## Molecular cloning and sequence of a cDNA coding for bovine lipoprotein lipase

Mototaka Senda\*, Kazuhiro Oka<sup>†</sup>, W. Virgil Brown<sup>†</sup>, Pradman K. Qasba<sup>‡</sup>, and Yasuhiro Furuichi<sup>\*§</sup>

\*Department of Molecular Genetics, Nippon Roche Research Center, Kamakura, 247 Japan; <sup>†</sup>Department of Medicine, Mount Sinai School of Medicine, New York, NY 10029; and <sup>‡</sup>Division of Cancer Biology and Diagnosis, National Cancer Institute, Rockville, MD 20851

Communicated by John J. Burns, March 9, 1987

ABSTRACT Lipoprotein lipase (LPL; triacylglycero-protein acylhydrolase, EC 3.1.1.34) was purified from bovine milk. Synthetic oligonucleotides were prepared, based on the amino acid sequences of three peptides obtained from partial digestion of purified LPL, and were used as probes to isolate cDNA clones for LPL mRNA from a bovine mammary gland. One of the clones, pLPL-49R2, contains an insert cDNA (49R2) of about 3.2 kilobases (kb) that hybridizes to all three probes and encodes a polypeptide that includes the NH<sub>2</sub>-terminal sequence of bovine LPL reported recently [Ben-Avram, C. M., Ben-Zeev, O., Lee, T. D., Hagga, K., Shively, J. E., Goers, J., Pedersen, M. E., Reeve, J. R. & Schotz, M. C. (1986) Proc. Natl. Acad. Sci. USA 83, 4185-4189]. Complete nucleotide sequence analysis revealed that cDNA insert 49R2 contains the entire coding region for LPL as well as a 3' untranslated region of about 1.6 kb. The predicted amino acid sequence indicates that bovine LPL is a hydrophilic protein consisting of 450 amino acids ( $M_r$  50,548) in its unglycosylated form. Blot hybridization analysis of poly(A)<sup>+</sup> mRNA from bovine mammary gland demonstrated that there are at least three sizes of LPL mRNAs-3.2, 2.5, and 1.7 kb-with the 2.5-kb mRNA being the most abundant. Restriction endonuclease mapping of other cDNA clones suggested that the variation in mRNA size results from differential utilization of polyadenylylation signals during mRNA processing.

Lipoprotein lipase (LPL; triacylglycero-protein acylhydrolase, EC 3.1.1.34) is a lipolytic enzyme involved in the metabolism of triacylglycerol-rich lipoprotein particles (1). It hydrolyzes the triacylglycerol cores of chyromicrons and low density lipoproteins, generating free fatty acids available for extrahepatic tissues. LPL is a glycosylated protein of apparent  $M_r$  55,000 that is synthesized and secreted from parenchymal cells of adipose tissue, heart, skeletal muscle, and mammary gland and is anchored on the luminal surface of the capillary endothelium by membrane-bound heparan sulfate (2, 3). The activity of LPL is markedly enhanced by apolipoprotein C-II present on the surface of triacylglycerol-rich lipoproteins (4). Recent findings (5) indicate that LPL is inhibited by tumor necrosis factor (cachectin). Detailed genetic analyses of several mutant mouse strains suggest that LPL is involved in the cause of atherosclerosis (6). LPL is also present in milk, but the reason for its presence is unknown (1).

Since the first description of canine LPL in 1943 (7), a sustained effort has been made to characterize its enzymatic activity and functions (for review, see refs. 1, 2, and 6). However, relatively little is known about the structure of LPL or the gene encoding LPL, or about the factors that govern its expression, despite clinical evidence that familial hypertriglyceridemia is a genetic disorder (8). Recently,

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.

Ben-Avram *et al.* (9) described the  $NH_2$ -terminal and partial amino acid sequence of bovine LPL and its homology to pancreatic lipase (10). We report here the isolation and characterization of a cDNA clone that contains the entire coding region of mature bovine LPL.

