The Mammalian Gene of Acetylcholinesterase-associated Collagen*

(Received for publication, September 16, 1996, and in revised form, March 18, 1997)

Eric Krejci‡§, Sébastien Thomine‡, Nicola Boschetti‡, Claire Legay‡, Janez Sketelj†, and Jean Massoulié‡

From the ‡Laboratoire de Neurobiologie Cellulaire et Moléculaire, CNRS URA 1857, Ecole Normale Supérieure, Paris, France and †Institute of Pathophysiology, School of Medicine, University of Ljubljana, 61105 Ljubljana, Slovenia

The collagen-tailed or asymmetric forms (A) represent a major component of acetylcholinesterase (AChE) in the neuromuscular junction of higher vertebrates. They are hetero-oligomeric molecules, in which tetramers of catalytic subunits of type T (AChET) are attached to the subunits of a triple-stranded collagen “tail.” We report the cloning of a rat AChE-associated collagen subunit, Q. We show that collagen tails are encoded by a single gene, COLQ. The ColQ subunits form homotrimers and readily form collagen-tailed AChE, when coexpressed with rat AChET. We found that the same ColQ subunits are incorporated, in vivo, in asymmetric forms of both AChE and butyrylcholinesterase. A splice variant from the COLQ gene encodes a proline-rich AChE attachment domain without the collagen domain but does not represent the membrane anchor of the brain tetramer. The COLQ gene is expressed in cholinergic tissues, brain, muscle, and heart, and also in noncholinergic tissues such as lung and testis.

Acetylcholinesterase (AChE, EC 3.1.1.7) is highly concentrated at vertebrate neuromuscular junctions. This enzyme is encoded by a single gene, and adult mammalian muscles express a single splice variant, corresponding to the catalytic subunit of type T (AChET), (1, 2). At the post-translational level, however, quaternary interactions introduce a considerable diversity of molecular forms that are characterized by distinct localizations in cellular structures. These molecules include amphiphilic monomers (G1), and dimers (G2), nonamphiphilic tetramers (G4), as well as hetero-oligomeric structures in which tetramers of catalytic subunits are disulfide-linked with a hydrophobic “tail” (20 kDa) in the membrane-bound G4 forms (3, 4) or with a collagenous “tail” in the collagen-tailed or asymmetric (A) forms. The latter molecules consist of one, two, or three tetramers (A1, A2, A3), which are disulfide-linked to the strands of the triple helical collagen tail (see Fig. 1A). G4 and G2 forms appear to remain mostly intracellular and represent precursors of more complex molecules. The G4,na form is secreted and hydrophobic-tailed tetramers (G4,na) are attached to the plasma membrane. The collagen-tailed molecules are tethered in the basal lamina, and are largely responsible for the high concentration of AChE at the neuromuscular junction.

To understand the biosynthesis of the various AChE forms and its regulation, it is necessary to analyze the association of AChET catalytic subunits with anchoring subunits, particularly the collagen subunits, which have been named Q, according to the nomenclature of AChE-associated proteins (5). Cloning and expression of the collagen tail subunit of the asymmetric AChE forms from Torpedo electric organ (tQ) allowed us to show that the structural and catalytic subunits assemble into collagen-tailed molecules when coexpressed in COS cells (6). The primary sequence of the Q subunit comprises an N-terminal region, QN, a collagen domain, and a C-terminal domain, QC. We showed that the QC domain is able to recruit monomers, producing a tailed tetramer (7, 8) and defining a proline-rich attachment domain (PRAD) that is sufficient for interaction with AChET (9).

In the present study, we cloned cDNAs encoding Q subunits from the rat, using cross-hybridization with Torpedo probes. We asked three major questions: Do several genes encode different Q subunits? Are collagenous and noncollagenous AChE-associated proteins, particularly the 20-kDa membrane anchor, generated by the same gene(s)? Is the expression of Q gene(s) exclusively restricted to cholinergic tissues?

EXPERIMENTAL PROCEDURES

Unless otherwise indicated, reagents were purchased from Prolabo (Paris, France), Sigma, or Appligene (Illkirch, France), and enzymes from New England Biolabs (Ozyme, France).

