

of these two lipoproteins, and to Kenneth Watt for the extended sequencer runs that yielded the amino acid sequence data upon which our oligonucleotide probes were patterned.

# REFERENCES

- Barker, W. C., & Dayhoff, M. O. (1977) *J. Comp. Biochem. Physiol.* 57B, 309-315.
- Bassiri, R. M., Dvorak, J., & Utiger, R. D. (1979) in *Methods of Hormone Radioimmunoassay* (Jaffey, B. M., & Behrman, H. R., Eds.) pp 46-47, Academic Press, New York.
- Biggin, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80, 3963-3965.
- Boguski, M. S., Elshourbagy, N., Taylor, J. M., & Gordon, J. I. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81, 5021-5025.
- Boguski, M. S., Elshourbagy, N., Taylor, J. M., & Gordon, J. I. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82, 992-996.
- Bohonus, V. L., Doolittle, R. F., Pontes, M., & Strong, D. D. (1986) *Biochemistry* 25, 6512-6516.
- Burnette, W. H. (1981) *Anal. Biochem.* 112, 195-203.
- Chapman, M. J. (1979) *J. Lipid Res.* 21, 789-853.
- Chapman, M. J., Goldstein, S., Mills, G. L., & Leger, C. (1978) *Biochemistry* 17, 4455-4464.
- Chou, P. Y., & Fasman, G. D. (1978) *Annu. Rev. Biochem.* 47, 251-276.
- Doolittle, R. F. (1981) *Science (Washington, D.C.)* 214, 149-159.
- Doolittle, R. F. (1984) in *The Plasma Proteins* (Putnam, F. W., Ed.) Vol. IV, pp 317-360, Academic Press, New York.
- Doolittle, R. F., Cottrell, B. A., & Riley, M. (1976) *Biochim. Biophys. Acta* 453, 439-452.
- Feng, D. F., Johnson, M. S., & Doolittle, R. F. (1985) *J. Mol. Evol.* 21, 112-125.

- Fitch, W. M. (1977) *Genetics* 86, 623-644.
- Garnier, J., Osguthorpe, D. J., & Robson, B. (1978) *J. Mol. Biol.* 120, 97-120.
- George, D. G., Barker, W. C., & Hunt, L. T. (1986) *Nucleic Acids Res.* 14, 11-15.
- Kuyas, C., Riley, M., Bubis, J., & Doolittle, R. F. (1983) *Fed. Proc., Fed. Am. Soc. Exp. Biol.* 42, 2085.
- Larsson, A., & Fänge, R. (1977) *J. Comp. Biochem. Physiol.* 57B, 191-196.
- Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning, A Laboratory Manual*, pp 1-545, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560.
- McLachlan, A. D. (1977) *Nature (London)* 267, 465-466.
- Messing, J. (1983) *Methods Enzymol.* 101, 20-78.
- Nelson, G., & Shore, V. G. (1974) *J. Biol. Chem.* 249, 536-542.
- Norgard, M. V., Emigholz, K., & Monahan, J. J. (1979) *J. Bacteriol.* 138, 270-272.
- Rubin, C. M., & Schmid, C. W. (1980) *Nucleic Acids Res.* 8, 4613-4619.
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463-5467.
- Strong, D. D., Moore, M., Cottrell, B. A., Bohonus, V. L., Pontes, M., Evans, B., Riley, M., & Doolittle, R. F. (1985) *Biochemistry* 24, 92-101.
- Thompson, J. A., Blakesley, R. W., Doran, K., Hough, C. J., & Wells, R. D. (1983) *Methods Enzymol.* 100, 368-399.
- von Heijne, G. (1986) *Nucleic Acids Res.* 14, 4683-4690.
- Wallace, R. B., Johnson, M. J., Hirose, T., Miyabe, T., Kawashima, E. H., & Itakura, K. (1981) *Nucleic Acids Res.* 9, 879-894.

## Isolation of Full-Length Putative Rat Lysophospholipase cDNA Using Improved Methods for mRNA Isolation and cDNA Cloning<sup>†</sup>

Jang H. Han,<sup>†</sup> Christian Stratowa,<sup>§</sup> and William J. Rutter\*

Hormone Research Institute and Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143

Received August 14, 1986; Revised Manuscript Received October 30, 1986

**ABSTRACT:** We have cloned a full-length putative rat pancreatic lysophospholipase cDNA by an improved mRNA isolation method and cDNA cloning strategy. These new methods allow the construction of a cDNA library from the adult rat pancreas in which the majority of recombinant clones contained complete sequences for the corresponding mRNAs. A previously recognized but unidentified long and relatively rare cDNA clone containing the entire sequence from the cap site at the 5' end to the poly(A) tail at the 3' end of the mRNA was isolated by single-step screening of the library. The size, amino acid composition, and the activity of the protein expressed in heterologous cells strongly suggest this mRNA codes for lysophospholipase [Van den Bosch, H., Aarsman, A. J., DeJong, G. N., & Van Deenen, L. M. (1973) *Biochim. Biophys. Acta* 296, 94-104].

**T**he molecular cloning of cDNAs has played a key role in elucidating the sequence, organization, and expression of eu-

caryotic genes. In addition, it has provided the coding sequences for synthesizing specific proteins in heterologous systems. The methodology usually requires screening a library for the cDNA of interest and splicing partial-length cDNAs to provide the full-length cDNA. The availability of an efficient method for obtaining full-length cDNAs would greatly improve the efficacy of various screening methods of cDNA libraries and of course eliminate the need for splicing partial-length cDNAs to obtain full-length coding regions.

<sup>†</sup> This research was supported by NIH Grants AM21344, GM28520, and AM32994 to W.J.R.

\* Address correspondence to this author at the Hormone Research Institute, University of California, San Francisco.

<sup>†</sup> Present address: Department of Virus and Cell Biology, Merck Sharp & Dohme Research Labs, West Point, PA 19486.

<sup>§</sup> Present address: Ernst-Boehringer-Institut für Arzneimittelforschung, Dr. Boehringer-Gasse 5-11, A-1120 Vienna, Austria.

Several years ago, we developed a method for isolating functional RNAs from the pancreas and other sources rich in RNase (Chirgwin et al., 1979). This method allowed the isolation of insulin cDNA (Ullrich et al., 1977) and subsequently the cDNAs of more than 15 pancreas-specific exocrine and endocrine cDNAs [listed in Han et al. (1986)]; it has also been used broadly by others. The method employed, however, usually leads to the isolation of partial cDNAs, and the isolation of the full-length cDNAs is progressively more difficult with increasing size. Two-dimensional gel analysis of the proteins of the rat exocrine pancreas by ourselves (Van Nest et al., 1980) as well as by others (Schick et al., 1984) has revealed more than 20 distinct proteins, several of which still remain unidentified. Several of these are large proteins with molecular weights greater than amylase (55 000).

We describe in this paper an improved method that produces cDNAs corresponding to full-length mRNAs. The procedures involve the isolation of intact mRNA by a refinement of the original guanidinium thiocyanate method (Chirgwin et al., 1979) and an optimization of the Okayama-Berg strategy for cDNA cloning (Okayama & Berg, 1982). Thus, instead of using plasmid vectors, we employ a new  $\lambda$  vector ( $\lambda$ gt18) and synthetic oligonucleotides as primers and adaptors. From this library, we have isolated in a single screening several full-length cDNAs which extend to the multiple cap sites of a previously unidentified pancreatic mRNA which is both relatively rare (0.05% of total message) and long [2.1 kilobases (kb)]. By expressing the cloned cDNA in heterologous cells, we have shown that it encodes lysophospholipase (Van den Bosch et al., 1973).

