Zebrafish acetylcholinesterase is encoded by a single gene localized on linkage group 7. Gene structure and polymorphism; molecular forms and expression pattern during development

Christelle Bertrand, Arnaud Chatonnet, Christina Takke, Yilin Yan, John H Postlethwait, Jean-Pierre Toutant, Xavier Cousin

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 1, Issue of January 5, pp. 464–474, 2001

Printed in U.S.A.

Zebrafish Acetylcholinesterase Is Encoded by a Single Gene Localized on Linkage Group 7

GENE STRUCTURE AND POLYMORPHISM; MOLECULAR FORMS AND EXPRESSION PATTERN DURING DEVELOPMENT*

Received for publication, July 17, 2000, and in revised form, September 20, 2000
Published, JBC Papers in Press, October 2, 2000, DOI 10.1074/jbc.M006308200

Christelle Bertrand‡, Arnaud Chatonnet‡, Christina Takke‡, YiLin Yan§, John Postlethwait§, Jean-Pierre Toutant‡, and Xavier Cousin‡¶

From the ‡Différenciation Cellulaire et Croissance, INRA, 2 Place Viala, 34060 Montpellier Cedex, France and the §Institute of Neuroscience, University of Oregon, Eugene, Oregon 97403

We cloned and sequenced the acetylcholinesterase gene and cDNA of zebrafish, *Danio rerio*. We found a single gene (*ache*) located on linkage group LG7. The relative organization of *ache*, *eng2*, and *shh* genes is conserved between zebrafish and mammals and defines a synteny. Restriction fragment length polymorphism analysis was allowed to identify several allelic variations. We also identified two transposable elements in non-coding regions of the gene. Compared with other vertebrate acetylcholinesterase genes, *ache* gene contains no alternative splicing at 5′ or 3′ ends where only a T exon is present. The translated sequence is 60–80% identical to acetylcholinesterases of the vertebrates and exhibits an extra loop specific to teleosts. Analysis of molecular forms showed a transition, at the time of hatching, from the globular G4 form to asymmetric A12 form that becomes prominent in adults. *In situ* hybridization and enzymatic activity detection on whole embryos confirmed early expression of the acetylcholinesterase gene in nervous and muscular tissues. We found no butyrylcholinesterase gene or activity in *Danio*. These findings make zebrafish a promising model to study function of acetylcholinesterase during development and regulation of molecular forms assembly *in vivo*.

The role of acetylcholinesterase (AChE,[1] EC 3.1.1.7) in synaptic transmission is clearly demonstrated by the effects of inhibitors of this enzyme. In addition to hydrolysis of the neurotransmitter acetylcholine at synapses, AChE was shown to be involved in non-cholinergic functions, influencing differentiation and neuronal outgrowth (reviewed in Ref. 1). During development in vertebrates, AChE appears long before synapses are functional, and its role in this context is not clear (2, 3). Successive expression of butyrylcholinesterase (BChE, EC 3.1.1.8) and AChE was associated with transition from the proliferation state to differentiation in embryonic chick somites (2).

Zebrafish is a useful model for vertebrate early development and is amenable to experimental modulation of the cholinergic system, AChE activity, or *ache* gene expression. In this fish, AChE is expressed in neurons long before axons reach their target (4–7). When embryos were bathed with organophosphate (diisopropylfluorophosphate), AChE was totally inhibited, and somitogenesis was altered (8). It is to note, however, that effect on development may not be related to AChE inhibition since diisopropylfluorophosphate inhibits many hydrolases.

In vertebrates, AChE is characterized by a large set of molecular forms. During development and tissue differentiation their proportion varies. This polymorphism is due to alternative mRNA processing at the 3′ end of the gene resulting in proteins containing different C-terminal peptides. Exon H encodes a hydrophobic peptide that is cleaved upon glycolipid addition. Exon T encodes a peptide highly conserved among species. Soluble monomers (G1), dimers (G2), and tetramers (G4) are composed of T subunits. In addition these tetramers may be associated with structural subunits, a collagenic tail in neuromuscular junctions, or a membrane protein in brain (9). In some cases, the genomic sequence following the last common exon is retained into the mRNA. Physiological significance of this readthrough transcript is not known (9).

In order to study AChE functions during development of *Danio rerio*, we first characterized its early expression. In this study, we cloned and sequenced the zebrafish *ache* gene and cDNA. We showed that zebrafish has a single *ache* gene encoding only T subunits and has no *bche* gene. We located the *ache* gene in the zebrafish genome on LG7 near *eng2* and *shh*. It contains transposable elements in non-coding regions and also presents allelic variations. We show that the fully mature pattern of AChE molecular forms is reached only after 1 week of development, in free swimming larvae, whereas the AChE expression pattern analysis showed an early expression from the 5- to 7-somite stage (12 h).

EXPERIMENTAL PROCEDURES

Materials—D. rerio adults and embryos were from our facility. Fish are maintained at 28 °C on a 13-h light/11-h dark cycle. The AB strain

* This research work was supported by grants from Association Française contre les Myopathies (to X. C., J.-P. T., and A. C.) and National Institutes of Health Grant P01HD22486 (to J.-H. P. and Y. Y.). 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.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AJ251640 (for ache) and AF003943 (for esterase).

¶ To whom correspondence should be addressed: Différenciation Cellulaire et Croissance, INRA, 2 Place Viala, 34060 Montpellier Cedex, France. Tel.: 33 4 99 61 28 14; Fax: 33 4 67 54 56 94; E-mail: cousin@ensam.inra.fr.

[1] The abbreviations used are: AChE, acetylcholinesterase; BChE, butyrylcholinesterase; ASCh, acetylthiocholine; BSCh, butyrylthiocholine; BW284c51, 1,5-bis(4-allyldimethylammoniumphenyl)-pentan-3-one dibromide; bp, base pair(s); kb, kilobase pair(s); LS, low salt, HS, high salt, LST, low salt Triton; HST, high salt Triton; MOP, Mother of pearl; PCR, polymerase chain reaction; RT-PCR, reverse transcription-polymerase chain reaction; nt, nucleotides; EST, expressed sequence tag.

was a gift of Dr. Bricaud (Université des Sciences et Techniques du Languedoc, Montpellier, France), and the ABO strain was a gift of Dr. Strähle (Institut de Génétique et de Biologie moléculaire et Cellulaire, Strasbourg, France). Embryos were collected from spontaneous spawnings and staged according to Westerfield (10).

All molecular biology procedures including genomic DNA and RNA isolation, Northern and Southern blots, PCR amplification followed standard techniques (10, 11) or the manufacturer's protocols. Chemicals are from Sigma, Fluka, or Aldrich, and enzymes are from Promega, New England Biolabs, and Roche Molecular Biochemicals. Sequencing was performed with the Big Dye kit from PerkinElmer Life Sciences following the supplier's protocol.

cDNA Cloning, Library Screening, and Gene Isolation—Degenerate primers P1, P2, and P3 were designed from conserved sequences of AChEs. Reverse transcription was performed with Expand RT (Roche Molecular Biochemicals) using primer (P3) 5′ GAACTC(A/G)AT(C/T)T-CATAGCC(A/G)TG 3′ on total adult RNAs. A first cDNA fragment was isolated by PCR using primers (P1) 5′ TT(T/C)C(C/A)(G/A)GGTTCiGA-GATGTGGAA 3′ and (P2) 5′-GC(T/G)(G/C)(G/C)iCCiGCACTCTC(C/T)-CC(A/G)AA-3′. Rapid amplification of cDNA ends experiments were performed to find the 5′- and 3′-untranslated regions and to locate the transcription start site.

A zebrafish genomic library cloned in Lambda Fix II vector was purchased from Stratagene. Approximately 1×10^6 clones were first screened with the P1–P2 cDNA fragment. Clone 14 was fully sequenced, and although it covered 8 kb of genomic AChE sequence, it did not contain non-coding exon 1. We performed a second screen with a 220-bp probe covering exon 1. Two additional clones 8N and 14N were isolated and sequenced, and they both contained the transcriptional start site and promoter region.

Sequence analysis using the blast algorithm (12) was performed at NCBI. For identification of transposable elements, blastn (nucleotide search *versus* a nucleotide data base) was performed *versus* the "non-human, non-mouse EST" or *versus* the non-redundant GenBank™ nucleotide data base. Blastx analysis (translated nucleotide sequence search *versus* a protein sequence data base) was performed *versus* the GenBank™ non-redundant protein data base. Cholinesterase sequence analysis was performed at ESTHER data base.

Genomic Mapping—The polymorphism found in the fourth intron (see "Results") in genomic clones was used to search for segregation in the MOP haploid mapping cross-panel (13). The allele-specific oligonucleotides, indicated in Fig. 1, (Forward) 5′ GAGGAAGTCATAACAGA-AGTGAGAAT 3′ and (Reverse) 5′ CGCAGACAAGGCATTTCCTTGA-TAA 3′ amplified a fragment of 300 bp for one allele or 1100 bp for the other allele. The mapping panel was previously genotyped for over 800 PCR-based markers (13, 14). The strain distribution patterns were analyzed using MapManager. The sequences of zebrafish loci, including simple sequence length (14) were compared with sequences of human and mouse genes in GenBank™ using the blastx algorithm (12). The map locations of human *ACHE* and mouse *Ache* genes were found in On Line Mendelian Inheritance in Man, GeneMap 99, and Mouse Genome Data base. Human/mouse comparative mapping was accomplished at the NCBI HomoloGene site.

Gene Polymorphism Analysis—For Southern blot experiments, about 30 μg of DNA were digested with each restriction enzyme and loaded on 0.8% agarose gels. After migration, the DNA was transferred to nylon membranes (Hybond-N+, Amersham Pharmacia Biotech) overnight in 0.4 N NaOH, 1 M NaCl. Probes were synthesized with [^{32}P]dCTPs and Megaprime kit (Amersham Pharmacia Biotech) and purified on G-50 columns (Amersham Pharmacia Biotech). Hybridizations were done in 6× SSC, 0.25% powdered milk, 0.5% SDS, 100 μg/ml salmon sperm DNA at 65 °C with 2×10^6 dpm/ml. After a 2-min wash in 6× SSC, 0.1% SDS at room temperature and two times for 20 min in 0.2× SSC, 1% SDS at 65 °C, membranes were exposed 24 h on a phosphoactivable screen and analyzed on a StormImager (Molecular Dynamics).

PCRs analyses were performed on genomic DNA from the two parent fish and nine 4-day-old offspring from one cross. For a single embryo, 5 μl of genomic DNA were used in PCR at a total volume of 50 μl. 5 μl of PCR products were digested with 10 units of *Eco*RI during at least 3 h at 37 °C. The entire reaction was loaded on 0.8% agarose gel.

