PHARMACOLOGY LETTERS

Accelerated Communication

CLONING AND SEQUENCING OF A
HUMAN LIVER CARBOXYLESTERASE ISOENZYME

Rochelle M. Long[1,3], Maria R. Calabrese[1], Brian M. Martin[2], and Lance R. Pohl[3]

[1]Department of Pharmacology and Toxicology, School of Pharmacy, University of
Maryland, Baltimore, MD 21201, [2]Clinical Neurosciences Branch,
National Institute of Mental Health, Bethesda, MD 20892, and
[3]Laboratory of Chemical Pharmacology, National Heart, Lung, and Blood
Institute, National Institutes of Health, Bethesda, MD 20892

Abstract A human liver lambda gt11 library was screened with antibodies raised
to a purified rat liver carboxylesterase, and several clones were isolated and
sequenced. The longest cDNA contained an open reading frame of 507 amino acids
that represented 92% of the sequence of a mature carboxylesterase protein. This
sequence possessed many structural features that are highly conserved among
rabbit and rat liver carboxylesterase proteins, including Ser, His, and Asp
residues that comprise the active site, two pairs of Cys residues that may
participate in disulfide bond formation, and one Asn-Xxx-Thr site for N-linked
carbohydrate addition. When the clone was used to probe human liver genomic DNA
that had been digested with various restriction enzymes, many hybridizing bands
of differing intensities were observed. The results suggest that the carboxyl-
esterases exist as several isoenzymes in humans, and that they are encoded by
multiple genes.

Introduction

The mammalian carboxylesterases (E.C. 3.1.1.1) comprise an important family
of isoenzymes that are biochemically, immunologically, and genetically distinct
proteins (1-7). Carboxylesterases are responsible for the hydrolysis of ester,
thioester, and amide bonds of a wide variety of therapeutic drugs and xenobiotic
chemicals, and can either activate or terminate the biological actions of these
substrates (1,2). Recently, it was discovered that purified carboxylesterase
enzymes can function as lipases that specifically catalyze the hydrolysis of
medium chain mono- and diacylglycerols, and fatty acyl esters of carnitine and
coenzyme A (3,4). Thus, there may be an important physiological role for the
carboxylesterases, perhaps in transmembrane signalling, in lipid transport
mechanisms, or in the maintenance of membrane structure. This ability of the
carboxylesterases to utilize such a broad range of substrates may well result
from the existence of multiple isoenzymic forms, each with unique structural
constraints that determine substrate specificity. However, structure/function
relationships are currently not known for the carboxylesterases, because there
is very little structural information available regarding the individual
isoenzymes. It is probable that the major distinguishing features among the
carboxylesterase isoenzymes will be revealed only when the primary amino acid
sequences are known, and this information can best be obtained through molecular
cloning. In this report, we present the cloning and sequencing of a comple-
mentary DNA (cDNA) encoding a human liver carboxylesterase, the first of the
human isoenzymes to be cloned.

Corresponding Author: Rochelle M. Long, Ph.D., Pharmacological Sciences Program,
Westwood Bldg. Rm. 919, NIGMS, NIH, Bethesda, MD 20892

Methods

Lambda gt11 Library Screening. A lambda gt11 library that had been constructed using messenger RNA prepared from human liver (a gift from Dr. Frank J. Gonzalez, NCI, NIH) was screened with antibodies raised to a purified rat liver carboxylesterase (8). Blocking and washing conditions were those as described previously for immunoblotting experiments (9). Clones were confirmed to be esterase-related by hybridization to oligonucleotide probes prepared from sequences obtained from tryptic-digest peptides generated from a purified rat liver carboxylesterase (10). A second library screening was conducted using a partial length clone, and plaque lifting and washing conditions as described for nucleotide probes (11).

cDNA Cloning and Sequencing. Lambda gt11 insert cDNAs were isolated by digestion with Eco RI and cloned into pGEM 7f(+) vector (Promega, Madison, WI). Clones were prepared by the alkaline lysis method (11), denatured, and sequenced by the double-stranded method using SP6 and T7 primers (Promega) and the Sequenase Kit from United States Biochemical (Cleveland, OH). Sequences were extended using primers synthesized to the 3'-most known regions of sequence. Overlapping clones were assembled and translated with the Microgenie computer program (Beckman, Palo Alto, CA).