#### MATERIALS AND METHODS

**Chemicals and Enzymes.** Guanidinium thiocyanate was purchased from Fluka, guanidine hydrochloride from Sigma, oligo(dT)-cellulose from Collaborative Research Inc., reverse transcriptase from Life Sciences, Inc., terminal deoxynucleotidyl transferase, RNase H, *Escherichia coli* DNA ligase, polynucleotide kinase, and unlabeled nucleotides from P-L Pharmacia, and endonuclease-free DNA polymerase, restriction enzymes, T4 DNA ligase, and *dam* methylase either from Boehringer Mannheim or from New England Biolabs. The RNase inhibitor RNasin and the in vitro  $\lambda$  DNA packaging extract were obtained from either Promega Biotech or Vector Cloning System, DNase I and RNase A from Worthington, and  $^{32}$ P-labeled nucleotides and 1-[ $^{14}$ C]palmitoyllyso-3-phosphatidylcholine from Amersham. The oligonucleotides were synthesized and supplied by Jennifer Barnett in this laboratory.

**Isolation of RNA. (A) Solutions.** Guanidinium thiocyanate stock (5 M) is prepared by mixing 60 g of guanidinium thiocyanate with 5 mL of 1 M tris(hydroxymethyl)amino-methane hydrochloride (Tris-HCl), pH 7.5 (final concentration 50 mM), 25 mL of 0.1 M ethylenediaminetetraacetic acid (EDTA) (25 mM final concentration), and 8 mL of  $\beta$ -mercaptoethanol (8% final concentration, added just prior to use) in a total volume of 100 mL; 6 M guanidine hydrochloride is prepared by mixing 57 g of the salt with 25 mL of 0.1 EDTA, pH 7.0 (25 mM), and 75  $\mu$ L of  $\beta$ -mercaptoethanol (10 mM final concentration, added just prior to use) in a total volume of 100 mL. Both solutions are filtered through 0.45- $\mu$ m Millipore filters before use.

**(B) Low-Temperature Guanidinium Thiocyanate Extraction Procedure.** Freshly removed tissue (about 1 g) was placed in 20 mL of 5 M guanidinium thiocyanate (precooled on ice for 5 min; longer times produce crystallization) in a graduated precooled polypropylene tube (50-mL volume) and immedi-

ately homogenized with an ice-chilled polytron (probe 1 in. in diameter, Brinkmann) for 20 s at the maximum speed. To prevent loss of tissue homogenate, the tube was loosely covered with parafilm during the homogenization. The contents were poured into a polypropylene tube (50-mL volume, which fits a HB4 Sorvall centrifuge rotor without rubber cushion) containing 0.3 volume of cold ethanol ( $-20^{\circ}\text{C}$ ), mixed, and immediately centrifuged in the HB4 rotor (precooled to  $-10^{\circ}\text{C}$ ) at 10 000 rpm for 5 min. During this period, the polytron probe was cleaned with 0.2% sodium dodecyl sulfate, rinsed with water, and chilled on ice. After the centrifugation, a protein film at the top of the fluid was removed, and the remaining supernatant was discarded by aspiration while the tube was maintained at  $-10^{\circ}\text{C}$ . Ten milliliters of ice-cold (0  $^{\circ}\text{C}$ ) 5 M guanidinium thiocyanate was added to the pellet and then homogenized for 10 s as above. The suspension was centrifuged at 10 000 rpm (3 min, 0  $^{\circ}\text{C}$ ) to remove contaminating tissue debris. The clear supernatant was mixed with 0.025 volume of 1 M acetic acid and 0.75 volume of cold ethanol ( $-20^{\circ}\text{C}$ ), incubated at  $-20^{\circ}\text{C}$  for longer than 3 h, and then centrifuged at 6000 rpm for 10 min at  $-10^{\circ}\text{C}$ . The supernatant and any floating material were discarded by aspiration, and the pellet was resuspended in 2 mL of cold ( $-20^{\circ}\text{C}$ ) 6 M guanidine hydrochloride (Gdn-HCl) by repeated pipeting. In most cases, the pellet was easily resuspended; if not, it was briefly sonicated. The volume was adjusted to 10 mL of Gdn-HCl, transferred to two polypropylene tubes (15 mL), and vortexed. In each tube, 0.25 mL of 1 M acetic acid and 0.5 volume of cold ( $-20^{\circ}\text{C}$ ) ethanol were added with mixing and the contents incubated at  $-20^{\circ}\text{C}$  for more than 3 h and centrifuged as above. This Gdn-HCl step was repeated twice, and the RNA pellet was resuspended in progressively decreasing volumes of Gdn-HCl (finally 5 mL). A small aliquot (less than 50  $\mu$ L) of Gdn-HCl solution containing RNA was then tested for RNase activity. The aliquot was precipitated with ethanol, dissolved in diethyl pyrocarbonate (DEP) treated 50 mM sodium acetate (pH 7.0), and incubated for 30 min at  $37^{\circ}\text{C}$ . The RNA was analyzed by electrophoresis in a 1.5% nondenaturing agarose gel (TBE buffer: 89 mM each of Tris and boric acid and 10 mM EDTA, 80 V, 1 h). The presence of 28S and 18S rRNA bands in a ratio of  $\sim 2:1$  would indicate intact RNA. If the ratio was less than 1:1, further cycles of Gdn-HCl-ethanol precipitation were performed to remove the residual RNase activity, after which the entire RNA pellet was dissolved in 5 mL of DEP-treated water. Complete dissolution of the RNA pellet may sometimes require extraction with additional aliquots of water. The RNA-containing solution was mixed with 0.1 volume of 2 M sodium acetate and 2.5 volumes of ethanol and stored at  $-70^{\circ}\text{C}$  until use. Poly(A) RNA was isolated by a standard method (Aviv & Leder, 1972).

**(C) RNA Isolation by Other Methods.** For comparison, RNA was also isolated by other methods using guanidinium thiocyanate such as CsCl ultracentrifugation (Chirgwin et al., 1979), ethanol precipitation (Chirgwin et al., 1979), and the LiCl precipitation method (Cathala et al., 1985).

**Construction of  $\lambda$ gt18 and  $\lambda$ gt19 (*lac5 nin5 c1857 S100 spi*<sup>-</sup>).**  $\lambda$ gt18 and  $\lambda$ gt19 (Figure 1) were constructed from  $\lambda$ gt11 (Young & Davis, 1983a). The two unwanted *Sal*I sites within the parent phage were removed by digesting  $\lambda$ gt11 DNA with *Sal*I and religating the DNA with a 10-fold molar excess of an oligonucleotide which has *Sal*I-cohesive ends but cannot generate at a *Sal*I site

TCGAGGGCC  
CCCGGAGCT

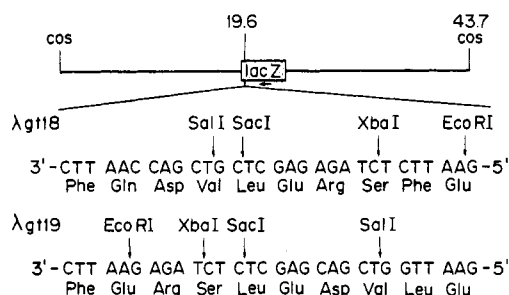
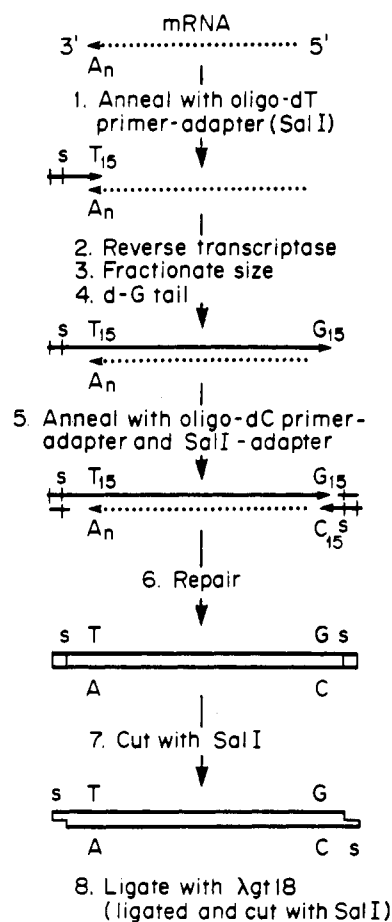


FIGURE 1: Map of bacteriophage  $\lambda$ gt18 and  $\lambda$ gt19 vectors in the region of polylinker. Restriction cleavage sites and the amino acids encoded are shown below the map. Other restriction sites are identical with that of  $\lambda$ gt11 (Young & Davis, 1983a) except for the deleted *SalI* sites in the new vectors.