Whole Mount in Situ Hybridization and Histochemistry on Embryos—Embryos were fixed in 4% paraformaldehyde in fixation buffer (0.15 M $CaCl_2$, 4% sucrose in 0.1 M $NaPO_4$, pH 7.4) for 12–16 h at 4 °C, washed twice for 5 min in PBST (1× phosphate-buffered saline, 0.1% Tween 20), pH 7.4, and conserved in methanol at −20 °C. Before staining, embryos were rehydrated 10 min in solutions containing 75, 50, and 25% methanol in PBST and washed 5 min in PBST. A cDNA, covering the entire coding sequence from the start codon to the stop codon, was obtained by high fidelity RT-PCR and cloned in the pBS-SK+ vector. Sense and reverse constructs were used to synthesize *in situ* hybridization antisense and sense mRNA probes. Fluorescein or digoxigenin-labeled RNA probes were synthesized with T7 polymerase and NTP-labeling mix (Roche Molecular Biochemicals). Hybridization was performed according to the method described (10). Staining was performed with 4-nitro blue tetrazolium chloride and 5-bromo-4-chloro-3-indolyl-phosphate disodium (Roche Molecular Biochemicals).

AChE activity was detected on fixed embryos (6–8 h in fixation buffer at room temperature) with a method adapted from Karnovsky and Roots (15). Embryos were incubated 4–5 h in 60 mM sodium acetate buffer, pH 6.4, 5 mM sodium citrate, 4.7 mM $CuSO_4$, 0.5 mM $K_3(Fe(CN)_6)$, and 1.7 mM acetylthiocholine iodide and washed extensively with PBST before observation.

Analysis of Molecular Forms—AChE proteins were extracted from adult zebrafish or embryos. Tissues were homogenized sequentially in 5–10 volumes of low salt (LS), high salt (HS), low salt-Triton (LST) buffers, or a unique extraction in high salt-Triton (HST) buffer. LS buffer contained 20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 0.1 mg/ml bacitracin, and aprotinin (7.5×10^{-3} trypsin inhibiting unit/ml, Sigma). HS buffer was LS + 1 M NaCl. LST and HST buffers contained the same components as LS or HS + 1% Triton X-100 (Sigma).

The AChE activity of extracts and gradient fractions was assayed according to the spectrophotometric method of Ellman *et al.* (16) using acetylthiocholine iodide or butyrylthiocholine iodide (Sigma) as substrates. Protein content of samples was estimated following the manufacturer's instructions (Bio-Rad Protein Assay).

Molecular forms were analyzed on non-denaturing polyacrylamide gel. Electrophoresis was performed in a Bio-Rad apparatus (5 × 8-cm gels of 7.5% bisacrylamide). Gels and running buffers contained 50 mM Tris glycine, pH 8.9, with detergents 0.2% Triton X-100 and 0.2% deoxycholate. Gels were washed in distilled water and stained with acetylthiocholine as substrate according to Karnovsky and Roots (15). Protein extracts were incubated for 1.5 h at 37 °C in the presence of 0.25 milliunits/phosphatidylinositol phospholipase C (Roche Molecular Biochemicals) or 0.5 mg/ml proteinase K (Sigma) before loading. For collagenase treatment, HS extracts were diluted six times to a final concentration of 150 mM NaCl, 10 mM $CaCl_2$, and 50 mM Tris-HCl buffer, pH 8.0. Collagenase (Clostridiopeptidase A, Worthington) was added to a final concentration of 125 μg/ml and incubated at 30 °C. The reaction was stopped by addition of NaCl to 0.5 M final and stored on ice prior to centrifugation on sucrose gradient.

Sedimentation analysis was performed in 11 ml of 5–20% sucrose gradients in LS, HS, LST, or HST buffers. Samples (250 μl) were loaded and centrifuged for 18 h at 40,000 rpm ($200{,}000 \times g$) at 4 °C in a SW41 rotor. Forty fractions were collected from the bottom of each gradient and assayed for AChE activity at pH 7. Alkaline phosphatase from calf intestine (Roche Molecular Biochemicals) ($s_{20,w}$ = 6.1) and β-galactosidase from *Escherichia coli* (Roche Molecular Biochemicals) ($s_{20,w}$ = 16) were used as internal markers of migration.

Expression of Recombinant AChE in Drosophila S2 Cells and in Zebrafish Embryos—A mini-gene containing the entire coding sequence of AChE was assembled from three genomic fragments and cloned in the expression vector pMT/V5-His (Invitrogen). The gene extends from 117 bases upstream of exon 2 (183 upstream of start codon) to the stop codon. The final construct was totally sequenced before transfection in *Drosophila* S2 cells. This mini-gene was also introduced in pcDNA3 vector for *in vivo* overexpression in zebrafish embryos.

Transfection of 50 μg of plasmid and expression in *Drosophila* S2 cells (3×10^6 cells/ml) was performed with the *Drosophila* Expression System (Invitrogen). After induction of expression with copper, no serum was added to the cell growth medium. AChE was recovered from the cell medium.

Microinjections into 1–2-cell stage embryos were performed with 200 ng/μl AChE-pcDNA3 plasmid in 0.5 M KCl, 2.5% rhodamine-B isothiocyanate dextran (Sigma). About 500 pl of solution were injected into each embryo using a Transjector 5246 (Eppendorf).

Determination of Kinetic Parameters—In kinetic experiments we used recombinant AChE secreted in growth medium or native AChE from whole zebrafish extracts. Activity was assayed in Ellman reaction buffer (16) containing 100 mM sodium phosphate, pH 7.4, 0.5 mM 5,5′-dithiobis(2-nitrobenzoic acid), 0.01% bovine serum albumin. AChE extracts were incubated for 15 min with buffer reaction before addition of various concentrations of acetylthiocholine iodide at 25 °C. The reaction was monitored in a spectrophotometer at 412 nm. K_m and K_{ss} values were determined according to the Haldan equation with 0.01–100 mM acetylthiocholine. Apparent first order rate constant, k_{cat}, was calculated as described previously (17). The constants of inhibition, K_i and

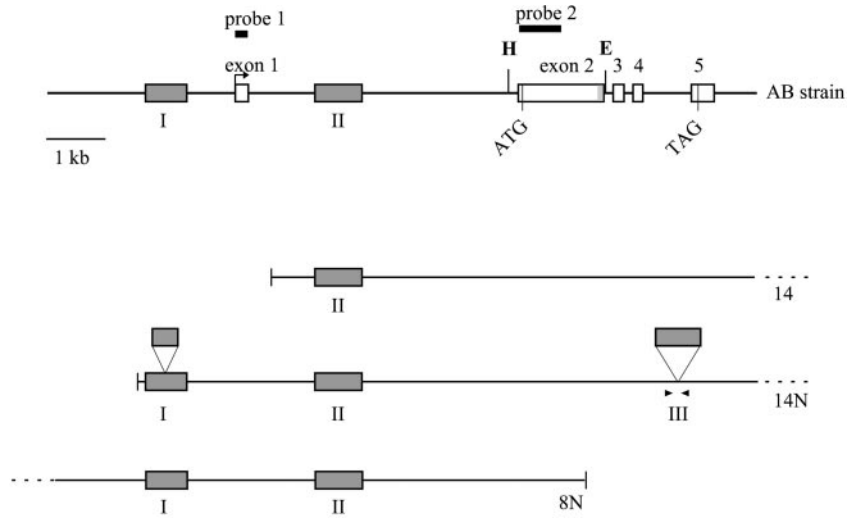

FIG. 1. **Structure of zebrafish *ache* gene and genomic clones.** The *scheme* represents 12 kb of sequenced genomic DNA containing the zebrafish *ache* gene. The five *ache* exons are shown as *open boxes,* and the extra-coding region specific of teleosts is *hatched.* ATG and TAG point out initiation and stop codons. *H* and *E* indicate positions of polymorphic *Hin*dIII and *Eco*RI restriction sites (see also Fig. 4). Three isolated genomic clones (14, 14N, and 8N) are shown *below* the deduced gene structure of AB strain. *Shaded boxes I, II,* and *III* indicate repeated and transposable elements that are further described in Fig. 3. Insertions found in clone 14N are indicated. *Arrowheads* on both sides of clone 14N 3′ insertion indicate location of primers used for gene mapping. Probes used during this work are represented as *thick lines above* the gene structure.

αK_i, were determined for several inhibitors with acetylthiocholine from 0.06 to 0.7 mM. Each experiment was repeated 3–5 times using 5 concentrations of ASCh and 4–5 concentrations of inhibitor. Results were identical for recombinant or native AChE, with a standard error less than 10%.

RESULTS

A Single Acetylcholinesterase in Zebrafish—Cholinesterase activity was first detected, in whole embryos, using a histochemical reaction adapted from Karnovsky and Roots (15). In 24-h embryos, strong activity appeared with acetylthiocholine (ASCh) within a few hours of staining, whereas there was no staining with butyrylthiocholine (BSCh). Eserine (10^{-5} M) inhibited all ASCh hydrolysis and staining of embryos.

Cholinesterase activity was also measured in protein extracts. The level of BSCh hydrolysis in 24-, 48-, and 72-h embryo extracts was less than 1% of the rate of ASCh hydrolysis. This value is similar to the ratio determined for AChE of lamprey and hagfish, which have only an AChE (18). A low V_{max}BSCh/V_{max}ASCh ratio is consistent with the properties of AChE (19). We also analyzed cholinesterase activity in adult in total extracts and in brain and heart extracts. In all tissues there was no residual ASCh or BSCh hydrolysis activity when a specific inhibitor of AChE, BW284c51, was used at 10^{-6} M. We conclude that a single AChE was responsible for ASCh and BSCh hydrolysis in zebrafish. Cholinesterase activity in zebrafish heart is sensitive to BW284c51, contrary to the cholinesterase activity found in the heart of *Torpedo*, mainly originating from BChE (20). In addition, an excess of the substrate ASCh inhibits activity of the enzyme, showing a specific property of AChEs *versus* BChEs.

We failed to detect *bche* gene by RT-PCR with degenerated oligonucleotides specific for BChEs. We did find, however, the 3′-coding region of an esterase (GenBank™ accession number AF003943) 67% homologous to carboxyl ester lipase from *Salmo salar* (GenBank™ accession number L23929). This indicates that we used conditions that were relaxed enough to allow us to detect hybridization of oligonucleotides to a putative zebrafish *bche* gene if it had been present.

