Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 3359–3363, May 1987
Genetics

Molecular cloning and characterization of esterase-6, a serine hydrolase of *Drosophila*

(peptide sequencing/synthetic oligonucleotide/*in situ* hybridization/transcript analysis/cDNA sequencing)

J. G. OAKESHOTT*†, C. COLLET‡, R. W. PHILLIS‡, K. M. NIELSEN‡, R. J. RUSSELL*†, G. K. CHAMBERS§¶, V. ROSS*, AND R. C. RICHMOND‡||

*Research School of Biological Sciences, Australian National University, P.O. Box 475 Canberra City, ACT 2601 Australia; ‡Department of Biology and Institute for Molecular and Cellular Biology, Indiana University, Bloomington, IN 47405; and §Museum of Comparative Zoology, The Agassiz Museum, Harvard University, Cambridge, MA 02138

Communicated by Bruce Wallace, January 30, 1987

ABSTRACT **The *Est-6* gene of *Drosophila melanogaster* was cloned by screening libraries with synthetic oligonucleotides corresponding to tryptic peptides from purified esterase-6 (Est-6) protein. cDNA clones were isolated that hybridized *in situ* to the site of *Est-6* on chromosome 3 at 69A1. Inserts in putative *Est-6* cDNA clones were 1.85 kilobases (kb) long, and blot hybridization analysis of electrophoretically fractionated RNA, using a cDNA clone as a probe, revealed two transcripts, of 1.68 and 1.83 kb. The two transcripts showed the same developmental profile as the Est-6 protein. Neither transcript was detected in an Est-6-null line. The cDNA fragment was homologous to a 2.3-kb *Eco*RI–*Bam*HI fragment in genomic clones, and this region was interrupted by the 8-kb *B104* transposable element in the Est-6-null line. Conceptual translation of the cDNA sequence revealed a protein of 548 residues with 19% sequence similarity to acetylcholinesterase from the *Torpedo* ray.**

Esterase-6 (Est-6; carboxylic-ester hydrolase, E.C. 3.1.1.1), the major β-carboxylesterase of *Drosophila melanogaster*, is produced as a monomer in the anterior ejaculatory duct of the male reproductive system and plays an important role in its reproductive biology (1, 2). There is considerable polymorphism at the *Est-6* locus in natural populations for activity and electrophoretic mobility of the enzyme, and at least some of this variation is subject to natural selection (3). Moreover, there are qualitative differences between *D. melanogaster* and its sibling species in the structure and expression of Est-6 (4).

This paper describes the cloning and initial characterization of the *Est-6* gene of *D. melanogaster*. This opens the way for a study of the molecular basis of the evolution of *Est-6* structural and regulatory sequences in *Drosophila*, although this report focuses on the more distant origins and relationships of the locus. In particular, the *Est-6* cDNA sequence has been analyzed for sequence similarity with members of the serine hydrolase family of enzymes.

The serine hydrolases are presently defined as a functionally related class of hydrolytic enzymes containing a serine residue in their active site (5). The class comprises the relatively well-characterized serine protease multigene family (6) as well as various carboxyl-, choline-, and aliesterases (7, 8). Despite their functional and structural similarities, it is not clear whether any of the various esterases are genetically related to each other or to the serine proteases, or whether their similarities are the outcome of convergent evolution.

The serine protease genes that have been cloned and sequenced show extensive regions of sequence similarity, including a characteristic Asp-Ser-(His/Gly) tripeptide at the active site (6, 9). Nucleotide sequence data have only been reported for one eukaryotic esterase, acetylcholinesterase from the ray *Torpedo californica* (8). This enzyme shows very little sequence similarity to the serine proteases and contains a Glu-Ser-Ala tripeptide at the active site.

The sequences of up to 20 amino acids around the active site have been reported for several esterases (refs. 11 and 12; see also Table 1). Each of these peptides has the Glu-Ser-Ala tripeptide found in acetylcholinesterase and also shares at least six of eight residues immediately around and including the active-site serine. Three of these eight residues are also found in the consensus sequence of the serine proteases. Thus there is limited evidence for a relationship among the eukaryotic esterases, but only equivocal evidence that they are related to the serine proteases.