The removal of two *SalI* sites resulted in a phage which gives rise to small plaques. This phage was passed for five generations in *E. coli* (LE392) until a relatively large plaque-forming phage was obtained according to the procedure by Stahl et al. (1975). A convenient polylinker containing *SalI*, *SacI*, *XbaI*, and *EcoRI* restriction sites

AATTGGTCGACGAGCTCTCTAGAG  
CCAGCTGCTCGAGAGATCTCTTAA

## A. Two Orientation Cloning



was then introduced at the *EcoRI* site of the changed  $\lambda$ gt11 DNA. Because of these changes, these phages are likely to be *Spi<sup>-</sup>* and *chi*.  $\lambda$ gt18 and  $\lambda$ gt19 are identical except for the orientation of the polylinker.

**cDNA Cloning with the  $\lambda$ gt18 (or  $\lambda$ gt19) Vectors.** The steps for cDNA synthesis and cloning in either one or two orientations in the  $\lambda$ gt18 (or  $\lambda$ gt19) vector are outlined in Figure 2. All restriction endonuclease digestions and ligations were performed as suggested by the enzyme suppliers. The isolated  $\lambda$ gt18 DNA was preligated and cut either with *SalI* for two-orientation cloning or with *SalI* and *XbaI* for one-orientation cloning after methylation by *dam* methylase. (The methylase treatment protects only the second *XbaI* site present in the right arm of the vector.) After digestion, the DNA was phenol-extracted and ethanol-precipitated twice in the presence of 2 M ammonium acetate as described (Maniatis et al., 1982).

For two-orientation cloning, about 20 pmol of mRNA was annealed with 100 pmol of oligo(dT) primer-adaptor (GA-ATTCGTCGACT<sub>15</sub>) in 20  $\mu$ L of 10 mM KCl by heating at 70 °C for 5 min and then chilling on ice. The first-strand cDNA was synthesized in a volume of 200  $\mu$ L, and sizes above 600 bases were selected by alkaline gel electrophoresis as described (Maniatis et al., 1982). A dG tail (~15 bases) was

## B. One Orientation Cloning

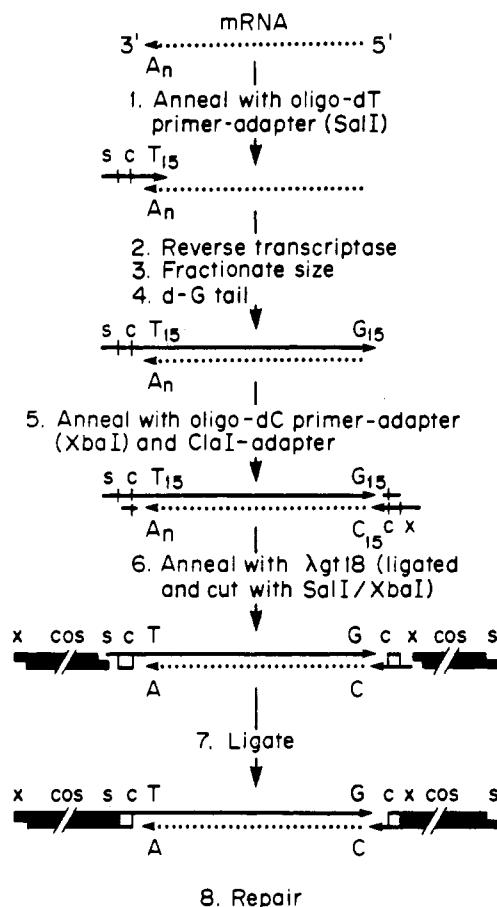


FIGURE 2: Steps in the construction of phage cDNA recombinant. The experimental details of the procedure are presented under Materials and Methods and under Results.  $A_n$  indicates the poly(A) tail of mRNA, and C, G, and T indicate the oligonucleotide primer-adaptor which has a dC, dG, and dT homopolymer tail, respectively. The small letters c, s, and x indicate restriction sites *ClaI*, *SalI*, and *XbaI*, respectively.

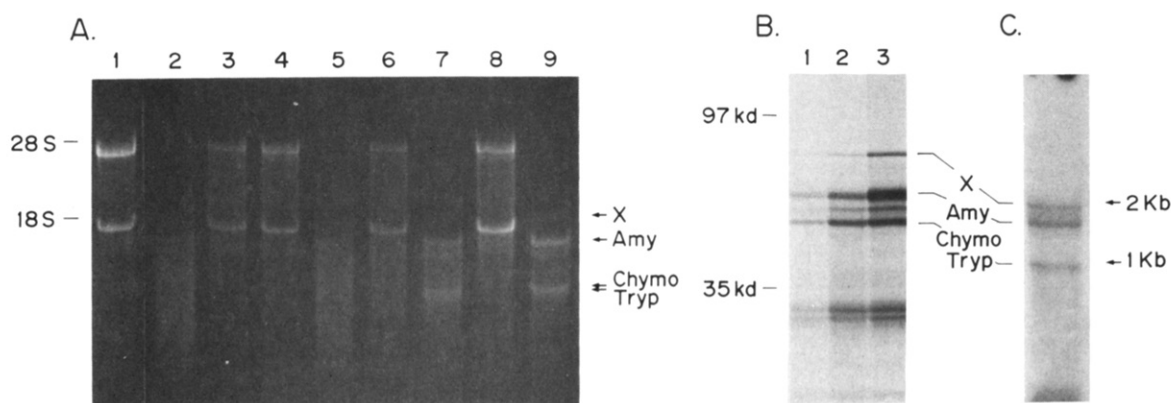


FIGURE 3: Comparison of different methods for RNA isolation using guanidinium thiocyanate. (A) An ethidium bromide staining pattern of total or poly(A<sup>+</sup>) RNA from rat. Lane 1, total RNA from adult spleen isolated by CsCl centrifugation method (Chirgwin et al., 1979). Lanes 2 and 3, total RNA from embryonic pancreas by CsCl centrifugation method (Chirgwin et al., 1979). The tissue homogenate in 6 M guanidinium thiocyanate was stored for 12 h at room temperature before centrifugation (lane 2) or immediately centrifuged (lane 3). Lanes 4–9, RNA isolated from adult pancreas. Lane 4, total RNA isolated by ethanol precipitation method (Chirgwin et al., 1979). Lane 5, poly(A<sup>+</sup>) RNA from lane 4. Lane 6, total RNA by LiCl method (Cathala, 1985). Lane 7, poly(A<sup>+</sup>) RNA from lane 6. Lane 8, total RNA by low-temperature method. Lane 9, poly(A<sup>+</sup>) RNA from lane 8. Five micrograms of RNA was denatured by heating at 65 °C for 10 min in 50% formamide and 2.2 M formaldehyde and electrophoresed on a 1.5% agarose gel containing formaldehyde. The gel was flushed with running water for 30 min to remove formaldehyde and stained with ethidium bromide. Positions of three known mRNAs are marked: Amy, amylase; Chymo, Chymotrypsin; and Tryp, trypsin. X indicates lysophospholipase mRNA. (B) Autoradiogram of [<sup>35</sup>S]methionine-labeled peptides synthesized from purified RNAs in a rabbit reticulocyte lysate (Pelham & Jackson, 1976). Poly(A<sup>+</sup>) RNA samples (1 µg per lane; all from adult rat pancreas) used for translation were the following: lane 1, the same RNA used for lane 5 in (A); lane 2, the same RNA used for lane 7 in (A); lane 3, the same RNA used for lane 9. Positions of three known proteins, amylase, chymotrypsin, and trypsin, are marked as above. (C) Autoradiogram of <sup>32</sup>P-labeled first-strand cDNA after size enrichment (larger than 0.5 kb). The positions of cDNAs of three proteins are marked as above. cDNA band marked X was eluted and used as a probe for screening cDNA library (see Results).