Preparation of Reinnervating Rat Soleus Muscle and Sternomastoid Muscle—Ten male Wistar albino rats weighing 180–250 g at the time of surgery were anesthetized by intraperitoneal injection of a mixture of 5 mg/kg xylazine (Bompoint, Bayer AG, Leverkusen, Germany) and 90 ng/kg ketamine (Retalar, Parke-Davis and Co., Berlin, Germany). The sciatic nerve on one side was exposed and crushed in the mid thigh region by a nonserrated hemostat for 30 s. The wound skin was closed, and the animals were left to recover for 21 days to allow for sciatic nerve regeneration and reinnervation of the soleus muscle. The animals were then sacrificed, and the reinnervating soleus muscle was rapidly isolated and frozen in liquid nitrogen.

Preparation and Screening of a cDNA Library—Total RNA was extracted from muscle (10), and a cDNA library was constructed in pCDM8. Briefly, double-stranded cDNA was obtained (11), using a cDNA synthesizing kit (Pharmacia). Then, BstXI adaptors were linked to the blunt-ended cDNA, after purification by Sephacyrl 400 chromatography, and the cDNAs were ligated in pCDM8, which had previously been cleaved by BstXI.

1.5 × 109 independent transformed MC1061/P12 bacteria were distributed in pools of 10,000–12,000 clones and amplified. 60 pools were grown, and for each pool the DNA was extracted, digested by Xhol, run on agarose gels, and blotted to nylon membranes that were hybridized

© 1997 by The American Society for Biochemistry and Molecular Biology, Inc.
with probes derived from the *Torpedo* cDNA (nucleotides 1–213 and 916–1416) (6). The membranes were incubated in 7% SDS, 0.5 mM sodium phosphate, pH 7.2, 2 mM EDTA at 50 °C with the radiolabeled probes and washed in 1% SDS, 0.2 mM sodium phosphate, and 2 mM EDTA at 45 °C. One pool was positive with the two probes. Sequential dilutions were used to isolate a positive clone. A second pool of this library with a probe derived from the rQ1 clone yielded five other clones with different structures, which will not be described in detail here.

Sequences were determined by the dideoxyxenucleotide method with the Sequenase kit (U. S. Biochemical Corp.) and analyzed by the Gene-works (Genoscope) program. Comparison of the protein and nucleic acid sequences with data banks were done with the FASTA and BLAST programs.

**Analysis of mRNA by Northern Blots—**We used the rat multiple tissue Northern blot from CLONTECH, prepared with 2 μg of poly(A)*+* RNA from heart, brain, spleen, lung, liver, skeletal muscle, kidney, and testis. The N-terminal probe (nucleotides 5–412 of rQ1 cDNA) was obtained by polymerase chain reaction with oligonucleotides 5′-GCTT GCCCTTCGATTGATA-3′ and 5′-GGGGCCACCTGAGGGCAGC-3′, the C-terminal probe (nucleotides 892–1412) with oligonucleotides 5′-GAAAGAGGATTTCCAGGGCC-3′ and 5′-TATCCGGAGGGTGGTGGA-3′. Radiolabeled probes of high specific activity were generated by random priming with the Rediprime kit (Amersham Corp.). Hybridization of RNA with the radioactive probes, RNase digestion, and inactivation were performed with the Hybspeed RNase protection assay kit (Ambion). The membranes were rinsed three times with 50 mM Tris/HCl, pH 7.4, 0.15 mM NaCl. For blocking, a solution of 20 mM Tris/HCl, pH 7.4, 0.15 mM NaCl (buffer A) and 3% bovine serum albumin was added to the blot, which was incubated for 4 h at room temperature. It was then incubated overnight at 4 °C with 1/200 diluted antiserum in buffer B (buffer A plus 0.05% Tween 20). The membranes were rinsed five times with buffer B and incubated for 2 h at room temperature with horseradish peroxidase-conjugated rabbit anti-mouse or anti-hen immunoglobulins diluted 1:1000 with buffer B. The strips were washed five times with buffer B and developed with a solution of 0.05% of diaminobenzidine in buffer A containing 1 μM H₂O₂.

**Results**

**Cloning of the Rat Collagen Subunit (rQ): A Modular Structure—**To increase the level of transcripts encoding the collagen tail, we used a rat soleus muscle which was in the process of reinnervation after a crush of the sciatic nerve; in this case, AChE collagen-tailed forms become highly expressed through-out nerve regeneration after a crush of the sciatic nerve; in this case, AChE collagen-tailed forms become highly expressed throughout the entire length of muscle fibers (22). We constructed a cDNA library containing 1.5 × 10⁶ independent clones from RNA isolated from the reinnervating soleus and screened it with probes corresponding to the QN and QC regions of the *Torpedo* Q1 subunit. We thus obtained one cDNA clone, rQ1, which cross-hybridized at low stringency with the two probes. The length of rQ1 was 2,731 nucleotides. It contains 45 nucleotides of the 5′-untranslated region, and 1,912 nucleotides of the 3′-untranslated region. The coding sequence (641 bp) is completely conserved between many species (21). The coding sequence contains 2731 nucleotides. It contains 45 nucleotides of the 5′-untranslated region, and 1912 nucleotides of the 3′-untranslated region. The coding sequence (46–1419) is entirely homologous to that of *Torpedo* Q1 (6); rQ1 and Q1 presented 60% identity at the nucleotide level and 52% identity at the amino acid level.