Sequence and Structure of AChE Gene and cDNA—By using primers P1 and P2, we first isolated a cDNA fragment that was used to screen a genomic library. We isolated three independent genomic clones (see "Experimental Procedures"). They overlap as shown in Fig. 1. The total sequence covered 12 kb of genomic sequence including 3.3 kb of 5′ potential promoter region. The *ache* gene contains 5 exons; the first is non-coding and is followed by a large intron of 5 kb. Compared with other vertebrate *ache* gene sequences, the zebrafish *ache* gene contains an insertion interrupting the large exon 2. This 226-nt insertion encodes 30 amino acids (Fig. 2) followed by a small intron of 136 bp (Fig. 1). The inserted amino acids were located on the protein surface as determined by three-dimensional modeling (21). The last exon is a T exon. We fully sequenced all intronic sequences upstream of the T exon where the H exon is located in other vertebrate *ache* genes. We did not find any sequence in genomic DNA that could correspond to a hydrophobic H exon. The start of transcription, a cytosine, was found 303 bp upstream from the Met initiation codon. A canonical TATA box is found 28 bp upstream of the transcription start site. Coding sequence covers 1905 bp, and at the 3′ end, a poly(A) site identified with 3′-rapid amplification of cDNA ends defined a non-coding sequence of 340 bp. An additional downstream poly(A) site may be also used as PCR performed with an oligonucleotide located after the first poly(A) signal also amplified AChE cDNA. The total length of sequenced cDNA is 2548 bp, and a unique mRNA was detected on Northern blots (data not shown). We never found 3′ alternative splicing or readthrough cDNA as seen in some vertebrate AChEs (Massoulié *et al.* (9)). Only T exon is present. No cryptic alternative splicing could be found similar to what was described for *Electrophorus electricus* AChE (21).

The 634-amino acid sequence of zebrafish AChE is 62% identical to mammalian AChE, 64% identical to *Torpedo,* and 80% identical to *Electrophorus*. We found at conserved positions all the elements specific of AChEs, detailed in Fig. 2. The protein contains 6 potential *N*-glycosylation sites, all identical in position to glycosylation sites in a ray-finned fish, the electric eel *Electrophorus*, but only two of them were conserved in the AChE of a cartilagenous fish, the electric ray *Torpedo*.

Transposable Elements in the Ache Gene—In the non-coding region of the zebrafish *ache* gene (Fig. 1 and 3), we identified three domains presenting high similarity to some otherwise unrelated fish genes and ESTs.

Domain I, located upstream from non-coding exon 1, is 1162 bp long. It is flanked by two inverted repeats of 35 nt showing

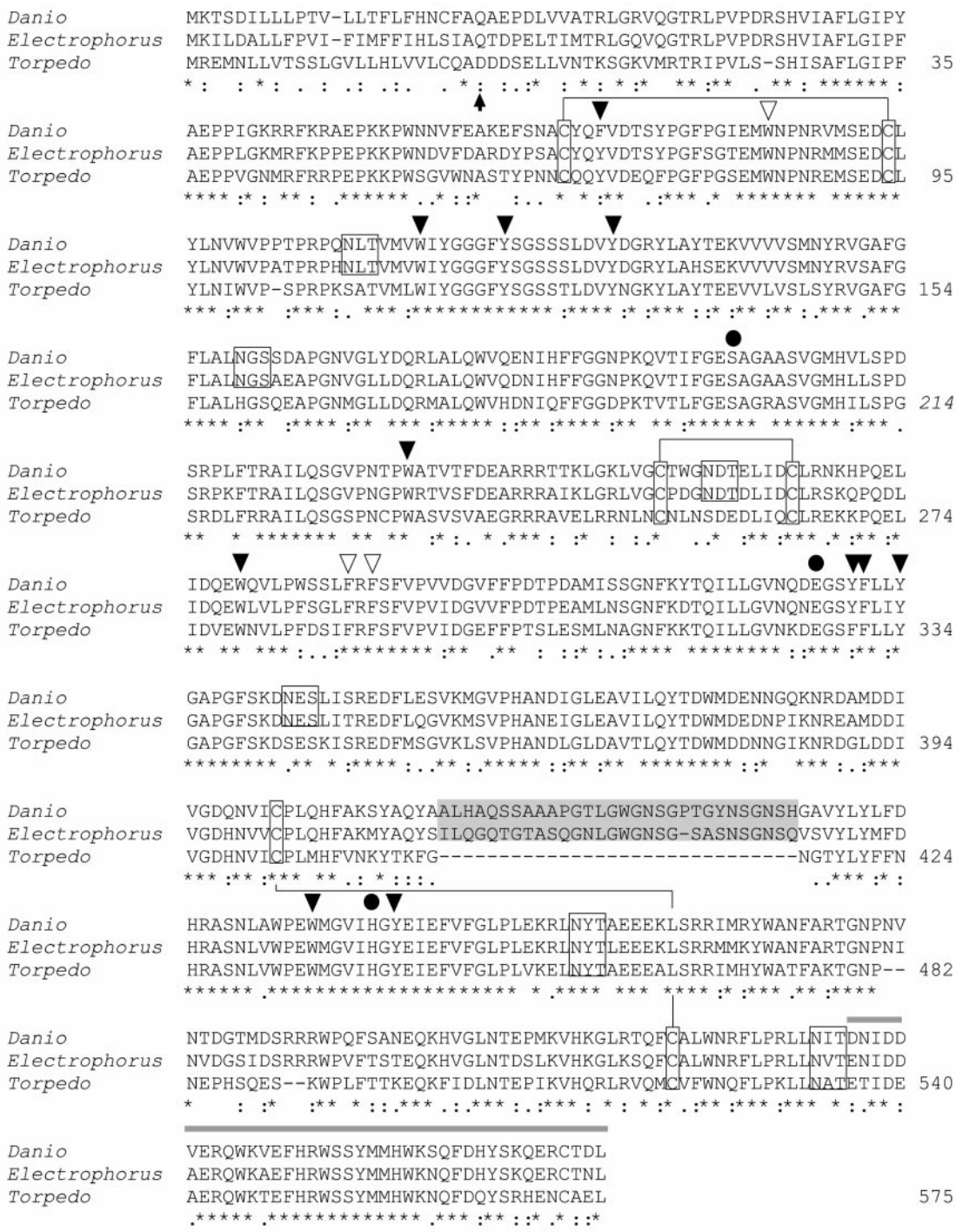

FIG. 2. **Alignment of fish AChE peptide sequences.** Common features of AChEs are indicated *above* the zebrafish sequence (*Danio*). Numbering is indicated according to *Torpedo* mature sequence. First residue of mature protein is marked by a *vertical arrow*. *Closed circles* indicate catalytic triad (Ser-200, Glu-327, and His-440), and *closed triangles* indicate the 14 aromatic residues lining the active site gorge. They include choline-binding site (Trp-84) and acyl pocket residues (Phe-288 and Phe-290) drawn as *open triangles*. 6 cysteines forming disulfide bridges are *boxed*. Glycosylation sites conserved in *Danio, Electrophorus,* and *Torpedo* are *boxed*. Extra-peptidic sequence, specific of teleost fish, is represented in a *shaded box*. *Asterisks below* the alignment indicate identical residues in all three sequences, and *semicolon* and *dots* show higher or lower similarity, respectively. Note the strong conservation of C-terminal sequences encoded by T exons (*thick line*).

80% of identity (see Fig. 8, in "Discussion"). Blastx analysis with the protein data base showed that almost all the sequence inside the inverted repeats was highly homologous to the *Tc1*-like transposase family. A translated sequence of transposase could be assembled by removing frameshifts and stop codons. We identified 3 boxes corresponding to the catalytic domain (for a review see Ref. 22). The second D box is shortened compared with the canonical sequence, and therein aspartic acid is changed to an asparagine. We also identified a G-rich region as well as the nuclear localization signal. The DNA binding domain is difficult to recognize in this sequence as such domains in transposases are structural elements with low sequence conservation (22). Blast analysis showed that more than 50 zebrafish ESTs showed homology with the transposase region. This is also the case of 7 non-coding regions of cloned genes from fish (4 from zebrafish and one from either medaka fish *Oryzia latipes*, tilapia *Haplochromis burtoni* or pufferfish *Fugu rubripes*). A 443-bp internal segment of this domain I is deleted

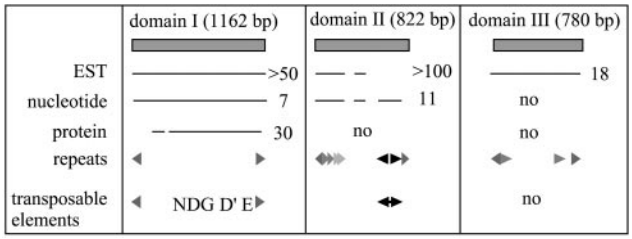

FIG. 3. **Description of transposable and repeated elements interspersed in the zebrafish *ache* gene.** Domains are *numbered* as in Fig. 1. *Below* each *shaded box*, representing the whole domain, *lines* indicate regions homologous with zebrafish EST, nucleotide, and protein data bases, respectively (details under "Experimental Procedures"). Values at *right* of each line indicate number of hits in blast analysis (E value $< 2{\cdot}10^{-4}$) and *no* indicates no homologue. *Pairs* of *arrowheads* in the *4th line* show localization of repeated elements. Transposon specific features are indicated in *last line.* In domain I, *letters* indicate identified domains of tranposase, nuclear localization signal (*N*) and G-rich domain (*G*). The *three boxes* defined by Plasterk *et al.* (22), corresponding to catalytic domains of Tc1 transposase, are also indicated (*D* and *E*). *D′* corresponds to the second mutated D box. For domain II *arrowheads* show the repeated elements of *angel* transposable element separated by a conserved motif (GGAAAAACAAA marked by a *line*). No homology with transposable element has been identified in domain III.

in the AB line and in one of the isolated genomic clones (Fig. 1).

Domain II (822 bp) is located inside intron 1. As in the case of the first domain, it is flanked by two inverted repeats (58 nt with 57% of identity). Domain II has no homology with coding sequences when compared with protein data base. It includes two sub-domains (Fig. 3). A first sub-domain shows a large number of blast homology hits with ESTs (>100), with non-coding region of 11 zebrafish genes, and contains several large direct repeats. A second sub-domain is characterized by two large inverted repeats separated by an A-rich domain. This second sub-domain is homologous to *Danio*-specific transposable element called *angel* (23). We identified only one of the T2 domains that form short inverted repeats at both ends of *angel* elements (23).

Domain III is located in intron 4 and was only found in genomic clone 14N (Fig. 1). A 16-nucleotide sequence (AGCCCCTTTCACACAG) is found in inverted orientation at both ends of the insertion. The central 0.5-kb domain was flanked by direct long repeats of about 60 bp. As for domains I and II, this sequence shows homology with a large number of zebrafish ESTs (Fig. 3).

