Molecular Biological Search for Human Genes Encoding Cholinesterases

Hermona Soreq and Averell Gnatt

Department of Biological Chemistry
The Life Sciences Institute
The Hebrew University Jerusalem, Israel

Contents

Abstract

Cholinesterases (ChEs) are highly polymorphic proteins, capable of rapidly hydrolyzing the neurotransmitter acetylcholine and involved in terminating neurotransmission in neuromuscular junctions and cholinergic synapses. In an attempt to delineate the structure and detailed properties of the human protein(s) and the gene(s) coding for the acetycholine hydrolyzing enzymes, a human cDNA coding for ChE was isolated by use of oligodeoxynucleotide screening of cDNA libraries. For this purpose, a method for increasing the effectiveness of oligonucleotide screening by introducing deoxyinosine in sites of codon ambiguity and using tetramethyl-ammonium salt washes to remove false-positive hybrids was employed. The resulting isolated 2.4-kilobase (kb) cholinesterase cDNA sequences encode for the entire mature secretory protein, preceded by an N-terminal signal peptide. The human ChE primary sequence shows almost no homology to other serine hydrolases, with the exception of a hexapeptide at the active site. In contrast, it displays extensive homology with acetycholinesterase form *Torpedo californica* and *Drosophila melanogaster* as well as with bovine thyroglobulin. These extensive homologies probably suggest the need of the entire coding sequence for the physiological function(s) fulfilled by the enzyme and further suggest a common, unique, ancestral gene for these cDNAs. In turn, the cDNA was used as a probe to isolate genomic DNA sequences for the 5'-region of the human ChE gene. The genomic DNA fragment encoding part of the 5'-region of ChEcDNA was detected by DNA blot hybridization, enriched 70-fold by gel electrophoresis and electroelution, cloned in λ phage and isolated. Sequencing of the cloned DNA revealed that it did indeed include part of the 5'-region of ChEcDNA, starting at an adjacent 5'-position to the nucleotides coding for the initiator methionine, and ending with an *Eco*RI restriction site inherent to the ChEcDNA sequence. The isolated fragment of the human cholinesterase gene is currently employed to complete the structural characterization of this and related genes.

Index Entries: Human genes, encoding cholinesterases; cholinesterase; protein polymorphism; neurotransmission; acetylcholine; butyrylcholine.

Introduction

Human Cholinesterases

Cholinesterase Classes and Their Physiological Significance.

Cholinesterases (ChEs) are scarce polymorphic proteins capable of rapidly degrading the neurotransmitter acetylcholine (Silver, 1974). Two major classes of the enzyme exist, which differ in the nature of their preferences for substrate: acetylcholine for acetylcholinesterase (AChE, true ChE, acetylcholine acetylhydrolase, EC 3.1.1.7) and butyrylcholine or propionylcholine for pseudocholinesterase (ψChE, pseudoChE, acylcholine acylhydrolase, EC 3.1.1.8). Both are generally distinguished by their susceptibility to inhibitors (Austin and Berry, 1953). Acetylcholinesterase is the major class in neuromuscular junctions ranging from fish to human and a key element in certain cholinergic synapses. At the neuromuscular junction and cholinergic synapse, it terminates the electrophysiological response to acetylcholine.

True AChE occurs in multiple molecular forms, exhibiting different sedimentation coefficients on sucrose gradients (Massoulie and Bon, 1982). Further polymorphism can be distinguished according to the interactions of such forms with nonionic detergents (Grassi et al., 1982). In mammalian nervous tissue there exist forms of AChE that are secreted, cytoplasmic, membrane-associated (Massoulie and Bon, 1982), and bound to the basal lamina (Dreyfus et

al., 1983). Heavier ChE forms have been shown to be associated with a collagen-like tail (Anglister and Silman, 1978). The hydrophobic properties of some forms might be explained by the finding of hydrophobic domains (possibly a C-terminal peptide) (Rosenberry and Scoggin, 1984) or tight interactions with phospholipid moieties (i.e., phosphatidyl inositol) (Futerman et al., 1985, 1986). Although such subcellular segregation suggests the existence of different domains in various ChE classes, pharmacological and kinetic evidence show similar catalytic properties (Vigny et al., 1978). ψChE is similarly polymorphic, and many forms of ψChE can be said to be homologous to the true AChE forms (Silman et al., 1979).

The existence of a ChE form that differs from AChE was shown by Alles and Hawes (1940), who found that human serum and red blood cell enzymes are qualitatively different. Mendel and Rudney (1943) have shown that the serum ChE (later designated as ψChE) hydrolyzes butyrylcholine and propionylcholine faster than the cell-bound ChE. This class is the principle one found in the serum, and, until recently, no known biological role has been found for it. However, it has recently been shown that the sole ChE in the *Torpedo marmorata* heart muscle is ψChE (Toutant et al., 1985), although in higher vertebrates both classes exist in the heart. Humans having nonfuntional ψChE in the serum show no known symptoms of illness (Hodgkin et al., 1965) and exhibit normal muscular and neuronal activity. Altogether this may indicate that in ancestral species each enzyme had specific localized roles, and, with the evolution of the two enzyme classes, the acetylcholine binding site in serum ψChE became less physiologically important.

Tissue and Cell Type Specificity

Although ChEs are abundant in nerve and muscle, they may be found in other tissue and cell types, including the erythrocyte (Silver, 1974), adrenal medulla, ovarian follicles (Karnovsky and Roots, 1964), and megakaryocytes (Burstein et al., 1980). Cholinesterases have been reported in a number of embryonic tissues (Drews, 1975). In addition, considreable levels of ChEs were detected in various neoplastic tissues, such as ovarian carinomas (Drews, 1975) and brain tumors of glial and mesenchymal origin (Ord and Thompson, 1952; Razon et al., 1984). The various ChEs localized in different tissues and cell types vary in their sedimentation properties, hydrophobicity, and glycosylation patterns (Massoulie and Bon, 1982; Razon et al., 1984; Meflah et al., 1984; Zakut et al., 1985). The physiological role of the enzymes in tissues other than brain or muscle remains unknown.

Regulation of Expression of Cholinesterase Genes

To comprehend the regulation of biosynthesis of ChEs fully, it is important to understand the genetic make-up of the ChE genes. In the nematode earthworm *Caenorhabditis elegans* two forms of ChEs were shown to exhibit different kinetic properties, though both enzyme classes (A and B) appear to be AChEs (Johnson and Russell, 1983). Two distinct genes, *ace*-1 and *ace*-2 (probably structural genes) were shown to be responsible for the expression of these classes (Johnson et al., 1981, Culloti et al., 1981). Mutations occurring in either the A class alone or the B class alone do not change the phenotype. If, however, mutations occur in both genes, a new, uncoordinated phenotype is observed (Culloti et al., 1981). The double mutations do not have a lethal effect, in contrast with mutations in the *Drosophila ace* locus (the only AChE locus in the *Drosophila*; for a review, *see* Hall, 1982). This observation led to the discovery of a third class of AChE, the C class, whose kinetic properties differ greatly from those of A and B, suggesting

the existence of a different active site (Johnson et al., 1981; Johnson and Russell, 1983; Kolson and Russel, 1985a,b).

In humans, two codominant alleles at a single locus, which has not been genetically mapped, are responsible for phenotypic variants of erythrocyte AChE (Coates and Simpson, 1972). In contrast, two genetically independent loci, namely E_1 and E_2, have been genetically linked to alterations in serum ψChE (Silver, 1974; Whittaker, 1980). Genetic linkage studies suggest that the generally expressed E_1 locus is situated on the long arm of chromosome no. 3 (Arias et al., 1985), in linkage with the transferrin (TF) gene (Sparkes et al., 1984), mapped at 3q31–q26.1 (Huerre et al., 1984), the ceruloplasmin gene, and the transferrin receptor gene (TFRC) (reviewed by Kidd and Gusella, 1985). Mutations in this ChE gene result in the appearance of the "atypic" and "silent" forms of serum ψChE. The E_2 gene is expressed in 8% of the caucasian population and is responsible for the production of the common C_5 variant of serum ψChE. This form has been suggested to be the result of the E_2 gene protein binding to the E_1 ψChE protein, which causes a change in the mobility of ψChE. It also increases the activity of ψChE by up to 48% (Simpson, 1966). Genetic linkage studies suggested a possible linkage between the E_2 gene and the α-haptoglobin gene (Lovrien et al., 1978), which maps in a region on chromosome no. 16, distal to the fragile site 16q22 (Simmers et al., 1986).

Recent *in situ* hybridization to human chromosomes, using cDNA probes encoding human ChE, revealed that two structural ChE genes exist in the areas where the E_1 and E_2 genes have been shown to reside (Soreq et al., submitted). This suggests that the E_2 gene might, by itself, encode for a catalytic subunit of ChE, which causes increased activity either related to its own enzymatic properties or by affecting the E_1 subunit conformation, creating an increase in the activity of the composite enzyme molecule.

Oocyte microinjection experiments (Soreq et al., 1984) and in vitro translation of ChEmRNAs (Sikorav et al., 1985; Schumacher et al., 1986) suggest that the polymorphism of the ChE proteins extends to the level of mRNA. Crossed immunoelectrophoretic analysis of ChEmRNA products indicates that various ChEmRNA species encode the biosynthesis of electrophoretically distinct ChE polypeptides in a tissue-specific manner (Djiegielewska et al., 1986; Soreq et al., 1986). However, it is not yet clear what the differences are between the various species of ChEmRNAs in general, and, particularly, in humans. Furthermore, it is impossible as yet to link particular ChE forms and specific ChEmRNA species to defined ChE genes included in the human genetic repertoire.

Why Search for Human Cholinesterase Genes?

Cholinesterase as a Neurobiological Model for Termination of Neurotransmission

The presence of an acetycholine-hydrolyzing protein seems to be an essential requirement in cholinergic synapses, as is evident from the genetic experiments described above, as well as from the lethal effects of cholinesterase inhibitors (*see* following sections for details). In addition, the structure of ChE molecular forms is related to the nature of the synapse. For example, neuromuscular junctions are rich in collagen-tailed asymmetric AChE, whereas muscarinic brain synapses mainly contain slightly hydrophobic AChE tetramers (Massoulie and Bon, 1982; Zakut et al., 1985). This divergent distribution of ChEs may be relevant to the physiological properties of particular synapses.

Recently accumulated evidence reveals that the primary structures of several major proteins play important roles in regulating the pace and mode of function of particular types of syn-

apses. For example, molecular cloning of the nicotinic (Sumikawa et al., 1982; Claudio et al., 1983; Noda et al., 1983) and muscarinic (Kubo et al., 1986) cholinergic receptors has shown that these two proteins, both of which bind acetylcholine, have completely different primary sequences. Production and mutagenesis of the synthetic nicotinic acetylcholine receptor from cloned DNA in heterologous expression systems, such as microinjected *Xenopus* oocytes (Mishina et al., 1984), has been performed by genetic engineering techniques. These studies linked many of the electrophysiological properties characteristic of the nicotinic synapse to primary sequence epitopes in the various subunits of the nicotinic acetylcholine receptor molecule. In contrast with this advanced stage of study of receptor proteins, little has been done to investigate the precise involvement of ChEs, functioning as the turning-off signal, in regulating cholinergic transmission. The extensive similarities between polymorphic cholinesterases suggests considerable homologies also at the level of nucleotides. However, the various ChEmRNAs coding for the different ChEs present in fast- and slow-twitch muscles (Jedrzejczyk et al., 1984) may carry form-specific differences contributing to the electrophysiological properties of such synapses. This suggestion and parallel questions may today be approached by gene transfection and site-directed mutagenesis studies to determine the significance of the structure of the ChE protein in termination of neurotransmission.

Cholinesterases as a Model System for Regulation of Protein Polymorphism

As mentioned previously, ChEs are extremely polymorphic proteins. They exist as monomers, dimers, and tetramers. Some forms are globular, and others, which are assymetric, are associated with a collagen-like tail (for a review of the forms, *see* Massoulie and Bon, 1982). In an attempt to understand the underlying causes of such polymorphism of proteins, high-level production of the various ChE forms in cell culture would be useful. It seems that the information localizing the asymmetric AChE within the quail muscle membrane is acquired in the Golgi apparatus (Rotundo, 1984), although the mechanism of this polymorphic determination is unknown. The polymorphism of ChEs could, in principle, be a result of different genes encoding various primary structures. However, this seems less likely than other possibilities because *in situ* chromosomal mapping (Soreq et al., submitted) and genomic DNA blot analysis using an isolated human ChEcDNA (Prody et al., 1987) suggest that the structural cholinestease genes do not exist in as many copies as the number of existing molecular forms. Other possibilities include posttranscriptional (*see*, for example, Amara et al., 1985) and/or posttranslational processing events, such as glycosylation, formation of intramolecular and intersubunit S—S bonds, distinct assembly patterns, and binding of the collagen-like tail. Since all of these processes do take place during ChE biosynthesis, the ChEs are appropriate to be used as a model system to study the regulation of protein polymorphism.

Clinical Potential to the Study of Cholinesterase Genes

Prolonged Apnea Following Succinylcholine Administration. Succinylcholine, which acts as a competitive analog of acetylcholine, is often used in surgery as a short-term muscle relaxant. Since the drug is hydrolyzed by ψChE, its administration into individuals carrying genetically abnormal ψChE causes prolonged apnea (Thompson and Whittaker, 1966). The most common variant with this problem is the atypical variant E^{a1} for which 3–6% of the Caucasian population is heterozygous and about 0.05% is homozygous (Kalow and Gunn, 1959).

This enzyme hydrolyzes acetylcholine, but not succinylcholine (Whittaker, 1980). Another variant, E^s, which causes the complete absence of catalytically active serum ψChE in homozygotes, is also associated with this clinical problem (Hodgkin et al., 1965). This type of "silent" enzyme cannot hydrolyze any ChE substrate, nor can it bind organophosphate compounds (Lockridge and La Du, 1986). High frequency of atypical and silent ψChE genes was reported among Iraqi and Iranian Jews (11.3% for heterozygotes and 0.08% for homozygotes, respectively) (Szeinberg et al., 1972). This could explain the high frequency of reports of prolonged apnea following surgery in Israel. It is likely that ψChE could be used intravenously to rid the body of the succinylcholine in cases of prolonged apnea. For such use large amounts of the purified functional human ChE would be necessary. Since human ChEs cannot be purified in sufficient quantities, a cloned product would be necessary for such a purpose.