**Results**

An alignment of the deduced peptide sequences shows that nonconserved regions alternate with remarkably well conserved regions, which define potentially functional domains (Fig. 1B). The N-terminal part of the sequence corresponds to a secretion signal peptide, which is predicted to contain 22 residues in rQ1, and 42 residues in Q1 (23). The predicted cleavage sites are located at the same position in the aligned rQ1 and Q1.
sequences, and this position was recently confirmed in the case of tQ1 (8). The N-terminal extremity of the mature protein shows only 10 conserved residues out of 28, but it is followed by a very conserved PRAD, containing 15/17 identical residues. The PRAD sequence includes two adjacent cysteines that establish disulfide bonds with AChET subunits (24) and stretches of five and three prolines. This peptidic domain is sufficient for binding an AChET tetramer (9). It is separated from the collagen domain by 30 residues, showing essentially no conservation between tQ1 and rQ1, except for the presence of a cysteine residue, which may participate in the stabilization of the triple helix through intersubunit disulfide bonds (tQ1 contains two cysteines in this region, only one of which is conserved in rQ1).

The collagen domain has approximately the same length in both sequences; it consists of 63 triplets of amino acids GXY in rat, compared with 60 in Torpedo, with an interruption of 10 residues in tQ1, 7 in rQ1, at the same position in both sequences. Two internal motifs, which have been proposed to represent binding sites for heparan sulfate proteoglycans in tQ1 (25), are totally conserved, RKGR (128–131) and KRGK (234–237).

The C-terminal noncollagenous domain contains three conserved regions of unequal length. The first conserved region consists of 14 residues, beginning essentially no conservation between tQ1 and rQ1, except for the presence of a cysteine residue, which may participate in the stabilization of the triple helix through intersubunit disulfide bonds (tQ1 contains two cysteines in this region, only one of which is conserved in rQ1).

The collagen domain has approximately the same length in both sequences; it consists of 63 triplets of amino acids GXY in

---

**Fig. 1. Organization of the collagen tail of AChE, ColQ.** **A**, schematic structure of the quaternary association of catalytic AChE subunits and collagen Q subunits in the collagen-tailed AChE (Q) sequence. Identical amino acids are boxed; vertical lines indicate the limits of exons. Conserved motifs are indicated by marks below the sequence: Hatched box for the PRAD, black boxes for putative HSP binding domains, a wavy line for the collagen triple helical domain, gray boxes for C-terminal subdomains.
known proteins. Separated from this domain by a short linker, the third conserved region covers the C-terminal 81 residues of rQ1 (83 in tQ1). It contains 10 conserved cysteines, and presents an imperfect internal repeat, as previously noted in the case of tQ1 (6).

Another cDNA clone, rQR, was isolated from the soleus muscle library and also obtained by a rapid amplification of cDNA ends procedure. In this transcript the PRAD-encoding sequence is followed by a short sequence (355 nucleotides) that is not included in rQ1 (see Fig. 3). As described in a subsequent section, analysis of the genomic sequence in the mouse showed that this unrelated sequence corresponds to the "readthrough" of the following intron. The rQR transcript was terminated by a poly(A) sequence, despite the absence of a canonical polyadenylation signal.

Collagen Q Is Encoded by a Single Gene; Partial Genomic Organization—Rat genomic DNA, cleaved by restriction enzymes, was hybridized in Southern blots at low stringency with probes corresponding to the QC domain. We observed a single strong signal (not shown), suggesting the absence of closely related genes, since C-terminal sequences are generally well conserved among members of a collagen family (26). We also screened a human cosmid genomic library with two probes corresponding to the rat QC domain, at low stringency. We obtained five independent cosmids, all of which overlapped in the same genomic region. In addition, we cloned a 35-kb cosmid containing part of the mouse Q gene, by hybridization with a probe covering rat QN domain. This cosmid covers less than a third of the coding sequence. Fig. 3 shows the exon-intron boundaries of the exons encoding the PRAD and the collagen domain. A single exon of 131 base pairs (EPRAD) en-