added to the 3' end according to Land et al. (1981). The tailed cDNA was annealed with a 2-fold molar excess of oligo(dC) primer-adaptor (GAATTCGTCGACC<sub>15</sub>) and *SalI* adaptor (GTCGACG) as above, and second-strand cDNA was synthesized by DNA polymerase I in a volume of 100 µL by the method of the Okayama and Berg (1982). The double-stranded cDNA was cut with *SalI*, phenol-extracted, and ethanol-precipitated twice in the presence of 2 M ammonium acetate. Varying amounts of cDNA (100, 200, and 500 ng) were separately ligated to 1 µg of *SalI*-cut λgt18 (or λgt19) DNA in a 10-µL volume with 1 unit of T4 ligase.

For one-orientation cloning, first-strand cDNA was prepared as above except that a different oligo(dT) primer-adaptor (TCGACATCGATGT<sub>15</sub>) was used as a primer for the first-strand synthesis. The dG-tailed first-strand cDNA was annealed with 2-fold molar excess of oligo(dC) primer-adaptor (CTAGAATCGATGC<sub>15</sub>) and *Clal* adaptor (CATCGATT and CATCGATG) as above and ligated in a 25-µL volume with 5 units of *E. coli* ligase to λgt18 DNA previously cut with *SalI* and *XbaI*. The ligated DNA was then converted into double-stranded cDNA as above, phenol-extracted, ethanol-precipitated as above, and packaged into infectious phage as suggested by the supplier of the package extract. The phage were propagated in *E. coli* strain (LE392 or K803) as described by Young and Davis (1983b).

**Sequencing.** The cDNA insert from λrLPL3 was isolated, cleaved with several restriction endonucleases, subcloned into M13mp18, and sequenced by the method of Sanger et al. (1977); ambiguous sequences were confirmed by the method of Maxam and Gilbert (1980). Direct λ sequencing was performed by a modified Sanger method (Chen & Seeburg, 1985) using two synthetic oligonucleotides (TTGACACCA-GACCAACTGG and CCATATGGGGATTGGTGGC-GAC) as primers. Direct mRNA sequencing and primer extension were carried out by the chain termination method using a synthetic primer (GAGGCCAAGAAACA-GAACCTC) as previously described (MacDonald et al., 1982b).

**Expression and Enzyme Assay.** The full-length cDNA (rLPL) was isolated from a λ clone (λrLPL3) and subcloned

into M13mp18 (M13rLPL). The GC tract at the 5' end of the full-length cDNA was removed by site-directed mutagenesis (Craik, 1985). The resultant cDNA was subcloned into pECE<sub>1</sub>, a mammalian cell expression vector utilizing the SV40 early promoter (Ellis et al., 1986). Ten micrograms of pECE<sub>1</sub>rLPL DNA was cotransfected into CHO cells (1 × 10<sup>6</sup>) with 10 µg of pSV<sub>2</sub>CAT, a vector in which the coding sequence of chloramphenicol acetyltransferase (CAT) is under the control of the same SV40 early promoter (Gorman et al., 1982), using the calcium phosphate method (Graham & van der Eb, 1973). For a control, CHO cells were transfected with 10 µg of pSV<sub>2</sub>CAT alone. After 48-h incubation, the cells were harvested, washed with saline, sonicated in 100 µL of 20 mM sodium phosphate buffer (pH 7.0), and centrifuged at 10000g for 5 min at 4 °C. The supernatant was frozen until the activity was assayed. The lysophospholipase activity was assayed according to van den Bosch (Van den Bosch et al., 1973) using a modified Dole procedure (Depierre, 1977).

## RESULTS

**Isolation of Intact mRNA.** The efficacy of our RNA isolation procedure is illustrated in Figure 3 by comparing the RNAs isolated from three different rat tissues (adult spleen, 16-day embryonic pancreas, and adult pancreas, exhibiting intermediate, high, and extremely high RNase activity, respectively) using variations of the guanidinium thiocyanate method including ultracentrifugation (Chirgwin et al., 1979), ethanol precipitation (Chirgwin et al., 1979), LiCl precipitation (Cathala et al., 1985), and the method described here. All methods are acceptable for isolating RNA from tissues with intermediate levels of RNase activity. As illustrated in lane 1, intact 28S and 18S ribosomal RNAs are evident when the RNA is isolated by any of the methods (data not shown). The deficiency of the ultracentrifuge method is evident when the RNA is isolated from tissues with high RNase activity such as 16-day rat embryonic pancreas (lanes 2 and 3) and is even more obvious when the RNA is isolated from tissues with extremely high RNase levels such as adult pancreas. In this case, degradation (as in lane 2) is always observed (data not shown). The three remaining methods produce satisfactory

28S and 18S rRNA bands (lanes 4, 6, and 8, respectively); however, differences in the quality of the isolated mRNA are evident when the poly(A<sup>+</sup>) RNAs are visualized by gel electrophoresis. The three prominent pancreatic mRNA species, amylase, chymotrypsin, and trypsin, are not easily recognizable with the ethanol precipitation method (lane 5) but are evident with the LiCl method (lane 7) and with the present method (lane 9). The intensity of the amylase mRNA band relative to the chymotrypsin or trypsin mRNA band (lane 8) is significantly higher with the latter method. The superiority of the low-temperature method is further demonstrated by *in vitro* translation of the poly(A<sup>+</sup>) RNA (Figure 2B). The low-temperature method (lane 3) gives a higher yield of translation products per unit of poly(A<sup>+</sup>) RNA input and also shows at least a 5-fold higher amylase:chymotrypsin or trypsin ratio as compared to the other two methods (lanes 1 and 2).

**Construction of a Full-Length cDNA Library.** The strategy and experimental procedures for preparing and cloning cDNA in either one or both orientations in  $\lambda$  phage are outlined in Figure 2. The key features are (a) a synthetic oligonucleotide acts both as primer and as adaptor, (b) the size-selected reverse transcripts are efficiently converted into full-length duplex cDNAs by tailing/priming and repair synthesis, (c) a  $\lambda$  phage vector is used for efficient cloning of cDNA and screening, and (d) the cDNA can be cloned in a defined orientation.

Expression vectors  $\lambda$ gt18 and  $\lambda$ gt19 were constructed from  $\lambda$ gt11 by deleting the restriction sites for *SalI* in the phage and then inserting a synthetic polylinker containing restriction sites for *SalI*, *SacI*, *XbaI*, and *EcoRI* in the 3' end of *lacZ* gene (Figure 1).  $\lambda$ gt19 contains the polylinker in the opposite orientation of  $\lambda$ gt18. By use of two restriction sites, single-orientation cloning can be performed, thus maximizing the chance of expression of the cloned cDNA as a fusion protein with *lacZ*. If the *SalI* primer-adaptor is used for first- and second-strand synthesis, and the resultant material is digested with *SalI* to generate cohesive ends, then two-orientation cloning occurs. Eucaryotic DNA contains few *SalI* sites; thus, possible internal cleavage of the cDNA during the digestion is minimal, but it is not eliminated.