Southern Blotting Analysis. DNA was prepared and digested to completion with Eco RI, Stu I, Bgl I, Sca I, Hind III, or Pvu II according to the manufacturer's directions. The fragments were separated on 0.7% agarose gels and transferred to nylon membrane by capillary blotting with 0.4 M NaOH, 0.6 M NaCl. Prehybridization and hybridization were performed as described by Church and Gilbert (12). The human cDNA used as a probe was labelled with [^{32}P]dCTP (3000 Ci/mmole, Amersham, Arlington Hts., IL) using an Oligolabelling Kit (Pharmacia, Piscataway, NJ). Blots were washed for 20 minutes twice each in 2X SSC at 25°C, 2X SSC at 65°C, and then 0.2X SSC at 65°C.

Results

We now report the nearly-complete sequence of a human liver carboxyl-esterase isoenzyme as deduced from its cDNA (Figure 1). The longest clone isolated from the library screening was 1725 bp and encoded 507 amino acids. The clone contained an open reading frame terminating with a stop codon (TGA), followed by untranslated region that included a polyadenylation signal (AATAAA) and a poly A tail. The clone started at what is likely to be nucleotide 127 and corresponded to 92% of the sequence of a mature, processed carboxylesterase protein, as deduced by alignment to the carboxylesterase sequences from other species (10, 13-16). Repeated screening attempts with human cDNA libraries obtained from independent sources (including one designed to be enriched in longer clones, a "5'-Stretch Library" from Clontech, Palo Alto, CA) did not yield any full-length clones or partial clones overlapping at the 5' end.