Our strategy for cloning *Est-6* involved the sequencing of tryptic peptides from purified Est-6 protein, the synthesis of corresponding oligonucleotides, and the screening of a *Drosophila* cDNA library with the oligonucleotide probes. The identity of putative *Est-6* clones was confirmed by *in situ* hybridization to polytene chromosomes, developmental RNA blot hybridization analysis, DNA and RNA blot hybridization analyses of an Est-6-null strain, and DNA sequencing.

MATERIALS AND METHODS

Protein Purification, Peptide Sequencing, and Oligonucleotide Synthesis. Est-6 protein was purified from 3- to 5-day-old males of an Oregon R strain of *D. melanogaster*, using the procedures of Mane *et al.* (1). Peptide mapping was carried out according to the methods of Chambers *et al.* (13), and the amino acid compositions of 56 tryptic peptides were determined using a Beckman 121 MB amino acid analyzer after hydrolysis with 6 M HCl for 18 hr at 110°C *in vacuo*. Six peptides with the lowest predicted codon degeneracy were sequenced using an Applied Biosystems 470A gas-phase amino acid sequencer.

Oligonucleotides were synthesized, using an Applied Biosystems DNA synthesizer, to correspond to all or a portion of three sequenced peptides which would provide minimum codon degeneracy (14). Deoxyinosine was inserted to further reduce redundancies in two of the three oligonucleotides (15). As is clear in Fig. 1, it was still necessary to prepare each oligonucleotide as an equal mixture of 32 or 64 related sequences to cover the remaining degeneracies.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Est-6, esterase-6 enzyme; *Est-6*, esterase-6 locus.

†Present address: CSIRO Division of Entomology, Black Mountain, Box 1700 Canberra City, ACT 2601, Australia.
¶Present address: Department of Biochemistry, Victoria University of Wellington, Private Bag, Wellington, New Zealand.
||To whom reprint requests should be addressed.

Nucleic Acid Preparations. Genomic DNA from adult flies, bacteriophage λ DNA, and plasmid DNA were prepared and analyzed by standard procedures (16). Total cellular RNA was extracted from several different life stages of the Canton-S strain of *D. melanogaster* and from adults of the Est-6-null, Dm 100 strain (17), by the guanidinium thiocyanate/cesium chloride procedure of Chirgwin *et al.* (18). Poly(A)$^+$ RNA was fractionated from total RNA preparations by affinity chromatography using oligo(dT)-cellulose (16). RNA was denatured in glyoxal and dimethyl sulfoxide prior to electrophoresis in horizontal 1% agarose gels prepared in 10 mM sodium phosphate (pH 7.0).

Hybridizations. A *D. melanogaster* adult male cDNA library (in λgt10; a gift of T. Kornberg and L. Kauvar, University of California, San Francisco) was screened with the oligonucleotide probes. Replicate filters of the library were screened by hybridization under low-stringency conditions (37°C, 6× SSC; 1× SSC is 0.15 M NaCl/15 mM sodium citrate, pH 7.0) with 5′-end-labeled oligonucleotides (16, 19); about 2×10^7 plaque-forming units were screened.

A library of genomic DNA from *D. melanogaster* (a partial *Sau*3A digest of the Dm 145 Est-6^s strain cloned into λ EMBL4) was subsequently screened with putative *Est-6* sequences isolated during the oligonucleotide screening above. Double-stranded DNA probes were prepared by nick-translation, and filter-hybridization was carried out under standard conditions (16).

DNA from positive clones was labeled with biotin by nick-translation (20) and hybridized *in situ* to salivary gland polytene chromosomes of Oregon R third-instar larvae (21). Hybridization was detected using streptavidin-horseradish peroxidase (Bethesda Research Laboratories) stained with 3,3′-diaminobenzidine.

For DNA (Southern) blot hybridizations, DNA was transferred from agarose gels to nitrocellulose filters and hybridized as described by Maniatis *et al.* (16).

