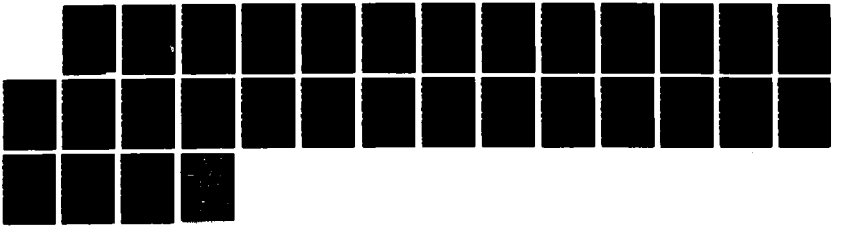
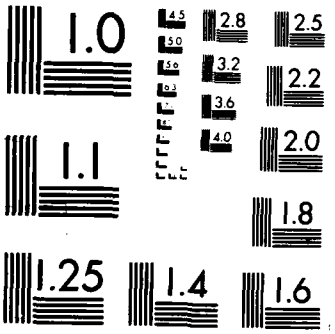
AD-A192 751 ISOLATION OF GENOMIC CLONE FOR HUMAN CHOLINESTERASE(U) 1/1
MICHIGAN UNIV ANN ARBOR DEPT OF PHARMACOLOGY
O LOCKRIDGE DEC 86 DAMD17-86-C-6037

UNCLASSIFIED F/G 6/1 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3

AD_____

ISOLATION OF GENOMIC CLONE FOR HUMAN CHOLINESTERASE

ANNUAL SUMMARY REPORT

OKSANA LOCKRIDGE

December 1986

Supported by

U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Fort Detrick, Frederick, Maryland 21701-5012

Contract No. DAMD17-86-C-6037

University of Michigan
Pharmacology Dept., Med. Sci. I
Ann Arbor, Michigan 48109-0010

SECURITY CLASSIFICATION OF THIS PAGE

ADA192757

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188
Exp Date Jun 30, 1986

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION University of Michigan Pharmacology Dept, Med. Sci. I 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Ann Arbor, Michigan 48109-0010	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING / SPONSORING ORGANIZATION U.S. Army Medical Research & Development Command 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER Contract No. DAMD17-86-C-6037
8c. ADDRESS (City, State, and ZIP Code) Fort Detrick Frederick, Maryland 21701-5012	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 62734A; PROJECT NO. 3M1-62734A875; TASK NO. AA; WORK UNIT ACCESSION NO. 523

11. TITLE (Include Security Classification)
Isolation of Genomic Clone for Human Cholinesterase

12. PERSONAL AUTHOR(S)
Oksana Lockridge, Ph.D.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Annual Report	FROM 15Nov85 TO 15Nov86	1986 December	33

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Cholinesterase Genes; Pseudocholinesterase; Genetic Engineering; Recombinant DNA; Gene Sequencing.
06	13		
06	01		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The goal of this work is to isolate the gene for human serum cholinesterase (EC 3.1.1.8). We have succeeded in isolating a partial clone that codes for amino acids phenylalanine 195 through leucine 574, and contains 302 bp of the 3' untranslated region. This clone was found by screening a human brain (cDNA) library with oligonucleotide probes. The probes corresponded to the amino sequence of human serum cholinesterase. The coding region of our 1440 bp clone matched the amino acid sequence of human serum cholinesterase perfectly. This suggested that human cholinesterase was identical in brain and in serum. We are continuing to search the brain cDNA library for other cDNA clones that will extend our present clone in the 5' direction so as to have the complete coding region of cholinesterase. Keywords:

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Mrs. Virginia Miller	22b. TELEPHONE (Include Area Code) 301/663-7325 22c. OFFICE SYMBOL SGRD-RMI-S

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted
All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

SUMMARY

The goal of this project is to isolate the gene for human serum cholinesterase (EC 3.1.1.8). We have isolated a cDNA clone from a human brain basal ganglia cDNA library, that codes for a portion of the cholinesterase enzyme. The clone contains 1440 bp, of which 1138 bp are in the coding region, and 302 bp are in the 3' untranslated region. The coding region corresponds perfectly to the amino acid sequence of human serum cholinesterase starting at phenylalanine 195 to leucine 574 at the carboxy terminus. This confirms that our cDNA clone codes for cholinesterase rather than for acetylcholinesterase (EC 3.1.1.7). Furthermore it leads to the conclusion that the amino acid sequence of cholinesterase from two different tissues, from human serum and human brain, is identical.

The strategy for finding the cDNA clone depended on our knowledge of the amino acid sequence of human serum cholinesterase. Our laboratory had determined the complete amino acid sequence, and had designed 12 oligonucleotide probes corresponding to regions containing minimum codon ambiguity. These oligonucleotide probes were successfully used to find the cDNA clone for cholinesterase. We are continuing to search the brain cDNA library for other cDNA clones that will extend our present clone in the 5'direction so as to have the complete coding region of cholinesterase.

FOREWORD

Citations of commercial organizations and trade names in this report do not constitute an official Department of the Army endorsement or approval of the products or services of these organizations.

For the protection of human subjects the investigator has adhered to policies of applicable Federal Law 45CFR46.

The investigator(s) have abided by the National Institutes of Health Guidelines for Research Involving Recombinant DNA Molecules (April 1982) and the Administrative Practices Supplements.

TABLE OF CONTENTS

INTRODUCTION

The goal of this project is to isolate the gene for human serum cholinesterase. Our strategy for finding the gene was based on our previous results, where we had determined the complete amino acid sequence of human serum cholinesterase. Knowledge of the amino acid sequence allowed us to design oligonucleotide probes that would selectively hybridize with the gene for human serum cholinesterase. We have used these probes to screen a human brain cDNA library and have found one positive clone. The DNA sequence of the positive clone corresponded to a portion of the amino acid sequence of human serum cholinesterase, starting at phenylalanine 195 and extending past leucine 574 at the carboxyl terminus, into the 3' untranslated region.