The protein that was encoded possessed many structural features that are highly conserved among the carboxylesterases sequenced from rabbit liver as proteins (13,14) and from rat liver as cDNA clones (10,15,16). These features included Ser, His, and Asp residues that comprise the active site and may form a putative charge relay system (17); two pairs of Cys residues that can partici-pate in intramolecular disulfide bond formation, by analogy to the cholinester-ases (18); and one Asn-Xxx-Thr site for N-linked carbohydrate addition, that is common to the carboxyl- and cholinesterases (Figure 2). There was also one free Cys residue that occurred only in the human carboxylesterase clone, that may be available to participate in intermolecular disulfide bridge formation. The hydropathy profile (not shown) revealed no large, very hydrophobic regions (potential membrane-spanning domains) in the protein, consistent with biochemical evidence that carboxylesterases are not integral membrane proteins (16).

```
GGA CCC CTG AGG TTT ACT CCA CCG CAG CCT GCA GAA CCA TGG AGC TTT GTG AAG    180
Gly Pro Leu Arg Phe Thr Pro Pro Gln Pro Ala Glu Pro Trp Ser Phe Val Lys     60

AAT GCC ACC TCG TAC CCT CCT ATG TGC ACC CAA GAT CCC AAG GCG GGG CAG TTA    234
Asn Ala Thr Ser Tyr Pro Pro Met Cys Thr Gln Asp Pro Lys Ala Gly Gln Leu     78

CTC TCA GAG CTA TTT ACA AAC CGA AAG GAG AAC ATT CCT CTC AAG CTT TCT GAA    288
Leu Ser Glu Leu Phe Thr Asn Arg Lys Glu Asn Ile Pro Leu Lys Leu Ser Glu     96

GAC TGT CTT TAC CTC AAT ATT TAC ACT CCT GCT GAC TTG ACC AAG AAA AAC AGG    342
Asp Cys Leu Tyr Leu Asn Ile Tyr Thr Pro Ala Asp Leu Thr Lys Lys Asn Arg    114

CTG CCG GTG ATG GTG TGG ATC CAC GGA GGG GGG CTG ATG GTG GGT GCG GCA TCA    396
Leu Pro Val Met Val Trp Ile His Gly Gly Gly Leu Met Val Gly Ala Ala Ser    132

ACC TAT GAT GGG CTG GCC CTT GCT GCC CAT GAA AAC GTG GTG GTG GTG ACC ATT    450
Thr Tyr Asp Gly Leu Ala Leu Ala Ala His Glu Asn Val Val Val Val Thr Ile    150

CAA TAT CGC CTG GGC ATC TGG GGA TTC TTC AGC ACA GGG GAT GAA CAC AGC CGG    504
Gln Tyr Arg Leu Gly Ile Trp Gly Phe Phe Ser Thr Gly Asp Glu His Ser Arg    168

GGG AAC TGG GGT CAC CTG GAC CAG GTG GCT GCC CTG CGC TGG GTC CAG GAC AAC    558
Gly Asn Trp Gly His Leu Asp Gln Val Ala Ala Leu Arg Trp Val Gln Asp Asn    186

ATT GCC AGC TTT GGA GGG AAC CCA GGC TCT GTG ACC ATC TTT GGA GAG TCA GCG    612
Ile Ala Ser Phe Gly Gly Asn Pro Gly Ser Val Thr Ile Phe Gly Glu Ser Ala    204

GGA GGA GAA AGT GTC TCT GTT CTT GTT TTG TCT CCA TTG GCC AAG AAC CTC TTC    666
Gly Gly Glu Ser Val Ser Val Leu Val Leu Ser Pro Leu Ala Lys Asn Leu Phe    222

CAC CGG GCC ATT TCT GAG AGT GGC GTG GCC CTC ACT TCT GTT CTG GTG AAG AAA    720
His Arg Ala Ile Ser Glu Ser Gly Val Ala Leu Thr Ser Val Leu Val Lys Lys    240

GGT GAT GTC AAG CCC TTG GCT GAG CAA ATT GCT ATC ACT GCT GGG TGC AAA ACC    774
Gly Asp Val Lys Pro Leu Ala Glu Gln Ile Ala Ile Thr Ala Gly Cys Lys Thr    258