For RNA blot hybridizations, RNA was transferred to Zeta-Probe membranes (Bio-Rad) and prehybridized and hybridized at 60°C in 1.5× SSPE (1× SSPE is 0.18 M NaCl/0.01 M sodium phosphate, pH 7.0/1 mM EDTA)/0.5% skim milk powder/0.5% $NaDodSO_4$/50% (vol/vol) formamide containing carrier DNA (500 μg/ml), yeast carrier RNA (500 μg/ml), and poly(A) (20 μg/ml). Membranes were washed at 65°C with 0.1× SSC/0.1% $NaDodSO_4$. Single-stranded RNA probes were prepared by subcloning putative *Est-6* fragments into pGemini vectors (Promega Biotec) and synthesizing radiolabeled transcripts by use of SP6 RNA polymerase (22).

DNA Sequencing. Both strands of the putative *Est-6* insert were sequenced by the dideoxy chain-termination method (23). *Eco*RI–*Pst* I and *Eco*RI–*Bam*HI fragments and partial *Sau*3A and *Taq* I digests were subcloned into appropriately cut M13mp18 or M13mp19 vectors.

RESULTS

Isolation of Clones. Amino acid sequences were determined for six tryptic peptides from purified Est-6 protein, and oligonucleotides were synthesized to correspond to five or more consecutive amino acids in three of the peptides (ETP-5, -13, and -19; Fig. 1). The oligonucleotides were used to screen a library of *D. melanogaster* adult male cDNA.

Two putative *Est-6* clones were isolated that hybridized to all three oligonucleotide probes and contained identical 1.85-kilobase (kb) inserts. The chromosomal site to which the insert hybridized was defined more precisely by *in situ* hybridization to polytene chromosomes prepared from larvae heterozygous for standard chromosomes (Oregon R) and one of three deficiencies in the *Est-6* region: *vin*^{6} (*68C8-11–69A4-5*) and *BK9* (*68E2–69B1*), which delete *Est-6*, and *vin*^{5} (*68A3–69A1*), which does not (ref. 24; R.C.R. and R.W.P., unpublished data). The putative *Est-6* clones hybridized to the *vin*^{5} but not *vin*^{6} or *BK9* chromosomes, and the region of hybridization was localized to 69A1 (data not shown), in agreement with the location of *Est-6* determined by the deficiency mapping of Akam *et al.* (24). Southern hybridization of the inserts in the *Est-6* clones to genomic DNA showed that they were homologous to the same single-copy genomic sequence (data not shown).

The 1.85-kb insert in the cDNA clone was used as a probe to isolate three overlapping clones spanning 35 kb of genomic DNA from the *D. melanogaster* Dm 145 library. The cloned genomic DNA only hybridized to 69A1 on polytene chromosomes, and Southern analysis indicated that homology to the insert was confined to a 2.3-kb region bounded by *Eco*RI and *Bam*HI sites (Fig. 2). The similarity of the *Pst* I and *Bam*HI restriction sites in the 5′ region of the cDNA and genomic clones suggests that there are no or only small introns in this region. We have no evidence to exclude the possibility of one or more small introns in the 3′ portion of *Est-6*.

RNA Blot Analysis. After agarose gel electrophoresis, blot-transfers of RNA from several life stages of the Canton-S strain of *D. melanogaster* and from adult males of the Est-6-null strain were hybridized to single-stranded RNA probes derived from each strand of the 1.85-kb cDNA fragment (Fig. 3). No specific hybridization was detected with one strand, but two poly(A)$^+$ transcripts, of about 1.68

ETP 5 Ser-Asp-Thr-Asp-Asp-Pro-Leu-Leu-Val-Gln-Leu-Pro-Gln-Gly-Lys
Oligo: CT^G_A TGX CT^G_A CT^G_A GG

ETP 13 Glu-Asp-Phe-Gly-Gln-Leu-Ala-Arg
Oligo: CT^C_T CTI AAI CCI GT^C_T IA^C_I CG^C_I IC^C_I

ETP 16 Val-Gly-Asp-Glu-Ser-Ala-Glu-Asp-Ser-(Ser-Gly)-Lys-Lys

ETP 19 Ser-Pro-Ala-Tyr-Ala-Tyr-Val-Tyr-Asp-Asn-Pro-Ala-Glu-Lys
Oligo: CG^I_C ATI CG^I_C ATI CA^I_C ATI CTI TTI GG^I_C CG^I_C CT^I_C TT

ETP 22 Asp-Thr-Lys

ETP 28 Ala-Phe-Glu-Leu-Gly-Arg

FIG. 1. Amino acid sequences of tryptic peptides from purified Est-6. The sequences of three synthetic oligonucleotides are given beneath the corresponding peptides. I, deoxyinosine.