---

**Fig. 2.** The sedimentation of AChE A forms in sucrose gradients is shifted by antibodies against the collagen subunit. Sedimentation profiles of AChE activity in extracts from A. *Xenopus* oocytes injected with 25 ng of AChE T and 2.5 ng of rQ1 mRNA and B. soleus muscle. In both cases, the extracts were centrifuged in sucrose gradients containing 0.4 M NaCl and 1% Brij-96, alone (open circles), or after incubation with 1/250 volume of IgY chicken antibodies directed against the N-terminal domain of rQ1 (anti-rQ35–51) (black circles). The sedimentation coefficients were deduced by linear relationship from the position of internal marker proteins, E. coli β-galactosidase (16 S), and alkaline phosphatase (6.1 S). The AChE activity of each fraction is plotted on an arbitrary scale, as a function of the calculated S value.

---

2 E. Krejci, unpublished results.
FIG. 3. Partial structure of the mouse COLQ gene. A, the map presents the BamHI restriction sites (B) and the position of the exons: EPRAD is hatched, the readthrough sequence of rQ1 in white, the collagen exons are black. B, the exons are shown in uppercase letters with the deduced amino acid sequence, the numbering refers to the rQ1 sequence, and the coding differences between mouse and rat noted. The extremities of the introns (30 base pairs on each side) are shown in lowercase letters, their approximate size is indicated.
codes the PRAD and flanking sequences, and the next 3’ exon (Ecoll 1, 96 nucleotides) extends into the beginning of the collagen domain. Exons encoding the N-terminal part of the collagen domain are small (27–72 nucleotides), and each one encodes a multiple of GXY triplets, starting with the first base of the glycine codon (GGN). The few coding differences between the rat and the mouse sequences are indicated on Fig. 3.

The Asymmetric Forms of AChE and BChE Possess the Same Collagen Tail—Another way to investigate the possible diversity of cholinesterase-associated collagens is to examine whether BChE asymmetric forms cross-react with antibodies raised against the nonconserved peptide that precedes the PRAD, anti-rQ135–51. Because collagen-tailed forms of BChE are relatively abundant in Torpedo heart (27), and because rQ1 appears to be expressed at a high level in rat heart, we thought that it might be possible to characterize asymmetric BChE in this tissue. We performed sequential extractions of rat heart ventricle in low salt and in high salt concentrations, as described previously (17), to obtain an extract enriched in AChE and BChE forms, was diluted to 0.4 M NaCl and incubated at room temperature without or with 1/200 volume of IgY anti-rQ135–51 antibodies, prior to sedimentation in sucrose gradients. For each fraction, AChE and BChE activities were measured by the colorimetric method of Ellman in the presence of appropriate inhibitors and plotted as described in Fig. 2.

A high salt extract from rat heart ventricle, enriched in asymmetric gradient, 16.5 S for AChE A12 and 18 S for BChE A12. This high salt extract was analyzed by sedimentation in sucrose gradients, followed by assays of AChE and BChE activities, performed in the presence of the specific inhibitors tetraisopropyl pyrophosphoramide and BW284C51, respectively (Fig. 4). The patterns obtained showed the presence of asymmetric forms of both AChE and BChE, with distinct sedimentation coefficients in the fractions of the same sucrose gradient, 16.5 S for AChE A12 and 18 S for BChE A12. This difference in sedimentation, in complete agreement with previous analyses in rat muscle (28), established that the AChE and BChE molecules were distinct and independent. The sedimentation coefficients of all asymmetric forms, BChE as well as AChE, were increased by about 2 S after incubation with the anti-rQ135–51 antibodies, while globular forms were not affected. Although the BChE A4 and A8 forms were not entirely resolved from the G4 peak, these molecules were also clearly shifted by the antibodies.