ACC ACC TCT GCT GTC ATG GTT CAC TGC CTG CGA CAG AAG ACG GAA GAG GAG CTC    828
Thr Thr Ser Ala Val Met Val His Cys Leu Arg Gln Lys Thr Glu Glu Glu Leu    276

TTG GAG ACG ACA TTG AAA ATG AAA TTC TTA TCT CTG GAC TTA CAG GGA GAC CCC    882
Leu Glu Thr Thr Leu Lys Met Lys Phe Leu Ser Leu Asp Leu Gln Gly Asp Pro    294

AGA GAG AGT CAA CCC CTT CTG GGC ACT GTG ATT GAT GGG ATG CTG CTG CTG AAA    936
Arg Glu Ser Gln Pro Leu Leu Gly Thr Val Ile Asp Gly Met Leu Leu Leu Lys    312

ACA CCT GAA GAG CTT CAA GCT GAA AGG AAT TTC CAC ACT GTC CCC TAC ATG GTC    990
Thr Pro Glu Glu Leu Gln Ala Glu Arg Asn Phe His Thr Val Pro Tyr Met Val    330

GGA ATT AAC AAG CAG GAG TTT GGC TGG TTG ATT CCA ATG CAG TTG ATG AGC TAT   1044
Gly Ile Asn Lys Gln Glu Phe Gly Trp Leu Ile Pro Met Gln Leu Met Ser Tyr    348

CCA CTC TCC GAA GGG CAA CTG GAC CAG AAG ACA GCC ATG TCA CTC CTG TGG AAG   1098
Pro Leu Ser Glu Gly Gln Leu Asp Gln Lys Thr Ala Met Ser Leu Leu Trp Lys    366

TCC TAT CCC CTT GTT TGC ATT GCT AAG GAA CTG ATT CCA GAA GCC ACT GAG AAA   1152
Ser Tyr Pro Leu Val Cys Ile Ala Lys Glu Leu Ile Pro Glu Ala Thr Glu Lys    384

TAC TTA GGA GGA ACA GAC GAC ACT GTC AAA AAG AAA GAC CTG TTC CTG GAC TTG   1206
Tyr Leu Gly Gly Thr Asp Asp Thr Val Lys Lys Lys Asp Leu Phe Leu Asp Leu    402

ATA GCA GAT GTG ATG TTT GGT GTC CCA TCT GTG ATT GTG GCC CGG AAC CAC AGA   1260
Ile Ala Asp Val Met Phe Gly Val Pro Ser Val Ile Val Ala Arg Asn His Arg    420

GAT GCT GGA GCA CCC ACC TAC ATG TAT GAG TTT CAG TAC CGT CCA AGC TTC TCA   1314
Asp Ala Gly Ala Pro Thr Tyr MET Tyr Glu Phe Gln Tyr Arg Pro Ser Phe Ser    438

TCA GAC ATG AAA CCC AAG ACG GTG ATA GGA GAC CAC GGG GAT GAG CTC TTC TCC   1368
Ser Asp Met Lys Pro Lys Thr Val Ile Gly Asp His Gly Asp Glu Leu Phe Ser    456

GTC TTT GGG GCC CCA TTT TTA AAA GAG GGT GCC TCA GAA GAG GAG ATC AGA CTT   1422
Val Phe Gly Ala Pro Phe Leu Lys Glu Gly Ala Ser Glu Glu Glu Ile Arg Leu    474

AGC AAG ATG GTG ATG AAA TTC TGG GCC AAC TTT GCT CGC AAT GGA AAC CCC AAT   1476
Ser Lys Met Val Met Lys Phe Trp Ala Asn Phe Ala Arg Asn Gly Asn Pro Asn    492

GGG GAA GGG CTG CCC CAC TGG CCA GAG TAC AAC CAG AAG GAA GGG TAT CTG CAG   1530
Gly Glu Gly Leu Pro His Trp Pro Glu Tyr Asn Gln Lys Glu Gly Tyr Leu Gln    510

ATT GGT GCC AAC ACC CAG GCG GCC CAG AAG CTG AAG GAC AAA GAA GTA GCT TTC   1584
Ile Gly Ala Asn Thr Gln Ala Ala Gln Lys Leu Lys Asp Lys Glu Val Ala Phe    528

TGG ACC AAC CTC TTT GCC AAG AAG GCA GTG GAG AAG CCA CCC CAG ACA GAA CAC   1638
Trp Thr Asn Leu Phe Ala Lys Lys Ala Val Glu Lys Pro Pro Gln Thr Glu His    546

ATA GAG CTG TGA ATG AAG ATC CAG CCG GCC TTG GGA GCC TGG AGG AGC AAA GAC   1692
Ile Glu Leu ---


TGG GGT CTT TTG CGA AAG GGA TTG CAG GTT CAG AAG GCA TCT TAC CAT GGC TGG   1746
GGA ATT GTC TGG TGG TGG GGG GCA GGG GAC AGA GGC CAT GAA GGA GCA AGT TTT   1800
GTA TTT GTG ACC TCA GCT TTG GGA ATA AAG GAT CTT TTG AAG GCC AAA AAA       1851
```