Isolation of the clone for human serum cholinesterase will allow large scale production of this enzyme via recombinant DNA technology. This is of military importance because cholinesterase has the property of being highly reactive with organophosphate esters of the type used in nerve gas. Cholinesterase could be used as a protective agent, to protect personnel against the toxic effects of nerve gas. It could also be used to clean up spills in situations where a chemical agent such as sodium hydroxide would be too corrosive for practical purposes.

METHODS

Oligonucleotide probes. Twelve oligonucleotide probes were made. The probes are listed in Table 1. The probes were based on the amino acid sequence of human serum cholinesterase (1). Regions of the amino acid sequence representing minimum codon ambiguity were chosen. The location of each probe within the complete amino acid sequence is illustrated in Table 2. Probes 6 and 7 were synthesized by the University of Michigan DNA synthesizing facility. All other probes were synthesized by PL-Biochemical Corp., Milwaukee, WI. Probes 1, 3, and 4 represent the coding strand of the amino acid sequence. All other probes represent the reverse complement.

Three different principles were used in designing these probes. Probes 1-5 are mixed oligonucleotides containing all possible DNA sequences corresponding to the selected amino acids. Probes 6 and 7 are unique sequence probes; in every position where more than one nucleotide could have been used for a codon, a single choice was made. Unique sequence probes are longer than mixed probes to compensate for the likelihood of a mismatch. Probes 8-12 contain inosine at every position where more than one nucleotide could be chosen.

The following variables were considered in designing our unique sequence probes (2). The amino acid sequence of cholinesterase was examined to find the longest region corresponding to the least codon ambiguity. Two such regions were found. The region EWGNNAFFYYFEH corresponded to 39 nucleotides having an 83.3% probability of being homologous to the DNA sequence. This probability was calculated from the observations that all 13 amino acids in this region had fixed bases in positions 1 and 2 of their codons, and that the 3rd base was fixed for 1 amino acid (W), but could be 1 of 2 choices for 5 amino aicds (E,N,F,Y,H), and could be 1 of 4 choices for 2 amino acids (G,A). A second region was WPEWMGVMHGYEIEFVFG corresponding to 53 nucleotides having an 85.0% probability of being homologous to the DNA sequence. Codon degeneracy was limited to the 3rd base in each codon; 4 amino acids had no degenerate codons, 7 amino acids had a choice of 2 codons, and 6 amino acids had a choice of 4 codons. These two regions were located within 6 amino acid residues of each other. This proximity increased the chance that a clone would be represented by both regions, and therefore that a positive recombinant clone might be positive with both probes. This is important for eliminating false positives.

To further increase the likelihood of homology, codon usage was considered. Codon usage in human blood coagulation factor VIII (3) was analyzed to determine which codons were used most

frequently. Human blood coagulation factor VIII was selected as a model because, like cholinesterase, it is a glycoprotein that is synthesized and secreted by the liver. Furthermore, it is present in low concentration in plasma, suggesting that its mRNA belongs to the category of "weakly expressed" mRNA (4). Grantham et al (4) found a correlation between codon usage and whether an mRNA was highly or weakly expressed. Codon usage in factor VIII showed a preference for Ala (GCT), Asp (GAT), Glu (GAA), Gly (GGA), Lys (AAA), and Val (GTC). The other amino acids had no clear codon choices. G was infrequently used in the 3rd positions of Pro, Ser, Thr, and Ala. Codons for those amino acids having no clear preference in codon usage were selected by following the rule that a G:T mismatch, though less stable than a G:C pair still contributes to hybridization stability whereas an A:C mismatch does not (2). Thus G was chosen for the 3rd position when the choice was G or A, and T was chosen for the 3rd position when the choice was T or C. Following this rule, we chose T for the 3rd positions of Asn, His, Phe, Tyr, and Ile. We also chose T for the 3rd position of Pro, because G was infrequently used.

The use of deoxyinosine at positions corresponding to ambiguous nucleotides was introduced by Ohtsuka et al (5). They synthesized 23mer and 26mer probes, each containing five deoxyinosines, and successfully used these to find a genomic clone. Deoxyinosine does not contribute to the stability of a nucleotide pair, but neither does it destabilize. For these reasons we designed a series of inosine probes. These are probes #8,9,10,11, and 12 in Table 1.

Probes were labeled at the 5' position with gamma ^{32}P-ATP and T4 polynucleotide kinase.

cDNA library. A cDNA library from the basal ganglia of a 1-day old human brain was from Dr. Robert A. Lazzarini (6). This library can now be purchased from the American Type Culture Collection, Rockville, MD. The cDNA was in vector lambda gt11 to which it was linked via Eco R1 linkers. The library had first been screened by Dr. Charles McTiernan in the laboratory of Dr. Terrone Rosenberry at Case Western Reserve University Medical School. Dr. McTiernan had screened the library with our 53mer unique sequence probe and had found several positive clones. Upon sequencing one of these clones, he found that it matched our amino acid sequence for human serum cholinesterase, which had not yet been published. Because of this collaboration, we knew that this particular library contained clones for human cholinesterase.

Screening the cDNA library. The in situ plaque hybridization method of Benton & Davis (7) as modified by Maniatis et al (8) was used to screen 150,000 plaques spread on 20 plates, containing 47 ug/ml ampicillin. The host was E. coli strain Y1088 (9). Four

lifts were made from each 150 mm plate onto four separate nitrocellulose filters (Millipore, HATF 137 50, pore size 0.45 um). Each filter was hybridized with a different probe or set of probes. Pre-hybridization was in 0.25% Carnation instant dry milk, 6 x SSC (6 x SSC is 0.9 M NaCl, 0.09 M NaCitrate, pH 7.0), at 60° for 2 to 16 h (10). Hybridization was in 50 ml of 0.25% Carnation instant dry milk, 6 x SSC containing 120 uCi of probe, for 12 to 16 h. The hybridization and washing temperatures were specific for each probe. The 53mer probe was hybridized at 46°; filters were washed at room temperature in 6 x SSC, 0.05% pyrophosphate, and then at 46° for 10 min in 2 x SSC, 0.1% SDS. The 39mer probe was hybridized at 36°; filters were washed at room temperature in 6 x SSC, 0.05% pyrophosphate and then at 46° for 10 min in 6 x SSC, 0.05% pyrophosphate. Probes #8, #10, and #12 in Table 1, were used together. They were hybridized at 36°; filters were washed at room temperature in 6 x SSC, 0.05% pyrophosphate, followed by 10 min at 43° in the same buffer. Probes #4, #9, and #11 were used together. They were hybridized at 36°; filters were washed at room temperature in 6 x SSC, 0.05% pyrophosphate, followed by 10 min at 41° in the same buffer.