Evidence for Polymorphism in Ache Gene—A commercial genomic library made with DNA from several adults was screened. Three independent clones were isolated. They differed from one other by two large insertions and multiple point mutations (Fig. 1).

The whole domain III, previously described in intron 4, was only found in the clone 14N. This insertional polymorphism was used to map *ache* on zebrafish genetic map (see below). In addition, a 443-bp segment located inside domain I is deleted in fish from our stock, in AB strain, and in one genomic clone.

Southern analysis of genomic DNA was first performed on fish from our stock using probes 1 and 2 (see Fig. 1). Fifteen restriction enzymes were used, and some results are presented on Fig. 4*A*. For most restriction enzymes, or combinations, the size of the bands matched the expected size deduced from the gene restriction map, but an extra high molecular weight band was detected when *Eco*RI or *Hin*dIII were used. Sizes of these second bands correspond to a genomic sequence in which one *Hin*dIII and one *Eco*RI site, previously located in intron 1 and 2, are missing (Fig. 1). The same experiment performed on AB and ABO fish showed that AB was polymorphic, whereas ABO genomic DNA had only the short *Hin*dIII and *Eco*RI fragments.

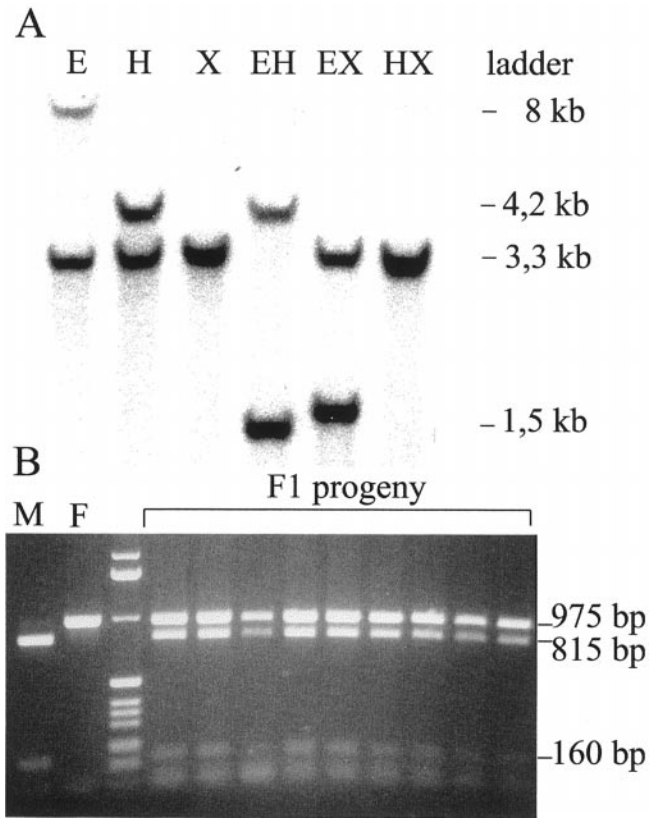

FIG. 4. **Restriction fragment polymorphism reveals two *ache* alleles in the AB strain.** *A,* genomic DNA was digested with enzymes indicated *above* each lane (*E*, *Eco*RI; *H*, *Hin*dIII; *X*, *Xba*I) and Southern blot was hybridized with probe 2 (see Fig. 1). Two bands visible in *Eco*RI and *Hin*dIII lanes revealed the existence of polymorphic sites. *B,* the genotype of AB fish was checked using PCRs, with primers surrounding *Eco*RI allelic site, on genomic DNA from one adult couple and 9 F1 progeny. The amplified 975-bp fragment was then digested with *Eco*RI. Parents are homozygotes for resistant (a unique 975-bp band in female) or sensitive (two bands of 160 and 815 bp in male) forms. All F1 progeny are heterozygotes as shown by the presence of the restriction bands and undigested fragment.

To ensure that the extra bands were due to allelic polymorphism rather than the presence of a gene duplicate, we analyzed genetic transmission of the two forms. We performed PCRs on genomic DNA with primers surrounding the alternative *Eco*RI site, on adult breeding pairs, and their progeny. Amplified fragments were digested with *Eco*RI to determine the genotype of each fish. When one parent was homozygous for one variant and the other parent was homozygous for the alternative variant (Fig. 4*B*), all F1 progeny were heterozygotes and presented bands both resistant and sensitive to *Eco*RI digestion. We also observed mendelian transmission of this variation of restriction site by crossing one homozygote and one heterozygote (not shown). By using the restriction polymorphism, we have cloned and sequenced the *Hin*dIII-*Eco*RI-resistant allele from the end of intron 1 to the 3′-untranslated regions. Most of the 20 nucleotide variations found in the whole gene are located in introns, and the four polymorphisms occurring in the coding sequence are silent. Sequences confirmed that the *Hin*dIII site and the *Eco*RI site located in introns 1 and 2, respectively, are lost. Mendelian transmission of the restriction polymorphism and conservation of the coding sequence of the two forms allowed us to conclude that these differences correspond to two different alleles of a unique *ache* gene.

Mapping of Ache on the Zebrafish Genetic Map—We took advantage of the length of polymorphism in the fourth intron to

TABLE I
ache single gene mapped on linkage group 7

The polymorphism found in the fourth intron in genomic clones and allele-specific oligonucleotides (see Fig. 1) was used to search for segregation in the MOP haploid mapping cross-panel (see "Experimental Procedures"). *ache* was mapped on distal LG7, between previously mapped genes *fkd5* and *islet 3*. *ache* is located about 1.6 cM from each. Mat, number with maternal genotype; Pat, number with paternal genotype; X, number of recombinants; N, total number of segregants scored for the two adjacent loci; Map, distance between two adjacent loci in centimorgans (cM); 95% = 95% confidence interval for map distance between adjacent markers; LOD, logarithm of the odds score for linkage of adjacent markers.

Locus	Mat	Pat	X	N	Map	SE	95%		LOD
z4706	45	49							
			5	59	8.47	3.63	2.8	18.7	10.3
fkd5	31	29							
			1	60	1.67	1.65	0.0	8.9	15.9
ache	46	39							
			1	62	1.61	1.60	0.0	8.7	16.4
islet3	31	34							
			6	63	9.52	3.70	3.6	19.6	10.4
z3445	49	43							
			2	39	5.13	3.53	0.6	17.3	8.3
z1182	19	21							
			3	40	7.50	4.16	1.6	20.4	7.4
cyce	42	54							
			3	46	0.00	0.00	0.0	7.7	13.8
z1059	19	27							
			0	46	0.00	0.00	0.0	7.7	13.8
shh	43	52							
			0	44	0.00	0.00	0.0	8.0	13.2
eng2	18	27							

locate the *ache* gene with the specific allele amplification method and segregation on the MOP cross-haploid mapping panel (13). According to usual nomenclature, symbols used are *ache* for zebrafish, *ACHE* for human, and *Ache* for murine genes.

On the MOP panel, *ache* mapped to LG07 between z4706 and z3445 (Table I). Intercalation between these markers places *ache* at about LG07_39.5 on the MGH sex-averaged diploid map (14). As shown in Table I, a few centimorgans away on LG07 reside *shh* and *eng2*, co-localizing with z1059 at LG07_52.3 on the MGH map. In human, *ACHE* is located at Hsa7q22, and *SHH* and *EN2*, the orthologues of *shh* and *eng2* (24, 25), reside in Hsa7q36. The mouse orthologues of these three loci are also syntenic, located on chromosome 5 (*Ache*, Mmu5_80.0; *Shh*, Mmu5_12.0; *En2*, Mmu5_15.0). Interspersed between *ache* and *shh* in zebrafish is *islet3*, whose human orthologue is uncertain, and *cyclin E*, whose human orthologue is located at Hsa19q12-q13. These results suggest that *ache* belongs to a conserved chromosome segment that originally included *ache*, *shh*, and *eng2* in the last common ancestor of zebrafish, human, and mouse.

Pattern of Expression of AChE mRNA and Protein in Embryos—In situ hybridization with labeled antisense mRNAs was performed at several developmental stages. In all cases, overnight or longer incubations are required to detect staining of AChE mRNAs in embryos. Controls performed with sense RNA probe showed only staining of notochord (not shown).

ache mRNA was first detected in the trunk, in discrete regions of paraxial mesodermal segmental plate at 12 h of development (6-somite stage). At this stage there was usually no signal in the two or three most recently formed somites. In contrast, in older embryos a weak diffused staining was always found, even in unsegmented presomitic mesoderm (Fig. 5*A*). Expression, probably located in myoblasts, proceeds in a rostrocaudal sequence according to the state of differentiation of the somites. Initial narrow staining progressively enlarged in the differentiating somite. At all stages, the ventral part of anterior somites showed more intense staining than the dorsal part (Fig. 5*A*). This could be due to a higher rate of development of ventral myotome as accurately described in mouse embryos (26). In the spinal cord, mRNAs are also present in primary motoneurons as seen in Fig. 5*B*, in a 24-h embryo.

In the brain, mRNAs first appear in a symmetrical cluster at presumptive midbrain-hindbrain boundary from the 10-somite stage (14 h) (not shown). Two hours later, additional stainings are detected in three large bilateral clusters, in anterior telencephalon, on the floor of the diencephalon and of the mesencephalon. At 24 h, these clusters co-locate with the axonal tract of the anterior commissure, the postoptic commissure, and the ventral longitudinal tract (6, 7, 27). We can also detect messengers in hindbrain, in ventrolateral clusters at regularly repeated intervals (not shown). Small groups of cells, segmentally reiterated, expand in longitudinal columns. This corresponds to the location of differentiating reticulospinal neurons in the seven hindbrain rhombomeres as previously shown with zn1-antibody staining (4).

The method of Karnovsky and Roots (15) was used to detect AChE activity in whole embryos, in order to complete the pattern of expression described for neuronal (5, 6) and brain (7) differentiation in zebrafish and also to correlate this pattern with the mRNA expression.