Figure 1.
Nucleotide sequence of the cDNA and deduced amino acid sequence of a Human Liver Carboxylesterase. Active site residues are double underlined, a potential site for glycosylation is in bold type, and cysteines are underlined. AATAAA is found at positions 1824 - 1829.

The -COOH terminus of the protein was -His-Ile-Glu-Leu; thus only the last 2 residues are identical to the consensus sequence proposed by Pelham (-Lys-Asp-Glu-Leu) for proteins that are inserted through the endoplasmic reticulum via the signal peptide mechanism, yet are destined not to be secreted, that are retained (or salvaged) at that luminal site (19). The other carboxylesterases from rat and rabbit liver also conform only minimally to this mammalian "consensus" sequence.

NH$_2$ ... 100 ... 200 ... 300 ... 400 ... 500 ... COOH

Human
CHO site $_{61}$, Asp$_{97}$, Ser$_{203}$, His$_{450}$
Cys$_{69}$–Cys$_{98}$, Cys$_{256}$–Cys$_{267}$, Cys$_{372}$

Rat
CHO site $_{61}$, Asp$_{97}$, Ser$_{203}$, CHO site$_{256}$ CHO site$_{283}$, CHO site$_{357}$, CHO site$_{458}$, His$_{435}$
Cys$_{69}$–Cys$_{98}$, Cys$_{255}$–Cys$_{266}$

Rat ES-10
CHO site $_{61}$, Asp$_{97}$, Ser$_{203}$, CHO site$_{471}$, His$_{448}$
Cys$_{69}$–Cys$_{98}$, Cys$_{255}$–Cys$_{266}$

Rabbit 1
CHO site $_{61}$, Asp$_{97}$, Ser$_{195}$, CHO site$_{363}$, His$_{441}$
Cys$_{69}$–Cys$_{98}$, Cys$_{247}$–Cys$_{258}$

Rabbit 2
Asp$_{95}$, Ser$_{201}$, CHO site$_{249}$, His$_{430}$
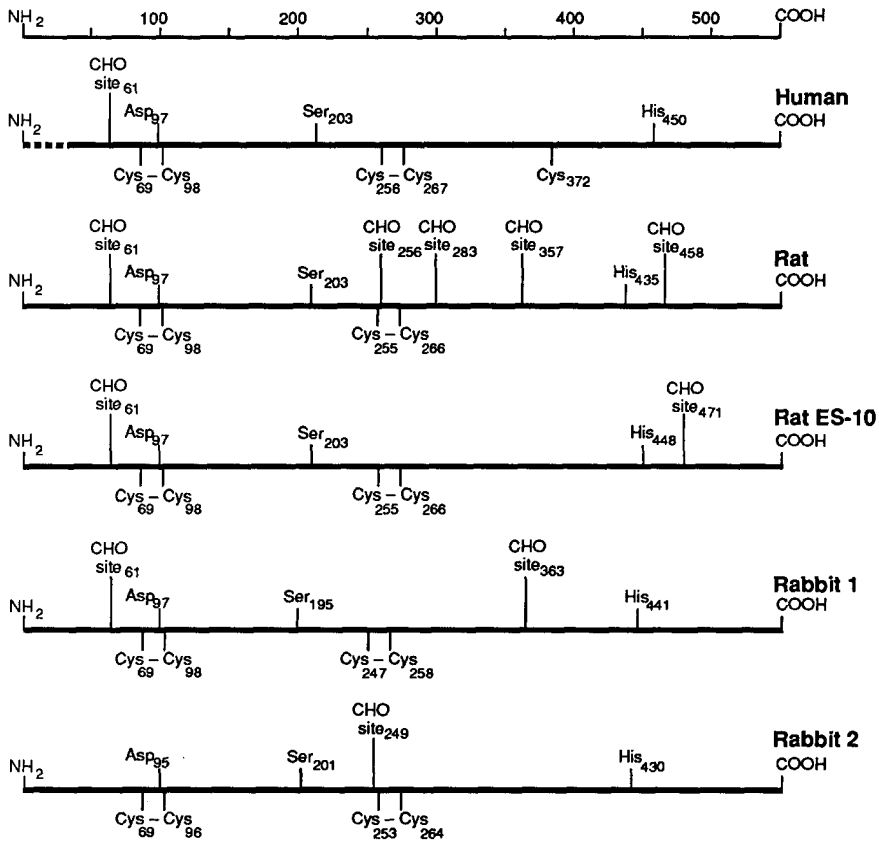Cys$_{69}$–Cys$_{96}$, Cys$_{253}$–Cys$_{264}$

Figure 2.
Comparison between Human, Rat, and Rabbit Liver Carboxylesterases.
Amino acid sequences were aligned using the proteins deduced from the respective cloned human (this paper), rat (form unidentified in Ref. 10, identified as E1 in Ref. 15), and rat form ES-10 (16) cDNAs, and the amino acid sequences obtained from purified proteins identified as rabbit form 1 (13) and form 2 (14).