Purification of DNA. This procedure is described in detail because it is unpublished. It was adapted from David Law's thesis (11). The advantage of this method is that it avoids a long centrifugation in cesium chloride.

PREPARATION OF HOST CELLS. A single colony of Y1088 was inoculated into 50 ml of NZYM medium supplemented with 0.25 ml of 1% thymidine. NZYM medium was made by dissolving 10 g casein enzymatic hydrolysate from Sigma, 5 g NaCl, 5 g bacto-yeast extract (Difco Labs), 2 g $MgSO_4.7H_2O$ in water to a total volume of 1 liter and adjusting the pH to 7.5 with NaOH. Bacteria were grown for 11 h at 37°C.

ADSORBTION OF PHAGE TO HOST CELLS. To 1 ml of the fresh overnight culture of Y1088, were added 5 ml of phage containing 5.6 x 10^6 pfu, 1 ml of 50 mM calcium chloride, and 15 ul of 1 M $MgSO_4$. It was important to infect the bacteria with 5 x 10^6 phage. If significantly fewer or significantly greater numbers of phage were used, the yield of DNA decreased drastically. The phage were allowed to adsorb to the bacteria for 10 min at 37°.

PHAGE GROWTH. The phage-Y1088 mixture was added to 1 liter of NZYM medium supplemented with 5 ml of 1% thymidine. The medium was in a 2 liter Fernbach flask. It was shaken at 200 rpm on a shaking table at 37° for 18 h. After 18 h the appearance of the medium was cloudy with white swirls and small white pieces of debris. 58 g NaCl and 20 ml chloroform were added to further lyse the bacteria. The medium was shaken an additional 15 min. It was centrifuged at 8000 rpm for 10 min at 4°C to remove cell debris.

PEG PRECIPITATION. To the clear supernatant was added 100 g of polyethylene glycol (PEG 8000 from Sigma) and 12 ml of 1 M $MgSO_4$. This was slowly stirred at 4° for 1 h, and then centrifuged for 10 min at 4° at 8000 rpm. Two centrifuge cups were used and pellets were stacked on top of each other.

CHLOROFORM EXTRACTION OF PEG PELLET. Pellets were suspended in SM. SM was prepared by adding 5.8 g NaCl, 2 g $MgSO_4.7H_2O$, 50 ml of 1 M TrisCl pH 7.5, and 5 ml of 2% gelatin to a total volume of 1 liter in water and autoclaving. 5 ml of SM was added to each centrifuge bottle and the suspension was gently homogenized in a glass homogenizer. The pooled suspension was poured into one 30 ml corex tube. 10 ml of chloroform was added, and the liquids mixed by drawing into a pasteur pipette. This was centrifuged at 8000 rpm at 4° for 15 min. The chloroform layer went to the bottom, and the aqueous layer to the top of the tube. The aqueous layer was transferred into 4 thick walled SW50.1 centrifuge tubes, and centrifuged in an SW50.1 swinging bucket rotor (Beckman Co.) at 20,000 rpm for 2 h at 4°. A small yellowish precipitate contained the DNA.

PROTEINASE K AND RNAse DIGESTION. The supernatant was discarded. 150 ul of SM was added to each tube and the pellet suspended by shaking at room temperature for 1 h. The total volume was adjusted to 2 ml with SM. EDTA was added to a final concentration of 20 mM; SDS to 0.5%; proteinase K (Sigma) to 50 ug/ml; and RNAse to 20 ug/ml. The mixture was incubated 1 to 12 h at 65°.

PHENOL/CHLOROFORM EXTRACTION. The solution was extracted with 1 volume of phenol, followed by 1 volume of phenol/chloroform (1:1), and finally 1 volume of chloroform. The DNA stayed in the aqueous layer.

ETHANOL PRECIPITATION. The aqueous layer was divided into 3 microcentrifuge tubes (1.5 ml capacity) so that each tube contained about 700 ul. Added 70 ul of 3 M NaAcetate pH 4.6 and 700ul cold isopropanol. Huge pellets came down after a 3 min centrifugation in the microfuge. The pellet was washed with 1 ml of 70% cold ethanol. The DNA pellet was air dried for 3 to 5 h.

ESTIMATION OF DNA YIELD. The pellet was incubated overnight in 400 ul of 20 mM TrisCl pH 7.5, 0.1 mM EDTA at room temperature. The resulting solution was viscous. Absorbance of a 1:100 diluted sample was measured at 260 nm and 280 nm. The DNA concentration was calculated from absorbance at 260 nm, where an OD of 1 corresponded to 50 ug/ml for double stranded DNA. The 260/280 absorbance ratio of pure DNA was 1.8. Contamination by RNA gave a ratio higher than 1.8. Contamination by protein or phenol gave a ratio lower than 1.8. The yield was 2 to 5 mg of DNA.

Restriction mapping. The purified DNA, still attached to lambda gt11 arms, was digested with restriction enzymes purchased from Boehringer Mannheim, BRL, or Amersham. The digested DNA was electrophoresed on 0.8% agarose (BRL, ultrapure), stained with ethidium bromide, photographed, and transferred to nitrocellulose by the method of Southern(12). The DNA bands containing the human cDNA were identified by hybridizing with various probes from the list in Table 1.