The methods which employ viral reverse transcriptase for first-strand synthesis usually generate partial copies of mRNA. The double-stranded cDNAs made from these partial copies are generally small [less than 500 base pairs (bp)] and are readily incorporated into the vector, thus lowering the yield of full-length cDNAs in the library. Generally, size fractionation is more efficient when the mRNA is primed with a short synthetic oligonucleotide as compared to priming with an oligo(dT)-tailed plasmid vector as employed by Okayama and Berg (1982). However, the cloning efficiency should be higher in methods utilizing plasmid vector primers. By using a synthetic oligonucleotide containing a preformed restriction site and homopolymer tail that serves as both primer and adaptor, we eliminate the need for linker attachment and subsequent digestion with restriction enzymes to create cohesive ends.

The efficiency of second-strand synthesis was measured under several different conditions as shown in Figure 4. The best conditions involve dG tailing of the 5' end of the cDNA in the mRNA/cDNA hybrid, and the priming of the second-strand synthesis with oligo(dC) followed by repair synthesis with DNA polymerase and DNA ligase (lane 3), thus forming duplex cDNA. This combination of priming (Land et al., 1981) and repair synthesis by nick translation (Okayama & Berg, 1982) generates largely complete copies of the first strand as illustrated by the identical banding pattern in lane 1 (first strand) and lane 2 (second strand). The mRNA/

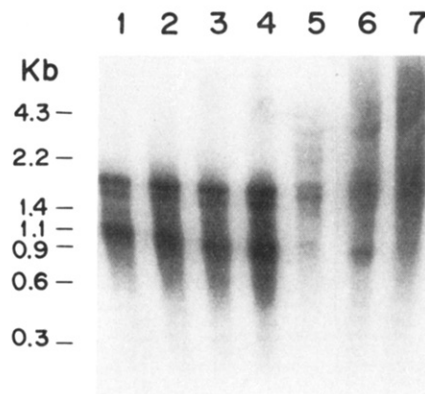


FIGURE 4: Analysis of second-strand cDNA synthesis. The conversion of unlabeled first-strand cDNA (10 ng) into second-strand cDNA was monitored by incorporation of [ $\alpha$ -<sup>32</sup>P]dCTP into the second strand. cDNA/mRNA cDNA was first size fractionated by alkaline gel (lanes 1–5) or by neutral gel electrophoresis (lanes 6 and 7) and then tailed with dG (lane 2–5). After the conversion, all samples were denatured by heating for 10 min in 30 mM NaOH at 70 °C before electrophoresis on a 0.7% alkaline gel. Lane 1,  $\gamma$ -<sup>32</sup>P-5'-end-labeled first strand; lane 2, second strand synthesized with oligo(dC) primer, DNA polymerase, DNA ligase, and RNase H; lane 3, same as lane 2 except for no DNA ligase; lane 4, same as lane 2 except for no RNase H; lane 5, same as lane 2 except for no oligo(dC) primer; lane 6, second strand synthesized with DNA polymerase, ligase, and RNase H; and lane 7, same as lane 6 except for no RNase H.

cDNA hybrids can be converted into duplex cDNA by repair synthesis alone, without tailing/priming, as shown in lane 6. However, this procedure generates slowly migrating bands (top band, lane 6) probably resulting from self-primed second-strand synthesis. Further, the second-strand cDNAs migrate as blurry bands (second and third bands, lane 6), indicating a heterogeneous size distribution which may arise from internal priming at different locations along the first-strand cDNA, resulting in the loss of 5'-mRNA sequence. The tailing/priming reaction should abolish both problems and increase the yield of full-length cDNAs in the library.

The data in Figure 4 (lanes 6 vs. 7) illustrate the requirement of RNase H for repair synthesis (second strand) in the Okayama–Berg method. If the mRNA/cDNA hybrid is sized by electrophoresis on an alkaline gel as in the present method, RNase H is no longer required (lane 4). During electrophoresis in an alkaline gel, the mRNA is partially digested as evidenced by a more than 10-fold increase in kinase-mediated  $\gamma$ -<sup>32</sup>P incorporation into the hybrid after electrophoresis (data not shown). In addition, the weak bands, 1.6 kilobases (kb) (amylase) and 1.0 kilobase pair (kbp) (chymotrypsin), of second-strand cDNA seen in lane 5 in the absence of oligo(C) primer probably result from partial internal RNA priming. Omission of DNA ligase has no apparent effect on second-strand synthesis in the cDNA size range of 1–2 kbp (lane 3); however, we add the enzyme in all reactions to ensure complete ligation of the cDNA to the  $\lambda$  vector.

The cloning efficiency depends greatly on the recovery of the mRNA/cDNA hybrid in the size-selection step. Using electroelution, we routinely obtain  $5 \times 10^5$  recombinant clones from 1  $\mu$ g of poly(A<sup>+</sup>) RNA for both one- and two-orientation cloning.

**Characterization of the cDNA Library.** To test the quality of the cDNA library, we screened with a single-strand size-selected <sup>32</sup>P-labeled cDNA corresponding to the largest mRNA species recognizable in rat adult pancreas (Figure 3A, lane 9, also denoted as “X” in Figure 3C). Ten positive clones were obtained from screening about 20 000 recombinant phages. Digestion of DNAs isolated from these 10 clones with *SalI* yielded a single band for each DNA; 8 out of the 10