The overall homologies of human liver carboxylesterase to the carboxylesterases from other species are: 69% to rat liver carboxylesterase (10,15), 76% to rat liver carboxylesterase form ES-10 (16), 77% to rabbit liver carboxylesterase form 1 (13), and 48% to rabbit liver carboxylesterase form 2 (14).

When the human liver cDNA was used to probe human genomic DNA digested with various restriction enzymes, the Southern blot revealed "complex" patterns with multiple bands at higher molecular weights (Figure 3). These bands were clearly of varying intensity, *i.e.* imperfect matches upon hybridization. Because the hybridizations and washings were conducted at relatively stringent conditions, the bands represent highly similar but distinct sequences. These results suggest there are multiple carboxylesterase genes. When a comparison is made between species, it is apparent that fewer bands were visible on the Southern blot of human genomic DNA (Figure 3) than were seen on the Southern blot of rat genomic DNA (10). This suggests that there are fewer carboxylesterase genes present in humans than in rats.
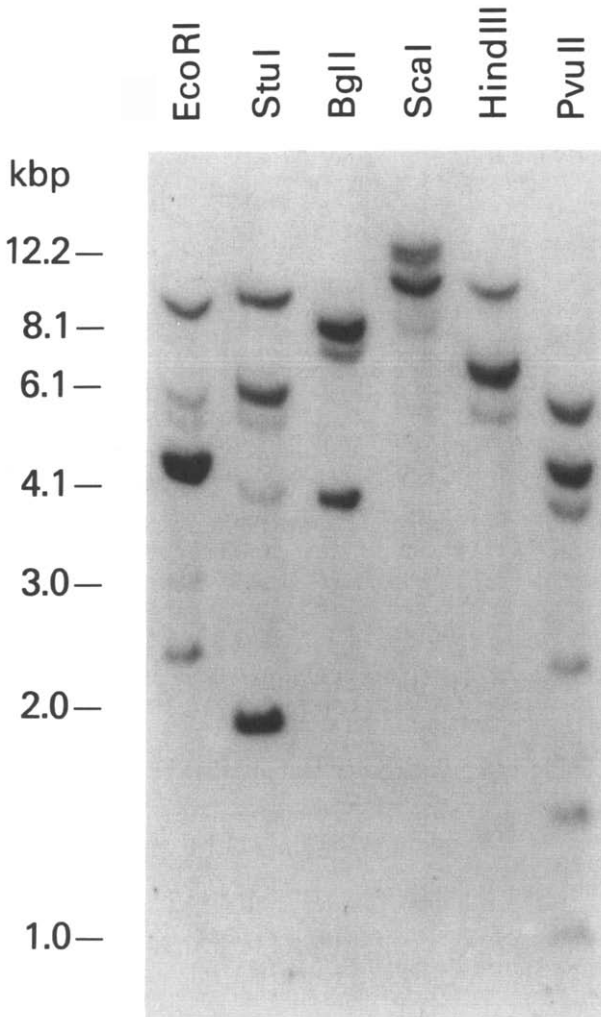
Figure 3.
Southern blot of Human Liver Genomic DNA. Human liver DNA was digested to completion with restriction enzymes, separated by electrophoresis, transferred, and hybridized as described in Methods.