DNA sequencing. The 2.4 kb fragment produced by digestion of the lambda gt11 clone with Kpn 1 and Eco R1 was transfected into M13mp19. The 2.4 kb clone was partially sequenced by the dideoxy chain termination method of Sanger (13). Subclones of the 2.4 kb fragment were prepared by isolating the Kpn 1-Eco R1 fragment from agarose gel by electroelution, and then digesting with Sau 3A1. The Sau 3A1 digest was ligated into the isoschizomeric Bam H1 site of M13mp18. These subfragments were prepared because the 2.4 kb clone was too long to sequence to completion when it was intact. 60 clear plaques were selected at random for preparation of the single stranded recombinant M13 to be used for DNA sequencing. All 60 subclones were sequenced by the dideoxy method of Sanger.

Computer. We used the Nucleic Acid Query Program of the Protein Identification Resource , National Biomedical Research Foundation, Georgetown University Medical Center, 3900 Reservoir Road, N.W., Washington, D.C. 20007 phone 202 625-2121. This program allowed us to store and align our DNA sequence files, and to compare the DNA sequence of our clone to those in the computer databank.

Table 1. List of oligonucleotide probes.

	Trivial name	Length	Amino Acids	Location in protein	Probe Sequence
1.	8mix	17mer	WPEWMG	430-435	5'TGGCCNCAATGGATGGG 3' G
2.	8mix	17mer	WPEWMG	430-435	Reverse complement 5'CCCATCCATTGNGGCCA 3' C
3.	16mix	17mer	WKNQFN	557-562	5'TGGAAAAACCAATTCAA 3' G T G T
4.	32mix	17mer	KEFQEG	348-353	5'AAAGAATTCCAAGAAGG 3' G G T G G
5.	4mix	17mer	MMDWKN	554-559	Reverse complement 5'TTTTTCCAGTCCATCAT 3' C A
6.	39mer	39mer	EWGNNAFFYYFEH	411-423	Reverse complement 5'ATGTTCAAAATAATAAAAAAAAGCATTATTTCCCCATTC 3'
7.	53mer	53mer	WPEWMGVMHGYEIEFVFG	430-447	Reverse complement 5'CCAAAGACAAATTCAATTTCATATCCATGCATGACTCCCATCCATTCAGGCCA 3'
8.	AEW	29mer	AEWEWKAGFH	539-548	Reverse complement 5'TGIAAICCIGCITTCCAITCCCAITCIGC 3'
9.	WNN	29mer	WNNYMMDWKN	550-559	Reverse complement 5'TTITTCCAITCCATCATITAITTITTCCA 3'
10.	GQF	23mer	GQFKKTQI	310-317	Reverse complement 5'ATITGIGTITTITTIAAITGICC 3'
11.	GVN	29mer	GVNKDEGTAF	320-329	Reverse complement 5'AAIGCIGTICCITCITCITTITTIACICC 3'
12.	VYG	20mer	VYGAPGF	331-337	Reverse complement 5'AAICCIGGIGCICCITAIAC 3'

Single letter code for amino acids: A=alanine, C=cysteine, D=aspartic acid, E=glutamic acid, F=phenylalanine, G=glycine, H=histidine, I= isoleucine, K=lysine, L=leucine, M=methionine, N=asparagine, P= proline, Q=glutamine, R=arginine, S=serine, T=threonine, V=valine, W=tryptophan, Y=tyrosine. Probe sequence: N=A,C,G or T.

Table 2. Location of oligonucleotide probes in the amino acid sequence of human serum cholinesterase.

```
          .         .         .         .        50         .
EDDIIIATKNGKVRGMNLTVFGGTVTAFLGIPYAQPPLGRLRFKKPQSLTKWSDIWNATK
          .         .         .       100         .         .
YANSCCQNIDQSFPGFHGSEMWNPNTDLSEDCLYLNVWIPAPKPKNATVLIWIYGGGFQT
          .         .       150         .         .         .
GTSSLHVYDGKFLARVERVIVVSMNYRVGALGFLALPGNPEAPGNMGLFDQQLALQWVQK
          .       200         .         .         .         .
NIAAFGGNPKSVTLFGESAGAASVSLHLLSPGSHSLFTRAILQSGSFNAPWAVTSLYEAR
       250         .         .         .         .       300
NRTLNLAKLTGCSRENETEIIKCLRNKDPQEILLNEAFVVPYGTPLSVNFGPTVDGDFLT
          .         .         .         .       350         .
DMPDILLELGQFKKTQILVGVNKDEGTAFLVYGAPGFSKDNNSIITRKEFQEGLKIFFPG
          --------  ----------  ------          ------
            #10         #11       #12            #4

          .         .         .       400         .         .
VSEFGKESILFHYTDWVDDQRPENYREALGDVVGDYNFICPALEFTKKFSEWGNNAFFYY
                                                  ----------
                                                     #6
          .         .       450         .         .         .
FEHRSSKLPWPEWMGVMHGYEIEFVFGLPLERRDNYTKAEEILSRSIVKRWANFAKYGNP
---       ------#1,#2
          ------------------#7

          .       500         .         .         .         .
NETQNNSTSWPVFKSTEQKYLTLNTESTRIMTKLRAQQCRFWTSFFPKVLEMTGNIDEAE
                                                            --
       550         .         .    574
WEWKAGFHRWNNYMMDWKNQFNDYTSKKESCVGL
-------- ----------#9
 #8           ------#5
                 ------#3
```

Refer to Table 1 for the probe sequence corresponding to each probe #. A dot is above every tenth amino acid.

RESULTS

One positive clone was found out of 150,000 screened. The positive clone hybridized with at least 4 different probes. It hybridized with the 53mer probe (#7 in Table 1), the 39mer probe (#6 in Table 1), with one or more probes in the set containing #8, #10, and #12 , and with one or more probes in the set containing #4, #11, and #9. The positive clone was called Z3. Clone Z3 was plaque purified by replating it two more times. On the second replating, 4% of the plaques hybridizd with the 53mer probe, and on the third replating 100% of the plaques hybridized.