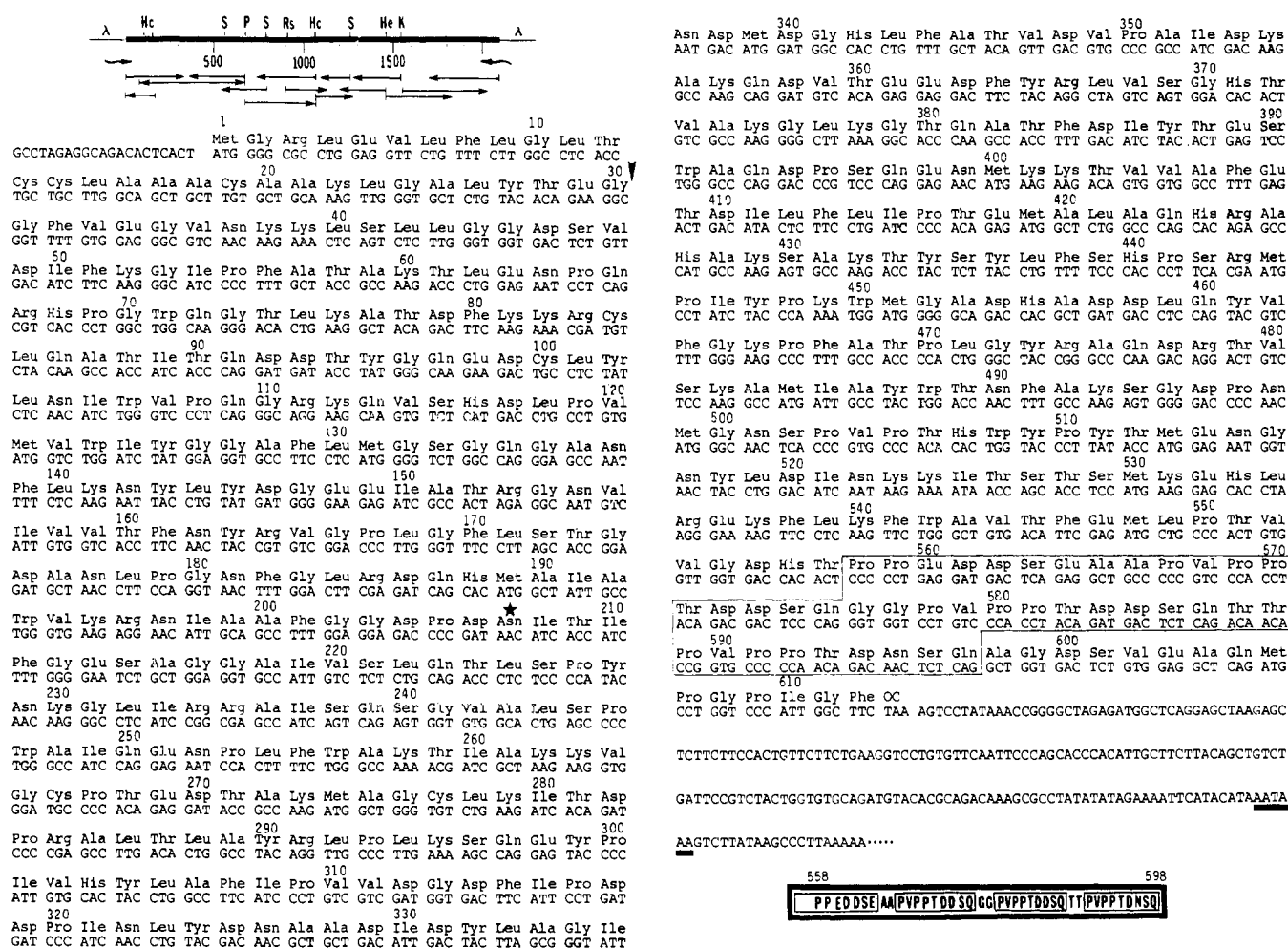


FIGURE 5: Sequence of the full-length lysophospholipase cDNA. The M13 sequencing strategy is shown above the nucleotide sequence and predicted amino acid sequence of putative lysophospholipase. The direction of each sequencing reaction is shown by the horizontal arrows starting at the restriction sites shown (Hc, *HincII*; He, *HaeIII*; P, *PstI*; RI, *EcoRI*; Rs, *RsaI*; S, *Sau3A*). The 5'- and 3'-end sequences of the insert were confirmed from direct  $\lambda$  sequencing (wavy arrows). Amino acids are numbered from the initiator Met residue. The possible cleavage site for the signal peptide is indicated by a vertical arrow. The conserved sequence used as the polyadenylation signal in the 3'-noncoding region is underlined. The site for potential asparagine-linked glycosylation is indicated by a star. The four direct repeats of amino acid sequences are boxed and further represented schematically below the sequences.

clones contained cDNA inserts of about 2000 bp; the other 2 cDNA clones contained inserts of 700 and 1200 bp, respectively. Southern analysis confirmed that all cDNA inserts from the 10 clones are homologous to the nucleotide sequence of the DNA probe used (data not shown). A Northern analysis indicated the size of the relevant mRNA was about the same (~2 kb) as the large cDNAs (data not shown). Sequence analysis of these clones showed that they were virtually full-length (see Figure 7 and below). Further evidence that this cDNA library contains largely full-length clones was obtained by analysis of the clones of three other known rat pancreatic mRNAs. The library was screened with an oligonucleotide probe specific for carboxypeptidase A2 and cDNA probes for amylase and trypsin. Ten positive clones specific for each probe were randomly picked. In each case, 8 of the 10 clones were similar in size to the corresponding full-length mRNAs (estimated by Northern analysis). In the case of carboxypeptidase A2, the 5' end of the cDNA was sequenced and found to terminate 30 nucleotides downstream from the TATAAA sequence of the corresponding gene. We presume this represents the bonafide cap site (the only other upstream purine residues are at positions 18, 19, and 20; S. Gardell, unpublished results).

**Isolation and Complete Nucleotide Sequence of Putative Lysophospholipase mRNA.** One of the eight putative full-

length clones,  $\lambda$ rLPL3, was randomly selected for detailed sequence analysis. As shown in Figure 5, the cDNA is 2067 nucleotides long and contains a long open-reading frame (1839 nucleotides) with 21 additional nucleotides at the 5' end and 207 nucleotides at the 3' end including the poly(A) tail. The single-coding frame prescribes a protein of 612 amino acids with a molecular weight of 67 107. The size and amino acid composition of the predicted protein are remarkably similar to those of the reported bovine lysophospholipase (Table I; see Discussion). The nucleotide sequence of actATGg around the first methionine agrees with the conserved sequence of  $^{\text{a}}\text{xx-AUGg}$  around the initiation codon (Kozak, 1984). The polyadenylation signal AAUAAA is found 15 nucleotides before poly(A) addition in the 3'-noncoding region.

**5' End of the Putative Lysophospholipase mRNA.** The nucleotide sequences at the 5' end of the lysophospholipase mRNA and cloned cDNA were determined by primer-extension analysis and direct  $\lambda$  sequencing, respectively. The analysis of the extended synthetic oligonucleotide primer revealed that the 5'-end sequence of the mRNAs is heterogeneous as indicated by the stars in Figure 6A. About 50% of the mRNA species begin at the position indicated as star 5, 20% at stars 2 and 3, and 5% at stars 1 and 4 as determined by densitometric scanning of the first lane in Figure 6A. The first nucleotide of each mRNA could not be identified since

Table I: Comparison of the Amino Acid Composition of Bovine Pancreatic Lysophospholipase with That Predicted by the Isolated Pancreatic cDNA Sequence

amino acid	bovine pancreatic lysophospholipase	putative rat pancreatic lysophospholipase without predicted signal peptide
Ala	48 (8.2)	51 (8.8)
Arg	18 (3.1)	17 (2.9)
Asn	70 (12.0)	69 (11.9)
Asp		
Cys	5 (0.9)	4 (0.7)
Gln	45 (7.7)	46 (7.9)
Glu		
Gly	48 (8.2)	48 (8.2)
His	10 (1.7)	13 (2.2)
Ile	29 (5.0)	32 (5.5)
Leu	51 (8.8)	42 (7.2)
Lys	35 (6.0)	34 (5.8)
Met	11 (1.9)	15 (2.6)
Phe	25 (4.3)	27 (4.6)
Pro	33 (5.7)	44 (7.5)
Ser	37 (6.4)	29 (5.0)
Thr	42 (7.2)	42 (7.2)
Trp	12 (2.1)	11 (1.9)
Tyr	30 (5.2)	23 (4.0)
Val	33 (5.7)	35 (6.0)
no. of amino acids	582	582
mol wt <sup>a</sup>	67000	63400

<sup>a</sup>The molecular weight and amino acid composition of bovine pancreatic lysophospholipase are from van den Bosch et al. (1973). The values for the rat enzyme are obtained from the amino acid sequence derived from the nucleotide sequence of the cDNA minus the first 30 amino acids (putative signal peptide). The mole fractions are listed in parentheses.