Organophosphate Poisoning. Complete inhibition of ChEs, for example, by the administration of organophosphorous (OP) poisons, is lethal (Koelle, 1972). This inhibition is achieved by formation of a stable stoichiometric (1:1) convalent conjugate with the active site serine (Aldridge and Reiner, 1972), and may be followed by a parallel competing reaction, termed "aging." This in turn transforms the inhibited ChE into a form that cannot be regenerated by the commonly used reactivators (Aldridge and Reiner, 1972), such as active-site-directed nucleophiles (e.g., quaternary oximes), which detach the phosphoryl moiety from the hydroxyl group of the active site serine (Hobbiger, 1963). The aging process is believed to involve dealkylation of the covalently bound OP group (Aldridge and Reiner, 1972) and renders therapy of intoxication by certain organophophates, such as sarin, diisopropylfluorophosphate (DFP), and Soman, exceedingly difficult (Loomis, 1963). The cloned ψChE protein could potentially be produced in large quantities and used prophylactically (e.g., in people who handle agricultural OP insecticides) and in place of oximes after intoxication. This prospect is strengthened by the recent observation that injection of highly purified fetal bovine serum AChE into mice protects them against organophosphate poisoning and by the positive effects of injected purified serum cholinesterase on patients suffering from alkyl phosphate poisoning (Klose and Gustensohn, 1976).

Open Neural Tube in Human Embryos. Neural tube defects in human embryos are biochemically characterized by secretion of a 10S tetrameric form of AChE into the amniotic fluid (Bonham and Atack, 1983). The detection of alterations in the level and isoform composition of AChE is currently carried out by sucrose gradient sedimentation of amniotic fluid, followed by enzymatic assays of ChE activity (Atack et al., 1983). In an alternative, less-quantitave method, amniotic fluid samples are separated by gel electrophoresis, and AChE activity is detected by specific staining (Brock and Bader, 1983). It would be particularly desirable to have a simple, specific procedure to determine the level of specific ChE forms in the amniotic fluid. This could be made possible by amino acid sequence determination, molecular cloning, and expression of the peptide regions specific to this cloned protein in sufficient quantities as to allow production of type-specific antibodies. Such antibodies could possibly be used in a radioimmunoassay to detect the existence of the enzyme isotype and thereby, the open neural tube itself.

Detection of ChE in Other Disorders. Modifications of human brain AChE have been reported in several neurological or genetic disorders, such as Alzheimer's disease (Coyle et al., 1983) and Down's syndrome (Yates et al., 1980). In the brains of patients with presenile dementia of the Alzheimer type (SDAT; about 5% of the

population above 65), the levels of AChE in cholinergic brain areas drop by about 50% (Atack et al., 1983). This particularly refers to soluble AChE tetramers (Fishman et al., 1986). However, it is not clear at what step through the pathway of gene expression this decrease is controlled. Probes for ChE genes and good antibodies for the various forms of AChE affected in such disorders will help to shed more light on this issue.

Technical Approaches and Methodology

The Use of Synthetic Oligodeoxynucleotides as Probes and/or Primers for Sequencing

Peptide sequencing information is often used to devise synthetic oligodeoxynucleotide probes, which may in turn be employed in various ways for hybridization with particular cDNA sequences (Wallace et al., 1981). Throughout the studies described in this review, oligodeoxynucleotides were prepared by phosphoramidite chemistry, first based on published peptide sequencing information (Lockridge, 1984) and then based on DNA sequences derived in our laboratory for ChEcDNA (Prody et al., 1986, 1987). The sequences and origin of part of these probes are detailed below, in Table 1. For example, the three probe mixtures designated OPSYN include all of the possibilities that could encode for the active site ChE hexapeptide (Lockridge, 1984). It was initially decided to limit the active site probe length to this hexapeptide since this was a consensus sequence identical in human ψChE and *Torpedo* AChE (Lockridge, 1984; McPhee-Quigley et al., 1985). Further publications revealed that the limitation of these probes' lengths was crucial to the success of the screening project since the human ψChE active site sequence that was published in 1986 (Lockridge and La Du) differed from that of 1984 (Lockridge, 1984) by 4 out of 29 amino acids.

Because of codon ambiguity, a total of 384 possibilities existed for the composition of the particular 17-long active site oligonucleotide. Therefore, it was essential to divide the OPSYN mixture of probes into three groups, to ensure that the specific activity of each individual sequence would be high enough to be detected in an autoradiogram. The ChEcDNA-derived oligonucleotide probes prepared later had the advantage of representing a single sequence each, which allowed a much higher specific activity in their use for hybridization experiments and eliminated the problem of false-positive hybrids.

To be used as probes in hybridization reactions, the synthetic oligodeoxynucleotides were end-labeled at their 5' end, with $\gamma^{32}P$, using the enzyme polynucleotide kinase (Prody et al., 1986). Hybridization and washing of hybrids using such probes was carried out with special caution because of the short nucleotide chains (Wallace et al., 1981); this also increases the probability of false-positive hybrids. To minimize such errors, 3*M* tetramethylammonium chloride [$(CH_3)_4NCl$] was employed to discriminate against short GC-rich hybrids in a base composition-independent manner (Wood et al., 1985).

Two cDNA clones were isolated from the fetal brain library by screening with the OPSYN II probe (Gnatt, 1986). These could represent members of the cholinesterase family or be false-positives. To distinguish between these possibilities and to characterize the oligodeoxynucleotide-cDNA hybridization, DNA sequencing was performed by the Sanger dideoxy-sequencing technique (Sanger, 1977), with the single-stranded vectors M13mp10, M13mp11, and M13mp19 (Messing, 1983). Similar techniques were employed for the characterization of other clones, including those

TABLE 1

Oligodeoxynucleotides[a]

Name	DNA sequences 5'—3'	Encoded polypeptide	Origin
Pseudo-C-term (–)80	A A C AGCCCNAC CA CT TC G G T	1864–1881	1
Opsyno (–) 151	A A C CT CT G G T A AC IGCIGCICCIGC TCICC AA IGI C G IGA T	742–770	1
Opsyn I (–)	A A C G C C A CC GC CT TC CC AA G A T G G T T	742–758	2
Opsyn II(–)	A A C G C C A CC GC GA TC CC AA G C T G G T T	None	2
Opsyn III(–)	A A C A C C A CC GC GA TC CC AA G T T G G T T	None	2
N-term (–) 200	TGTTGCAATTATGATGTCATCTTC	160–183	1
SP (+) 212	GGATTCTTAGCTTTGCC	613–629	1
TH2 (+) 214	GGATCAGAGATGTGGAA	391–407	1
SP (+) 216	TTGGAGAAAGTGCAGGA	743–759	1

(continued on next page)

TABLE 1 (*continued*)

Oligodeoxynucleotides[a]

Name	DNA sequence, 5'—>3'	Encoded polypeptide	Origin
N-term (+) 217	GATGACATCATAATTGC	163–179	1
SP (–) 220	AGCCCIACICAICTITC	1864–1880	1
SP (+) 226	GAIAGITGIGTIGGGCT	1864–1880	1
SP (+) 232	AAAGATGAAGGGACAGC	1126–1142	1
SP (+) 233	AATTATCAGTGCTCTGC	2187–2203	1
SP1424 (–) 250	AAAGGCATTATTTCCCC	1424–1440	1
SP1175 (–) 253	AGGAGCACCATAGACTA	1175–1191	1
SP1715 (–) 254	GCACGTAGTTTCGTCAT	1715–1732	1
SP2012 (–) 252	TCCTTCTGGCATTTGTG	2012–2028	1
308IV (–)	ACTTAACCAAGGCTGAA	Genomic	3
309IV (+)	TTCAGCCTTGGTTAAGT	Genomic	3
Nt-true 121	A A A C C C C GG TC TC GG CC TC G T G G T T T	None	4

[a]Oligonucleotides used as probes and/or primers for sequencing. Part of the synthetic oligodeoxynucleotides employed throughout this work as probes and/or primers are presented. (+): the mRNA (coding) strand; (–): the cDNA (noncoding) strand. Underlined nucleotides represent mistakes related to primary sequencing only from one strand, which were found after extensive sequencing. The polypeptides encoded by these oligodeoxynucleotides are numbered according to the nucleic acid residue numbers in the full-length ChEcDNA (*see* results). *Origin* refers to the source of information by which these sequences were determined. Probes labeled 1 were synthesized according to nucleotide sequences found in the isolated ChEcDNA. Those labeled 2 were deduced form peptide sequencing data published by Lockridge (1984). Those labeled 3 were determined from genomic nucleotide sequences, included in phages hybridizing with ChEcDNA probes, and those labeled 4 follow the sequence of Haas and Rosenberry for the N-terminal peptide of human erythrocyte AChE (1985).

coding for cholinesterase from fetal brain and liver (Prody et al., 1986,1987; *see* further sections for details).

Isolation and Characterization of Genomic DNA Fragments Hybridizing with Cholinesterase cDNA Probes

ChEcDNA was employed as a probe for the screening of genomic DNA libraries to detect and isolate DNA fragments from the human ChE gene(s). In addition, special attention was paid to the 5'-terminal region of such genes. In genomic DNA blots, *Eco*RI-digested DNA from all of the tissue origins checked to date gave seemingly identical patterns when hybridized with ChEcDNA probes (*see* further sections for details). When this hybridization was carried out using the 5'-terminal fragment of the human brain cDNA coding for ChE, one major band of approximately 4.7 ± 0.5 kilobase (kb) was visible and was, therefore, presumed to contain DNA fragments complementary to the 5'-terminal domain of the ChE gene. The DNA fragment containing the 5'-terminus is of particular interest to us since it should lead to the promotor region and possibly other sequences that might shed light on the regulation of ChE expression.

In order to isolate the 5'-terminal domain of the ChE gene, genomic DNA fragments cut with *Eco*RI were separated electrophoretically, and the electrophoresed DNA fragments in the size range of 3.7–4.7 kb were electroeluted and concentrated by chromatography on a DEAE cellulose minicolumn. The enriched DNA was ligated with dephosphorylated *Eco*RI-digested λgt10 DNA and packaged using the Gigapack Plus packaging kit (Stratagene). Phage colonies containing DNA sequences that hybridized with ^{32}P-labeled ChEcDNA were isolated and further characterized (Gnatt, 1986). The enrichment of the 5'-domain of the ChE gene and its cloning are presented schematically in Fig. 1.

Genomic double-stranded DNA fragments packaged in the λgt10 phages were subjected to direct sequencing by the dideoxy Sanger technique using the enzyme reverse transcriptase and synthetic oligodeoxynucleotide primers. Sequencing was carried out essentially as described by Zagursky et al., (1985), with minor modifications, or by use of the Klenow fragment of DNA polymerase. This was done using the M13 sequencing Kit (Amersham), with minor adjustments in the reaction conditions.

Current State of Experimental Observations

cDNA Screening and Sequence Analysis of Positively Hybridizing Phages

In an attempt to isolate ChEcDNA clones, four screens were performed on a fetal brain cDNA library using the OPSYN probes (Prody et al., 1986). The findings obtained in one of these screens are described below as an example for the procedures involved. The screen using the oligonucleotide probe mixture designated OPSYN II yielded two positively hybridizing phages, both of which did not form stable hybrids with the longer OPSYNO probe. (Prody et al., 1986; *see* Table 1 for the detailed sequences included in each of these probes.) In order to determine whether these cDNAs contained the active site hexapeptide and to examine the hybridization properties of the isolated cDNAs, OPSYN II hybridizing fragments derived from phages isolated from the OPSYN II screen were inserted into M13 single-stranded phages and their nucleotide sequence determined. Both sequences proved to contain open reading frames. The sequences of the OPSYN II hybridizing region of clone 7B and its translation in an open reading frame were as follows:

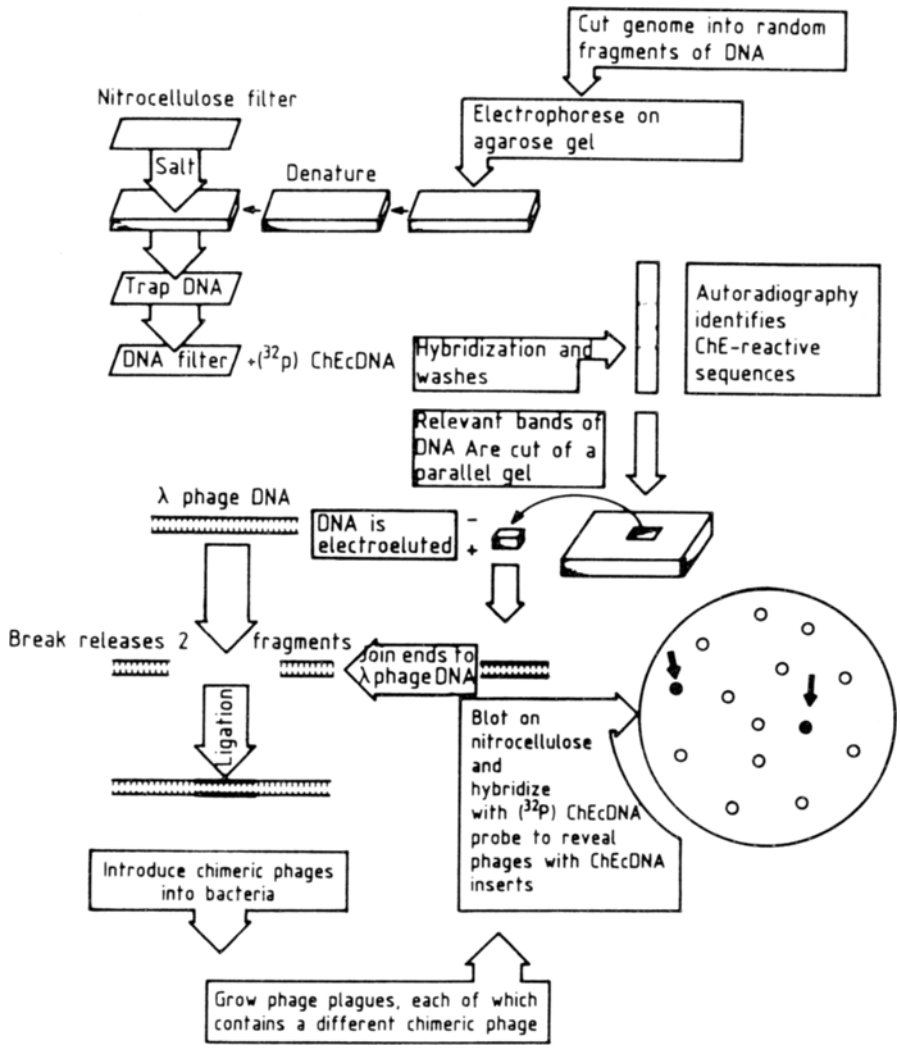


Fig. 1. Cloning and isolation of a 5'-terminal fragment from the human ChE gene. *See* text for details.

```
0
 AGG TGC GCA GGC CGA TTC TCC AAA
 Arg  Cys Ala  Gly  Arg  Phe Ser  Lys
25
GAA CTT GCA TTT AAG CTA AGG TGC
Glu  Lys  Ala  Phe  Lys  Leu  Arg  Cys
50
GCA  GGC CGA TTC TCC AAA GAA CTT
Ala  Gly  Arg  Phe Ser  Lys  Glu  Leu
75                  84
GCA TTT AAG CTT
Ala  Phe Lys  Leu
```

Thus, the mRNA from which this clone was reverse-transcribed codes for a protein that does not contain the ChE active site hexapeptide. Comparison of the nucleotide sequence in this clone with that of the OPSYN II probe revealed a base-pairing with a single mismatch, as shown in Table 2.