In the trunk the spatio-temporal expression of AChE protein matches mRNA expression in peripheral nervous system and in muscles, along differentiation of myoblasts. AChE activity is first detected in small clusters of cells on both sides near the spinal cord (Fig. 6*A*). No activity is detectable before 5–7 somites, at this stage staining appears in few cells in each segment. This staining should correspond to muscle precursors because it is found in myoblasts than in myofibers (8). The enzyme activity appears more intense in somites following a rostro-caudal sequence in a pattern similar to mRNA, and all somites show activity (Fig. 6, *B* and *E*). Shortly after 18 h, myocommata, the borders between somites, start to express AChE (see also at 24 h, Fig. 6*E*). This coincides with clustering of acetylcholine receptors at neuromuscular junctions (28) and the first spontaneous twitches of embryos. In addition, in the peripheral nervous system, AChE protein initially appears in presumptive cell bodies of primary sensory and motoneurons from 14 h (9–10-somite stage). By 24 h (Fig. 6*E*), AChE is found in the spinal cord in motoneurons, located in each hemisegment, in sensory neurons (Rohon-Beard cells), and in reticulospinal interneurons. As a general feature, AChE activity is strongly detected in cell bodies of neurons but only slightly in axons.

In the posterior brain activity appeared after 14 h. A bilateral large cluster, just anterior to hindbrain rudiment, indicates position of forming trigeminal ganglia. In the embryo at 16 h, three new symmetrically bilateral clusters are present as follows: dorso-rostral, ventro-rostral, and ventrocaudal clusters are shown on Fig. 6*B*. Shortly after new clusters appear in the anlage of the epiphysis, in the posterior commissure, and in differentiating rhombomeres of hindbrain. At 24 h (Fig. 6*C*) AChE is detected in all primary neurons in the brain (see also Ref. 5). Expression in cranial ganglia can be detected in whole embryo until about 36 h. We found additional stainings, yet undescribed, in a sensory system in anterior and posterior lateral line ganglia as seen in the embryo at 24 h (Fig. 6*D*). In addition the heart shows strong AChE activity from its early morphogenesis.

Identification and Localization of Molecular Forms of AChE—AChE protein in vertebrates exists in numerous molecular forms that can be identified by their sedimentation and hydrophobic properties after centrifugation on sucrose gradi-

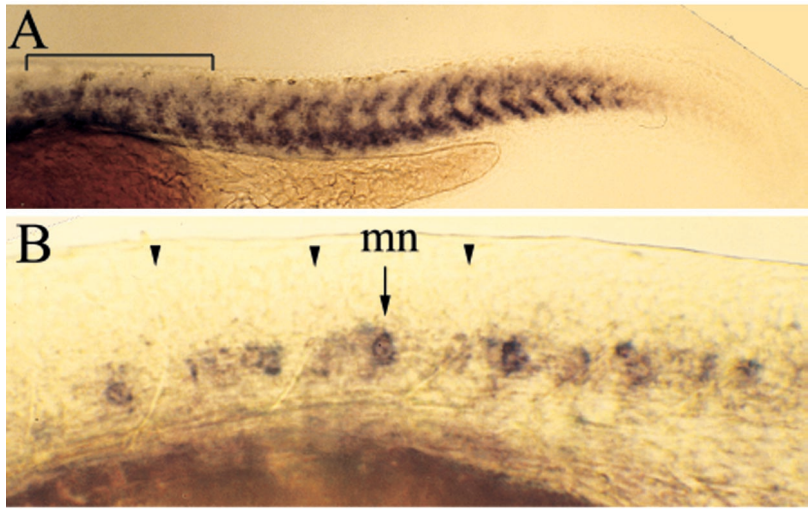

FIG. 5. **Pattern of AChE transcripts expression in 24-h embryos.** *In situ* hybridizations were performed with an antisense mRNA probe covering AChE cDNA. 24-h embryos are oriented in lateral view with anterior to the *left* and dorsal to the *top*. *A,* in the trunk narrow staining in caudal somites progressively enlarges. Expression is stronger in the ventral part of the somite compared with the dorsal part. *B,* a shorter incubation allows visualization of AChE mRNAs in cell bodies of primary motoneurons (*mn*) in the spinal cord. *Triangles* indicate borders of somites. *B* corresponds to *bracket* in *A*.

ents with or without non-denaturing detergent.

We first followed the evolution of different molecular forms repartition during zebrafish embryonic development (Fig. 7*A*). Until 48 h of development, in an HST protein extract, AChE sediments as soluble globular G4 (12 S) and a minor G2 form (6 S). The asymmetric form A12 (17.5 S) that contains a collagenic subunit appears after 48 h and becomes prominent after the 1st week of development.

In total extracts of adults (Fig. 7*B*), sequential extractions of proteins in LS, HS, and LST showed that asymmetric A12 and A8 (14.5 S), the major forms of HS extract (Fig. 7*B2*), were extremely predominant as 70% of total activity was found in HS extracts.

To ensure these 17.5 S forms correspond to collagen-tailed asymmetric forms, we performed collagenase digestion of HS extract (*inset* in Fig. 7*B2*). After treatment, two forms appear. In the heavier (19.5 S), the C-terminal domain of collagen tail which slows migration of A12 forms is removed and migration accelerated. In the lighter (15.5 S) digestion probably removes most of the collagen domain. We also observed appearance of tetramers (not shown). We conclude that 17.5 S corresponds to true A12 asymmetric forms. Tetramers and dimers extracted with LS or LST were found in similar amounts (Fig. 7, *B1* and *B3*). The sedimentation coefficient of dimers was reduced in the presence of non-denaturing detergent indicating their amphiphilic nature (Fig. 7, *B1* and *B3*). In addition, G2 forms were insensitive to phosphatidylinositol phospholipase C treatment, but their migration was slightly modified on non-denaturing electrophoresis when proteinase K was used (not shown). We conclude that these dimers have no glycolipid anchors and are type II amphiphilic dimers made of T subunits (9).

In isolated tissues of the adult (Fig. 7*C*), we found asymmetric forms prominent in muscles and a few G4 and G2. In contrast, amphiphilic tetramers, which aggregate in absence of detergent, were the major form in the brain (not shown). In the heart, we found G2 and G4 and also some A12 in HST extracts (Fig. 7*C*). We never identified glycolipid-anchored dimers. This is in agreement with the absence of H exon in cDNA and confirm what is observed in the other teleost, *E. electricus* (21).

In Vitro and in Vivo Expression of AChE, Catalytic Properties—Activity and molecular forms of expression of a recombinant AChE were checked *in vitro* and *in vivo*. Enzymatic assays indicated that AChE was produced in transfected *Drosophila* S2 cells and highly secreted into the growth medium. At 36 h, zebrafish embryos injected with the AChE mini-gene construct under control of cytomegalovirus promoter expressed recombinant AChE in all tissues as detected by *in situ* hybridization and activity staining in whole embryos (not shown). Four independent injection experiments were performed, and for each, two batches of 10–15 embryos were assayed for AChE activity in whole protein extracts. Depending on injections, AChE was overexpressed 5–25-fold compared with control embryos. This result was consistent with a variable amount of mRNAs injected in overexpression experiments and observed by *in situ* hybridization.

Fig. 7*D* shows that overexpressed AChE in embryos was composed of dimers and tetramers in similar amounts, whereas in non-injected embryos, prominent G4 and few G2 forms were detected. Recombinant AChE in S2 cell extracts was only found as G2 forms (Fig. 7*D*). These results suggest a regulation of tetramers assembly, different in *Drosophila* cells and in embryos.

The AChE of zebrafish has all residues characteristic of AChEs (Fig. 2). Kinetics parameters of AChE from zebrafish extracts or recombinant enzyme are close to values measured for ASCh with other vertebrates AChE, K_m = 230 μM and K_{ss} = 20 mM. Zebrafish AChE k_{cat} (1300/s) is lower than mammalian AChE k_{cat} (about 3000/s) and 10 times lower than the *E. electricus* k_{cat} despite the very high sequence homology between these two enzymes. Zebrafish enzyme is not sensitive to specific BChE inhibitor tetraisopropylphosphoramide. On the contrary, high inhibition is observed in the presence of active site inhibitors, serine (K_i = 14.4 μM) and edrophonium (K_i = 0.53 μM), or bis-quaternary inhibitor BW284c51 (K_i = 66 nM). K_i for propidium (16 μM), a peripheral site inhibitor, is 5–10-fold higher than K_i for mammalian or *Torpedo* AChEs. This could be due to the change of tyrosine 70 in phenylalanine since this residue is part of the propidium-binding site, located at the periphery of AChE.

DISCUSSION

AChE Is Encoded by a Unique Gene—The insertional polymorphism observed in intron 4 (see Fig. 1) allowed us to locate *ache* gene on zebrafish genome by segregation on MOP cross-haploid mapping panel (13). A unique *ache* gene was mapped in linkage group 7 within a few centimorgans of *shh* and *eng2* loci