Analysis of RNA Splicing at the Junction between Exons Encoding the PRAD and the Collagen Domain—The rQR cDNA clone corresponds to unspliced readthrough transcript, in which the PRAD exon is followed by the subsequent genomic sequence (Fig. 3), instead of the collagen exons. This transcript is polyadenylated and thus represents a mature mRNA, which encodes the PRAD without the collagen domain. By RNase protection assays with a probe covering EPRAD and the readthrough sequence rQ1 (Q1 probe), we showed that this mRNA represents about 5% of the transcripts in the soleus muscle and about 15% in the heart ventricle (Fig. 5). We also performed RNase protection assays with a probe covering EPRAD and the three following exons, encoding the N-terminal part of the collagen domain (Q1 probe). As shown in Fig. 5, we observed a major protected fragment corresponding to the structure of Q1 in both tissues; the fragment corresponding to EPRAD originating from rQ6 or other splicing events was nearly undetectable in this experiment but was visible in other experiments. Using reverse transcription-polymerase chain reaction analysis we failed to detect any transcript in which the PRAD-containing exon would be associated with downstream sequences differing from rQ6 or rQ1. We obtained similar results with brain mRNA (not shown).

Amphiphilic Tetramers from Brain Do Not Contain the PRAD of ColQ—Since the protein encoded by rQR contains the PRAD of ColQ—Another way to investigate the possible diversity of cholinesterase-associated collagens is to examine whether BChE asymmetric forms cross-react with antibodies raised against the nonconserved peptide that precedes the PRAD, anti-rQ135–51. Because collagen-tailed forms of BChE are relatively abundant in Torpedo heart (27), and because rQ1 appears to be expressed at a high level in rat heart, we thought that it might be possible to characterize asymmetric BChE in this tissue. We performed sequential extractions of rat heart ventricle in low salt and in high salt concentrations, as described previously (17), to obtain an extract enriched in AChE and BChE forms, was diluted to 0.4 M NaCl and incubated at room temperature without or with 1/200 volume of IgY anti-rQ135–51 antibodies, prior to sedimentation in sucrose gradients. For each fraction, AChE and BChE activities were measured by the colorimetric method of Ellman in the presence of appropriate inhibitors and plotted as described in Fig. 2.

A high salt extract from rat heart ventricle, enriched in asymmetric gradient, 16.5 S for AChE A12 and 18 S for BChE A12. This high salt extract was analyzed by sedimentation in sucrose gradients, followed by assays of AChE and BChE activities, performed in the presence of the specific inhibitors tetraisopropyl pyrophosphoramide and BW284C51, respectively (Fig. 4). The patterns obtained showed the presence of asymmetric forms of both AChE and BChE, with distinct sedimentation coefficients in the fractions of the same sucrose gradient, 16.5 S for AChE A12 and 18 S for BChE A12. This difference in sedimentation, in complete agreement with previous analyses in rat muscle (28), established that the AChE and BChE molecules were distinct and independent. The sedimentation coefficients of all asymmetric forms, BChE as well as AChE, were increased by about 2 S after incubation with the anti-rQ135–51 antibodies, while globular forms were not affected. Although the BChE A4 and A8 forms were not entirely resolved from the G4 peak, these molecules were also clearly shifted by the antibodies.

Analysis of RNA Splicing at the Junction between Exons Encoding the PRAD and the Collagen Domain—The rQR cDNA clone corresponds to unspliced readthrough transcript, in which the PRAD exon is followed by the subsequent genomic sequence (Fig. 3), instead of the collagen exons. This transcript is polyadenylated and thus represents a mature mRNA, which encodes the PRAD without the collagen domain. By RNase protection assays with a probe covering EPRAD and the readthrough sequence rQ6 (Q6 probe), we showed that this mRNA represents about 5% of the transcripts in the soleus muscle and about 15% in the heart ventricle (Fig. 5). We also performed RNase protection assays with a probe covering EPRAD and the three following exons, encoding the N-terminal part of the collagen domain (Q1 probe). As shown in Fig. 5, we observed a major protected fragment corresponding to the structure of Q1 in both tissues; the fragment corresponding to EPRAD originating from rQ6 or other splicing events was nearly undetectable in this experiment but was visible in other experiments. Using reverse transcription-polymerase chain reaction analysis we failed to detect any transcript in which the PRAD-containing exon would be associated with downstream sequences differing from rQ6 or rQ1. We obtained similar results with brain mRNA (not shown).

Amphiphilic Tetramers from Brain Do Not Contain the PRAD of ColQ—Since the protein encoded by rQR contains the PRAD of ColQ
nonsenselectrophoresis, we found that anti-rQ\textsubscript{35–51} antibodies recognize the heavy dimer of G\textsubscript{11} produced by coexpression of \textit{rQ}\textsubscript{A} and AChE\textsubscript{B} in oocytes but not the dimers that constitute the G\textsubscript{11} form extracted from bovine brain (Fig. 6). In addition to the labeling of the heavy dimer, which appears diffuse, we observed a labeled band corresponding to the association of a rQ\textsubscript{A} with a single AChE subunit. In the brain extract, a labeled band of high molecular weight could correspond to the A form. As positive controls, we verified that the antibodies shifted the sedimentation of bovine A forms in sucrose gradient, demonstrating that they interact with the bovine Q\textsubscript{A} sequence (not shown) and that the heavy dimer of purified brain G\textsubscript{11} is recognized by an anti-anchor antibody (29), producing a labeled band at the expected position in Western blot (not shown).