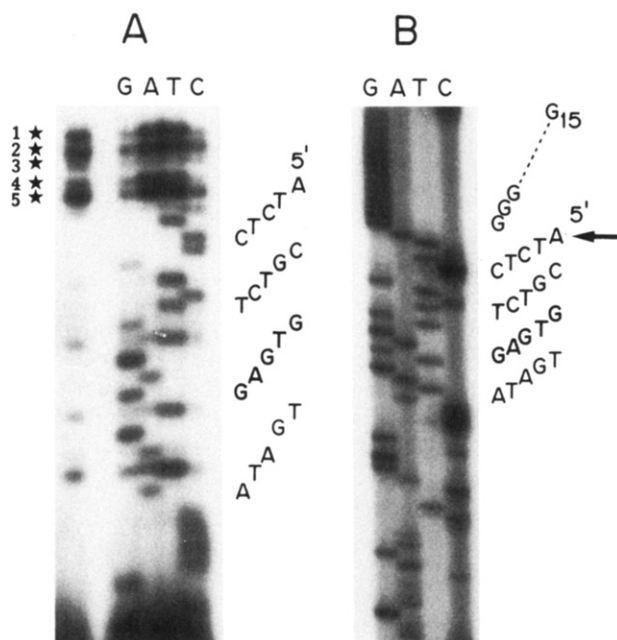


FIGURE 6: Comparison of 5'-end sequences of putative lysophospholipase mRNA and cDNA. The nucleotide sequence shown in (A) is the mRNA complement because the mRNA was used as a template for the chain termination sequencing reaction. The primer-extension analysis of the mRNA is shown in the left-hand lane. The sequence in (B) is obtained from direct  $\lambda$  sequencing (see Materials and Methods). Possible multiple 5' ends of the mRNA are marked by stars, and the end of the cloned sequence in  $\lambda$ LPL3 is indicated by an arrow in (B).

the elongated molecules terminate at these positions. Comparison of these extension results with the 5'-end sequences of the cDNA in  $\lambda$  phages indicates that star 5 corresponds to A, star 4 to T, star 3 to C, star 2 to C, and star 1 to G at the

Lysophospholipase mRNA ;		G <sup>m</sup> $\blacktriangledown\blacktriangledown\blacktriangledown\blacktriangledown$ GCCTAGAGGCAGA-----
		ppp
Number of $\lambda$ LPL clones obtained;		
1	G-----	18 -----GCCTAGAGGCAGA-----
1	G-----	14 -----CCTAGAGGCAGA-----
1	G-----	15 -----TAGAGGCAGA-----
4	G-----	15 -----AGAGGCAGA-----

FIGURE 7: Comparison of 5'-terminal sequences of rat pancreas lysophospholipase mRNA with 5'-end sequences of seven putative full-length lysophospholipase cDNAs determined by direct  $\lambda$  sequencing. Multiple-cap sites determined by primer extension analysis are marked by triangles on mRNA. The number of dG residues resulting from tailing are indicated before the 5' end of each cDNA sequence.

Table II: Relative Activity of Lysophospholipase and Chloramphenicol Acetyltransferase Expressed in CHO Cells<sup>a</sup>

	CHO transfected with pSV <sub>2</sub> CAT	CHO transfected with pECE <sub>1</sub> rLPL and pSV <sub>2</sub> CAT
lysophospholipase activity	1	12
CAT activity	1	1

<sup>a</sup> 300 nmol of 1-[<sup>14</sup>C]palmitoyllyso-3-phosphorylcholine ( $3 \times 10^{11}$  cpm/mol) was incubated for 30 min at 37 °C with 10  $\mu$ L of each CHO cell extract (50  $\mu$ g in protein) in a total volume of 100  $\mu$ L of 20 mM sodium phosphate buffer (pH 7.0). The reaction was stopped by adjusting the pH to 3.5. The <sup>14</sup>C-free fatty acid was extracted with petroleum ether, and the radioactivity in the organic phase was determined by scintillation spectrometry. An aliquot of the reaction was checked by thin-layer chromatography for free fatty acid. The CAT activity was determined as described (Gorman et al., 1982). Each enzyme activity detected in CHO transfected only with pSV<sub>2</sub>CAT was taken as 1 to obtain relative activity.

5'-terminus of the molecule (see Figure 7). Because of the 5'-end heterogeneity, it is not rigorously possible to assess whether the isolated cDNAs indeed represent full-length mRNA copies; however, the frequencies of the 5'-end sequence in the cDNAs appear to reflect those of primer-elongated molecules and are compatible with the view that each cDNA clone is full-length.

**Expression of cDNA and Demonstration of Lysophospholipase Activity.** In the absence of a peptide sequence for lysophospholipase (LPL) that could provide more definitive evidence for the identity of the cDNA sequence, we expressed the cloned cDNA in CHO cells using an SV40 expression vector (pECE<sub>1</sub>) (Ellis et al., 1986) and tested the biological activity of the protein. Cells were cotransfected with pECE<sub>1</sub>rLPL and plasmid pSV<sub>2</sub>CAT. In both vectors, the early SV40 promoter drives the expression of the lysophospholipase and CAT coding sequences, respectively. Control cells were transfected with pSV<sub>2</sub>CAT alone. Lysophospholipase activity was then tested in extracts by measuring the release of free fatty acid from the specific lysolecithin substrate. As shown in Table II, the fatty acid release was 12 times higher in cell extracts which were transfected with pECE<sub>1</sub>rLPL than in control cell extracts. The CAT activity, in contrast, was about the same in the two cell extracts, indicating the transfection was equally efficient in both cells. Thus, the cDNA sequence does indeed code for lysophospholipase.

## DISCUSSION

Several innovations in cDNA cloning methods have been made over the original procedures. Rabbitts first introduced a vector-priming method which allowed directional cloning of cDNA with improved efficiency (Rabbitts, 1976). Okayama and Berg (1982) refined the Rabbitts procedure and further introduced a more efficient method for second-strand cDNA synthesis which elicited a better yield of more complete copies

of a given mRNA (Okayama & Berg, 1982). More recently, Young and Davis (1983a) have developed a  $\lambda$  cDNA cloning method that exploits the high cloning and plaque screening efficiencies of the phage system. In this method, the cloned cDNA is expressed as a relatively stable fusion protein so screening of the library can be performed with an antibody to the desired protein.

The cloning strategy presented here employs the advantages of the  $\lambda$  system together with refinements for regenerating full-length cDNAs. The expression vector ( $\lambda$ gt18) has two unique cloning sites for the restriction enzymes *Sa*I and *Eco*RI, and an additional *Xba*I site which is of particular use if the  $\lambda$  DNA is *dam* methylated before treatment with restriction endonucleases. The polylinker allows directional cDNA cloning which increases the efficiency of producing positive clones. A key aspect of our cDNA cloning strategy is the use of versatile synthetic primer-adaptors that contain both homopolymer tails for the primed synthesis of cDNA and preformed restriction sites for cloning into the  $\lambda$  vector. These oligonucleotides (1) allow size selection after the first- or second-strand synthesis, (2) minimize the manipulation of cDNA necessary to create sticky ends for cloning, and (3) increase cloning efficiency. The successful use of a similar primer-adaptor for cDNA cloning in a plasmid vector (Coleclough & Erlitz, 1985) and in  $\lambda$ gt10 (Krawinkel & Zobelein, 1986) has recently been reported.

We have also improved the guanidinium thiocyanate method of mRNA isolation. Guanidinium thiocyanate combines the strong denaturing characteristic of guanidine with the chaotropic action of thiocyanate. This reagent efficiently solubilizes cell or tissue homogenates and denatures protein quickly. However, guanidinium thiocyanate, even in near-saturating concentrations, does not completely inactivate RNase in an RNase-rich homogenate such as that of adult rat pancreas. This has been the source of much variation in the original method (Chirgwin et al., 1979). We observed that guanidinium thiocyanate is effective at low temperature for isolation of mRNA. Thus, the RNase-rich tissue can be quickly homogenized in guanidinium thiocyanate solution at 0 °C and the RNA rapidly separated by ethanol precipitation at -10 °C. This step eliminates the bulk of RNase activity and makes the remaining steps less critical. The RNA preparations prepared by this simple method appear superior to those prepared by other methods. Using this method, we have been able to isolate virtually intact RNA species from various sources such as placenta, intestine, and several cultured cell lines.