## **MATERIALS AND METHODS**

**Characterization of Bovine Milk LPL.** *Purification.* LPL was purified from low-fat bovine skim milk according to the method described by Bengtsson and Olivecrona (11), with modifications. In brief, a 5000-fold purification of LPL was obtained by heparin-agarose column chromatography followed by phenyl-agarose column chromatography. The assay for LPL enzyme activity was carried out as described (12). Polyacrylamide gel electrophoresis according to Laemmli (13) was carried out in a slab gel, and the gels were stained with Coomassie brilliant blue R-250.

Digestion of LPL with trypsin and cyanogen bromide, and separation of peptides by reverse-phase HPLC. Purified LPL (50  $\mu$ g) was treated with trypsin (0.5  $\mu$ g) in 20 mM Tris/HCl buffer (pH 7) for 14 hr at 37°C. For chemical digestion, S-carboxymethylated LPL (50  $\mu$ g) was incubated with cyanogen bromide ( $\approx$ 50 mg) in 50  $\mu$ l of 70% formic acid 24 hr at room temperature. The digests were fractionated on a reverse-phase column (8.0 mm  $\times$  75 mm, Chemcosorbs ODS-H, Chemco Scientific, Osaka, Japan), using a 40-min linear gradient of 0–60% acetonitrile in 0.1% trifluoroacetic acid at a flow rate of 1 ml/min. The fractionated peptides were recovered by evaporation of the solvent.

Microsequence analysis. A gas-phase peptide/protein microsequencer [Applied Biosystems (Foster City, CA), model 470A] was used for automated Edman chemical degradations. The phenylthiohydantoin amino acid derivatives were identified by reverse-phase HPLC on an Aquasil SEQ-4K column (4.6 mm  $\times$  30 cm; Senshu, Tokyo) and were quantitated by an integration program on a Spectra Physics (San Jose, Ca) 4270 spectrophotometer.

Molecular Cloning of cDNA Encoding LPL. cDNA library. A cDNA library was prepared from poly(A)<sup>+</sup> mRNA obtained from lactating bovine mammary gland, using the Okayama–Berg procedure (14). Double-stranded cDNAs were ligated to plasmid pCD-X (15) and used to transform *Escherichia coli* MC1061 cells.

Oligonucleotide probes. Based on the amino acid sequence data, several oligodeoxyribonucleotide probes (21- to 23mers) were synthesized using an Applied Biosystems 381A DNA synthesizer. These probe DNAs were 5'-<sup>32</sup>P-labeled ( $\approx 10^8 \text{ cpm}/\mu g$ ) by polynucleotide kinase and were used in colony hybridization (16) to screen for bacterial clones that contain LPL-specific sequences. Hybridization was carried out overnight at 50°C in 6× SSC containing 5× Denhardt's

Abbreviation: LPL, lipoprotein lipase.

<sup>§</sup>To whom reprint requests should be addressed.

Step	Volume, ml	Protein, mg	Activity, μmol/hr	Specific activity, µmol/hr per mg of protein
Skim milk	1800	25,090	52,200	2.1
Heparin-agarose column	675		ar ann	10 100 5
chromatography	275	358	36,300	10,139.7
Phenyl-agarose column				
chromatography	19	0.49	5,472	11,167.3

LPL activity was assayed by the method of Baginsky and Brown (12) and expressed as  $\mu$ mol of fatty acid released during a 60-min incubation at 37°C.

solution, 0.5% NaDodSO<sub>4</sub>, and 20  $\mu$ g of salmon sperm DNA per ml. (SSC is standard saline citrate; 1× SSC is 0.15 M NaCl/0.015 M sodium citrate, pH 7.0. Denhardt's solution (1×) is 0.02% Ficoll/0.02% polyvinylpyrrolidone/0.02% bovine serum albumin.) Filters were washed in 6× SSC at 50°C for 2 hr with three changes of washing buffer.

DNA and RNA blot analyses. To verify that cDNA clones contained LPL-specific sequences, transblotting analysis by the Southern procedure (17) was carried out for individual plasmid DNAs, using 5'-<sup>32</sup>P-labeled synthetic DNA probes.