The sequence of the OPSYN II hybridizing region of clone 24 and its translation in an open reading frame were as follows:

0
AAG CAT AGC CAT CAG TTA GAA GTT
Lys His Ser His Gln Leu Glu Val
25
TTT TAT TTT TGG GGA TGT CGG CAG
Phe Tyr Phe Trp Gly Cys Arg Gln
50
GAG GAA TTT CCT TTA AAG GAG CAT
Glu Glu Phe Pro Leu Lys Glu His
75 96
ATA TAT ACG TCA GGA TTT GTC TTA
Ile Tyr Thr Ser Gly Phe Val Leu

In this case as well, the ChE-active site hexapeptide was not included in the protein encoded by the positively hybridizing cDNA, and the proposed base-pairing of the hybrids contained a single, unpaired base (Table 3).

Comparative Analysis of Cholinesterase cDNA of Various Genetic Origins

Primary Structure of Human Cholinesterase cDNA

In addition to the false-positive hybrids detailed above, the initial screening procedure described under Methods also resulted in the isolation of a single, true positive, in the form of a fetal brain cDNA clone, 765 nucleotides in length, designated FBChE12. This cDNA clone hybridized with both OPSYN and OPSYNO probes (Prody et al., 1986). The nucleotide sequence of FBChE12 that is complementary to probes OPSYN and OPSYNO corresponded

TABLE 2

Proposed Base Pairing Probe-DNA, Clone 7bxOPSYN II[a]

5'—>3' PROBE	C	C	G	G	C	C	G	A	T	T	C	T	C	C	A	A	A
3'—>5' DNA	G	T	C	C	G	G	C	T	A	A	G	A	G	G	T	T	T
		*															

[a]*See* text for details.

TABLE 3

Proposed Base Pairing Probe-DNA, Clone 24xOPSYN II[a]

5'—>3' PROBE	C	C	T	G	C	C	G	A	C	T	C	C	C	C	A	A	A
3'—>5' DNA	G	G	A	C	G	G	C	T	G	A	G	G	G	G	T	T	T
									T								
									*								

[a]*See* text for details.

exactly to the peptide sequence used to design these oligodeoxynucleotide probes (Fig. 2, amino acid residues encoded by nucleotides 742–759 and 742-771, respectively). FBChE12 was then used as a probe to screen the fetal brain and liver cDNA libraries. Four clones of 2.4 kb in length were isolated from the fetal liver library, and one of these, designated FL39, was further characterized in comparison with FBChE12. It was found that both clones contained an identical sequence of 693 nucleotides, with the 5'-end of the FL39 insert starting at nucleotide no. 73 of FBChE12 (Fig. 2), suggesting that both cDNAs were derived form similar mRNA transcripts (Prody et al., 1987). When the amino acids predicted from the FBChE12 and the FL39 sequence are aligned with the available peptide sequence of human ψChE, (Lockridge et al., 1987) the entire coding region for the mature enzyme is defined, starting at residue 1 (nucleotide 160), which corresponds to the N-terminal peptide, and ending at residue 574 (nucleotide 1881), which is the last amino acid residue in the C-terminal tryptic peptide of ψChE, as determined from amino acid sequencing (Lockridge and LaDu, 1986). This sequence also includes the active site tryptic peptide of human ψChE, which contains a serine residue that can be labled by diisopropylfluorophosphate (DFP) (Lockridge, 1984) (Fig. 2, circled). The polypeptide inferred by the FBChE12 and the FL39 sequences is identical to the ψChE polypeptide. In contrast, the amino acid composition of the Fl39-coded protein clearly differed from the parallel composition derived for erythrocyte AChE (Dutta-Chundhurry and Rosenberry, 1984). In addition, the N-terminus of the cholinesterase encoded by FBChE12 and Fl39 differs from the peptide reported for erythrocyte AChE (Haas and Rosenberry, 1985). Altogether, this proves that both FBChE12 and Fl39 code for ψChE. It should be noted that the amino acid sequence of nervous system ChEs has not been approached as yet, because of difficulties in purifying sufficient quantities of the active proteins. Thus, there is no indication at present regarding correlations between the isolated ChEcDNA and nervous system ChEs in humans.

The region upstream of the ψChE amino-terminal residue (nucleotides 88–147) in FBChE12 codes for 20 amino acids characteristic of leader peptides of membrane-associated and exported protein precursors (Heijne, 1985). The hydrophobic sequence in this region is rich in large, nonpolar amino acids. It is preceded by the tripeptide His–Ser–Lys, and terminates with Lys–Ser–His, both composed of polar amino acids. Further upstream, the cDNA sequence consists of a fully open reading frame without stop codons and includes a presumptive ribosome-binding site and an additional ATG triplet (Fig. 2), perhaps indicating that this mRNA is subjected to translational control at the level of initiation.

To examine further the molecular properties of the protein encoded by the human cholinesterase cDNA, we subjected it to hydrophobicity analysis according to Hopp and Woods (1981). The results of this analysis are presented in Fig. 3. The hydrophobicity pattern is consistent with a globular protein, in which a clearly hydrophobic region, that of the signal peptide, can be observed (Fig. 3).

Comparison of Human Cholinesterase cDNA to cDNAs of Other Proteins

The coding region in the cDNA and inferred amino acid sequence of the FL39 clone were compared to the parallel sequences recently published for cDNA clones coding for AChE from *Torpedo californica* electric organ (Schumacher et al., 1986) and from *Drosophila melanogaster* (Hall and Spierer, 1987). This analysis revealed considerable homologies between the corresponding parts of the cholinesterases from *T. californica, D. melanogaster,* and human,

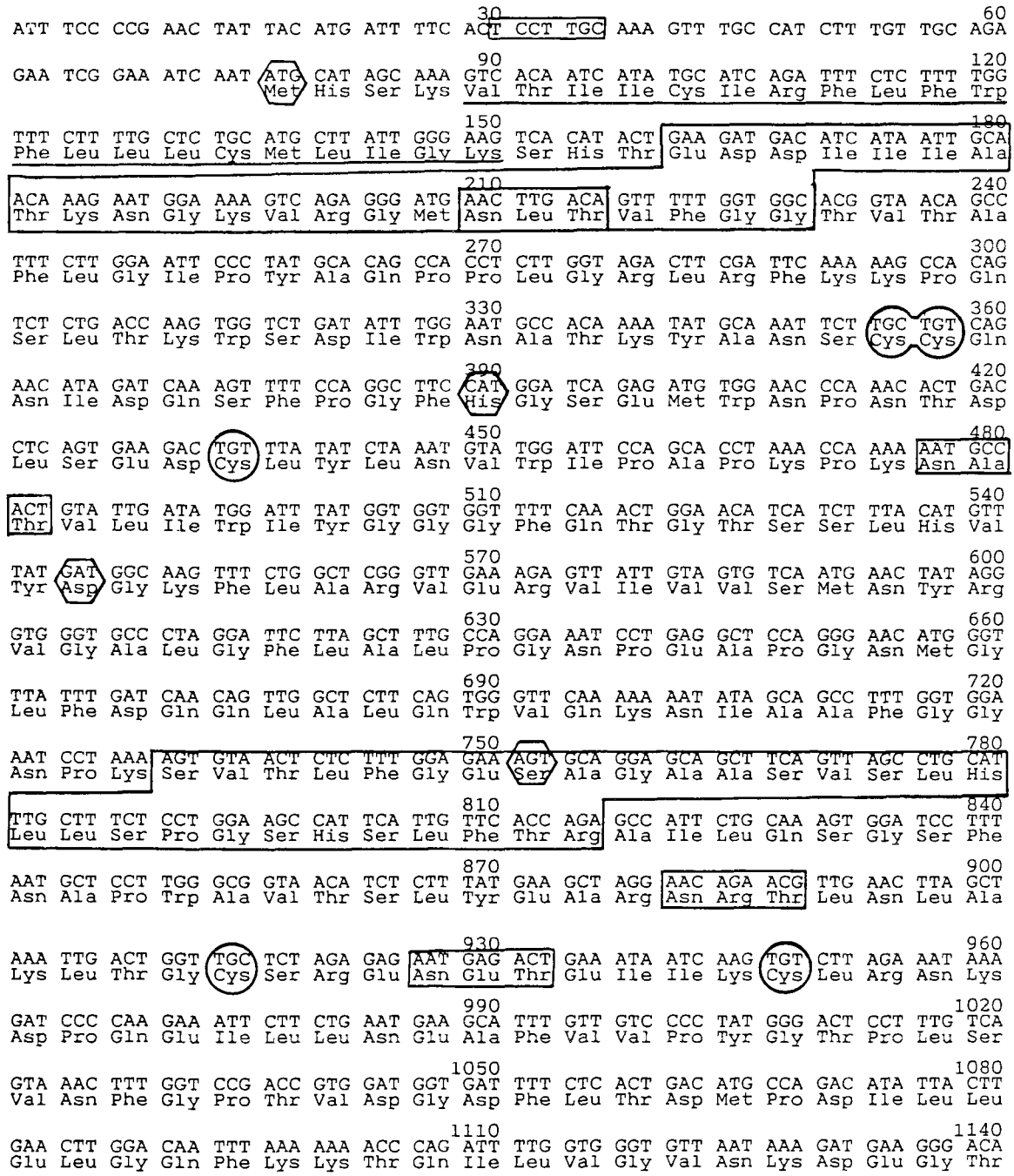
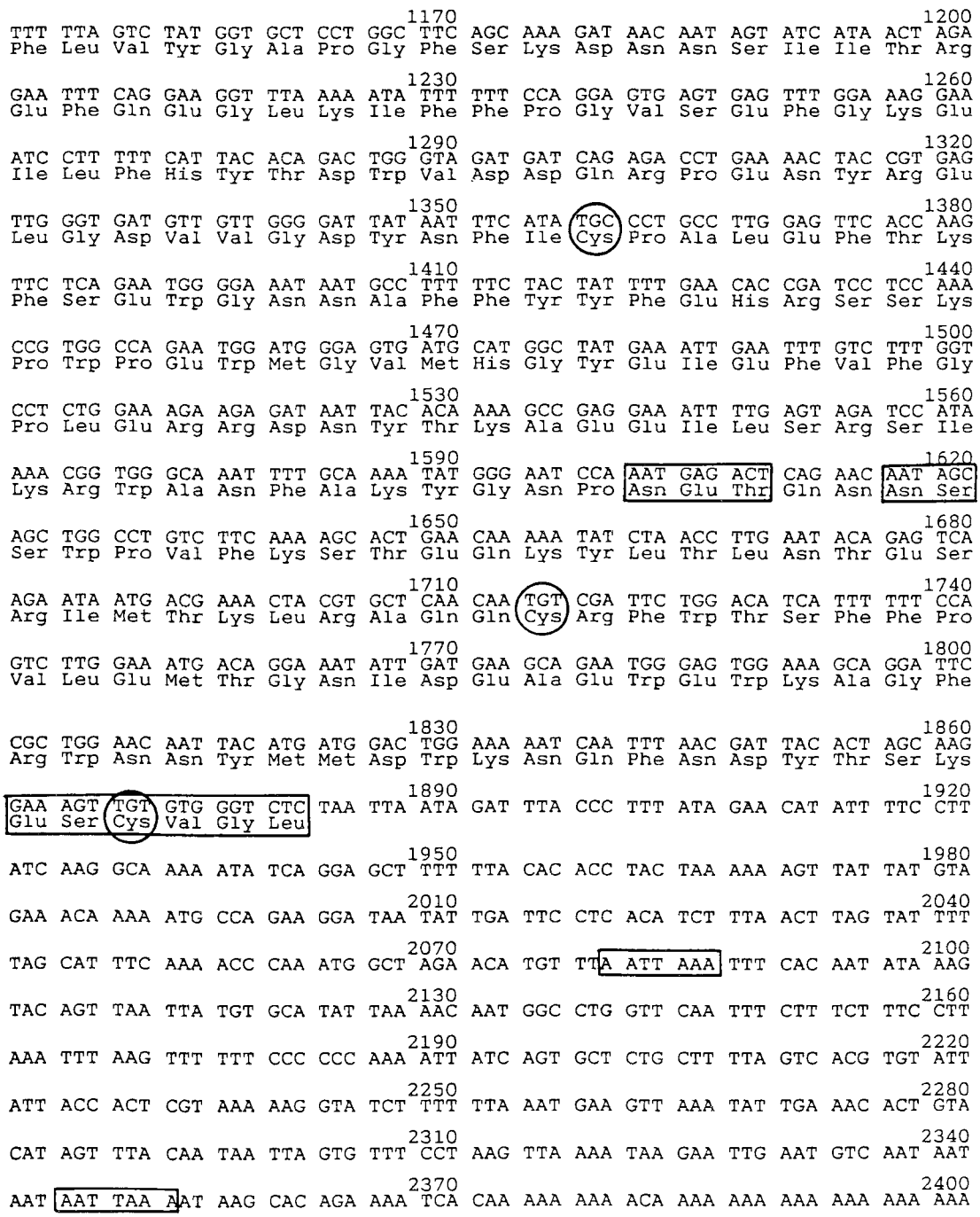


Fig. 2. Primary structure of the fetal human cholinesterase encoded by FBChE12 and FL39. The 2.4 kb composite nucleotide sequence of clones FBChE12 and FL39 (Prody et al., 1987) was translated into its encoded amino acid sequence. Nucleotides are numbered in the 5'–3' direction, and the predicted amino acids are shown below the corresponding nucleotide sequence. Boxed sequences indicate three amino acid sequences that were found to match peptices present in human ψChE, as shown by peptide sequencing (Lockridge, 1984). These are the N-terminal peptide (nucleotides 160–225), the active-site peptide (nucleotides 730–765, with a full circle indicating the active-site Ser residue no. 198), and the C-terminal peptide (nucleotides 1864–1881). The amino



(Fig. 2 *cont.*) acid sequence of the active-site peptide served as a basis for designing the oligodeoxynucleotide probes with which these cDNA clones were selected (Table 1). Also boxed are a presumptive ribosome binding site (nucleotides 30–36) and signal peptide (nucleotides 88–147), with three polar amino acid residues at both ends (Queen and Korn, 1980), as well as seven potential sites for N-linked glycosylation (starting at nucleotides 208, 475, 880, 925, 1180, 1600, and 1615), predicted by the sequence Asn–*X*–Thr/Ser, in which *X* represents any amino acid except proline (Bause, 1983). His 77 and Asp 129, which are the best candidates to be involved in the active site by comparision with other serine esterases (Dayhoff, 1978), are circled. The FL39 sequence also includes a long 3'-untranslated region, ending with a polyadenylation site and a poly(A)-tail.