Using our method for mRNA isolation and cDNA cloning, we have isolated a complete nucleotide sequence of a 2.1 kb mRNA by a single screening of our cDNA library. A computer search of our cloned 2.1 kb (lysophospholipase) cDNA for both nucleotide and deduced amino acid sequence homology in the data bank revealed no significant strong homologies. However, the size and amino acid composition of the predicted protein resemble that of bovine pancreatic lysophospholipase. As shown in Table I, the predicted amino acid composition of the putative rat lysophospholipase (without the putative signal peptide) is similar to that of the bovine enzyme. Furthermore, close inspection reveals that the tryptophan and cysteine residues are present in quite similar frequencies in both proteins. Tryptophan is evolutionarily conserved in low frequency in many proteins (Dayhoff et al., 1978), and cysteine is important for protein conformation. The detection of lysophospholipase activity in extracts from CHO cells transfected with LPL cDNA is persuasive evidence that the cloned cDNA represents lysophospholipase. However, amino acid sequence

comparison and/or demonstration of common antigenic determinants is required to definitively identify the cloned rat enzyme as lysophospholipase.

The isolated putative lysophospholipase cDNA exhibits structural features resembling the cDNAs of other rat pancreatic secretory proteins. The cDNAs (mRNAs) encoding trypsin (MacDonald et al., 1982a), elastases I and II (MacDonald et al., 1982b), chymotrypsin B (Bell et al., 1985), carboxypeptidase A (Quinto et al., 1982), and amylase (MacDonald et al., 1982c) all have short (11–21 nucleotides) 5'-noncoding leader sequences. Primer-extension analysis shows that the gene is transcribed at multiple start sites and that lysophospholipase mRNA also has a short leader of 17–21 nucleotides. The putative lysophospholipase mRNA, however, lacks a sequence within the 5'-noncoding region that is complementary to the conserved oligonucleotide sequence at the end of the 3' end of eucaryotic 18S rRNA. The putative lysophospholipase contains an N-terminal peptide which resembles a signal (Watson, 1984). We cannot infer from the sequence whether the enzyme has an activation peptide. The presence of a glycosylation site at amino acid residue 207 indicates the possibility of glycosylation. Another structurally interesting feature of this enzyme is 4 blocks of tandem amino acid repeats located 15 amino acids before the carboxy-terminal end of the protein. These repeats are regularly separated by a dipeptide of the same amino acid. The repeating unit consists of the consensus nanopeptide (PVPPTDDSQ), but the first two amino acid residues are deleted and the fifth and ninth residues are substituted with E in the first block (PPEDDSE). At present, the role of these repeats in the structure and function of the enzyme is not known.

These results illustrate the advantage of these methods in the routine construction of full-length cDNA libraries. Besides the improvements in the methodology for intact mRNA isolation and efficient cDNA cloning, these full-length expression libraries should improve the prospects for the efficient screening for individual proteins based on their biological or immunological properties.

#### ADDED IN PROOF

A recent report (Rogers et al., 1986) indicated that proteins with a short half-life have common structural domains rich in proline which may serve as a signal for selective proteolytic degradation (PEST hypothesis). All four tandem amino acid repeats found in the cloned lysophospholipase show remarkable structural similarity to the PEST regions demonstrated in many unstable proteins. Because of the presence of these repeats in lysophospholipase, the possibility that it will have a short half-life can now be tested.

#### ACKNOWLEDGMENTS

We thank Dr. D. Standring for reading the paper, Dr. M. Walker for helpful discussions concerning the expression of lysophospholipase cDNA in a mammalian cell, J. Barnett for synthesizing the oligonucleotides, J. Turner for preparing plasmid DNAs, and L. Spector for assistance in preparing the manuscript.

**Registry No.** DNA (rat pancreas phospholipase B messenger RNA complementary), 106587-90-0; phospholipase B (rat pancreas precursor protein moiety reduced), 106587-93-3; phospholipase B (rat pancreas protein moiety reduced), 106587-91-1; phospholipase B, 9001-85-8; guanidinium thiocyanate, 593-84-0.

#### REFERENCES

- Aviv, H., & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A.* 69, 1408–1412.



- Bell, G. I., Quinto, C., Quiroga, M., Valenzuela, P., Craik, C. S., & Rutter, W. J. (1984) *J. Biol. Chem.* 259, 14265-14270.
- Cathala, G., Savouret, J., Mendez, B., West, B., Karin, M., Martial, J., & Baxter, J. (1985) *DNA* 2, 329-335.
- Chen, E., & Seeburg, P. H. (1985) *DNA* 4, 165-170.
- Chirgwin, J., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry* 18, 5294-5399.
- Coleclough, C., & Erlitz, F. (1985) *Gene* 34, 305-314.
- Craik, C. S. (1985) *BioTechniques*, 12-19.
- Dayhoff, M. O., Hunt, L. T., & Hurst-Calderone, S. (1978) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., Ed.) Vol. 5, Suppl. 3, p 363, National Biomedical Research Foundation, Washington, D.C.
- DePierre, J. W. (1977) *Anal. Biochem.* 83, 82-90.
- Ellis, L., Clauser, E., Morgan, D., Edery, M., Roth, R., & Rutter, W. J. (1986) *Cell (Cambridge, Mass.)* 45, 721-732.
- Gorman, C., Moffat, L. F., & Howard, B. H. (1982) *Mol. Cell. Biol.* 2, 1044-1051.
- Graham, F. L., & van der Eb, A. J. (1973) *J. Virol.* 52, 456-467.
- Han, J. H., Rall, L., & Rutter, W. J. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 110-114.
- Kozak, M. (1984) *Nucleic Acids Res.* 12, 857-872.
- Krawinkel, U., & Zobelein, R. (1986) *Nucleic Acids Res.* 14, 1913.
- Land, H., Grez, M., Hansen, H., Lindermaier, W., & Schuetz, G. (1981) *Nucleic Acids Res.* 9, 2251-2266.
- MacDonald, R. J., Stry, S. J., & Swift, G. (1982a) *J. Biol. Chem.* 257, 9724-9732.
- MacDonald, R. J., Swift, G. H., Quinto, C., Swain, W., Pictet, R., Nikovits, W., Jr., & Rutter, W. J. (1982b) *Biochemistry* 21, 1453-1463.
- MacDonald, R. J., Crerar, M. M., Swain, W. F., Pictet, R. L., Thomas, G., & Rutter, W. J. (1982c) *Nature (London)* 287, 117-122.
- Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Maxam, A., & Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560.
- Okayama, H., & Berg, P. (1982) *Mol. Cell. Biol.* 2, 161-170.
- Pelham, H. F. B., & Jackson, R. J. (1976) *Eur. J. Biochem.* 67, 247-266.
- Quinto, C., Quiroga, M., Swain, W. F., Nikovits, W. C., Standring, D. N., Pictet, R. L., Valenzuela, P., & Rutter, W. J. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79, 31-35.
- Rabbitts, T. H. (1976) *Nature (London)* 260, 221-225.
- Rogers, S., Wells, R., & Rechsteiner, M. (1986) *Science (Washington, D.C.)* 234, 364-368.
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463-5467.
- Schick, J., Kera, H., & Scheele, G. (1984) *J. Cell Biol.* 99, 1569-1574.
- Stahl, F. W., Crasemann, J. M., & Stahl, M. M. (1975) *J. Mol. Biol.* 94, 203-212.
- Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischler, E., Rutter, W. J., & Goodman, H. M. (1977) *Science (Washington, D.C.)* 196, 1313-1319.
- Van den Bosch, H., Aarsman, A. J., DeJong, G. N., & Van Deenen, L. M. (1973) *Biochim. Biophys. Acta* 296, 94-104.
- Van Nest, G., MacDonald, R. J., Raman, R. K., & Rutter, W. J. (1980) *J. Cell Biol.* 86, 784-794.
- Watson, M. E. E. (1984) *Nucleic Acids Res.* 12, 5145-5164.
- Young, R. A., & Davis, R. W. (1983a) *Proc. Natl. Acad. Sci. U.S.A.* 80, 1194-1198.
- Young, R. A., & Davis, R. W. (1983b) *Science (Washington, D.C.)* 222, 778-782.