Expression of the \textit{COLQ} Gene in Rat Tissues—We analyzed the tissue distribution of \textit{rQ} transcripts by Northern blotting, using an N-terminal probe extending into the beginning of the collagen domain (N-terminal probe), and a probe corresponding to the Q\textsubscript{C} domain (C-terminal probe). As shown in Fig. 7, both N-terminal and C-terminal probes recognized a major transcript, of about 2.7 kb, corresponding to the size of \textit{rQ}\textsubscript{1}, in muscle, heart, and brain, as well as in noncholinergic tissues including lung, spleen, and testis, but not in the liver (Fig. 7, lane 5). Surprisingly, the \textit{rQ}\textsubscript{1} transcripts appeared less abundant in skeletal muscle than in lung or spleen, and were particularly abundant in heart. In addition to the 2.7-kb transcript, the two probes hybridized with smaller distinct bands but of much lower intensity; for example, the lung contains a 2.2-kb transcript that hybridizes with the C-terminal but not with the N-terminal probe.

DISCUSSION

A Single Gene Encodes the Collagen Tail of AChE and BChE: Collagen of Type Q, \textit{ColQ}—The asymmetric forms of AChE are hetero-oligomers in which AChE\textsubscript{B} catalytic subunits are associated with collagenous Q subunits. We report the cloning of a Q subunit from rat muscle, \textit{rQ}\textsubscript{1}, which is homologous to the previously cloned \textit{Torpedo} collagen subunit, now renamed \textit{tQ}\textsubscript{1} (6). The fact that \textit{rQ}\textsubscript{1} indeed encodes the AChE-associated collagen tail was demonstrated by the production of asymmetric forms with AChE\textsubscript{B} subunits in COS cells and in \textit{Xenopus} oocytes and by the fact that antibodies directed against a nonconserved peptide from \textit{rQ}\textsubscript{1} (anti-rQ\textsubscript{35–51}) interact with all asymmetric forms extracted from tissues.

Southern blotting of rat genomic DNA and screening of cDNA libraries containing mouse and human genomic fragments showed that Q subunits are encoded by a single gene, \textit{COLQ}. This conclusion is supported by the fact that the anti-rQ\textsubscript{35–51} antibodies recognize BChE as well as AChE A forms, indicating that the collagen tails of the two enzymes are generated from the same gene. A partial analysis of the collagen exons of the \textit{COLQ} gene reveals that they are organized as in the genes of true collagens. We found that collagen exons contain an integral number of codons, corresponding to multiples of GXY triplets, as in the case of several fibrillar collagens (30).

An alignment of the \textit{Torpedo} and rat sequences defines well conserved motifs and nonconserved regions. In the collagen domain, the regular repetition of the GXY triplet is interrupted once by a short noncollagenous motif (10 residues) that could correspond to a hinge between independent parts of the collagen. Note that the N- and C-terminal parts of the collagen domain contain a higher proportion of prolines (essentially triplets of amino acids GXP and GPP), compared with the central part. The two putative heparan sulfate binding sites (25) are localized at the two ends of the first part of the collagen domain preceding the hinge. In addition to the collagen domain itself, the conserved motifs include an N-terminal PRAD that binds an AChE tetramer (9), a short sequence that immediately follows the collagen domain (8 conserved residues out of 14), a highly hydrophilic motif containing a high proportion of charged amino acids (45 residues), and a C-terminal motif of 80 residues, containing 10 cysteines with an internal repeat. The C-terminal non-collagen sequences are usually well conserved within collagen families and may participate in the formation of trimers, prior to the organization of the triple helix. The conserved C-terminal motifs of ColQ have no homology with those of other collagens or of other known proteins; in addition, the length and the structure of the collagen domain have no equivalent in any known type of collagen (26). Therefore the \textit{Torpedo} and rat collagen subunits of cholinesterase collagen-tailed forms constitute a distinct type of collagen that we name collagen of type Q or \textit{ColQ}.