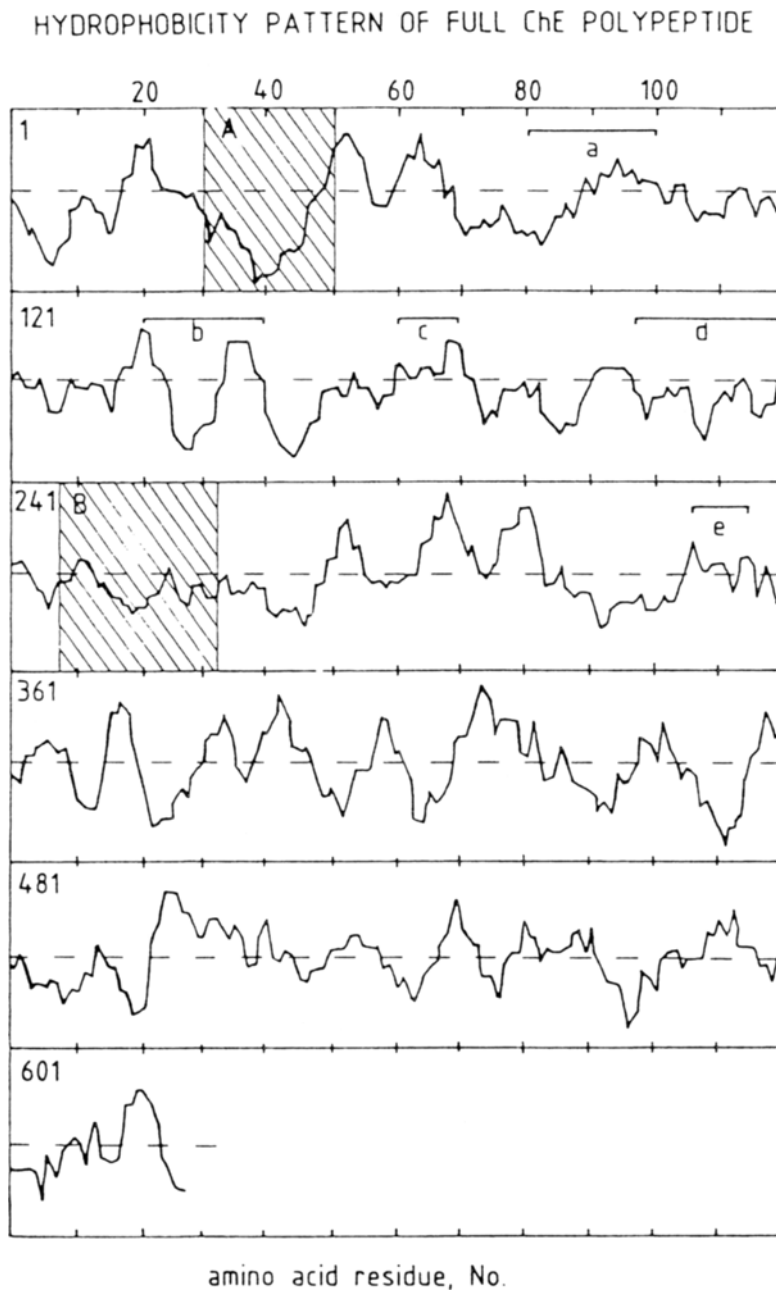


Fig. 3. Hydrophobicity pattern of the complete ChE protein. Represented is the prediction of the hydrophobic and hydrophilic regions of human ChE protein, using the algorithm of Hopp and Woods (1981). The baseline represents a hydrophilicity value of 0, increasing hydrophilicity is in the upward direction and increased hydrophobicity is in the downward direction. "A" represents the putative signal peptide, "B" the active site region, and a, b, c, d, and e represent regions that show high homology to the amino acid sequence of bovine thyroglobulin (*see* text for details). These are possible epitopes for autoimmune antibodies (*see* discussion).

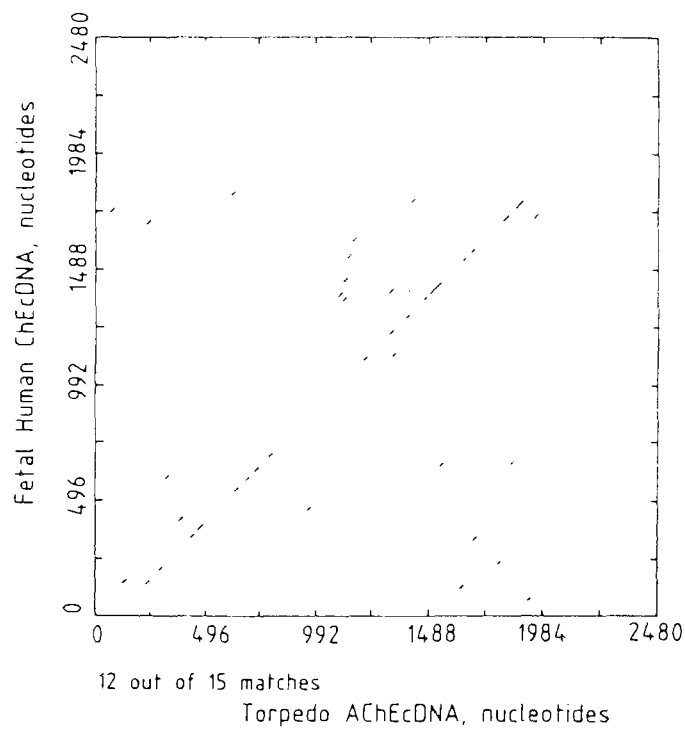


Fig. 4(a). Nucleotide and amino acid matrix homologies between the coding regions of cDNAs for cholinesterases from human, *Torpedo californica*, *Drosophila melanogaster*, and part of bovine thyroglobulin. Nucleotide (a) and amino acid sequence (b) data for FL 39 (prody et al., 1987) were compared with the parallel sequences published for a cDNA clone coding for AChE from *Torpedo* electric organ (Schumacher et al., 1986). Regions of homologies were searched for by the dot matrix approach (Maizel and Lenk, 1981) as modified by Unger and Sussman (personal communication). Match values that yielded clear homology regions and minimal background noise are presented (12 out of 15 matches for nucleotide sequence and 4 out of 5 matches for amino acid residues). Nucleotides are numbered in the 5' to 3' directions and amino acids in the N' to C' directions for the cDNAs in (a), (b), (c), and (d). The homologies start from around nucleotide 110 in FBChE12, a region that matches the beginning of the *Torpedo* cDNA clone. Note the presence of regions in which both the nucleotide sequence and the primary structure of amino acids are homologous (*see*, for example, nucleotides 1450–1500 in the human cholinestrase cDNA), as compared with regions with amino acid similarities but no nucleotide match (such as nucleotides 1000–1050 in the human cDNA) and with the short domains where the simlarities in both nucleotide and amino acid sequence were lower than the match frequency of choice (for example, nucleotides 310–340 in the human cDNA). Parallel analyses are presented for cDNAs coding for *Drosophila* AChE (Hall and Spierer, 1987) and for bovine thyroglobulin (Mercken et al., 1985).

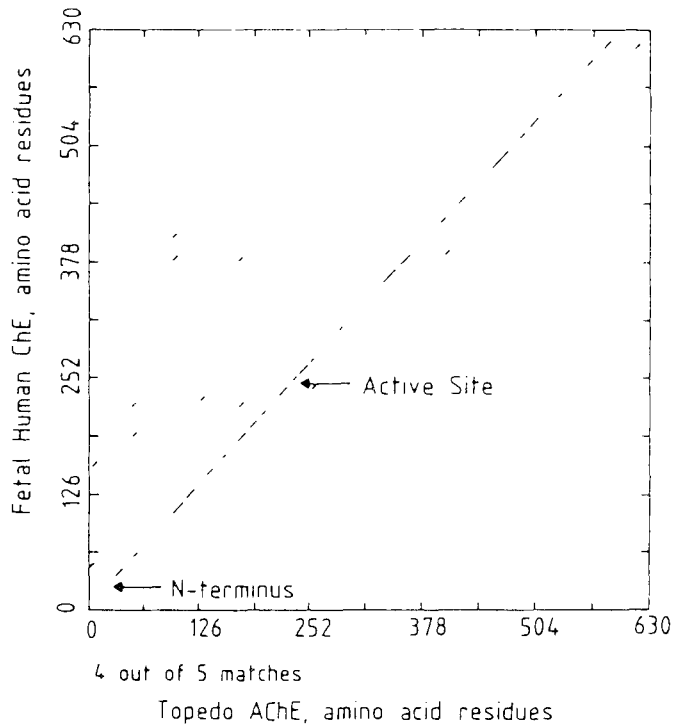


Figure 4b

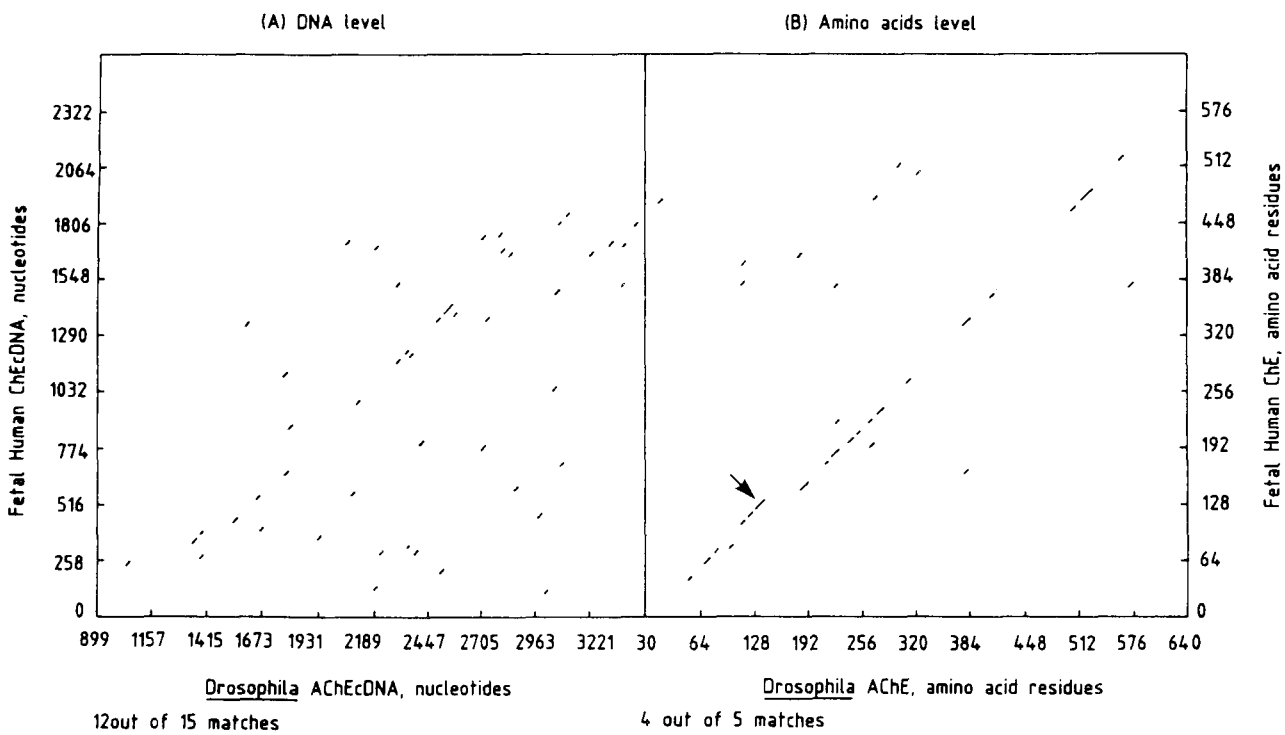


Figure 4c

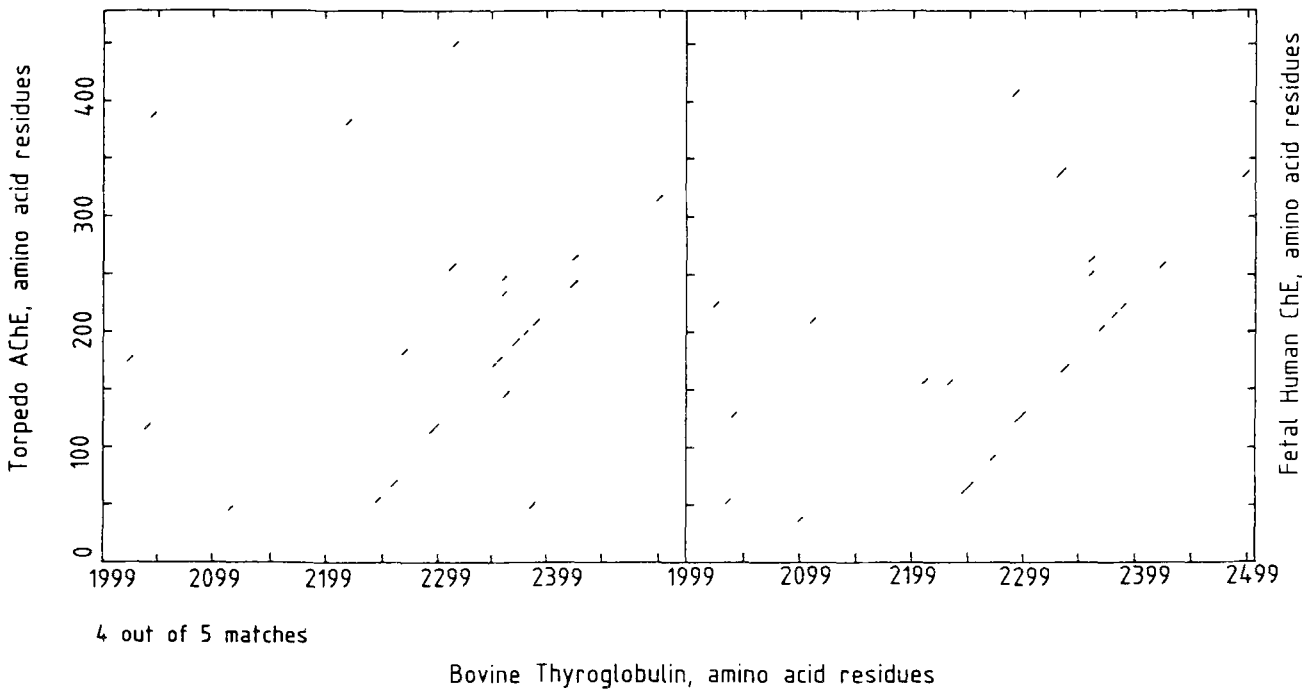


Figure 4d.

strongly suggesting that they have a common ancestral origin. A higher level of conservation was consistently found at the amino acid level than at the DNA level. Significant homology was also observed with the DNA and the amino acid sequence of bovine thyroglobulin (Merken et al., 1985). These homologies are presented in Fig. 4(a–d) as computer-derived matrices. The intrinsic differences between the cDNAs encoding these proteins have also been approached by analyzing their nucleotide composition and distribution of codons and dinucleotides, as shown in Table 4. In contrast with the pronounced species-specific differences in codon usage and nucleotide frequencies, the amino acid composition of the various cholinesterases remained surprisingly conserved, as displayed in Table 5.