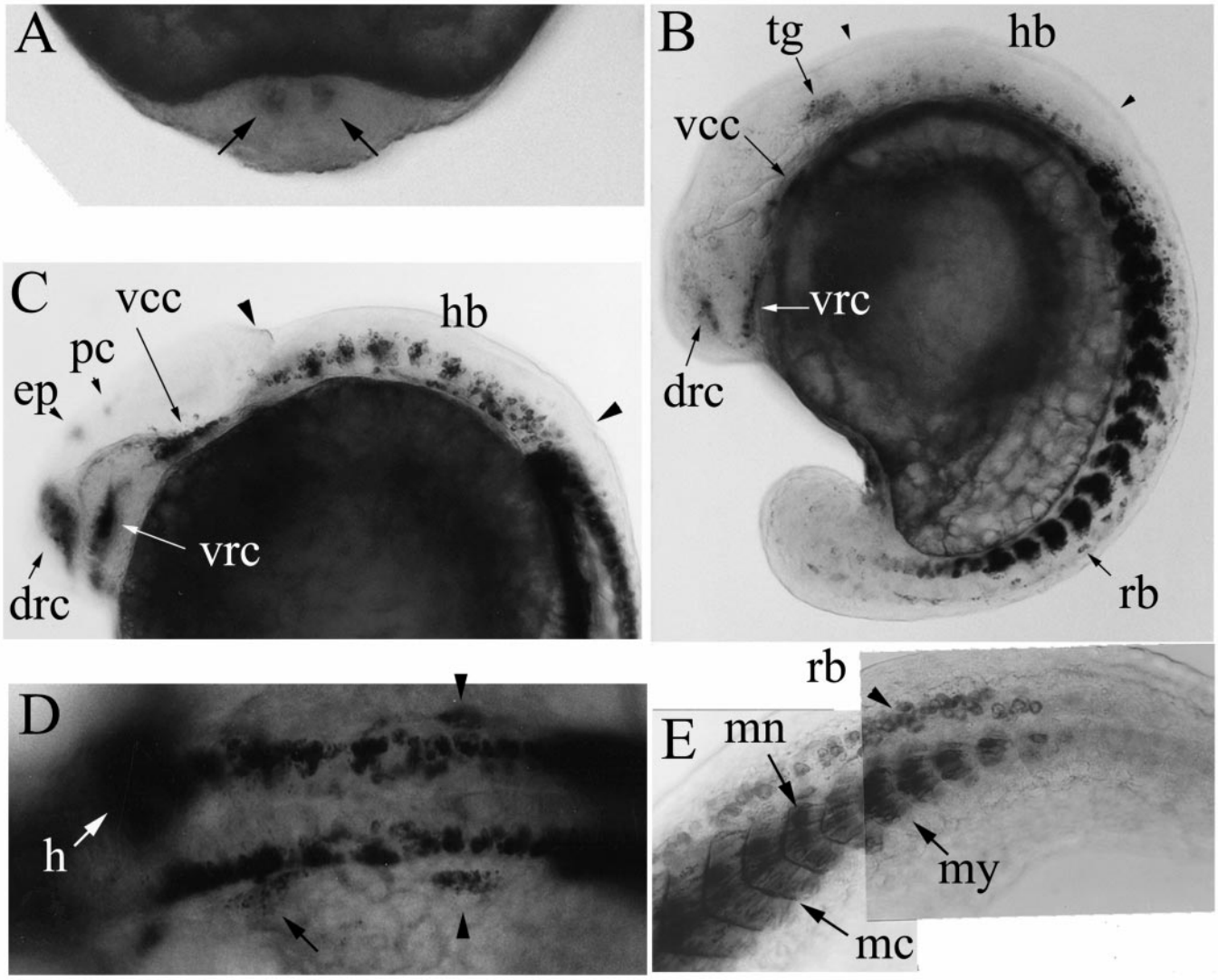

FIG. 6. **Detection of AChE activity in whole embryos.** Embryos are oriented in lateral view with anterior to the *left* and dorsal to the *top,* except for *A* and *D* which are dorsal views. *A,* appearance of AChE activity in a 5–7-somite embryo, seen in dorsal view, anterior to the top; on both sides of the spinal cord large cells are stained in all somites. *B,* 19-somite embryo; staining is strongly detected in somites and in Rohon-Beard cells (*rb*). In the brain, trigeminal ganglia (*tg*) were already stained, and three new clusters of few cells appear in the anterior part (*drc, vrc, vcc*). Small clusters are found in the hindbrain, indicated by *arrowheads*. *C–E,* details of 24-h embryo. In the head (*C*) rhombomeres are labeled in the hindbrain (*hb*). The three clusters *drc, vrc,* and *vcc* are found enlarged in the telencephalon, diencephalon, and mesencephalon. Epiphysis (*ep*) and posterior commissure (*pc*) stainings are now visible. In dorsal view of the head (*D*), labeling is clearly seen in anterior (*arrow*) and posterior (*arrowheads*) lateral line ganglia. Strong activity is also found in the heart (*h*). In the bud tail (*E*), AChE activity appears in myocommata (*mc*) between somites, at junction of myofibers (*my*). In each somite, heavily stained cells correspond to primary motoneurons (*mn*). Dorsally, numerous Rohon-Beard cells are also detected (*rb*). *vrc*, ventro-rostral cluster; *vcc,* ventro-caudal cluster; *drc*, dorso-rostral cluster.

(Table I). The synteny observed for these genes in zebrafish, human, and mouse could indicate relative organization of the genes in the common ancestor of teleosts and mammals living approximately 450 million years ago according to molecular time scale (29).

Many genes have been duplicated during evolution in zebrafish lineage (13, 30, 31). For example, zebrafish has two copies of the *EN2* and *SHH* genes of human, and they are called *eng2a* and *eng2b* (formerly called *eng3* (25)) and *shh* and *twhh* (32), respectively. The *eng2a* and *shh* loci are closely linked on LG7, and the *eng2b* and *twhh* loci are syntenic on LG19 (13). These are duplications of a portion of Hsa7q as shown by genetic mapping and phylogenetic analysis of HOX clusters (30). We hypothesize that subsequent to the divergence of the human and fish lineages, a chromosome inversion event in the zebrafish lineage rearranged this chromosome segment and separated *ache* from *shh* and *eng2. ache* and *Shh*/*En2* are also separated by apparent translocations and inversions on mouse chromosome 5: pieces orthologous to human chromosomes 4, 1, 12, and 22 intervene between them. In teleost lineage, the segment of chromosome containing *ache* was probably duplicated, and then the duplicate *ache* was lost in zebrafish. Loss of one or more copies of duplicated genes occurred frequently during evolution as also demonstrated by phylogenetic analysis in neurotrophin and Trk receptor families (33).

Zebrafish Have No BChE—In protein extracts of 24 and 48-h and 3-day embryos, or whole adult or heart, BSCh hydrolysis was always less than 1% of ASCh activity. Catalytic parameters measured in the presence of specific inhibitors of BChE (tetraisopropylphosphoramide) or AChE (BW284c51) correspond to a classical AChE. All attempts to identify BChE activity or BChE cDNA failed, showing that there is no BChE in zebrafish. Several duplications of cholinesterase genes occurred independently in different phylogenetic lineages as follows: 4 AChE genes in *Caenorhabditis elegans* (34) and 2 genes in *Amphioxus* (18). The duplication of an ancestral gene, giving

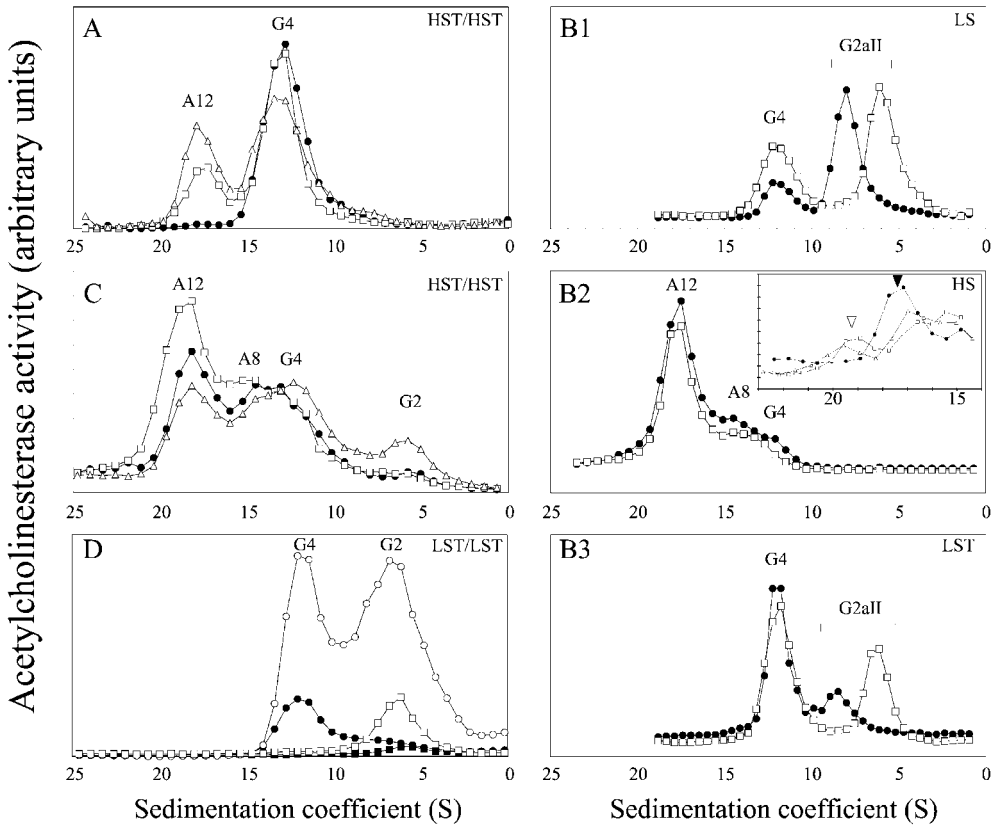

FIG. 7. **Evolution of AChE sedimentation of molecular forms during development and in adult tissues.** Comparison with molecular forms produced *in vitro* is shown. Extraction and sedimentation conditions are indicated in each panel (for example *HS*/*HST* indicates extraction in HS buffer and sedimentation in HST buffer). Buffers are described under "Experimental Procedures." Graphics represent AChE activity, measured in each fraction of gradient, as a function of sedimentation coefficient (*S*). *A,* protein extractions were performed in HST buffer in 48-h (●) and 96-h (□) embryos and 1-week larvae (△). At 48 h, the large peak of sedimentation indicates globular forms, G4 (12 S). Later, the peak at 17.5 S corresponds to appearance of the asymmetric forms A12 which will become prominent in adult (see *C*). *B,* analysis of molecular forms was performed in total adult with three sequential extractions in LS, HS, and LST (1, 2, and 3). In each panel presence (□) or absence of 1% Triton X-100 (●) in gradient is indicated. The shift of the G2 form in B1 and B3 indicates it is amphiphilic. Asymmetric forms A12 and A8 are only extracted with HS buffer. *Inset* in *B2* shows the effect of collagenase treatment of HS extract during 15 (□) or 30 min (△) compared with untreated enzyme (●). *Arrowheads* indicate the undigested form (*black*); heavier (*open*) and lighter (*shaded*) forms confirm the 17.5 S form is a collagen-tailed asymmetric form. *C,* molecular forms in adult. Extraction with HST buffer shows that A12 and A8 are prominent in total adult extract (□) as well as in muscles (●) with lower amounts of globular forms (G4 and G2). On the contrary, in the heart (△), most forms are globular (G4 and G2), and A12 are also present. *D,* recombinant molecular forms produced after injections of *ache* mini-gene in zebrafish embryos (○) or after transfection in *Drosophila* S2 cells (□) were analyzed in LST extracts. In S2 cells there is no activity in controls (■), and AChE is overexpressed as G2 form. In 36-h injected embryos two peaks of identical height correspond to overexpressed G4 and G2. In control embryos (●), at the same developmental stage, tetramers G4 are the major form.

rise to AChE and BChE, probably occurred before the split of cartilagenous fish, but after divergence of jawless fish lineage as BChE is found in *Torpedo*, birds, and mammals but is absent in hagfish and lamprey (reviewed in Ref. 18). The situation is less clear in bony fish because a BChE or pseudocholinesterase activity is found in some perciform fish as follows: flounder, sea bass, and surgeon fish (35, 36). A BChE activity was clearly demonstrated in gymnote brain (37). With no sequence information it is impossible to tell if they correspond to orthologues of tetrapod BChE. Our findings suggest that *bche* gene has been lost in zebrafish.

The Zebrafish ache Gene Contains Several Transposable Elements—Analysis of the non-coding region of the *ache* gene revealed several domains containing repeats and showing homology with transposable elements (Fig. 3).