\textit{ColQ} Is Expressed in Cholinergic and Noncholinergic Tissues—Northern blots showed that the \textit{rQ} gene is expressed in cholinergic tissues such as brain, muscle, and heart, as expected, but also in noncholinergic tissues, including lung and testis, but not liver. We found that all tissues expressing \textit{rQ}\textsubscript{1} contained transcripts that hybridized with both 5' and 3' probes and corresponded to the full length of \textit{rQ}\textsubscript{1}. However, shorter transcripts that hybridized only with a 5' or with a 3' probe occurred in a tissue-specific manner, \textit{e.g.} in the lung (3').
and in the testis (5'). The short 3' transcript cannot encode the binding domain (PRAD) nor the beginning of the collagen domain, so that the deduced protein would not associate with AChE footcatalytic subunits and may not be collagenous in structure. As for the short 5' transcript, it could correspond to an alternatively spliced transcript that would encode the PRAD, but not the collagen domain, as in the rQ7 cDNA clone (see next section), or to nonspecific hybridization with the PRAD-coding sequence, that contains a repetition of proline codons (CCN).

Extracts from rat heart ventricles were found to contain only a low level of collagen-tailed AChE forms (31). It was therefore surprising to find that heart contains the highest level of COLQ transcripts, which we found to correspond to rQ1 (not shown). This suggests that COLQ is not exclusively associated with accumulated AChE at synapses, but may also exist as an independent component of extracellular matrices. It should be recalled, however, that collagen-tailed forms of cholinesterases may have been underestimated or undetected, if present in some tissues in a nonextractable state (32, 33).

Diversity of AChE Anchor Molecules—The N-terminal non-collagen sequence contains two vicinal cysteines, as well as stretches of consecutive prolines, constituting the PRAD that binds a tetramer of AChE subunits (7–9). This very well conserved peptide motif of 17 amino acids is sufficient to interact with AChE subunits, organizing tetramers from monomers and dimers.

We found that the exon encoding the binding domain, ExPRAD, may be expressed without the collagen exons. This raised the possibility that the 20-kDa hydrophobic anchor of AChE (Gx)3,4 (3, 4) might be encoded by a splice variant of the COLQ gene. The same attachment domain, PRAD, would be used for associating AChE subunits with both collagenous and hydrophobic anchors, allowing the localization of the enzyme in the basal lamina and in membranes, respectively. One cDNA clone (rQ8) corresponds to a readthrough transcript, in which ExPRAD is not spliced to the first collagen exon (Ecoll 1) as in rQ1, but followed by the continuous "intronic" sequence that encodes a short peptide. This transcript is mature since its terminates with a poly(A) stretch. When co-expressed with AChE foot, the resulting protein was incorporated in a catalytic tetramer composed of a light dimer and a heavy dimer. The heavy dimer was labeled in Western blots by anti-rQN35–51 antibodies, showing that it binds a tetramer of AChE subunits (Gx)3,4, (3, 4) might be encoded by a splice variant of the COLQ gene.

This suggests that COLQ is not exclusively associated with AChE T subunits into tetramers requires the presence of a binding protein that is expressed at variable levels in the different cell types. It is intriguing that such cholinesterase-associated proteins may not be genetically or even structurally related to COLQ, since the hydrophobic anchor of membrane-bound AChE tetramers from brain does not contain the PRAD.

Acknowledgments—We thank Anne Le Goff and Boris Pecenko for expert technical assistance, Dr. Jacques Neyton for help with injection of the Xenopus oocytes, Dr. Suzanne Bon for transfections in the COS cells, Dr. Neva Crne-Finderle for help with reinervation experiments, Prof. U. Brodbeck for the trimethylammonium phenylenediamine affinity resin, and Jacqueline Pons for typing the manuscript.