Preliminary Characterization of the Structural Human Cholinesterase Genes

To identify the genomic DNA fragments encoding various regions of the ChE protein, segments of the molecularly cloned, full-length cholinesterase cDNA cut by the enzyme *Eco*RI were employed as probes of DNA blot hybridization of genomic DNA from postnatal and fetal brain, as well as from primary meningioma tumors. This analysis revealed a 10-kb DNA band that hybridized with the 3'-end of the cDNA, a 2.3-kb DNA fragment that appears to carry most of its central part, and a 4.3-kb fragment that codes for at least part of the 5'-terminal region of this cDNA, as shown in Fig. 5.

TABLE 4
Nucleotide Compositions and Frequencies in Various Cholinesterase cDNAs*

I_a *Human ChEcDNA mature protein encoding sequence*

The sequence contains 1722 nucleotides

Nucleotide	No.	%
A	535	(31.1)
C	308	(17.9)
G	388	(22.5)
T	491	(28.5)
A + T	1026	(59.6)
C + G	696	(40.4)

The dinucleotide frequency is

AA	195	(11.3)	CA	118	(6.9)	GA	139	(8.1)	TA	83	(4.8)
AC	84	(4.9)	CC	68	(4.0)	GC	60	(3.5)	TC	96	(5.6)
AG	117	(6.8)	CG	17	(1.0)	GG	105	(6.1)	TG	148	(8.6)
AT	139	(8.1)	CT	104	(6.0)	GT	84	(4.9)	TT	89	(9.5)

I_b *Distribution of codons in the open reading frame*

TTT	Phe	27(4.7)	TCT	Ser	7(1.2)	TAT	Tyr	14(2.4)	TGT	Cys	5(0.9)
TTC	Phe	12(2.1)	TCC	Ser	5(0.9)	TAC	Tyr	6(1.0)	TGC	Cys	3(0.5)
TTA	Leu	9(1.6)	TCA	Ser	9(1.6)	TAA	End	0(0.0)	TGA	End	0(0.0)
TTG	Leu	15(2.6)	TCG	Ser	0(0.0)	TAG	End	0(0.0)	TGG	Trp	18(3.1)
CTT	Leu	12(2.1)	CCT	Pro	12(2.1)	CAT	His	8(1.4)	CGT	Arg	2(0.3)
CTC	Leu	4(0.7)	CCC	Pro	3(0.5)	CAC	His	1(0.2)	CGC	Arg	1(0.2)
CTA	Leu	4(0.7)	CCA	Pro	13(2.3)	CAA	Gln	10(1.7)	CGA	Arg	3(0.5)
CTG	Leu	6(1.0)	CCG	Pro	2(0.3)	CAG	Gln	9(1.6)	CGG	Arg	2(0.3)
ATT	Ile	12(2.1)	ACT	Thr	13(2.3)	AAT	Asn	25(4.4)	AGT	Ser	9(1.6)
ATC	Ile	5(0.9)	ACC	Thr	6(1.0)	AAC	Asn	13(2.3)	AGC	Ser	7(1.2)
TAT	Ile	11(1.9)	ACA	Thr	15(2.6)	AAA	Lys	25(4.4)	AGA	Arg	13(2.3)
ATG	Met	11(1.9)	ACG	Thr	3(0.5)	AAG	Lys	8(1.4)	AGG	Arg	3(0.5)
GTT	Val	10(1.7)	GCT	Ala	11(2.3)	GAT	Asp	17(3.0)	GGT	Gly	17(3.0)
GTC	Val	6(1.0)	GCC	Ala	10(1.7)	GAC	Asp	7(1.2)	GGC	Gly	5(0.9)
GTA	Val	8(1.4)	GCA	Ala	12(2.1)	GAA	Glu	26(4.5)	GGA	Gly	18(3.1)
GTG	Val	8(1.4)	GCG	Ala	1(0.2)	GAG	Glu	11(1.9)	GGG	Gly	6(1.0)

(Table continued next page)

TABLE 4 (*continued*)

II_a *Torpedo californica ChEcDNA mature protein-encoding sequence*

The sequence contains 1725 nucleotides:

Nucleotide	*No.*	*%*
A	417	(24.2)
C	460	(26.7)
G	481	(27.9)
T	367	(21.3)
A + T	784	(45.4)
C + G	941	(54.6)

The dinucleotide frequency is

AA	101	(5.9)	CA	133	(7.7)	GA	148	(8.6)	TA	35	(2.0)
AC	121	(7.0)	CC	120	(7.0)	GC	96	(5.6)	TC	123	(7.1)
AG	130	(7.5)	CG	84	(4.9)	GG	146	(8.5)	TG	120	(7.0)
AT	65	(3.8)	CT	123	(7.1)	GT	90	(5.2)	TT	89	(5.2)

II_b *Distribution of codons in the open reading frame*

TTT	Phe	11(1.9)	TCT	Ser	10 (1.7)	TAT	Tyr	2 (0.3)	TGT	Cys	4 (0.7)
TTC	Phe	25(4.3)	TCC	Ser	7 (1.2)	TAC	Tyr	17 (3.0)	TGC	Cys	4 (0.7)
TTA	Leu	2(0.3)	TCA	Ser	5 (0.9)	TAA	End	0 (0.0)	TGA	End	0 (0.0)
TTG	Leu	8(1.4)	TCG	Ser	5 (0.9)	TAG	End	0 (0.0)	TGG	Trp	17 (3.0)
CTT	Leu	4(0.7)	CCT	Pro	7 (1.2)	CAT	His	2 (0.3)	CGT	Arg	0 (0.0)
CTC	Leu	18(3.1)	CCC	Pro	11 (1.9)	CAC	His	16 (2.8)	CGC	Arg	3 (0.5)
CTA	Leu	2(0.3)	CCA	Pro	5 (0.9)	CAA	Gln	1 (0.2)	CGA	Arg	4 (0.7)
CTG	Leu	17(3.0)	CCG	Pro	8 (1.4)	GAG	Gln	17 (3.0)	CGG	ARg	5 (0.9)
ATT	Ile	6(1.0)	ACT	Thr	3 (0.5)	AAT	Asn	8 (1.4)	AGT	Ser	5 (0.9)
ATC	Ile	11(1.9)	ACC	Thr	9 (1.6)	AAC	Asn	30 (5.2)	AGC	Ser	15 (2.6)
ATA	Ile	3(0.5)	ACA	Thr	7 (1.2)	AAA	Lys	7 (1.2)	AGA	Arg	6 (1.0)
ATG	Met	18(3.1)	ACG	Thr	4 (0.7)	AAG	Lys	19 (3.3)	AGG	Arg	9 (1.6)
GTT	Val	7(1.2)	GCT	Ala	6 (1.0)	GAT	Asp	5 (0.9)	GGT	Gly	5 (0.9)
GTC	Val	20(3.5)	GCC	Ala	9 (1.6)	GAC	Asp	23 (4.0)	GGC	Gly	16 (2.8)
GTA	Val	1(0.2)	GCA	Ala	4 (0.7)	GAA	Glu	11 (1.9)	GGA	Gly	12 (2.1)
GTG	Val	12(2.1)	GCG	Ala	6 (1.0)	GAG	Glu	29 (5.0)	GGG	Gly	12 (2.1)

(Table continued next page)

TABLE 4 (*continued*)

III$_a$ *Drosophila melanogaster ChEcDNA protein encoding sequence*

The sequence contains 1945 nucleotides:

Nucleotide	No.	%
A	406	(20.9)
C	576	(29.6)
G	567	(29.2)
T	396	(20.4)
A + T	802	(41.2)
C + G	1143	(58.8)

The dinucleotide frequency is

AA	89	(4.6)	CA	133	(6.8)	GA	136	(7.0)	TA	48	(2.5)
AC	109	(5.6)	CC	160	(8.2)	GC	176	(9.1)	TC	131	(6.7)
AG	94	(4.8)	CG	153	(7.9)	GG	174	(9.0)	TG	145	(7.5)
AT	114	(5.9)	CT	130	(6.7)	GT	81	(4.2)	TT	71	(3.7)

II$_b$ *Distribution of codons in the open reading frame*

TTT	Phe	6 (0.9)	TCT	Ser	1 (0.2)	TAT	Tyr	7 (1.1)	TGT	Cys	1 (0.2)
TTC	Phe	21 (3.2)	TCC	Ser	19 (2.9)	TAC	Tyr	19 (2.9)	TGC	Cys	10 (1.5)
TTA	Leu	1 (0.2)	TCA	Ser	2 (0.3)	TAA	End	0 (0.0)	TGA	End	0 (0.0)
TTG	Leu	8 (1.2)	TCG	Ser	11 (1.7)	TAG	End	0 (0.0)	TGG	Trp	15 (2.3)
CTT	Leu	2 (0.3)	CCT	Pro	5 (0.8)	CAT	His	4 (0.6)	CGT	Arg	4 (0.6)
CTC	Leu	9 (1.4)	CCC	Pro	19 (2.9)	CAC	His	11 (1.7)	CGC	Arg	11 (1.7)
CTA	Leu	2 (0.3)	CCA	Pro	3 (0.5)	CAA	Gln	3 (0.5)	CGA	Arg	5 (0.8)
CTG	Leu	32 (4.9)	CCG	Pro	13 (2.0)	CAG	Gln	20 (3.1)	CGG	Arg	4 (0.6)
ATT	Ile	7 (1.1)	ACT	Thr	4 (0.6)	AAT	Asn	15 (2.3)	AGT	Ser	5 (0.8)
ATC	Ile	23 (3.5)	ACC	Thr	17 (2.6)	AAC	Asn	18 (2.8)	AGC	Ser	6 (0.9)
ATA	Ile	4 (0.6)	ACA	Thr	4 (0.6)	AAA	Lys	5 (0.8)	AGA	Arg	1 (0.2)
ATG	Met	19 (2.9)	ACG	Thr	9 (1.4)	AAG	Lys	17 (2.6)	AGG	Arg	7 (1.1)
GTT	Val	2 (0.3)	GCT	Ala	11 (1.7)	GAT	Asp	19 (2.9)	GGT	Gly	8 (1.2)
GTC	Val	14 (2.2)	GCC	Ala	24 (3.7)	GAC	Asp	14 (2.2)	GGC	Gly	30 (4.6)
GTA	Val	2 (0.3)	GCA	Ala	6 (0.9)	GAA	Glu	7 (1.1)	GGA	Gly	14 (2.2)
GTG	Val	21 (3.2)	GCG	Ala	16 (2.5)	GAG	Glu	27 (4.2)	GGG	Gly	4 (0.6)

[a]Data were derived by computerized analysis of the specific cDNA sequences coding for human, *Torpedo*, and *Drosophila* ChEs (*see* text for details regarding the sources of these sequences).