Clone Z3 was cultured in 1 liter of Y1088 bacteria, and the phage DNA purified. To determine the size of the human DNA insert, clone Z3 was digested with Kpn 1, Sst 1, Eco R1 and with combinations of these enzymes. The digested DNA was electrophoresed on 0.8% agarose, transferred to nitrocellulose, and hybridized with the 53mer probe. Fig. 1 shows that digestion with Eco R1 alone did not yield the human DNA insert. Eco R1 had been expected to cut out the human cDNA because the cDNA had been ligated to the vector via Eco R1 linkers. This would have given a minimum of 3 fragments: the cDNA fragment, the 19.6 kb left arm and the 24.1 kb right arm of lambda gt11. The two large fragments obtained after digestion with Eco R1 suggested that Eco R1 cut at only one end of the cDNA. The digestion with Kpn 1 and Eco R1 yielded a 2.4 kb fragment of which 1 kb belonged to lambda phage and 1.4 kb was human. The Kpn 1 and Sst 1 digest yielded a 3.4 kb fragment which contained 2 kb of lambda phage DNA and 1.4 kb of human DNA. The digestion with Eco R1 and Sst 1 verified the result already apparent from other digestions regarding the orientation of the Eco R1 site. The functional Eco R1 site was 1 kb away from the Sst 1 site, while the faulty Eco R1 site was 1 kb away from the Kpn 1 site. Another conclusion from Fig. 1 was that the 1.4 kb cDNA contained no internal Eco R1, Kpn 1 or Sst 1 sites.

Clone Z3 was ligated into M13mp19 for DNA sequencing. Two types of primers were used: the 15 b primer complementary to M13, and a 17 b oligonucleotide, probe #3 in Table 1, complementary to the cDNA. DNA sequencing gave approximately 450 nucleotides of sequence. To complete the sequence, clone Z3 was subfragmented with Sau3A1 and ligated into the isoschizomeric Bam H1 site of M13mp18. For DNA sequencing 60 clear plaques were selected at random. These 60 subclones contained 11 fragments of cholinesterase. Fig. 2 shows the location of the 6 Sau 3A1 restriction sites in clone Z3. The functional Eco R1 site was at nucleotide 580, and the faulty Eco R1 site was at nucleotide 2025. The faulty Eco R1 site had a single base change. It had the sequence AAATTC instead of the correct GAATTC. Despite the faulty

Eco R1 site, the junction between human and lambda phage DNA was recognized by the lac Z sequence which is part of lambda gt 11.

Fig. 2 shows that most of the clone was sequenced in both directions. The Sau 3A1 subclones provided the greatest part of the sequence. The uncut Z3 clone extended the sequence at the 5' end and overlapped the 3' coding strand with the 3' untranslated region. The alignment of the Sau 3A1 subclones was determined by comparing their DNA sequences to the DNA sequence of human serum cholinesterase predicted from its amino acid sequence. The computer was used for this purpose.

Table 3 shows the nucleotide sequence of clone Z3 and the corresponding amino acid sequence. The amino acid sequence predicted by clone Z3 is a perfect match to the amino acid sequence of human serum cholinesterase which we had determined by Edman degradation (1).

Clone Z3 contains 1440 nucleotides. These correspond to the coding sequence of cholinesterase from phenylalanine 195 through leucine 574 at the carboxyl terminus, and extend 302 bases into the 3' untranslated region. The stop codon immediately after leucine 574 is TAA.