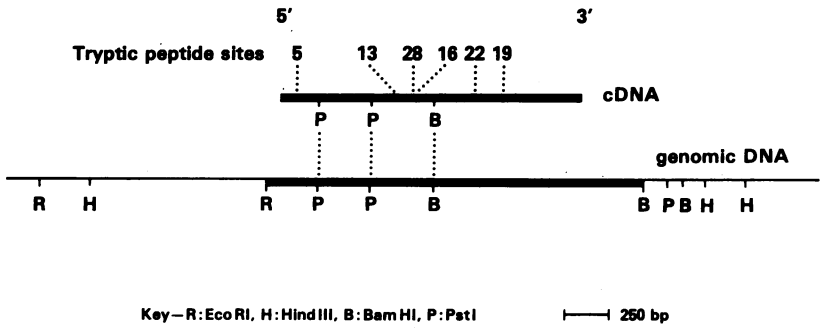

FIG. 2. Restriction map of the *Est-6* region showing the direction of transcription, the location of sequences encoding the six tryptic peptides in the cDNA, and alignment of some of the homologous restriction sites in the genomic and cDNA. The solid bar on the genomic DNA defines the region of homology to the cDNA. bp, Base pairs.

kb and 1.83 kb, were detected with the other strand. The larger of these transcripts is very similar in length to the size of the insert in the cDNA clones, suggesting that the cDNAs are, if not full-length, nearly full-length copies of *Est-6* message. The direction of transcription implied by the strand specificity of hybridization is indicated in Fig. 2 and confirmed by the location of the small poly(A) tract in the cDNA sequence (see below). The developmental profiles of the two transcripts are congruent with that of Est-6 protein (17). However, only one major protein form with Est-6 activity has been mapped to the *Est-6* locus (1), and it remains unclear whether both or only one of the transcripts is translated into this protein.

The 1.68- and 1.83-kb transcripts are both absent in adult males of the Est-6-null strain of *D. melanogaster* (Fig. 3). Subsequent Southern analysis of this strain (data not shown) indicates an insertion of the 8-kb *B104* transposable element (25) adjacent to the *Bam*HI site in the middle of the gene (Fig. 2).

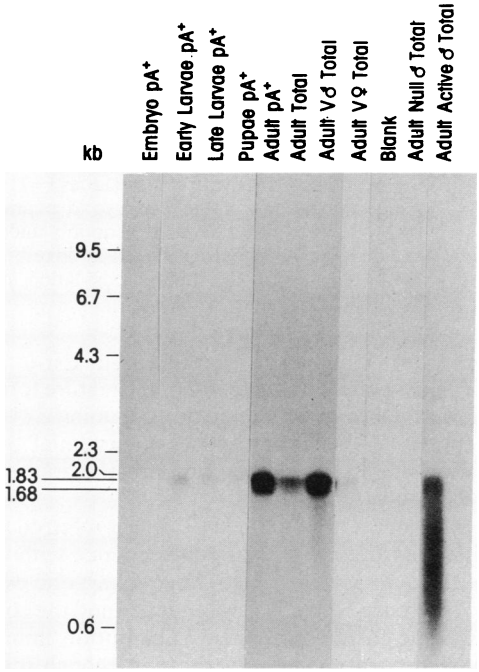

FIG. 3. Blot hybridization analysis of transcripts homologous to Est-6 cDNA. The first eight lanes give a developmental profile in wild-type *D. melanogaster* (pA^+, $poly(A)^+$ RNA; total, total RNA; V♂ and V♀, virgin male and female). The rightmost lanes compare total RNA of adult males from Est-6-null and wild-type (active) strains. Embryos had aged 0–16 hr after egg laying. Early larvae were early second instar, and late larvae were mid third instar. Pupae were collected 1–3 days after pupariation, and adults, 3–7 days after eclosion. All stages were reared at 22°C. Markers at left show positions and sizes (kb) of *Hin*dIII restriction fragments of bacteriophage λ DNA used as standards and of the 1.83- and 1.68-kb transcripts that hybridized to the probe.