TABLE 5

Amino Acid Composition and Primary Protein Properties in Various Cholinesterases[a]

I *Amino acid composition of the mature human ChE*

Amino acids	No. of residues	%	Amino acids	No. of residues	%
Ala	34	5.9	Leu	51	8.7
Arg	24	4.2	Lys	26	5.7
Asn	38	6.6	Met	18	1.9
Asp	24	4.2	Phe	36	6.8
Cys	8	1.4	Pro	31	5.2
Gln	19	3.3	Ser	23	6.4
Glu	37	6.4	Thr	17	6.4
Gly	46	8.0	Trp	19	3.1
His	9	1.6	Val	40	3.5
Ile	28	4.9			5.6
Acidic	(Asp + Glu)			68	10.6
Basic	(Arg + Lys)			53	9.9
Aromatic	(Phe + Trp + Tyr)			72	13.4
Hydrophobic	(Aromatic + Ile + Leu + Met + Val)			201	34.5

Mol wt = 65,087. Total amino acids = 574

II *Amino acid composition of the mature* T. californica *AChE*

Amino acids	No. of residues	%	Amino acids	No. of residues	%
Ala	25	4.3	Leu	54	8.9
Arg	27	4.7	Lys	22	4.5
Asn	38	6.6	Met	19	3.1
Asp	28	4.9	Phe	27	6.3
Cys	8	1.4	Pro	40	5.4
Gln	18	3.1	Ser	44	8.2
Glu	40	7.0	Thr	34	4.0
Gly	45	7.8	Trp	15	3.0
His	18	3.1	Tyr	26	3.3
Ile	20	3.5	Val	39	7.0
Acidic	(Asp + Glu)			67	11.8
Basic	(Arg + Lys)			54	9.2
Aromatic	(Phe + Trp + Tyr)			68	12.5
Hydrophobic	(Aromatic + Ile + Leu + Met + Val)			214	35.0

Mol wt = 65,596. Total amino acids = 575

(Table continued next page)

TABLE 5 (*continued*)

III *Amino acid composition of the* D. melanogaster *AChE*

Amino acids	No. of residue	%	Amino acids	No. of residues	%
Ala	57	8.8	Leu	54	8.3
Arg	32	4.9	Lys	22	3.4
Asn	33	5.1	Met	19	2.9
Asp	33	5.1	Phe	27	4.2
Cys	11	1.7	Pro	40	6.2
Gln	23	3.5	Ser	44	6.8
Glu	34	5.2	Thr	34	5.2
Gly	56	8.6	Trp	15	2.3
His	15	2.3	Tyr	26	4.0
Ile	34	5.2	Val	39	6.0
				67	10.3
Acidic	(Asp + Glu)			54	8.3
Basic	(Arg + Lys)			68	10.5
Aromatic	(Phe + Trp + Tyr)			214	33.0
Hydrophobic	(Aromatic + Ile + Leu + Met + Val)				

Mol wt = 71,662. Total amino acids = 648

[a]The amino acid sequences of each of the ChEs detailed under this Table were deduced from the published cDNA data and compared by computerized analysis. Note the high similarities in general amino acid composition as compared with distinct differences in nucleotide composition and dinucleotide frequencies and codon usage (Table 4).

from postnatal and fetal brain, as well as from primary meningioma tumors. This analysis revealed a 10-kb DNA band that hybridized with the 3'-end of the cDNA, a 2.3-kb DNA fragment that appears to carry most of its central part, and a 4.3-kb fragment that codes for at least part of the 5'-terminal region of this cDNA, as shown in Fig. 5.

To obtain the 5'-region of the ChE gene with its flanking region, DNA blot hybridization was performed with enzymatically restricted human genomic DNA and a ^{32}P-labeled fragment derived from the 5'-terminus of the cloned ChEcDNA. A 4.7 ± 0.5-kb long cDNA fragment was detected and enriched 70-fold by preparative restriction, gel electrophoresis and electroelution (*see* previous sections for details of the enrichment protocol). The enriched DNA fraction was ligated with λgt_{10} DNA and packaged in the lambda coat. The resultant genomic library was screened using ChEcDNA probes and positive phages containing genomic ChEDNA fragments were isolated. The DNA-sequencing analysis of one these phages revealed that the 4.2-kb DNA fragment inserted into the λgt_{10} DNA included in its 3'-terminus a stretch of 180 nucleotides similar to the sequence that has previously been found in the 5'-terminal part of ChEcDNA.

Our findings at present indicate the existence of at least one intervening sequence in the 5'-region of the isolated human cholinesterase gene included in a 4.2-kb DNA fragment that ends with an *Eco*RI site. A large portion of the coding sequence then appears to be included in a 2.5-kb fragment. This fragment contains an exon at least 585 nucleotides in length, which starts with an *Eco*RI site and ends with a *Bam*HI

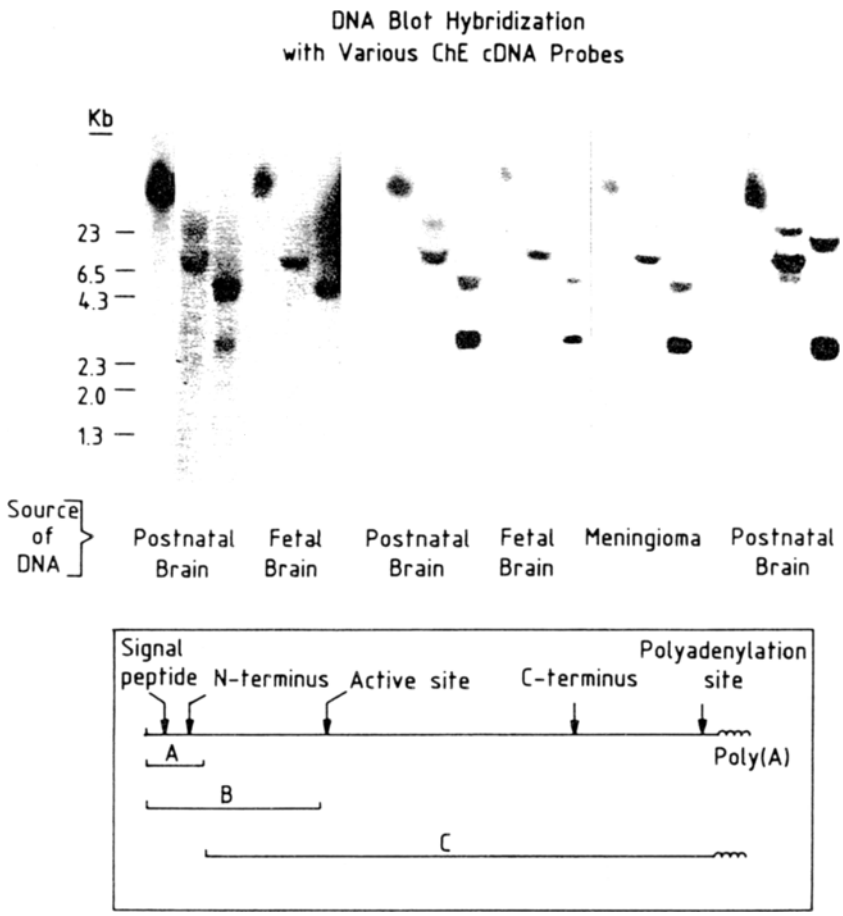


Fig. 5. Inset: Schematic drawing of the DNA probes employed. The full-length human ChEmRNA (upper line) contains 2500 nucleotides, is expressed in both fetal brain and liver (Prody et al., 1986, 1987), and includes sequences coding for a signal peptide and the N-terminal, active site, and C-terminal peptides found in human serum ψChE (Lockridge, 1984). Probe A represents a 250-nucleotides long *Eco*RI fragment of a cDNA insert isolated from a λgt10 library of fetal brain origin and spanning from the 5'-end region of the ChEcDNA through the N-terminal peptide that appears in the ψChE encoding sequences. Probe B represents the original FBChE12 cDNA clone isolated by use of oligodeoxynucleotides (Prody et al., 1986). This clone contains a 765-nucleotides long insert from which probe A was derived, begining at the same point as probe A, but reaching the active site region of the human seum ψChE sequences. Probe C represents a 2230-long cDNA fragment isolated form λt10 library of fetal liver origin. It contains a stretch of 585 nucleotides overlapping with probe B, and it spans from an *Eco*RI restriction site within probe B through the active site and C-terminal regions of ψChE as well as the 3'-untranslated region and polyadenylation site of ChEcDNA.

A, B, and C: ChEcDNA blot hybridization reveals various fragments of genomic DNA. Twenty micrograms of genomic DNA from human fetal brain, postnatal brain, or a primary meningioma tumor were restricted with the following enzymes: *Eco*RI (E), *Msp*I (M), and *Hpa*II (H), separated by agarose gel electrophoresis, and hybridized with a human ChEcDNA probe (B). The autoradiogram was washed in denaturing solution to "peel off" the probe and was then divided into two parts, which were then hybridized with probes A and C. Note the enhanced appearance of a stron 4.3-kb *Eco*RI-cut band with probe A but not c, whereas a 2.3 kb *Eco*Ri-cut band is only visible with probes B and C. A large (10-kb) *Eco*RI-cut band can only be detected with probe C.

site, both of which are inherent to the ChEcDNA sequence (Fig. 2). Finally, the C-terminal part of the cDNA seems to hybridize with another (10-kb) fragment of DNA. Figure 6 represents these findings in a schematic manner. It should be stated that molecular cloning and detailed sequence analysis of these DNA fragments will be required to establish whether they are indeed parts of the actively expressed ψChE gene or whether they are derived from genes coding for other cholinesterases or related proteins or from ψChE pseudogenes.

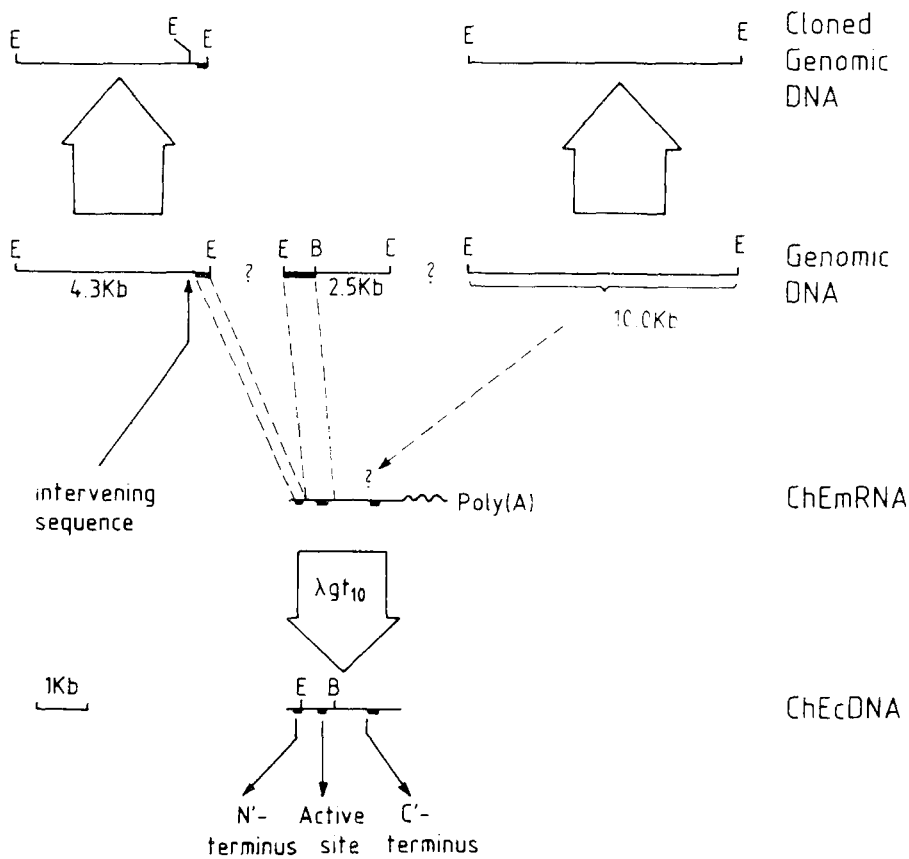


Fig. 6. Schematic drawing of genomic clones so far isolated. *See* text for details

Discussion

Oligonucleotide Stability in Screening of Libraries

Nonperfect Hybrids: Stability in the Presence of Tetramethylammonium Salts

Pools of oligodeoxynucleotides, ranging in length from 14–20 bases, are commonly used to screen libraries of cloned cDNA to isolate desired DNA sequences (Wallace et al., 1981). The oligonucleotide mixtures contain probes of equal size that represent all of the DNA posibilities encoding for a specific polypeptide. Because of the ambiguity of the codon, there might be many such possibilities for a given peptide, which requires the preparation of probe mixtures of high complexity. Such high complexity probes often contain sequences with varying G–C content, creating a problem in determining the stringency of hybridization and wash, since the stability of G–C pairs is higher than that of

A–T pairs (Wood et al., 1985). To overcome this difficulty, a selective method has been developed in which tetramethyl ammonium (TMA) salts are employed to allow the stringency of hybridization to be controlled as a sole function of probe length (Wood et al., 1985). The effect of TMA was suggested to cancel the stabilizing effect of G–C pairs on the dissociation temperature (t_d) of DNA hybrids. Thus, the length factor of the probe should remain as a single determining element (Melchior and Hippel, 1972). Since all of the unique probes in the pool are of identical size, the t_d of true positives should be identical under TMA washes and higher than the t_d of all the nonexact matches, regardless of their G–C content. This should exclude nonexact false positives, which remain stable under standard washing conditions.

Our findings indicate that the use of TMA is efficacious but not perfect, since some of the mismatched hybrids remain stable even following TMA washes (Prody et al., 1986; Gnatt, 1986). If it is simply assumed that the TMA washes cancel the G–C effect, both of the OPSYN II positive clones described above should not have been stable under the TMA washes used in the third screen. However, clone 7b was stable and included a single CxT mismatch, whereas clone 24 represented a perfect homology with an additional unpaired thymidine base. Futhermore, an additional OPSYN I positive clone proved to be stable to TMA washes, although it included three successive mismatches (Prody and Soreq, unpublished data).

Studies of others confirm the ability of oligodeoxynucleotides to form a double helix, with an additional unpaired adenine base stacked into the duplex (Patel et al., 1982), or an additional thymidine base forming a single base loop (Evans and Morgan, 1986). Mismatched base pairs in oligodeoxynucleotide duplexes have also been shown to exist and might play a role in mutagenesis (Patel et al., 1982; Brown et al., 1985; Dohet et al., 1985 and Hunter et al., 1986). However, in all of the reported cases of mismatched or additional unmatched base pairs in duplexes, the stability of hybrids was found to be lower than in perfectly matched duplexes. In contrast, clones 7b and 24 displayed high stability hybridizations with the mismatched oligodeoxynucleotide probes. In light of the present studies, we would like to suggest that sequence-specific elements also contribute to the stability of probe-DNA hybrids in the presence of TMA. These could possibly be structural effects related to interactions of the probe–DNA hybrids with the TMA itself. The stabilizing factor contributed by the specific sequences would thus be apparent in the presence of TMA, with each TMA–DNA complex having unique physicochemical properties. It must be noted that 78.4% of the positively hybridizing mismatched clones behaved as expected and were not TMA stable (Prody et al., 1986). The TMA treatment therefore appears to be an effective method for screening with large pools of oligonucleotide probes, but it does not exclude *all* of the mismatched oligonucleotides.

Use of Base Substitution in Designing of Oligonucleotide Probes

In place of mixtures of oligonucleotides that cover all of the alternatives allowed by codon amibiguity, it is possible to use nucleic acid analogs when polypeptide sequences are used to design probes for screening of cDNA libraries. In this approach, the nucleotide analogs may possibly substitute for more than one of the Watson-Crick bases, thus allowing the limitation of probe complexity in highly ambiguous mixtures. Deoxyinosine (dI) is the preferred analog for such substitutions and was successfully used in places of codon ambiguity to isolate a human cholecystokinin gene (Takahshi et al., 1985) and many others. Analysis of the thermal stability of oligodeoxyribonucleotide duplexes containig dI were examined (Martin et

al., 1985). The results suggest that the use of dI is useful at A/C, G/T, three- and fourfold ambiguities. It seems, though, that deoxyguanosine (dG) might be preferable in cases of T/C ambiguities, since dI is less stabilizing on the average. We can conclude from our results using the dI in the OPSYNO probe (Prody et al., 1986) that, in fact, it very specifically labeled the human ChEcDNA and was useful in verification of the "true positive." Since the dI might not contribute to the stability of the hybrids, it would be advantageous to use dI-containing probes longer than the 17-mers, such as those employed in the high-complexity OPSYN probes.