Upstream of *ache*, we identified an inactive transposon of the *Tc1*-like family (*domain I*, Fig. 1). Many mutations accumulated in transposase sequence since transposition occurred. Despite several frameshifts and stop codons, it is still possible to identify the original amino acid sequence. The transposase in *ache* shares less than 30% identity with *tes1* from hagfish (38) or other zebrafish transposases (39). This is in agreement with the fact that transposition is thought to have taken place 10 million years ago (22). As can be expected, no homology was found at the nucleotide level with these sequences. However, we identified several fish genes containing domains homologous to the *ache* transposon-like (80–95% identity in nucleotides). For three of these genes, the homologous domain has a different size, but the same inverted repeat flanks the central part (Fig. 8). This suggests that AChE transposon is member of a new *Tc1*-like transposon family in zebrafish.

Domains II and III are made of a different set of repeats, associated in domain II with *angel* transposable elements. This could indicate that insertion occurred by combining different repeated sequences. In *Danio* a retroelement inserted in the intron of the *elf-4E* gene is also composed by association of a DANA-transposable element with other repetitive sequences (40).

Since large numbers of ESTs present high similarity (>80%) with all three domains, it appears that repeated and transposable elements are very frequent in the zebrafish genome.

Repeated elements are also found in other cholinesterase genes. Two tandem repeats have been described in the non-coding region of the two alternative last exons of *Torpedo californica ache* gene (41). Two Alu sequences probably integrated by retrotransposition in human *BChE* gene have also

```
ache_s    1757 ATACATACACTTACCGGTCACTTTATTAGGTACACTTGT 1795
ache_as   2871 GAATATAAACTCACTGGCCACTTTATTAAGTACACCTTA 2833
odor_s   22732 GCAAATATGCTCACTGGCCACTTTATTAGGTACACGTTA 22770
odor_as  23404 ATATATACACGCAACGGCCACTTTATTAGGTACACCTAA 23367
spsy_s   20177 ACATTTACATTTAATGGCCACTTTATTAGGTGATCTGCT 20215
spsy_as  20990 GACAATACACTCACCGGCCACTATATTAGGTAAACTTTA 20952
oryz_s    7575 TTATATAAACTCACTGGCCACTTTATTTGGCACATCATG 7613
oryz_as   8339 CTGCATACACTCACTGGCCACTTCATTAAATACACCTTA 8301
consensus 75%    A ATA ACTCAC GGCCACTTTATTAGGTACAC T
```

FIG. 8. **Repeated element flanking AChE domain I is compared with transposable elements of other fish genes.** Pairs of inverted repeats identified at the 5′ end of zebrafish genes *ache* (*ache*), odorant cluster (*odor*, GenBank™ accession number AF112374), and spermine synthase (*spsy*, GenBank™ accession number DRAJ9633) and *O. latipes* gene guanylyl cyclase (*oryz,* GenBank™ accession number AB016081) are aligned. *s* and *as* indicate sense and antisense repeats. *Last line* shows 75% consensus. Position of repeated elements in each sequence are indicated.

been described: one in the promoter region (42) and the other one occurred in exon 2 leading to gene inactivation and a silent phenotype in a Japanese family (43). A short interspersed element was also found in the rabbit BChE intron 2 (44). Numerous repeats of Alu sequences are found upstream of human *ACHE* gene as well as in introns. All these retroposons are short interspersed repetitive elements that do not encode reverse transcriptase. The zebrafish *ache* gene is the first cholinesterase gene showing trace of ancient active transposition.

Intron Capture Occurred in Teleost Ache Genes Introducing Extra Coding Sequence—The zebrafish gene contains 5 exons and the first one is non-coding. No alternative splicing was found in 5′ or in 3′ ends contrary to what is found in the mammalian and *Torpedo* gene (reviewed in Ref. 9). Sequence alignment with the *Torpedo* gene showed a similar organization of exon-intron junctions in *Danio* with the exception of an insertion that splits the large coding exon 2 into two parts. This additional sequence (Fig. 2), also found in *Electrophorus ache* gene (21), introduces 30 residues in the protein and a small intron. The inserted peptide domain is a glycine-rich hydrophilic peptide similar in zebrafish and gymnote (94% of homology). Such coding insertions are not uncommon in AChE since some are also found in chick, *Drosophila,* and nematodes AChEs (9). Removal of insertions by genetic deletion in gymnote (21) does not modify catalytic properties or oligomerization of the enzyme. This is in agreement with the three-dimensional structure which shows that insertions are located at periphery of the protein, in loops between strands or helices, far away from the active site. In both teleosts, insertion takes place at the same position in *ache* exon 2 at position 415; in *Danio* it is located between consensus 5′-TG/GT-3′ nucleotides usually found at exon junctions. This suggests that the insertion may be the result of an intron capture as recently described in zebrafish apolipoprotein E gene (45). When compared with the *Torpedo ache* gene, an insertion also occurred in exon 2 at position 318 in tetrapod *ache* genes. In mammals, it corresponds to an additional intron introduced at 5′-AG/GT-3′ site. In chick AChE, the additional sequence introduced at 5′-TG/GA-3′ is not spliced and encodes a large hydrophilic domain. The intron capture mechanism may also have been responsible of these insertions. It is likely that in teleost lineage, after the divergence of tetrapods and teleost ancestors, a single insertion event of in frame intronic sequence insertion occurred and allowed addition of coding sequence and intron.

Assembly of Different Forms of AChE Is Regulated by Structural Proteins—In vertebrate AChEs the C terminus is encoded by alternatively spliced exons, T, H, or S exons. In insects only H cDNAs are found. In nematodes T exon is present in *ace-1,* whereas H exons are present in the three other *ache* genes (9). The absence of H exon appears to be frequent since it is lacking in *Electrophorus* (21), chick (46), or snake genes (47). In zebrafish AChE cDNA and gene, we identified only one kind of 3′ exon encoding peptide T. The T peptide is characterized by the succession of regularly spaced aromatic residues (see Fig. 2). In vertebrate AChEs it is responsible for the association with the proline-rich attachment domain of structural collagenic subunit and formation of homotetramers (48).

Repartition of molecular forms in adult, shown in Fig. 7, *B* and *C*, is consistent with what is observed in other vertebrates, asymmetric forms being major forms in muscle and amphiphilic tetramers the major form in brain.

As shown by sedimentation analyses, globular forms present at early developmental stages and asymmetric forms appearing later are both composed of T subunits (Fig. 7). Mechanisms driving molecular oligomerization of forms depend mainly on non-catalytic subunits. This has been shown by mRNA injections in *Xenopus* oocytes. Assembly of asymmetric forms is strictly correlated to collagenic tail amount (49). It is very likely that the same mechanism occurs *in vivo*. In zebrafish RT-PCRs performed on RNAs during embryonic development indicate that the cDNA encoding the collagenic tail appears around 48 h,[2] shortly before the appearance of asymmetric forms (Fig. 7*A*).

AChE produced by S2 cells transfected with the AChE minigene are dimers, whereas the injection of the same construct in embryos resulted in overexpression of equal amounts of tetramers and dimers (Fig. 7*D*). Molecular forms in control embryos and in AChE-overexpressing embryos differ widely. In the latter, dimers and tetramers are present in equal amounts, whereas in non-injected embryos we found only tetramers. It has been shown that oligomerization of tetramers could depend on proline-rich factors (9). We suggest that in the case of AChE overexpression the "assembler proline-rich protein" may be present in too low a concentration to assemble all AChE dimers in tetramers, and then excess dimers accumulated in embryos. It is possible that such assembler elements are missing in S2 cells. It should be noted that in *Drosophila*, *ache* gene has no T exon and tetramers never form. These findings indicate that zebrafish could be a suitable model to study regulation of molecular forms homo- and hetero-oligomerization.

Early Expression of mRNAs and Protein in Embryos—Tissue specificity and early expression of cholinesterases in mammals are due to specific transcription factor binding sites (50), alternative promoter usage (51), and intronic enhancers (52, 53). In the zebrafish *ache* 5′ non-coding region and first intron, no homology was detected with other *ache* genes. We identified a canonical TATA box that contrasts with other cholinesterase genes that are devoid of this regulating element. However, a TATA box has been identified in the recently described murine AChE neuronal promoter (51). A TATA repeat was also found 32 bp upstream from the *Torpedo* AChE transcription start site. Putative binding sites for AP1, Egr-1, MyoD, and TTF-2 transcription factors are also present.

AChE mRNA and activity are undetectable before the 5-somite stage, after which expression begins in recently formed somites. After about the 10-somite stage, AChE transcripts and proteins are present in every somite primordium, preceding boundary formation, and their expression increases along with differentiation.

Current knowledge of somitogenesis suggests that the first few rostral somites could be patterned by a different mechanism than the more caudal ones (54). Appearance of AChE simultaneously in the 5–7 early somites, and progressively in the later ones, is consistent with this hypothesis. It suggests that AChE expression could depend on factors controlling the

[2] X. Cousin, unpublished results.

pathway of somite anteroposterior determination (55) among which transcription factors Mesp, MyoD, or other members of basic helix-loop-helix proteins, and the Eph family of receptor tyrosine kinases and their ligand.

Such a regionalization of cholinesterases was also found during somitogenesis in other species. In chick embryo a rostrocaudal asymmetry is established with differential expression of BChE and then AChE in the rostral part of the somite (2). This suggests a conserved regulation of expression of AChEs in myotomes. Early expression could indicate that in zebrafish AChE covers both domains of expression of BChE and AChE found in chick somites.

In the nervous system, mRNA and then protein were found in several brain clusters and in spinal cord. Our results are consistent with previous studies showing that, in zebrafish, all primary neurons, including reticulospinal interneurons, sensory neurons, and primary motoneurons, expressed AChE before axonal outgrowth and neurite arborization (4–7). The pattern of AChE mRNAs expression in brain is very similar to expression of the growth cone component GAP-43 (56). The presence of AChE, before expression of GAP-43 (starting at 17–18 h) and growth of primary neurons, adds strength to the hypothesis of AChE involvement in the control of neuronal differentiation as recently suggested by *in vitro* experiments (57). However, in the mouse, neither collagenic tail gene inactivation (58), which prevented accumulation of all AChE at the synapse, nor *ache* gene inactivation (59) prevented embryonic development. In these mutant animals, showing delayed postnatal growth, BChE or other carboxylesterases may supply AChE function (58, 59). In *Drosophila,* which has no other cholinesterase than AChE, a mutation in AChE induced abnormal neuronal development (60). Similarly in zebrafish no *bche* gene was found, and there is no close relative of the active AChE. Identification of *ache* mutants in *D. rerio* will thus be a valuable tool to investigate the implication of AChE in the development of vertebrates and especially during regulation of morphogenesis of the nervous system.