Sequence Analysis. Fig. 4 presents the full sequence of the cDNA fragment and the inferred amino acid sequence derived from its only long open reading frame. The amino acid sequence encoded by the open reading frame probably begins at position 25 and ends at 1669. A 17-nucleotide poly(A) tail begins at 1820. The consensus poly(A) signal sequence, AAUAAA (26), begins at nucleotide 1799. The inferred amino acid sequence includes peptides similar to all six tryptic peptides sequenced from purified Est-6 protein. Only 4 of the 59 residues in the tryptic peptides differ from those deduced from the cDNA.

The first 21 residues (−21 through −1) of the inferred amino acid sequence probably define a signal peptide. This is suggested by the fact that one of the tryptic peptides (ETP 5) begins at residue 1, although the serine at this position does not constitute a site for cleavage by trypsin. Several characteristics of the putative signal peptide—its length, hydrophilic residues near the beginning and end, a hydrophobic core, and a final residue with a short side chain—are commonly found in signal peptides (27).

Mature Est-6 protein probably extends to residue 527, which precedes the first termination codon in this reading frame. This sequence would result in a nonglycosylated protein of M_r 59,380, in accordance with previous estimates (1). Potential N-linked glycosylation sites occur at residues −2, 21, 399, 435, and 485 in the inferred amino acid sequence (28).

Active-site peptides from a series of eukaryotic esterases all contain the sequence Gly-Glu-Ser-Ala (Table 1). This sequence is not found in its entirety in the inferred Est-6 sequence, although the tripeptide Glu-Ser-Ala is found at residues 242–244 within tryptic peptide ETP 16 (Fig. 4). However, the residues immediately surrounding this tripeptide show no homology with active-site sequences from the other esterases. On the other hand, a longer region with substantial homology to the active sites of other esterases is found at residues 186–193 (Table 1). Comparison of this region with a consensus sequence for the active site peptides from nine esterases shows only one mismatch in eight residues, a histidine for glutamic acid difference at residue 187. Although histidine is basic and glutamate is acidic, their side-chain hydration potentials are very similar (29). Moreover, similar hydropathies and secondary structures are predicted (30) for this region in Est-6 and *Torpedo* acetylcholinesterase.

DISCUSSION

Six lines of evidence show that we have cloned the *Est-6* gene. First, the cDNA clone originally isolated hybridized to all three *Est-6* oligonucleotide probes. Second, the cDNA clones and homologous genomic clones hybridized *in situ* to the cytological site of the *Est-6* gene. Third, the developmental profile of RNA complementary to the cDNA clone parallels that of Est-6 activity. Fourth, flies from an Est-6-null strain lacked this RNA and have an 8-kb insertion in homologous genomic DNA. Fifth, the cDNA fragment contains sequences that would encode peptides similar to all six