Discussion

Despite their potential importance in the metabolism of both endogenous and exogenous substrates, there is much that is not yet known about the individual carboxylesterases, including the primary structures, tissue distribution, physiological regulation, and substrate specificities of each of the individual isoenzymes. To date, experimental approaches to these problems have depended upon protein purification, and such studies have confirmed the presence of at least two distinct human carboxylesterase isoenzymes (20). However, this approach is a difficult if not impossible one to take, because there are many hydrolytic activities within cells, and if the "correct" substrate is not known results can be misleading. Thus, a molecular biological approach has been initiated to study the structure and function of the carboxylesterases. As complete sets of clones are developed, comparative studies of enzyme structure and catalytic function can commence through *in vitro* expression of the carboxylesterase clones, and the activity of each single isoenzyme can be characterized. The eventual goal of this work is to exploit this information for therapeutic purposes. For example, it may be possible to "tailor" the structure of an ester pro-drug, so that it can be hydrolyzed and trapped as a biologically active form in a specific tissue, based upon the existence of an isoenzyme unique to that tissue. This knowledge will also provide insight into the physiological role(s) played by these isoenzymes, particularly in lipid metabolism, and dysfunctions which may be involved in certain disease states.

Note

This sequence was submitted to Genbank and can be accessed under #M55509.

Acknowledgements

References

1. T. SATOH, Rev. in Biochem. Toxicol. 8 155-181 (1987).
2. E. HEYMANN, Metabolic Basis of Detoxication, 229-245, Academic Press, NY (1982).
3. R. MENTLEIN, S. HEILAND, and E. HEYMANN, Arch. Biochem. Biophys. 200 547-559 (1980).
4. M. HOSOKAWA, T. MAKI, and T. SATOH, Mol. Pharmacol. 31 579-584 (1987).
5. R. MENTLEIN, A. RONAI, M. ROBBI, E. HEYMANN, and O. VON DIEMLING, Biochim. Biophys. Acta 913 27-38 (1987).
6. R. MENTLEIN, G. REUTER, and E. HEYMANN, Arch. Biochem. Biophys. 240 801-810 (1985).
7. R. MENTLEIN, H. RIX-MATZEN, and E. HEYMANN, Biochim. Biophy. Acta 964 319-328 (1988).
8. H. SATOH, B.M. MARTIN, A. SCHULICK, D.D. CHRIST, J.G. KENNA, and L.R.POHL, Proc. Natl. Acad. Sci. 86 322-326 (1989).
9. J.G. KENNA, H. SATOH, D.D. CHRIST, and L.R. POHL, J. Pharmacol. Exp. Ther. 245 1103-1109 (1988).
10. R.M. LONG, H. SATOH, B.M. MARTIN, S. KIMURA, F.J. GONZALEZ, and L.R. POHL, Biochem. Biophys. Res. Comm. 156 866-873 (1988).
11. T. MANIATIS, E.F. FRITSCH, and J. SAMBROOK, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY (1982).

12. G.M. CHURCH and W. GILBERT, Proc. Natl. Acad. Sci. 81 1991-1995 (1984).
13. G. KORZA and J. OZOLS, J. Biol. Chem. 263 3486-3495 (1988).
14. J. OZOLS, J. Biol. Chem. 264 12533-12545 (1989).
15. Y. TAKAGI, K. MOROHASHI, S. KAWABATA, M. GO, AND T. OMURA, J. Biochem. 104 801-806 (1988).
16. M. ROBBI, H. BEAUFAY, and J.-N. OCTAVE, Biochem. J. 269 451-458 (1990).
17. R.C. AUGUSTEYN, J. DE JERSEY, E.C. WEBB, AND B. ZERNER, Biochim. Biophys. Acta 171 128-137 (1969).
18. O. LOCKRIDGE, S. ADKINS, and B. LA DU, J. Biol. Chem. 262 12945-12952 (1987).
19. S. MUNRO and H.R.B. PELHAM, Cell 48 899-907 (1988).
20. A.J. KETTERMAN, M.R. BOWLES, and S.M. POND, Int. J. Biochem. 21 1303-1312 (1989).