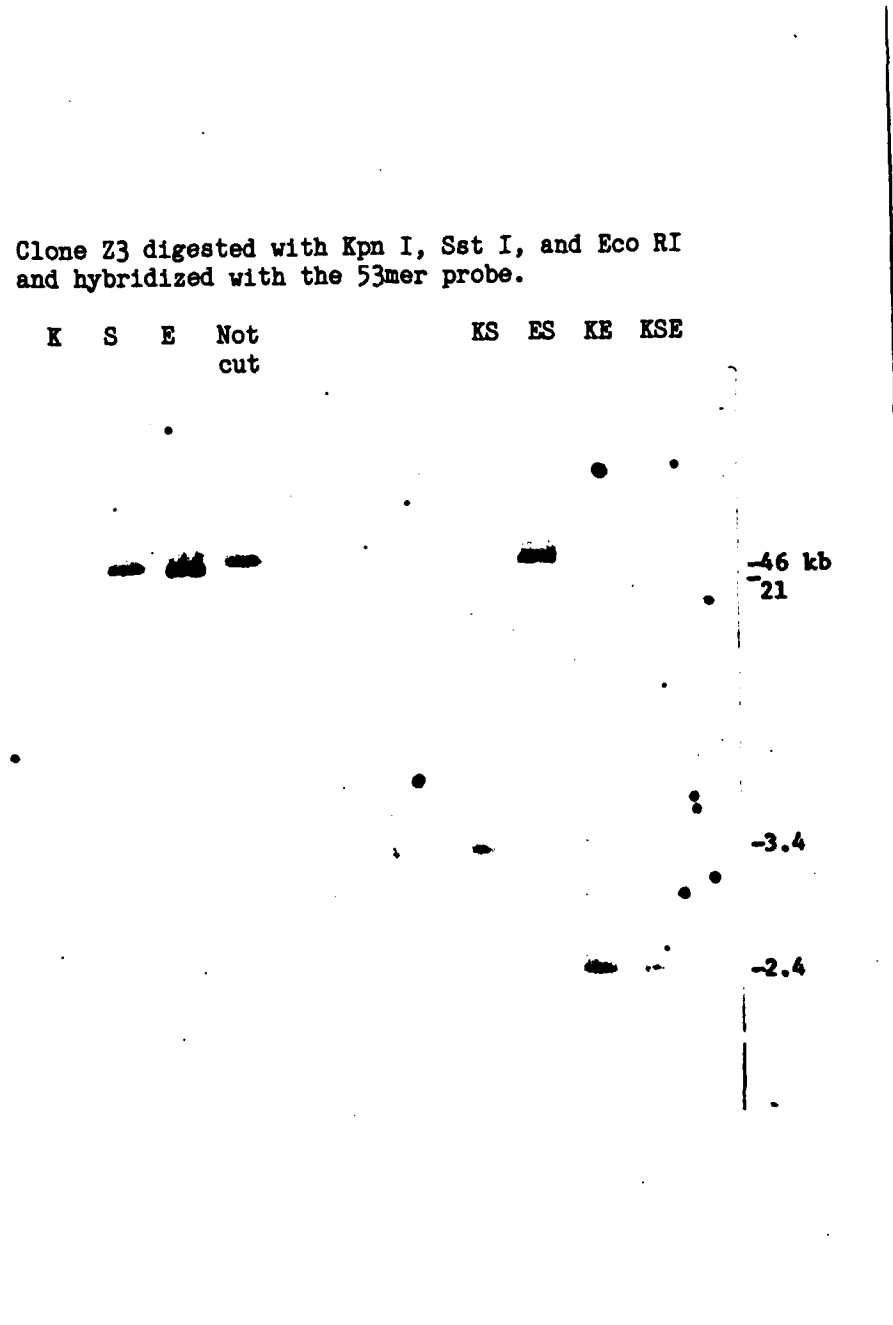


Fig.1. Southern blot of restriction digested clone Z3. Clone Z3 is in lambda gt11. Clone Z3 was digested with Kpn 1, Sst 1, Eco R1 and with combinations of these enzymes. The digests were electrophoresed on Ø.8% agarose, transferred to nitrocellulose, and hybridized with the 53mer probe.

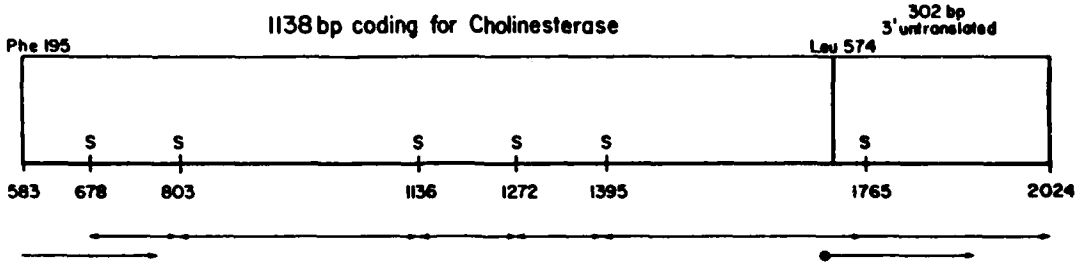


Fig.2. Restriction map of clone Z3 showing sequencing strategy. Clone Z3 contains 1440 bases. 1138 bases are in the coding region and these code for phenylalanine 195 through leucine 574 of human cholinesterase. 302 bases are in the 3' untranslated region. The 5' end of clone Z3 was defined by an Eco R1 site. The 3' end had a single base substitution in the Eco R1 site, and therefore was not cleaved by Eco R1. The location of six Sau 3A1 sites is indicated by S. Sau 3A1 subclones were sequenced in both directions. The uncut Z3 clone was sequenced using two different primers: a 15 b primer for M13, and a 17 b primer (probe #3 in Table 1) that corresponded to cholinesterase amino acids 557 to 562.

Table 3. DNA sequence of clone Z3 and derived amino acid sequence.

```
                        200
Phe Gly Glu Ser Ala Gly Ala Ala Ser Val Ser Leu His Leu Leu Ser
TTC GGA GAA AGT GCA GGA GCA GCT TCA GTT AGC CTG CAT TTG CTT TCT

Pro Gly Ser His Ser Leu Phe Thr Arg Ala Ile Leu Gln Ser Gly Ser
CCT GGA AGC CAT TCA TTG TTC ACC AGA GCC ATT CTG CAA AGT GGA TCC

Phe Asn Ala Pro Trp Ala Val Thr Ser Leu Tyr Glu Ala Arg Asn Arg
TTT AAT GCT CCT TGG GCG GTA ACA TCT CTT TAT GAA GCT AGG AAC AGA

Thr Leu Asn Leu Ala Lys Leu Thr Gly Cys Ser Arg Glu Asn Glu Thr
ACG TTG AAC TTA GCT AAA TTG ACT GGT TGC TCT AGA GAG AAT GAG ACT

Glu Ile Ile Lys Cys Leu Arg Asn Lys Asp Pro Gln Glu Ile Leu Leu
GAA ATA ATC AAG TGT CTT AGA AAT AAA GAT CCC CAA GAA ATT CTT CTG

Asn Glu Ala Phe Val Val Pro Tyr Gly Thr Pro Leu Ser Val Asn Phe
AAT GAA GCA TTT GTT GTC CCC TAT GGG ACT CCT TTG TCA GTA AAC TTT
                                    300
Gly Pro Thr Val Asp Gly Asp Phe Leu Thr Asp Met Pro Asp Ile Leu
GGT CCG ACC GTG GAT GGT GAT TTT CTC ACT GAC ATG CCA GAC ATA TTA

Leu Glu Leu Gly Gln Phe Lys Lys Thr Gln Ile Leu Val Gly Val Asn
CTT GAA CTT GGA CAA TTT AAA AAA ACC CAG ATT TTG GTG GGT GTT AAT

Lys Asp Glu Gly Thr Ala Phe Leu Val Tyr Gly Ala Pro Gly Phe Ser
AAA GAT GAA GGG ACA GCT TTT TTA GTC TAT GGT GCT CCT GGC TTC AGC

Lys Asp Asn Asn Ser Ile Ile Thr Arg Lys Glu Phe Gln Glu Gly Leu
AAA GAT AAC AAT AGT ATC ATA ACT AGA AAA GAA TTT CAG GAA GGT TTA

Lys Ile Phe Phe Pro Gly Val Ser Glu Phe Gly Lys Glu Ser Ile Leu
AAA ATA TTT TTT CCA GGA GTG AGT GAG TTT GGA AAG GAA TCC ATC CTT

Phe His Tyr Thr Asp Trp Val Asp Asp Gln Arg Pro Glu Asn Tyr Arg
TTT CAT TAC ACA GAA TGG GTA GAT GAT CAG AGA CCT GAA AAC TAC CGT
                                              400
Glu Ala Leu Gly Asp Val Val Gly Asp Tyr Asn Phe Ile Cys Pro Ala
GAG GCC TTG GGT GAT GTT GTT GGG GAT TAT AAT TTC ATA TGC CCT GCC

Leu Glu Phe Thr Lys Lys Phe Ser Glu Trp Gly Asn Asn Ala Phe Phe
TTG GAG TTC ACC AAG AAG TTC TCA GAA TGG GGA AAT AAT GCC TTT TTC

Tyr Tyr Phe Glu His Arg Ser Ser Lys Leu Pro Trp Pro Glu Trp Met
TAC TAT TTT GAA CAC CGA TCC TCC AAA CTT CCG TGG CCA GAA TGG ATG
```