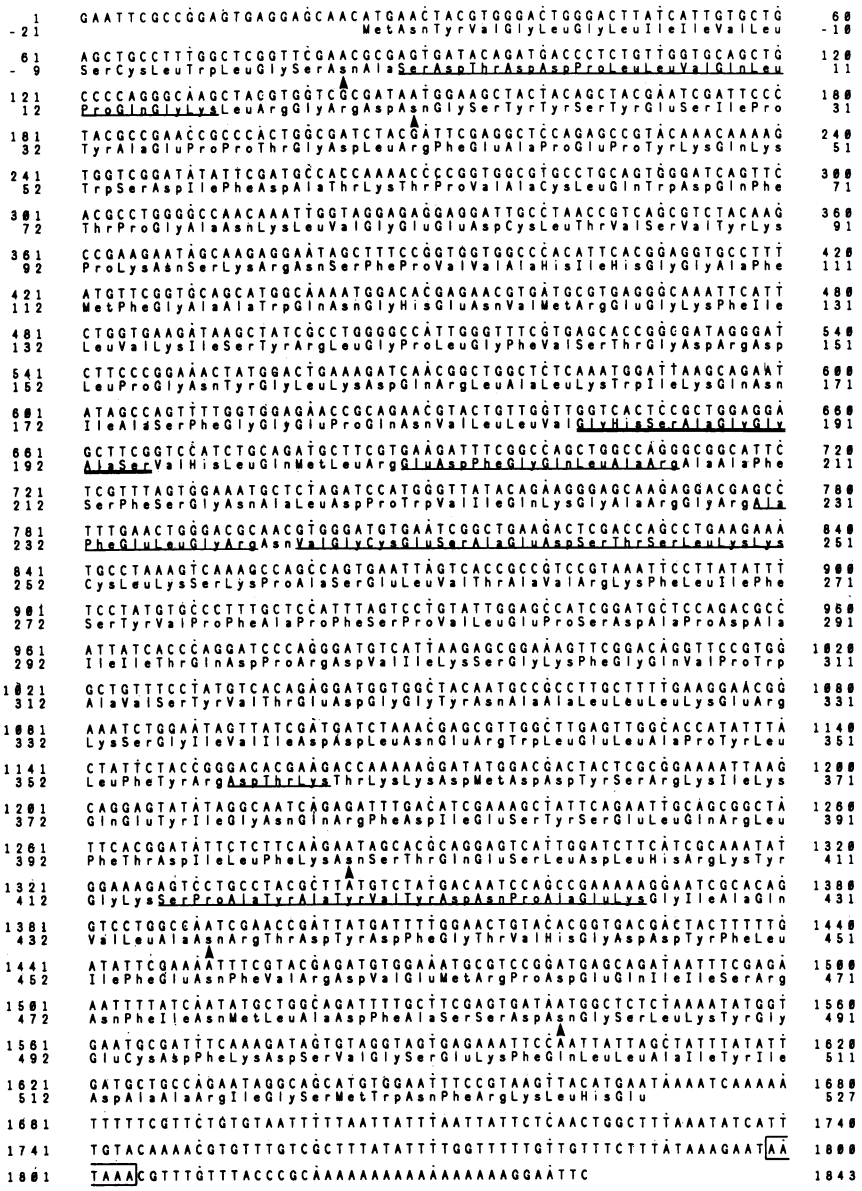

FIG. 4. Nucleotide and inferred amino acid sequences for Est-6 cDNA. The proposed signal peptide covers the first 21 residues (−21 to −1). The locations of the six tryptic peptides are underlined, the active site is doubly underlined, the poly(A) signal AATAAA is boxed, and the five potential N-linked glycosylation sites are indicated by arrowheads.

Est-6 tryptic peptides sequenced, although corresponding oligonucleotides had been prepared from only three of these peptides. Sixth, a region of the inferred amino acid sequence of Est-6 has pronounced homology to the active-site peptides of known esterases.

Developmental RNA blot analysis using the *Est-6* cDNA clone as a probe revealed two poly(A)$^+$ RNA species, of 1.68 and 1.83 kb. The shorter transcript is less abundant, although both have the same developmental profile. Both transcripts are likely to derive from the same gene because genomic Southern analysis at reduced hybridization stringency (55°C) did not reveal additional homologies. One explanation for the two transcripts involved the use of alternative polyadenylylation sites; in addition to the poly(A) signal sequence at nucleotide 1799, another consensus signal sequence occurs at position 1667 just prior to the termination codon. The distance between the two sequences is similar to the difference between the two transcript sizes, and there is precedent for a signal sequence prior to the termination codon in a eukaryotic system (31). Alternative splicing of a primary transcript or the use of two promoters could also account for the two message sizes.

The cDNA sequence contains a long open reading frame coding for a protein of 548 residues that includes the six tryptic peptides sequenced from purified Est-6 protein. Analysis of this sequence suggests that the mature form of the enzyme is derived from a proenzyme containing a 21-residue signal peptide (27). The presence of potential N-linked glycosylation sites in the enzyme supports previous evidence that Est-6 is a glycoprotein (1).