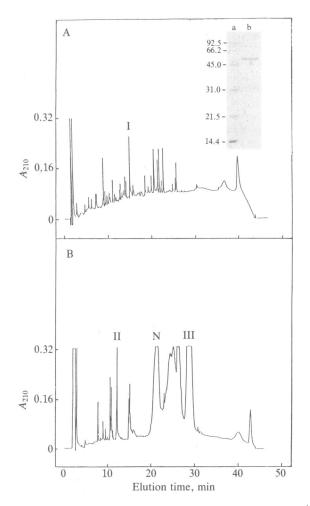


FIG. 1. Reverse-phase HPLC maps of peptides obtained by digestion of purified LPL (50  $\mu$ g) with trypsin (A) or with cyanogen bromide. Peak fractions I, II, and III were further characterized by ainito acid sequencing. Peak N contained the NH<sub>2</sub>-terminal peptide of LPL. The amino acid sequences of these peaks are given in the text. (*Inset*) Gel electrophoresis of purified bovine milk LPL. Purified enzyme was electrophoresed in a 12% polyacrylamide gel in the presence of 0.1% NaDodSO<sub>4</sub> and 0.1% 2-mercaptoethanol. The gel was stained with Coomassie brilliant blue. Lane a: molecular weight markers ( $M_{\rm r} \times 10^{-3}$  at left). Lane b:  $\approx 10 \ \mu$ g of LPL.

Similarly, RNA blot analysis (18) was performed for  $poly(A)^+$ mRNA prepared from bovine mammary gland (19) to detect LPL-specific mRNA species. For this, the *Pst* I fragment [541 base pairs (bp)] from 49R2 cDNA (bases 313–853 in Fig. 3) was <sup>32</sup>P-labeled by nick-translation and used as the probe.

Nucleotide sequencing. Sequencing of the cloned LPL cDNA was carried out by the M13 procedure of Sanger (20).

## RESULTS

Amino Acid Sequence Analysis of LPL. LPL was purified to homogeneity from bovine skim milk (Table 1; Fig. 1 A Inset). The purified LPL was then subjected to amino acid and peptide analyses as described in Materials and Methods. Our amino acid composition data were consistent with the data reported by Ben-Avram et al. (9). Also, our NH<sub>2</sub>-terminal sequence data (a stretch of 22 amino acid residues: Asp-Arg-Ile-Thr-Gly-Gly-Lys-Asp-Phe-Arg-Asp-Ile-Glu-Ser-Lys-Phe-Ala-Leu-Arg-Thr-Pro-Glu) agreed with the 19 residues determined by Ben-Avram et al. and, as discussed in a later section, are also consistent with the sequence predicted from the nucleotide sequence of the cloned LPL cDNA. These results not only identify the purified polypeptide as LPL but also confirm that the NH<sub>2</sub> terminus of bovine milk LPL is not modified, for example, by N-acetylation.

To obtain more information about the amino acid sequence, we digested the purified protein with pancreatic trypsin or cyanogen bromide, and the resulting peptides obtained as discrete peaks by HPLC were sequenced.

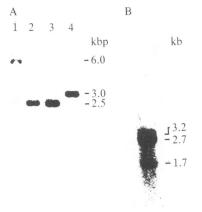


FIG. 2. Transfer blot analyses of cloned cDNA (A) and poly(A)<sup>+</sup> RNA (B) from bovine mammary gland. (A) cDNA clones pLPL-49R2 (lane 4), pLPL-49R3 (lane 2), and pLPL-54R1 (lane 3) were digested with restriction endonuclease *Bam*HI, electrophoresed in a 1% agarose gel, and hybridized with  $5'^{-32}$ P-labeled probe A after transblotting to nitrocellulose paper (18). Lane 1 shows undigested pLPL-49R2. (B) Poly(A)<sup>+</sup> RNA was electrophoresed in a 1% agarose/6% formaldehyde gel and blotted onto a Hybond N filter (Amersham). The RNA was fixed on the filter by exposure to UV (300-nm) light for 5 min and hybridized with a <sup>32</sup>P-labeled *Pst* I DNA fragment (541-bp fragment from 5'-proximal region).