Acknowledgments—We thank Dr. Bricaud (Université des Sciences et Techniques du Languedoc, Montpellier, France) for the AB strain and Dr. Strähle (Institut de Génétique et de Biologie moléculaire et Génétique, Strasbourg, France) for the ABO strain. Dr. Westerfield and Dr. Rosa are acknowledged for constant support.

REFERENCES

1. Brimijoin, S., and Koenigsberger, C. (1999) *Environ. Health Perspect.* **1,** 59–64
2. Layer, P. G., Alber, R., and Rathjen, F. G. (1988) *Development* **102,** 387–396
3. Layer, P. G. (1990) *BioEssays* **12,** 415–420
4. Hanneman, E., Trevarrow, B., Metcalfe, W. K., Kimmel, C. B., and Westerfield, M. (1988) *Development* **103,** 49–58
5. Hanneman, E., and Westerfield, M. (1989) *J. Comp. Neurol.* **284,** 350–361
6. Wilson, S. W., Ross, L. S., Parrett, T., and Easter, S. S., Jr. (1990) *Development* **108,** 121–145
7. Ross, L. S., Parrett, T., and Easter, S. S., Jr. (1992) *J. Neurosci.* **12,** 467–482
8. Hanneman, E. H. (1992) *J. Exp. Zool.* **263,** 41–53
9. Massoulié, J., Anselmet, A., Bon, S., Krejci, E., Legay, C., Morel, N., and Simon, S. (1998) *J. Physiol.* (*Paris*) **92,** 183–190
10. Westerfield, M. (1995) *The Zebrafish Book,* 3rd Ed., University of Oregon Press, Eugene, OR
11. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A laboratory Manual,* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
12. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402
13. Postlethwait, J. H., Yan Y.-L., Gates, M., Horne, S., Amores, A., Brownlie, A., Donovan, A., Egan, E., Ekker, M., Force, A., Gong, Z., Goutel, C., Fritz, A., Kelsh, R., Knapik, E., Liao, E., Paw, P., Ransom, D., Singer, A., Thomson, M., Abduljabbar, T. S., Yelick, P., Beier, D., Joly, J.-S., Larhammar, D., Rosa, F., Westerfield, M., Zon, L. I., Johnson, S. L., and Talbot, W. S. (1998) *Nat. Genet.* **18,** 345–349
14. Shimoda, N., Knapik, E. W., Ziniti, J., Sim, C., Yamada, E., Kaplan, S., Jackson, D., de Sauvage, F., Jacob, H., Fishman, M. C. (1999) *Genomics* **58,** 219–232
15. Karnovsky, M. J., and Roots, L. J. (1964) *J. Histochem. Cytochem.* **12,** 219–221
16. Ellman, G. L., Courtney, D. K., Andres, V., and Featherstone, R. M. (1961) *Biochem. Pharmacol.* **7,** 88–95
17. Cousin, X., Bon, S., Duval, N., Massoulié, J., and Bon, C. (1996) *J. Biol. Chem.* **271,** 15099–15108
18. Pezzementi, L., Sutherland, D., Sanders, M., Soong, W., Milner, D., McClellan, J. S., Sapp, M., Coblentz, W. B., Rulewicz, G., and Merritt, S. (1998) in *Structure and Function of Cholinesterases and Related Proteins* (Doctor, B. P., Taylor, P., Quinn, D. M., Rotundo, R. L., and Gentry M. K., eds) pp. 105–110, Plenum Publishing Corp., New York
19. Vellom, D. C., Radic, Z., Li, Y., Pickering, N. A., Camp, S., and Taylor, P. (1993) *Biochemistry* **32,** 12–17
20. Toutant, J. P., Massoulié, J., and Bon, S. (1985) *J. Neurochem.* **44,** 580–592
21. Simon, S., and Massoulié, J. (1997) *J. Biol. Chem.* **272,** 33045–33055
22. Plasterk, R. H., Izsvak, Z., and Ivics, Z. (1999) *Trends Genet.* **15,** 326–332
23. Izsvak, Z., Ivics, Z., Shimoda, N., Mohn, D., Okamoto, H., and Hackett, P. B. (1999) *J. Mol. Evol.* **48,** 13–21
24. Krauss, S., Concordet, J. P., and Ingham, P. W. (1993) *Cell* **75,** 1431–144
25. Ekker, M., Wegner, J., Akimenko, M. A., and Westerfield, M (1992) *Development* **16,** 1001–1010
26. Venters, S. J., Thorsteinsdottir, S., and Duxson, M. J. (1999) *Dev. Dyn.* **216,** 219–232
27. Kimmel, C. B., Hatta, K., and Eisen, J. S. (1991) *Development* **2,** (suppl.) 47–57
28. Liu, D. W. C., and Westerfield, M. (1992) *J. Neurosci.* **12,** 1859–1866
29. Kumar, S., and Hedges, S. B. (1998) *Nature* **392,** 917–920
30. Amores, A., Force, A., Yan, Y. L., Joly, L., Amemiya, C., Fritz, A., Ho, R. K., Langeland, J., Prince, V., Wang, Y. L., Westerfield, M., Ekker, M., and Postlethwait, J. H. (1998) *Science* **282,** 1711–1714
31. Nornes, S., Clarkson, M., Mikkola, I., Pedersen, M., Bardsley, A., Martinez, J. P., Krauss, S., and Johansen, T. (1998) *Mech. Dev.* **77,** 185–196
32. Ekker, S. C., Ungar, A. R., Greenstein, P., von Kessler, D. P., Porter, J. A., Moon, R. T., and Beachy, P. A. (1995) *Curr. Biol.* **5,** 944–955
33. Hallbook, F. (1999) *Curr. Opin. Neurobiol.* **9,** 616–621
34. Grauso, M., Culetto, E., Combes, D., Fedon, Y., Toutant, J. P., and Arpagaus, M. (1998) *FEBS Lett.* **424,** 279–284
35. Leibel, W. S. (1988) *J. Exp. Zool.* **247,** 209–223
36. Stieger, S., Gentinetta, R., and Brodbeck, U. (1989) *Eur. J. Biochem.* **181,** 633–642
37. Musset, F., Frobert, Y., Grassi, J., Vigny, M., Boulla, G., Bon, S., and Massoulie, J. (1987) *Biochimie* (*Paris*) **69,** 147–156
38. Heierhorst, J., Lederis, K., and Richter, D. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 6798–6802
39. Ivics, Z., Izsvak, Z., Minter, A., and Hackett, P. B. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 5008–5013
40. Izsvak, Z., Ivics, Z., Garcia-Estefania, D., Fahrenkrug, S. C., and Hackett, P. B. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 1077–1081
41. Maulet, Y., Camp, S., Gibney, G., Rachinsky, T. L., Ekstrom, T. J., and Taylor, P. (1990) *Neuron* **4,** 289–301
42. Jbilo, O., Toutant, J. P., Vatsis, K. P., Chatonnet, A., and Lockridge, O. (1994) *J. Biol. Chem.* **269,** 20829–20837
43. Muratani, K., Hada, T., Yamamoto, Y., Kaneko, T., Shigeto, Y., Ohue, T., Furuyama, J., and Higashino, K. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 11315–11319
44. Chatonnet, A., Lorca, T., Barakat, A., Aron, E., and Jbilo, O. (1991*) Cell. Mol. Neurobiol.* **11,** 119–130
45. Durliat, M., Andre, M., and Babin, P. J. (2000) *Eur. J. Biochem.* **267,** 549–559
46. Randall, W. R., Rimer, M., and Gough, N. R. (1994) *Biochim. Biophys. Acta* **1218,** 453–456
47. Cousin, X., Bon, S., Massoulié, J., and Bon, C. (1998) *J. Biol. Chem.* **273,** 9812–9820
48. Simon, S., Krejci, E., and Massoulié, J. (1998) *EMBO J.* **17,** 6178–6187
49. Krejci, E., Legay, C., Thomine, S., Sketelj, J., and Massoulié, J. (1999) *J. Neurosci.* **19,** 10672–10679
50. Mutero, A., Camp, S., and Taylor, P. (1995) *J. Biol. Chem.* **270,** 1866–1872
51. Atanasova, E., Chiappa, S., Wieben, E., and Brimijoin, S. (1999*) J. Biol. Chem.* **274,** 21078–21084
52. Luo, Z. D., Camp, S., Mutero, A., and Taylor, P. (1998) *J. Biol. Chem.* **273,** 28486–28495
53. Chan, R. Y., Boudreau-Lariviere, C., Angus, L. M., Mankal, F. A., and Jasmin, B. J. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 4627–4632
54. Sawada, A., Fritz, A., Jiang, Y.-J., Yamamoto, A., Yamasu, K., Kuroiwa, A., Saga, Y., and Takeda, H. (2000) *Development* **127,** 1691–1702
55. Durbin, L., Sordino, P., Barrios, A., Gering, M., Thisse, C., Thisse, B., Brennan, C., Green, A., Wilson, S., and Holder, N. (2000) *Development* **127** 1703–1713
56. Reinhard, E., Nedivi, E., Wegner, J., Skene, J. H. P., and Westerfield, M. (1994) *Development* **120,** 1767–1775
57. Bigbee, J. W., Sharma, K. V., Chan, E. L., and Bogler, O. (2000) *Pharmacol. Ther.* **86,** 29–48
58. Feng, G., Krejci, E., Molgo, J., Cunningham, J. M., Massoulié, J., and Sanes, J. R. (1999) *J. Cell Biol.* **144,** 1349–1360
59. Xie, W., Stribley, J., Chatonnet, A., Wilder, P. J., Rizzino, A., McComb, R. D., Taylor, P., Hinrichs, S. H., and Lockridge, O. (2000) *J. Pharmacol. Exp. Ther.* **293,** 896–902
60. Greenspan, R. J., Finn, J. A., Jr., and Hall, J. C. (1980) *J. Comp. Neurol.* **189,** 741–774

Zebrafish Acetylcholinesterase Is Encoded by a Single Gene Localized on Linkage Group 7: GENE STRUCTURE AND POLYMORPHISM; MOLECULAR FORMS AND EXPRESSION PATTERN DURING DEVELOPMENT

Christelle Bertrand, Arnaud Chatonnet, Christina Takke, YiLin Yan, John Postlethwait, Jean-Pierre Toutant and Xavier Cousin

J. Biol. Chem. 2001, 276:464-474.
doi: 10.1074/jbc.M006308200 originally published online October 2, 2000