The identity of the active site in Est-6 was inferred from the strong similarity of an eight-residue sequence with the consensus of the active sites in nine other eukaryotic esterases. The only difference between this sequence and the consensus was a histidine for glutamic acid substitution immediately adjacent to the central serine residue. Although this difference has little effect on predicted hydropathies and secondary structures, it may affect the net charge and substrate specificity of the region. A second feature of the Est-6 active site concerns the alanine residue at position 192. This residue also occurs in one aliesterase and three cholinesterases. The five other eukaryotic carboxylesterases characterized have glutamic acid or isoleucine at this position (Table 1). The functional significance of this difference is uncertain, but the evolutionary implications are that Est-6 may be more closely related to mammalian cholinesterases and aliesterases than to mammalian carboxylesterases.

Table 1. Comparison of the presumptive active-site peptide of Est-6 with known active sites of other serine hydrolases

Enzyme	Sequence
Esterases	
Drosophila Est-6	N V L L V G H S A G G A S V H L Q M L R
Torpedo acetylcholinesterase	T V T I F G E S A G G A S V G M H I L S
Eel acetylcholinesterase	G G E S S E G A A G
Horse aliesterase	F G E S A G A A S
Human cholinesterase	S V T L F G E S A G A A S V S L H L L S
Carboxylesterase	
Chicken	G E S A G G I S
Horse	G E S A G G(E S)
Pig and sheep	G E S A G G E S
Ox	G E S A G A E S
Consensus	G E S A G G A S
Selected proteases	
Bovine trypsin	G K D S C Q G D S G G P V V C S G K
Drosophila trypsin-like	G K D A C Q G D S G G P L V S G G V
Streptomyces trypsin-like	G V D T C Q G D S G G P M F R K D N
Hornet chymotrypsin	G E G A C H G D S G G P L V A N G V
Crab collagenase	G K G T C D G D S G G P L N Y D G L
Pig elastase	V R S G C Q G D S G G P L H C L V N

Data for *Torpedo* acetylcholinesterase and the *Drosophila* trypsin-like protein are taken from Schumacher *et al.* (8) and Davis *et al.* (9). All other data are from Dayhoff (11) and references therein. Standard one-letter amino acid symbols are used.

The relationship of Est-6 to the mammalian cholinesterases is likely to be a distant one. A comparison of *Torpedo* acetylcholinesterase (8) and Est-6 reveals only 19% overall amino acid sequence similarity. There are two local regions of greater similarity; one 20-residue sequence from positions 30 to 49 in Est-6 has 65% sequence similarity to residues 33–52 in acetylcholinesterase (8); a second region, of 151 residues, which contains the presumptive active site and runs from positions 132 to 282, shows 43% similarity with residues 143–295 in *Torpedo* acetylcholinesterase.

Searches of the National Biomedical Research Foundation** and Genbank†† data bases revealed only one other protein, bovine thyroglobulin (10), with significant sequence similarity to Est-6. The specific regions of Est-6 responsible for the similarity are the same as those similar to acetylcholinesterase, although in this case the active site is not conserved.

None of several serine proteases in the data bases show significant overall sequence similarity to Est-6 or acetylcholinesterase. There is direct evidence for limited local similarity around the active site (Table 1), but it is clear that Est-6 and acetylcholinesterase are at most only distant relatives of the serine protease multigene family.

**Protein Identification Resource (1987) Protein Sequence Database (Natl. Biomed. Res. Found., Washington, DC), Release 12.0.

††National Institutes of Health (1986) Genetic Sequence Databank: GenBank (Research Systems Div., Bolt, Beranek, and Newman, Cambridge, MA), Tape Release 44.0.

Dr. Matthew White purified Est-6. We thank Sanya Collet and Donna Rigby for expert technical assistance. James Brady, Peter Cooke, Nick Deacon, and Anne Game provided advice and assistance. Dr. Douglas Cavener made valuable comments for improvement of the manuscript. This work was supported in part by grants from the National Science Foundation (BSR 83-14910 and INT 84-13637) and the National Institutes of Health (AG02035). R.W.P. is the recipient of a National Institutes of Health Predoctoral Fellowship in Genetics (GM7757).