Codon Usage Specificity as a Tool for Designing Probes

An alternative approach in the use of oligonucleotide probes is to exploit the fact that different taxa have been found to display different codon preferences. (Chen et al., 1982). Instead of synthesizing a pool of probes, one can synthesize a single or a few probes whose codons are taxa specific (Lathe 1985). This method has been used successfully in the search for several cDNAs. However, it involves a "statistical risk." Even if, statistically, a specific codon is in genral preferable, it might occur that an upredicted codon is used in the particular site for which the probe is synthesized. This would cause the synthesis of an incorrect probe. Therefore, it was decided in our case to use inosine probes and oligodeoxynucleotide pools; in fact, the codon usage of the particular OPSYNO sequence that served for the preparation of our oligodeoxynucleotide probes appears to be nonfavorable according to the general codon usage tables for human genes (Chen et al., 1982). Thus, in this specific case the statistical risk was also a practical one, and the choice of probe mixtures and deoxyinosine was therfore justified.

Human Cholinesterase Homology to Other Proteins

The primary sequence of human ChE as encoded by the isolated cDNA can clearly be distinguished from those of other serine hydrolases (Dayhoff, 1978), although they share common amino acids in the immediate vicinity of the OP-binding serine (Prody et al., 1986). Thus, the genes coding for cholinesterases most probably have arisen form a unique gene. As shown above, ChEs from various species, such as the *T. californica, D. melanogaster,* and human, share extensive sequence homology throughout the polypeptide sequence of the enzyme proteins. This conservation suggests that most of the regions in the ChEs are necessary for biological functions and that the ancestral gene for ChE has developed very early in evolution with essentially the same properties as observed in most species today. One may postulate several important domains within the ChE protein. These include one for binding of the collagen-like "tail," another responsible for membrane binding, and, clearly, the active site and the anionic site domains (Massoulie and Bon, 1982). In addition, sites of contact between subunits and S—S bonds may be postulated. Since all of these properties should be displayed by ChEs in all species, it seems logical to expect to find such high homology between ChEs of geneticallly remote species. The exceptional, additional amino acids found in the *Drosophila* ChE might hence be necessary for the *Drosphila* enzyme alone, but probably do not create any steric hindrance of activity.

To our surprise, homology searches with both human ChE and the *T. californica* protein sequences determined from their nucleic acid cDNA sequences show extensive homology with bovine thyroglobulin, as already noted by Schumacher et al. (1986). This suggests that these two proteins contain, in part, a shared common ancestral origin. The accepted role of thyroglobulin is that of a carrier protein

(Merken et al., 1985), however, this homology suggests that there may exist another, as yet undetermined, function in common for both proteins. The divergence of the ancestral gene is intriguing, considering the different functions of these proteins. Furthermore, the conservation of the primary structure in the thyroglobulin protein may actually become a "physiological nuisance" in cases of hyperthyroidism (such as Grave's opthalmopathy). In this disorder, there exists an over-production of the hormone thyroxin. It was suggested (Swillens et al., 1986) that antibodies raised against thyroglobulin, the thyroxine precursor, could cross-react wth ChEs because of the homology in their primary sequence. These new antibodies, recognizing epitopes shared by both proteins, would then cause an autoimmune effect, responsible for at least some of the symptoms observed in hyperthyroidism. The location of primary sequences of human ChE showing high homology to bovine thyroglobulin (presented in Fig. 4d) is in the area of the active site of the human ChE. This implies that antithyroglobulin antibodies cross-reacting with the ChE protein would be likely to inhibit its enzymatic activity. It remains to be shown, though, that the thyrogolobulin in fact reaches the vascular system in this disorder and if, at all, such autoimmune antibodies exist (Ludgate et al., 1986).

Preliminary Findings on the Structure of the Human ChE Genes

Composition and Structure of the Cholinesterase cDNA

The full-length cholinesterase cDNA fragment contains 2400 nucleotides, with a high (63.3%) content of pyrimidines and a very low content (0.9%) of C–G pairs, a dinucleotide frequency characteristic of the human genome (Lathe, 1985). Its nucleotide sequence displays high levels of homology to the cDNA clones coding for *Torpedo* AChE (Schumacher et al., 1986) and for *Drosophila* AChE (Hall and Spierer, 1987). Futhermore, it resembles bovine thyroglobulin cDNA sequence (Merken et al., 1985), with a similarly open reading frame preceding the signal peptide. The fractional codon utilization of this cDNA is compatible with other human protein-coding sequences (Lathe, 1985) and differs considerably from that found in the *Drosophila* and *Torpedo* AChEcDNA clones, yet its inferred protein displays 53 and 35% matches with the parallel protein sequences of *Torpedo* and *Drosophila* AChE, respectively. In the few domains where the amino acid sequence was not conserved (for example, the region encoded by nucleotides 310–340 in FBChE12), there were also no similarities in the nature of the corresponding residues, as recently classified by Doolittle (1985). This confirms our previous assumption (Prody et al., 1987) that most of the sequence inferred by these cDNA clones is required as such for maintaining the yet undefined biochemical properties characteristic of ψChE.

Genomic DNA Fragments Hybridizing with ChEcDNA Probes

Human DNA restricted with the *Eco*RI enzyme includes fragments of 4.2, 2.3, and 10 kb in length hybridizing with the 5'-terminal, active site region, and 3'-terminal parts of the cDNA encoding the cholinesterase protein. The 2.3-kb fragment is not homologous to the 5' region of the gene since it does not hybridize with a 5'-terminal probe. When human DNA is restricted with *Eco*RI and *Bam*HI together, this 2.3-kb band does not appear, and instead a fragment of 585 bases can be seen. The cDNA restriction pattern reveals a similar size fragment using both restriction enzymes. Therefore, it seems likely that this 2.3-kb fragment contains a 585-nucleotide-long fragment that appears in exactly the same form in the cDNA. The 10-kb fragment

contains sequences homologous to the 3'-terminal area. Finally, the 4.2-kb fragment was enriched from genomic DNA and cloned into λgt10. Its sequence analysis revealed that it indeed contained a portion of the 5' region of the cDNA in addition to at least one intervening sequence, separating the 5'-regulatory region of the ChEcDNA from the nucleotides coding for the initiator methionine (Gnatt, 1986). The existence of an intervening sequence at this important region of the ChEcDNA sequence may reflect an alternative splicing phenomenon, which could regulate the production of various ChEmRNAs from a single gene (Amara et al., 1985). Further characterization of genomic and cDNA sequences encoding for human ChEs would be required to examine these possibilities.

Acknowledgments

We are grateful to Prof. H. Zakut, Drs. C. Prody, P. Dreyfus, and D. Zevin-Sonkin, Mrs. R. Zamir, and Mrs. R. Zisling for their contributions toward this work and to Prof. A. H. Mehler for critically reviewing this manuscript. This study was supported by the US Army Medical Reasearch and Development Command (Contract DAMD 17-85C-5025, to H. S.).

References

Aldridge W. N. and Reiner E. (1972) *Enzyme Inhibitors as Substrates*. North-Holland, Amsterdam.

Alles G. A. and Hawes R. C. (1940) Cholinesterases in the blood of man. *J. Biol. Chem.* **133**, 375–390.

Amara S. G., Arriza J. S., Leff S. E., Swanson L. W., Evans R. M., and Rosenfeld, M. G. (1985) Neuropeptide homologous to calcitonin-gene related peptide.

Anglister L. and Silman I. (1978) Molecular structure of elongated forms of electric eel acetylcholinesterase. *J. Mol. Biol.* **125**, 293–311.

Atack J. R., Perry E. K., Bonham J. R., Perry R. H., Tomlinson B. E., Blessed G., and Fairbairn A. (1983) Molecular forms of acetylcholinesterase in senile dementia of Alzheimer Type: Selective loss of the intermediate (10S) form. *Neurosci. Lett.* **40**, 199–204.

Arias S., Rolo M., and Gonzalez N. (1985) Gene dosage effect present in trisomy 3q25.2-qter for serum cholinesterase (ChE1) and absent for transferrin (TF) and ceruloplasmin (CP). *Cytagenet. Cell Genet.* **40**, 571

Austin L. and Berry W. K. (1953) Two selective inhibitors of cholinesterase. *Biochem. J.* **54**, 695–700.

Bause E. (1983) Structural requirements of *N*-glycosylation of proteins. *Biochem. J.* **209**, 331–336.

Bonham J. R. and Atack J. R. (1983) A neural tube defect specific form of acetylcholinesterase in amniotic fluid. *Clin. Chem. Acta* **135**, 233–237.

Brock D. J. H. and Bader P. (1983) The use of commercial antisera in resolving the cholinesterase bands of human amniotic fluids. *Clin. Chem. Acta* **127**, 419–422.

Brown T., Kennard O., Kneale G., and Rabinovitch D. (1985) High-resolution structure of a DNA helix containing mismatched base pairs. *Nature* **315**, 604–606.

Burstein, S. A., Adamson, J. W., and Waker, L. A. (1980) Megakaryocytopoiesis in culture: Modulation by cholinergic mechanisms. *J. Cell Physiol.* **103**, 201–208.

Chen H. R., Dayhoff M. O., Baker W. C., Hunt L. T., Yen L. S., George D. G., and Orcutt B. C. (1982) Nucleic acid sequence data-base IV. *DNA* **1**, 365–374.

Claudio T., Ballivet M., Patrick J., and Heinemann S. (1983) Nucleotide and deduced amino acid sequences of *Torpedo californica* acetylcholine receptor α-subunit. *Proc. Natl. Acad. Sci. USA* **80**, 1111–1114.

Coates P. M. and Simpson N. E. (1972) Genetic variation in human erythrocyte acetylcholinesterase. *Science* **175**, 1466–1477.

Coyle J. T., Price D. L., and DeLong M. R. (1983) alzheimer's disease: A disorder of cortical cholinergic innervation. *Science* **219**, 1184–1190.

Culotti J. C., Von Ehrenstein G., Culotti M. R., and Russell R. L. (1981) A second class of Acetylcholinesterase-deficient mutants of the nematode *Caenorhabditis Elegans*. *Genetics* **97**, 281–305.

Dayhoff M. O. (1978) Atlas of Protein Sequence and structure, vol. 5,, suppl. 3, National Biomedical Research Foundation.

Dohet C., Wagner R., and Radman M. (1985) Repaire of defined single base-paire mismatches in *Escherichia coli. Proc. Natl. Acad. Sci. USA* **82**, 503–505.

Doolittle R. F. (1985) The Proteins. *Sci. Am.* **253**, 74–83.

Drews E. (1975) Cholinesterase in embryonic development. *Prog. Histochem. Cytochem.* **7**, 1–52.

Dreyfus P. A., Rieger F., and Pincon-Raymond M. (1983) Acetylcholinesterase of mammalian neuromuscular junctions: Presence of tailed asymmetric acetylcholinesterase in synaptic basal lamina and sarcolemma. *Proc. Natl. Acad. Sci. USA* **80**, 6698–6702.

Dutta-Choudhury T. A. and Rosenberry T. (1984) Human erythrocyte acetylcholinesterase is an amphipathic protein whose short membrane-binding domain is removed by papain digestion *J. Biol. Chem.* **259**, 5653–5660.

Dzjiegielewska K. M., Saunders N. R., Schejter R. J, Zakut H., Zevin-Sonkin D., Zisling R., and Soreq H. (1986) Synthesis of plasma proteins in fetal, adult, and neoplastic human brain tissue. *Dev. biol.* **115**, 93–104

Evans D. H. and Morgan A. R. (1986) Characterization of imperfect DNA duplexes containing unpaired bases and non-Watson-Crick base pairs. *Nucl. Acids Res.* **14**, 4267–4280.

Fishman E. B., Siek G. C., MacCallum R. D., Bird E. D., Volicer L., and Marquis J. K. (1986) Distribution of the molecular forms of acetylcholinesterase in human brain: Alterations in Dementia of the Alzheimer type. *Ann. Neurol.*, in press.

Futerman A. H., Low M. G., and Silman I. (1985) A hydrophobic dimer of acetylcholinesterase from *Torpedo californica* electric organ is solubilized by phsophatidylinostiol-specific phospholipase C. *Neurosci. Lett.* **40**, 58–89.

Futerman A. H., Barton P. L., Fiorini R. M., Low M. G., Sherman W. R., and Silman I. (1986) the involvement of phosphatidylinositol in the anchoring of hydrophobic forms of acetylcholinesterase to the plasma membrane, in *Molecular basis of nerve activity* (Changeux, Hucho, Maelicke, and Neumann, eds.), Walter de Gruyter, Berlin, New York, pp. 635–650.

Gnatt A. (1986) *The isolation and characterization of human cholinesterase cDNA and genomic sequences.* M. Sc. Thesis, The Weizmann Institute of Science.

Grassi J., Vigny M., and Massoulie J. (1982) Molecular forms of acetylcholinesterase in bovine caudate nucleus and surperior cervical ganglion: Solubility properties and hydrophobic character. *J. Neurochem.* **38**, 457–469.

Haas, R.and Rosenberry T. L. (1985) Quantitative identification of N-terminal amino acids in proteins by radiolabeled reductive methylation and amino acid analysis: Application to human erythrocyte acetylcholinesterase. *Anal. Bio. Chem.* **148**, 74– 77.

Hall J. C. (1982) Genetics of the nervous system in *Drosophila. Q. Rev. Biophy.* **15**, 3–479.

Hall J. C. and Spierer P. (1987) The ACE locus of *Drosophila melanogaster*: Structural gene for Acetylcholinesterase with an unusual 5' leader. *EMBO J.* 34, in press.

Heijne G. V. (1985) Signal sequence: the limits of variation. *J. Mol. Biol.* **184**, 99–105.

Hobbiger F. (1963) Reactivation of phosphorylated acetylcholinesterase, in *Cholinesterases and Anticholinesterase Agents*, (Koelle G. B., ed.), (Springer, Berlin), pp. 921–988.