REFERENCES

1. Li, Y., Camp, S., Rachinsky, T. L., Getman, D., and Taylor, P. (1991) J. Biol. Chem. 266, 23083–23090
2. Legay, C., Hutuch, M., Massoulie, J., and Changeux, J. P. (1995) Eur. J. Neurosci. 7, 1803–1809
3. Gennari, K., Brunner, J., and Brodbeck, U. (1987) J. Neurochem. 49, 12–18
4. Inestrosa, N. C., Roberts, W. L., Marshall, T. L., and Rosenberry, T. L. (1987) J. Biol. Chem. 262, 4441–4444
5. Massoulie, J., Sussman, J. L., Doctor, B. P., Soreq, H., Velan, B., Cugy, M., Rotundo, R., Shafferman, A., Silman, I., and Taylor, P. (1991) in Cholinesterases (Massoulie, J., Bacou, F., Barnard, E., Chatenet, A., Doctor, B. P., and Quin, M. D., eds) pp. 285–288, American Chemical Society, Washington, D. C.
6. Kreici, E., Coussen, F., Durval, N., Chatel, J. M., Legay, C., Puyve, M., Vandekerchkove, J., Cartaud, J., Bon, S., and Massoulie, J. (1991) EMBO J. 10, 1285–1293
7. Durval, N., Kreici, E., Grassi, J., Coussen, F., Massoulie, J., and Bon, S. (1992) EMBO J. 11, 3255–3261
8. Bon, S., and Massoulie, J. (1997) J. Biol. Chem. 272, 3007–3015
9. Bon, S., Coussen, F., and Massoulie, J. (1997) J. Biol. Chem. 272, 3016–3021
10. Chomczynski, P., and Sacchi, N. (1987) Anal. Biochem. 162, 156–159
11. Rubes, J. P., and Grham, B. J. (1983) Gene (Amst.) 25, 263–269
12. Miller, S. A., Dykes, D. D., and Polesky, H. F. (1988) Nucleic Acids Res. 16, 1215
13. Henikoff, S. (1987) Methods Enzymol. 155, 156–165
14. Legay, C., Bon, S., Vernier, P., Coussen, F., and Massoulie, J. (1993) J. Neurochem. 60, 337–346
15. Krieg, P. A., and Melton, D. A. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 2351–2355
16. Soreq, H., Parvari, R., and Silman, I. (1982) Proc. Natl. Acad. Sci. U. S. A. 79, 830–834
17. Bon, S., Cartaud, J., and Massoulie, J. (1978) Eur. J. Biochem. 85, 1–14
18. Brodbeck, U., Gentinetta, R., and Ott, P. (1981) in Membrane Proteins: A Laboratory Manual (Azzi, A., Brodbeck, U., and Zahler, P., eds) pp. 85–96, Springer-Verlag, Berlin
19. Ellman, G. L., Courtney, D. K., Andrews, V., and Featherstone, R. M. (1961) Biochem. Pharmacol. 7, 88–95
20. Laemmli, U. K. (1970) Nature 270, 680–685
21. Massoulie, J., Pezzemignt, L., Bon, S., Kreici, E., and Vallette, F. M. (1993) Prog. Neurobiol. 41, 31–91
22. Crne-Finderle, N., Thipke, J., and Sketlej, J. (1995) J. Neurosci. Res. 41, 745–752
23. von Hejne, G. (1986) Nucleic Acids Res. 14, 4683–4696
24. Lee, S. L., and Taylor, P. (1982) J. Biol. Chem. 257, 12922–12931
25. Deprez, P. N., and Inestrosa, N. C. (1995) J. Biol. Chem. 270, 11043–11046
26. van der Rest, M., and Garrone, R. (1991) FASEB J. 5, 2814–2823
27. Toutant, J. P., and Massoulie, J. (1988) Handb. Exp. Pharmacol. 86, 225–265
28. Vigny, M., Gisiger, V., and Massoulie, J. (1978) Proc. Natl. Acad. Sci. U. S. A. 75, 2588–2592
29. Liao, J., Norgard-Pedersen, B., and Brodbeck, U. (1993) J. Neurochem. 61, 1127–1134
30. Pockop, D. J., and Kivirikko, K. I. (1995) Annu. Rev. Biochem. 64, 403–434
31. Nyqvist-Battie, C., Hodges-Savola, C., and Fernandez, H. L. (1987) J. Mol. Cell. Cardiol. 19, 935–943
32. Younkin, S. G., Rosenstell, C., Collins, P. L., and Rosenberry, T. L. (1982) J. Biol. Chem. 257, 13630–13637
33. Rossi, S. G., and Rotundo, R. L. (1993) J. Biol. Chem. 268, 19152–19159
34. Cater, J. L., and Bingham, S. (1981) J. Neurochem. 36, 1018–1025
35. Anglister, L. (1991) J. Cell. Biol. 115, 755–764
36. Coussen, F., Bonnerot, C., and Massoulie, J. (1995) Eur. J. Cell Biol. 67, 254–260
37. Velan, B., Kronman, C., Grosfeld, H., Leitner, M., Gozes, Y., Flashner, Y., Sery, T., Cohen, S., Ben-Aziz, R., Seidman, S., Shafferman, A., and Soreq, H. (1991) Cell. Mol. Neurobiol. 11, 143–156