Peptides in peaks I, II, and III (Fig. 1) contained the sequences Ile-Thr-Gly-Leu-Asp-Pro-Ala-Gly-Pro-Asn-Phe--Glu-Tyr-Ala-Glu-, (Met)-Gly-Tyr-Glu-Ile-Asn-Lys-Val-Arg--Ala-Lys-Arg-Ser-Ser-Lys-Met-, and (Met)-Ala-Asp-Glu-Phe--Asn-Tyr-Pro-Leu-, respectively. Mixed oligodeoxyribonucleotide probes potentially coding for these peptides were prepared, based on codon degeneracy and frequency of eukaryotic codon usage (21): probe B (a 20-mer corresponding to a peptide in peak I) was a 48-fold-degenerate mixture of the sequence 5'-ATCACRGGSCTSGAYCCVGC-3'; probe C (a 21-mer corresponding to a peptide in peak II) was a 16-fold-degenerate mixture of the sequence 5'-ATGGGY-TAYGARATYAACAAG-3'; and probe A (a 23-mer corresponding to a peptide in peak III) was a 128-fold-degenerate mixture of the sequence 5'-ATGGCNGAYGARTTYAART-AYCC-3'. (Y is C or T; R is A or G; S is C or G; V is A, C, or G; and N is A, C, G, or T.)

Isolation and Characterization of cDNA Clones. About

49R3 [.] 54R1 10 20 30 10 30 (G), TEGEEGEAGAEAGGATTACAGGAGGAGAAGGATTTTAGAGACATTGAAAGTAAATTT D R I E G K <u>P F R D I E S K F</u> 90 110 100 120 GCTCTCAGGACTCCCGAAGACACAGCTGAGGACACTTGCCACCTCATTCCTGGAGTGACC A I R T P E D T A E D T C II L I P G V T 130 1.10 160 260 200 280 290 300 AGGGAACTGGACTGGTGGTGGACGGGCGCAGGGGCCAGCAGCAGCATAT R E P D S N V I V D V L S R A Q Q H Y CCAGTGTCIGCAGGGTACACCAAGCTGGTGGGACAGGATGTGGCCAAGTTTATGAACTGG ΑΤΟΘΟΟΘΑΤΩΑΛΕΤΙΑΑΘΤΑΤΟΟΟΟΤΟΘΟΟΑΤΟΤΟΟΛΤΟΤΟΤΟΟΛΤΑΟΛΟΟΟΤΤΟΘΟ  $(\mathbf{A}) = (\mathbf{111})_{110}$ 490 500 510 520 530 540 TTAGATCCAGCTGGACCTAACTTCGAGTATGCAGAAGCTCCAAGTCCCCTTTCTCCTGAT L D P A G P N F E Y A E A P S R I S P N 550(B) = (11500)GATGUGGATTTTGIAGACGTTTTAUAUACACTTUACCAGGGGGTUAUCAGGTUGAAGTATU DADIVUVIL HTFTRGSPGRSI 620 GGAATCCAGAAACCAGTAGGGCATGTGGACATTTACCCTAACGGAGGCACTTTCCAACCA G I U K P V G II V D I Y P N G G T F U P GGATGTAACATTGGGGAAGCTCTCCCGTGTGAATGCAGGAGGAGGGCCTTGGAGATGTGGA G C H I G E A L R V I A E R G L G D V D 790 800 810 820 830 840 GANGANANTCCANGTANGGCUTACCGGTGCGANTCCANAGANGCCTTTGAGANAGGTCTC E E N P S K A Y R C N S K E A F E K G L

600,000 transformants, derived from a cDNA library constructed from lactating bovine mammary gland poly(A)<sup>+</sup> RNA in Okayama-Berg vectors, were screened by using labeled oligonucleotides as hybridization probes. <sup>32</sup>P-labeled probe A was used as the primary probe to screen all transformants, and probes B and C were used as second and third sorting probes to select LPL-specific cDNA clones. Three cDNA clones were identified that reacted with probes A, B, and C, and one clone that reacted only with probes B and C. The cDNA inserts of these plasmids were excised by digestion with endonuclease BamHI. Their sizes ranged from 1.7 to 3.2 kilobases (kb). The plasmid with the longest insert (3.2 kb) was designated pLPL-49R2; the two with mediumsized inserts (2.5 kb) were called pLPL-49R3 and pLPL-54R1. Upon Southern blot analysis, labeled probe A reacted with all inserts examined (Fig. 2A). In anticipation that cDNA insert 49R2 might contain the entire coding region for LPL, we analyzed this clone by restriction enzyme mapping and