Hodgkin W. E., Giblett E. R., Levine H., Bauer W., and Motulsky A. G. (1985) Complete pseudocholinesterase deficiency: Genetic and immunologic characterization. *J. Clin. Invest.* **44**, 486–497.

Hopp T. P. and Woods K. R. (1981) Prediction of protein antigenic determinants from amino acid sequences. *Prot. Natl. Acad. Sci. USA* **78**, 38. 3824–3828.

Huerre C., Uzan G., Greschik K. H., Weil D., Levin M., Hors-Cayla M. C., Boue J., Kahn A., and Junien C. (1984) The structural gene for transferrin TF maps to 3q21-3qter. *Ann. Genet.* **40**, 107–127.

Hunter W. N., Brown T., Anand N. N., and Kennard O. (1986) Structure of an adenine–cytosine base pair in DNA and its design implications . *Nucl. Acids Res.* **13**, 8927–8938.

Jedrzejczyk J., Silman I., Lai J., and Barnard E. A. (1984) Molecular forms of acetylcholinesterase in synaptic and extrasynaptic regions of avian tonic muscle. *Neurosci. Lett.* **46**, 283–289.

Johnson C. D. and Russell R. L. (1983) Multiple molecular forms of acetylcholinesterase in the nematode *Caenorhabditis elegans. J. Neurochem.* **41**, 30–36.

Johnson C. D., Duckett J. G., Culotti J. G., Herman R. K., Meneely P. M., and Russell R. L. (1981) An

acetylcholinesterase-deficient mutant of the nematode *Caenorhabditis elegans*. *Genetics* **97**, 261–279.

Kalow W. and Gunn D. R. (1959) Some statistical data on atypical cholinesterase of human serum. *Ann. Hum. Genet* (Lond.) **23**, 239–248.

Karnovsky M. J. and Roots L. (1964) A direct-coloring' thiocholine method for cholinesterases. *J. Histochem. Cytochem.* **12**, 219–221.

Kidd K. K. and Gusella J. (1985) Report of the committee on the genetic constitution of chromosomes 3 and 4. *Cytogenet. Cell Genet* **40**, 107–127.

Klose R. and Gustensohn G. (1976) Treatment of alkyl phosphate poisoning with purified serum cholinesterase. *Prakl. Asasthe.* **11**, 1–7.

Koelle G. B. (1972) Anticholinesterase agents, in (Goodman L. S. and Gilman A., eds.), Fifth ed., McMillan, NY, pp. 445–466.

Kolson D. L. and Russell R. L. (1985a) New acetylcholinesterase deficient mutants of the nematode *Caenorhabditis elegans*. *J. Neurogenet* **2**, 69–91.

Kolson D. L. and Russell R. L. (1985b) A novel class of acetylcholinesterase, revealed by mutation in the nematode *Caenorhabitis elegans*. *J. Neurogenet.* **2**, 93–110.

Kubo T., Fukuda K., Milkami A., Maeda A., Takahashi H., Mishina M., Hatga T., Haga K., Ichiyama A., Kanagawa K., Kojima M., Matsuo H., Hirose T., and Numa S. (1986) Cloning, sequencing and expression of complementary DNA encoding the muscarinic acetylcholine receptor. *Nature* **323**, 411–416.

Lathe R. (1985) Synthetic oligonucleotide probes deduced from amino acid sequence data: Theoretical and practical considerations. *J. Mol. Biol.* **183**, 1–12.

Lockridge O. (1984) Amino acid composition and sequence of human serum cholinesterase: A progress report, in *Cholinesterase Fundamental and Applied Aspects* (Brzin M., Barnard E. A., and Sket eds.) Walter de Gruyter, NY pp. 5–12.

Lockridge O., Bartels C. G., Vaughan T. A., Wong C. K., Norton S. E., and Johnson L. L. (1987) Complete amino acid sequence of human serum cholinesterase *J. Biol. Chem.* **262**, 549– 557.

Lockridge O. and La Du B. N. (1986) Amino acid sequence of the active site of human serum cholinesterase from usual, atypical and atypical-silent genotypes. *Biochem. Genet.* **24**, 485–498.

Loomis T. A. (1963) Distribution and excretion of pyridine-2-aldoxime methiodide (PAM); atropine and PAM in sarin poisoning *Toxicol. Appl. Pharamacol.* **5**, 489–499.

Lovrien E. W., Magenis R. E., Rivas M. L., Lamvik N., Rowe S., Wood J., and Hemmerling J. (1978) Serum cholinesterase E2 linkage analysis: Possible evidence for localization to chromosome 16. *Cytogenet. Cell. Genet.* **22**, 324–326.

Ludgate, M., Swillens S., Mercken L., and Vassart G. (1986) Homology between thyroglobulin and acetylcholinesterase: An explanation for pathogenesis of ginnes opthamology? *Lancet* **II**, 219.

Maizel J. V. and Lenk R. R. (1981) Enhanced graphic matrix analysis of nucleic acid and protein sequences. *Proc. Natl. Acad. Sci. USA* **78**, 7665–7669.

Martin F. H., Castro M., Aboul-ela, Grave's and Tinoco I., Jr. (1985) Base pairing involving deoxyinosine: Implications for probe design. *Nucl. Acids Res.* **13**, 8927– 8938.

Massoulie J. and Bon S. (1982) The molecular forms of cholinesterase and acetylcholinesterase in vertebrates. *Ann. Rev. Neurosci.* 5, 57–106.

McPhee-Quigley K., Taylor P., and Taylor S. (1985) Primary structures of the catalytic subunits from two molecular forms of acetylcholinesterase: A comparison of NH_2-terminal and active center sequences. *J. Biol. Chem.* **260**, 12185– 12189.

Meflah K., Bernard S., and Massoulie J. (1984) Interactions with lectins indicate differences in the carbohydrate composition of the membrane-bound enzymes acetylcholinesterase and 5'-nucleotidase in different cell types. *Biochimie* **66**, 59–69.

Melchior E. B. and Hippel H. V. (1972) Alteration of the relative stability of dA-dT and dG-dC base pairs in DNA. *Proc. Natl. Acad. Sci. USA* **70**, 298–302.

Mendel B., and Rudney H. (1943) Studies on cholinesterase. A specific test for true pseudocholinesterase. *Biochem. J.* **37**, 59–63.

Merken L., Simons M. J., Swillens, S., Massaer M., and Vassart G. (1985) Primary structure of bovine thyroglobulin deduced from the sequence of its 8,431-base complementary DNA. *Nature* **316**, 647–651.

Messing J. (1983) new M13 vectors for cloning, in *Methods in Enzymology-Recombinant DNA Techniques* vol. 101, part c, pp. 20–78.

Mushina M., Kurosaki T., Tobimatsu T., Morimoto Y., Noda M., and Numa S., (1984) Expression of functional acetylcholine receptor from cloned cDNAs. *Nature* **307**, 604–608.

Noda M., Takahashi H., Tanabe T., Toyasato M., Kikyotani S., Furutani Y., Hirosi T., Tasashima H., Inayama S., Niyata T., and Numa S. (1983) Structural homology of *Torpedo californica* acetylcholine receptor subunits. *Nature* 69. **302,** 528–531.

Ord G. M. and Thompson R. H. S. (1952) Pseudocholinesterase activity in the central nervous system. *Biochem. J.* **51,** 245– 251.

Patel D. J., Kozlowski S. A., Marky L. A., Broka C., Dallas J., Itakura K., and Breslauer K. J. (1982) Structure, dynamics, and energetics of deoxyguanosine-thymidine Wobble base pair formation in the self-compleme d(CGTGAATTCGCG) duplex in solution. *Biochemistry* **21,** 437–444.

Prody C., Zevin-Sonkin D., Gnatt A., Koch R., Zisling R., Goldberg O., and Soreq H. (1987) Use of synthetic oligodeoxynucleotide probes for the isolation of a human cholinesterase cDNA clone. *J. Neurosci. Res.* **16,** 25–35.

Prody C., Gnatt A., Zevin-Sonkin D., Goldberg O., and Soreq H. (1987) Isolation and characterization of full length cDNA clones coding for cholinesterase: from fetal human tissues *Proc. Natl. Acad. Sci. USA* , in press.

Queen C. L. and Korn L. J. (1980) Computer Analysis of Nucleic Acids and Proteins *Meth. Enz. Mole* **65,** 595–609.

Razon N., Soreq H., Roth E., Bartal A., and Silman I. (1984) Characterization of levels and forms of cholinesterases in human primary brain tumors. *Exp. Neurol.* **84,** 681–695.

Rosenberry T. L. and Scoggin D. M. (1984) Human erythrocyte acetylcholinesterase is an amphipathic protein whose short membrane-binding domain is removed by papain digestion. *J. Biol. Chem.* **250,** 5643– 5652.

Rotundo R. L. (1984) Asymmetric acetylcholinesterase is assembled in the Golgi apparatus. *Proc. Natl. Acad. Sci. USA* **81,** 479–483.

Sanger G., Nicklen S., and Coulson A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. sci. USA* **74,** 5463–5468.

Schumacher M., Camp S., Maulet Y., Newton M., MacPhee-Quigley K., Taylor S. S., Friedmann T., and Taylor P. (1986) Primary structure of *Torpedo Californica* acetylcholinesterase deduced from its cDNA sequence. *Nature* **319,** 407–409.

Sikorav J. L., Frassi J., and Bon S. (1985) Synthesis in vitro of precursors of the catalytic subunits of acetylcholinesterase from *Torpedo marmorata* and *Electrophorus electricus. Eur. J. Biochem.* **145,** 519–524.

Silman I., DiGiamberardino L., Lyles J., Couraud J. Y., and Barnard E., (1979) Parallel regulation of acetylcholinesterase and Pseudocholinesterase in normal, denervated, and dystrophic chicken skeletal muscle. *Nature* **28,** 160–161.

Silver A. (1974) *The Biology of Cholinesterases.* North-Holland, Amsterdam.

Simmers R. N., Stupans I., and Sutherland G. R. (1986) Localization of the human haptoglobin genes distal to the fragile site at 16922 using in situ hybridization *Cytogenet. Cell Genet.* **41,** 838–41.

Simpson N. E. (1966) Factors influencing cholinesterase activity in a Brazilian population *Am. J. Hum. Gen.* **18,** 29–35.

Soreq H., Dzjiegielewskla K. M., Zevin-Sonkin D., and Zakut H. (1986) The use of mRNA translation *in vitro* and *in ovo* followed by crossed immunoelectrophoretic antoradiography to study the biosynthesis of human cholinesterases. *Cell Molec. Neurobiol.* **6,** 227–237.

Soreq H., Zamir R., Zevin-sonkin D., and Zakut H. (1987) Human cholinesterase genes localized by hybridization to chromosomes 3 and 16, submitted.

Soreq H., Zevin-Sonkin D., and Razon N. (1984) Expression of cholinesterase gene(s) in human brain tissues: Translational evidence for multiple mRNA species. *EMBO J.* **3,** 1371–1375.

Sparkes R. S., Field L. L., Sparkes M. C., Crist M., Spence M. A., James K., and Garry P. J. (1984) Genetic linkage studies of transferrin, pseudocholinesterase, and chromosome 1 loci. *Hum. Hered.* **34,** 96–100.

Sumikawa K., Houghton M., Smith J. C., Bell L., Richards B. M., and Barnard E. A. (1982) The Molecular cloning and characterization of cDNA coding for the α-subunit of the acetylcholine receptor. *Nucl. Acids Res.* **10,** 5809–5818.

Swillens S., Ludgate M., Mercken L., Dumont J. E., and Vassart G. (1986) analysis of sequence and structure homologies between thyroglobulin and acetylchoinesterase: Possible functional and clinical significance. *Biochem. Biophys. Res. Commun.* **137,** 142–148.

Szeinberg A., Pipano S., Assa M., Medalie J. H., and Neufeld H. N. (1972) High frequency of atypical

pseudocholinesterase gene among Iraqi and Iranian Jews. *Clin. Genet.* 3, 123–127.

Takahashi Y., Kato K., Hayashizaki Y., Wakabayashi T., Ohtsuka E., Matsuki S., Ikehara M., and Matsubara K. (1985) Molecular cloning of the human cholecystokinin gene by use of a synthetic probe containing deoxyinosine. *Proc. Natl. Acad. Sci. USA* **82**, 1931–1935.

Thompson J. C. and Whittaker M. (1966) A study of the pseudocholinesterase in 78 cases of apnea following suxamethonium. *Acta Genet.* **16**, 206–215.

Toutant J. P., Massoulie J., and Bon s. (1985) Polymorphism of pseudocholinesterase in *Torpedo marmorata* tissues: Comparative study of the catalytic and molecular properties of this enzyme with acetylcholinesterase. *J. Neurochem.* **44**, 580–592.

Vigny M., Gisiger V., and Massoulie J. (1978) "Nonspecific" cholinesterase and acetylcholinesterase in rat tissues: Molecular forms, structural, and catalytic properties and significance of the two enzyme systems. *Proc. Natl. Acad. Sci. USA* **75**, 2588–2592.

Wallace R. B., Johnson M. J., Hirose T., Miyake T., Kawashima E. H., and Itakura K. (1981) The use of synthetic oligonucleotides as hybridization probes. II. Hybridization of oligonucleotides of mixed sequence to rabbit beta-globin *DNA. Nucl. Acids. Res.* **9**, 879–894.

Whittaker M. (1980) Plasma cholinesterase and the anaesthetist. *Anaesthesia* **35**, 174–197.

Wood W. I., Gitschier J., Laskey L. A., and Lawn R. M. (1985) Base composition-independent hybridization in tetramethylammonium chloride: A method for oligonucleotide screening of highly complex gene libraries. *Proc. Natl. Acad. Sci. USA* **82**, 1585–1588.

Yates C. M., Simpson J., Maloney A. F. J., Gordon A., and Reid A. H. (1980) Alzheimer-like cholinergic deficiency in Down's syndrome. *Lancet* **2**, 979–980.

Zagursky R. J., Bumeister K., Lomax N., and Berman M. L. (1985) Rapid and easy sequencing of large linear double-stranded DNA and supercoiled plasmid DNA. *Gene Anal. Techn.* **2**, 89–94.

Zakut H., Matzkel A., Schejter E., Avni A., and Soreq H. (1985) Polymorphism of acetylcholinesterase in discrete regions of the developing human fetal brain. *J. Neurochem.* **45**, 382–389.