Table 3. continued

```
Gly Val Met His Gly Tyr Glu Ile Glu Phe Val Phe Gly Leu Pro Leu
GGA GTG ATG CAT GGC TAT GAA ATT GAA TTT GTC TTT GGT TTA CCT CTG

Glu Arg Arg Asp Asn Tyr Thr Lys Ala Glu Glu Ile Leu Ser Arg Ser
GAA AGA AGA GAT AAT TAC ACA AAA GCC GAG GAA ATT TTG AGT AGA TCC

Ile Val Lys Arg Trp Ala Asn Phe Ala Lys Tyr Gly Asn Pro Asn Glu
ATA GTG AAA CGT TGG GCA AAT TTT GCA AAA TAT GGG AAT CCA AAT GAG

Thr Gln Asn Asn Ser Thr Ser Trp Pro Val Phe Lys Ser Thr Glu Gln
ACT CAG AAC AAT AGC ACA AGC TGG CCT GTC TTC AAA AGC ACT GAA CAA
    500
Lys Tyr Leu Thr Leu Asn Thr Glu Ser Thr Arg Ile Met Thr Lys Leu
AAA TAT CTA ACC TTG AAT ACA GAG TCA ACA AGA ATA ATG ACG AAA CTA

Arg Ala Gln Gln Cys Arg Phe Trp Thr Ser Phe Phe Pro Lys Val Leu
CGT GCT CAA CAA TGT CGA TTC TGG ACA TCA TTT TTT CCA AAA GTC TTG

Glu Met Thr Gly Asn Ile Asp Glu Ala Glu Trp Glu Trp Lys Ala Gly
GAA ATG ACA GGA AAT ATT GAT GAA GCA GAA TGG GAG TGG AAA GCA GGA

Phe His Arg Trp Asn Asn Tyr Met Met Asp Trp Lys Asn Gln Phe Asn
TTC CAT CGC TGG AAC AAT TAC ATG ATG GAC TGG AAA AAT CAA TTT AAC
                                          574
Asp Tyr Thr Ser Lys Lys Glu Ser Cys Val Gly Leu
GAT TAC ACT AGC AAG AAA GAA AGT TGT GTG GGT CTC taa tta ata gat
tta ccc ttt ata gaa cat att ttc ctt tag atc aag gca aaa ata tca
gga gct ttt tta cac acc tac taa aaa agt tat tat gta gct gaa aca
aaa atg cca gaa gga taa tat tga ttc ctc aca tct tta act tag tat
ttt acc tag cat ttc aaa acc caa atg gct aga aca tgt tta att aaa
ttt cac aat ata aag ttc tac agt taa tta tgt gca tat taa aac aat
ggc ctg gtt caa ttc ttt ctt tcc tta ata aat tta agt ttt tcc ccc
ca
```

Amino acids are numbered according to their location in the complete sequence of human serum cholinesterase.

DISCUSSION

We have succeeded in isolating a cDNA clone for human cholinesterase. Our strategy was based on our knowledge of the complete amino acid sequence of human serum cholinesterase. The amino acid sequence was determined in our laboratory and is not yet published, though it is in press (1). The amino acid sequence allowed us to prepare oligonucleotide probes to use for screening gene libraries. Our most successful probe was the unique sequence 53mer probe. It could be hybridized at a fairly high stringency of 46°C. The inosine probes, #8-#12 in Table 1, were the least useful because they hybridized at lower temperatures than the mixed oligonucleotide probes, #1-#5 in Table 1. Our 12 probes are clustered near the carboxy terminus portion of cholinesterase (see Table 2). This was partly by chance because the first probes synthesized (#1,2,3,4,6,7) had unknown locations within the complete sequence. Probes #5,8,9,10,11,12 were designed to be near the carboxy terminus because we expected this region to be better represented in a cDNA library.

We screened 4 different human cDNA libraries: a liver, muscle, lymphocyte cell line GM2131A, and a brain library. We found a cholinesterase clone only in the brain basal ganglia library, despite the fact that each of the other tissues has been reported to synthesize cholinesterase (14,15,16). It is possible that the other cDNA libraries also contained cholinesterase and that we missed the cholinesterase clone by chance.

Clone Z3 contained 1440 nucleotides. The amino acid sequence of this clone corresponded perfectly to amino acids 195-574 of human serum cholinesterase. It was especially reassuring to find asparagines in positions 241, 256, 341, 455, 481, and 486 because our amino acid sequencing paper had reported that these asparagines were glycosylated. Glycosylated asparagines do not show up during amino acid sequencing and therefore identification was by indirect arguments. Our cloning results support the conclusion that these are glycosylated asparagines.

The perfect match of the amino acid sequence predicted from clone Z3 with the amino acid sequence of human serum cholinesterase leads to the conclusion that human brain cholinesterase is identical to human serum cholinesterase, at least in the region of amino acids 195 to 574. This is the first evidence that cholinesterases from two different tissues are identical.

The only published nucleotide sequence of a cholinesterase is the sequence of acetylcholinesterase from the electric organ of Torpedo californica (17). No human acetylcholinesterase sequence

is available for comparison, nor any mammalian acetylcholinesterase. Acetylcholinesterase is considered to be distinct from cholinesterase, though the structural basis for the observed activity differences is unknown. Comparison of the amino acid sequences of acetylcholinesterase from Torpedo and cholinesterase from human serum shows an identity of 53.8%. Comparison of the nucleotide sequences shows an identity of approximately 55%. These are strikingly high degrees of identity when it is considered that the species are far apart in evolution, and that the two proteins are thought to be products of different genes. One might speculate that human acetylcholinesterase will be even more similar to human serum cholinesterase than Torpedo acetylcholinesterase, so that the identity may turn out to be as high as 70 to 80%.