1. Mane, S. D., Tepper, C. S. & Richmond, R. C. (1983) *Biochem. Genet.* **21**, 1019–1040.
2. Scott, D. S. (1986) *Evolution* **40**, 1084–1091.
3. Anderson, P. R. & Oakeshott, J. G. (1984) *Nature (London)* **308**, 729–731.
4. Morton, R. A. & Singh, R. S. (1985) *Biochem. Genet.* **23**, 959–973.
5. Heymann, E. (1980) in *Enzymatic Basis of Detoxification*, ed. Jakoby, W. B. (Academic, New York), pp. 291–323.
6. Rogers, J. (1985) *Nature (London)* **315**, 458–459.
7. Augustinsson, K. B. (1968) in *Homologous Enzymes and Biochemical Evolution*, eds. Thai, N. V. & Roche, J. (Gordon & Breach, New York), pp. 299–311.
8. Schumacher, M., Camp, S., Maulet, Y., Newton, M., MacPhee-Quigley, K., Taylor, S. S., Friedmann, T. & Taylor, P. (1986) *Nature (London)* **319**, 407–409.
9. Davis, C. A., Riddell, D. C., Higgins, M. J., Holden, J. J. A. & White, B. N. (1985) *Nucleic Acids Res.* **13**, 6605–6619.
10. Merken, L., Simmons, M. J., Swillens, S., Massaer, M. & Vassart, G. (1985) *Nature (London)* **316**, 647–650.
11. Dayhoff, M. O., ed. (1972) *Atlas of Protein Sequence and Structure*, (National Biomedical Research Foundation, Washington, DC), Vol. 5.
12. MacPhee-Quigley, K., Taylor, P. & Taylor, S. (1985) *J. Biol. Chem.* **260**, 12185–12189.
13. Chambers, G. K., Felton, A. A., Ramshaw, J. A. M., Rigby, D. L. & Sullivan, D. T. (1985) *Biochem. Genet.* **23**, 801–814.
14. Matteuci, M. D. & Caruthers, M. H. (1981) *J. Am. Chem. Soc.* **103**, 3185–3191.
15. Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y. & Matsubara, K. (1985) *J. Biol. Chem.* **260**, 2605–2608.
16. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).
17. Sheehan, K. B., Richmond, R. C. & Cochrane, B. J. (1979) *Insect Biochem.* **9**, 443–450.
18. Chirgwin, J. M., Przybyla, A. E., McDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
19. Woods, D. (1984) in *Focus* (Bethesda Research Laboratories, Rockville, MD), Vol. 6, No. 3, pp. 1–3.
20. Langer-Safer, P. R., Levine, M. & Ward, D. C. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4381–4385.
21. Pardue, M. L. & Gall, J. G. (1975) *Methods Cell. Biol.* **10**, 1–25.
22. Melton, D. A., Kreig, P. A., Rebagliati, M. R., Maniatis, T., Zinn, T., Zinn, K. & Green, M. R. (1984) *Nucleic Acids Res.* **12**, 7035–7056.
23. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
24. Akam, M. E., Roberts, D. B., Richards, G. P. & Ashburner, M. (1978) *Cell* **13**, 215–225.
25. Scherer, G., Tschudi, C., Perera, J. & Delius, H. (1982) *J. Mol. Biol.* **157**, 435–451.
26. Birnstiel, M. L., Busslinger, M. & Strub, K. (1985) *Cell* **41**, 349–359.
27. Carne, T. & Scheele, G. (1982) *J. Biol. Chem.* **257**, 4133–4140.
28. Bause, E. (1983) *Biochem. J.* **209**, 331–336.
29. Wolfenden, R. V., Cullis, P. M. & Southgate, C. C. F. (1979) *Science* **206**, 575–577.
30. Kyte, J. & Doolittle, R. J. (1982) *J. Mol. Biol.* **157**, 105–132.
31. Bishop, D. F., Calhoun, D. H., Bernstein, H. S., Hantzopoulos, M. Q. & Desnick, R. J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4859–4863.