Comparison of the 53mer probe region in Table 4 shows that the 18 amino acids in this probe are identical in Torpedo acetylcholinesterase and human cholinesterase except for the substitution of one isoleucine for methionine, which is a single base change. However, the Torpedo nucleotide sequence differs from the human sequence in 10 nucleotides. The 53mer probe turned out to be 94.3% correct for human cholinesterase but only 81% correct for Torpedo acetylcholinesterase. This suggests that frequency of codon usage in a particular species is an important consideration in designing a probe. Furthermore, it suggests that the amino acid sequence is more conserved than the nucleotide sequence.

FUTURE PLANS

We plan to use a portion of clone Z3 as a probe to rescreen the library. Approximately 120 bp from the 5' end of clone Z3 will be used to find the overlapping cDNA clone. The purpose will be to find the rest of the coding region for cholinesterase and to find the 5' untranslated region.

Table 4. Amino acid and nucleotide sequences of the 53mer probe, human serum cholinesterase and Torpedo acetylcholinesterase in the region corresponding to the 53mer probe.

Amino acid sequence

53mer	W	P	E	W	M	G	V	M	H	G	Y	E	I	E	F	V	F	G
ChE	W	P	E	W	M	G	V	M	H	G	Y	E	I	E	F	V	F	G
AChE	W	P	E	W	M	G	V	I	H	G	Y	E	I	E	F	V	F	G

Nucleotide sequence

		#mismatch
53mer	CCAAAGACAAATTCAATTTCATATCCATGCATGACTCCCATCCATTCAGGCCA	3
	* * *
ChE	CCAAAGACAAATTCAATTTCATAGCCATGCATCACTCCCATCCATTCTGGCCA	0
	* * * * * * * * * * *
AChE	CCGAAGACGAACTCAATCTCATAGCCGTGGATGACGCCCATCCACTCCGGCCA	10

53mer=unique sequence probe used for finding the ChE clone
ChE=human serum cholinesterase; amino acids 430 to 447
AChE=acetylcholinesterase from Torpedo californica;
amino acids 432-449

LITERATURE CITED

1. O.Lockridge, C.F.Bartels, T.A.Vaughan, C.K.Wong, S.E.Norton and L.L.Johnson. J.Biol.Chem. in press Jan.15, 1987 issue. Complete amino acid sequence of human serum cholinesterase.

2. M.Jaye, H.de la Salle, F.Schamber, A.Balland, V.Kohli, A. Findeli, P.Tolstoshev and J.P.Lecocq. Nucleic Acids Res. 11, 2325-2335 (1983). Isolation of a human anti-haemophilic factor IX cDNA clone using a unique 52-base synthetic oligonucleotide probe deduced from the amino acid sequence of bovine factor IX.

3. W.I.Wood, D.J.Capon, C.C.Simonsen, D.L.Eaton, J.Gitschier, B.Keyt, P.H.Seeburg, D.H.Smith, P.Hollingshead, K.L.Wion, E.Delwart, E.G.D.Tuddenham, G.A.Vehar and R.M.Lawn. Nature 312, 330-337 (1984). Expression of active human factor VIII from recombinant DNA clones.

4. R.Grantham, C.Gautier, M.Gouy, M.Jacobzone and R.Mercier. Nucleic Acids Res. 9, r43-r74 (1981). Codon catalog usage is a genome strategy modulated for gene expressivity.

5. E.Ohtsuka, S.Matsuki, M.Ikehara, Y.Takahashi and K.Matsubara. J.Biol.Chem. 260, 2605-2608 (1985). An alternative approach to deoxyoligonucleotides as hybridization probes by insertion of deoxyinosine at ambiguous codon positions.

6. J.Kamholz, F.DeFerra, C.Puckett and R.Lazzarini. Proc. Natl.Acad.Sci. USA 83, 4962-4966 (1986). Identification of three forms of human myelin basic protein by cDNA cloning.

7. W.D.Benton and R.W.Davis. Science 196, 180-182 (1977). Screening lambda gt recombinant clones by hybridization to single plaques in situ.

8. T.Maniatis, E.F.Fritsch and J.Sambrook. Molecular Cloning, A Laboratory Manual. Cold Spring Harbor, New York, 1982, p.322.

9. R.A.Young and R.W.Davis. Science 222, 778-782 (1983). Yeast RNA polymerase II genes: isolation with antibody probes.

10. D.A.Johnson, J.W.Gautsch, J.R.Sportsman and J.H.Elder. Gene Anal.Techn. 1, 3-8 (1984). Improved technique utilizing nonfat dry milk for analysis of proteins and nucleic acids transferred to nitrocellulose.

11. D.J.Law. Ph.D. Thesis, 1981, Wayne State Univ., Detroit, MI. Developmental regulation of laccase accumulation in Aspergillus nidulans and attempts to clone the laccase gene.

12. E.M.Southern. J.Mol.Biol. 98, 503-517 (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis.

13. F.Sanger, A.R.Coulson, B.J.Barrell, A.J.H.Smith and B.A. Roe. J.Mol.Biol. 13, 161-178 (1980). Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing.

14. A.Silver. The Biology of Cholinesterases. North-Holland Publishing Co., Amsterdam, 1974, p.440.

15. S.C.Sung. Comp.Biochem.Physiol. 73C, 33-36 (1982). Acetylcholinesterase and butyrylcholinesterase released from normal and dystrophic muscles by treatment with proteolytic enzymes.

16. H.M.Rubinstein, T.Lubrano, H.L.Mathews, C.F.Lange, S.Silberman, E.M.Adams and J.Minowada. Biochem.J. 224, 1171-1175 (1984). A lymphocyte cell line that makes serum cholinesterase instead of acetylcholinesterase.

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

END

DATE

FILMED

6-88

DTIC