850 860 870 880 890 900	
TGCCTGAGCTGCAGAAAGAACCGTTGCAACAACATGGGCTACGAGATCAACAAGGTCAGA C L S C R K N R C H N <u>H G Y E I H K V R</u>	
910 920 930 940 $950^{\circ} = (11)_{950}$	
GCCANANGANGCAGCAAGATGTACCTGAAGACTCGTTCTCAGATGCCTTACAAAGTCTTC A K R S S K <u>H Y</u> L K T R S Q H P Y K V F	
970 980 990 1000 1010 1020 CATTACCANGTANAGATACATTACTGGANGTAGAGAGTANTACATCACCANGCGAGCC H Y Q V K L H F S G T E S H T Y T N Q A	
1030 1040 1050 1060 1070 1080	
TTCGAGATCTCTCTGTATGGCACTGTGGGTGAGAGGGGAGAACATCCCTTTTACCCTGCCT F E I S L Y G T V A E S E N I P F T L P	
$ \begin{array}{cccc} GAAGTTTCCACAAATAAGACATACTCCTTTTCACCGGAGGTGGATATTGGAGAA\\ \mathsf{E &V &S &T & \overset{H}{\overset{H}{\overset{K}{\overset{T}}{\overset{T}{\overset{T}}{\overset{T}{\overset{T}{\overset{T}{\overset{T}}{\overset{T}{\overset{T}{\overset{T}{\overset{T}{\overset{T}}{\overset{T}{\overset{T}{\overset{T}{\overset{T}{\overset{T}{\overset{T}{\overset{T}}}}}}}}}$	
1150 1160 1170 1180 1190 1200	
TTGCTAATGTTGAAACTCAAATGGATTAGTGATTCCTACTTCAGCTGGTCCAACTGGTGG L L H L K L K V I S D S Y F S V S H V V	
1210 1220 1230 1240 1250 1260	
AGCAGCCCCGGCTTTGATATTGGGAAGATCAGAGTAAAGGCAGGAGAGAGA	
1270 1280 1290 1300 1310 1320	
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	
1330 1340 1350 1360 1370 1380	
TTTGTGANATGCCATGACAAGTCCCTGAATAGAAAGTCTGGCTGAAACTGGGCAAAGCTA F V K C II D K S L N R K S G $\Rightarrow$	
CGGAAGAAAG AACAGCATAT GAATTCTATG AAGAATGAAG TAACTTTTAC AAAAGATGCO	2
CAGTGCTTTA GATGGTGAAA TGTGGATTTT CCGGAGTATT AACCCCAGCT CTAGCCTTA	
TAGTTATTTA GGAGACAGTC TCAAATACTA AAACTAATTC AATTTGAGGT GTATAGTGG	
CAAATAGCAC ATCTTCCAAC ATTAAAAAAA TAACAGATAT GAAAAGCACT GCATTCTGT	
TITITGAAAAA ATATGAGTTA TITAAAATGA TAAAATAATC AGATCTCTTC ATGTAGTAA	
TITAGCCATA GCGTAAAATT GATTAAAATC TCATTTTTAA TTGGGATTCT GCGTTTTCT.	
GACTGATAAC CTICTCAGAG TTTTCTCCGA GTCTAAATAT AGGAAGTAAG TTTTTTTGC	
CONTRACTOR CACCOUNTER CONTRACTOR ANALYCICATA CONTRACTOR	A
GCGTGAGTAG GACCCGTTTA CCTATCAATC AAAATCCCTA CTTTCTTGGA ATTACTCTC	A C
TCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA	A C T
TCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAACCCCC TCTACGGTTG GTTGCATCA	A C T
ΤCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAAACCCCC TCTACGGTTG GTTGCATCA AACTAAGTTA CCAATTAAAG GAGATATATA AAGTTGAGAT CAATTAAATC TT <u>AATAAA</u> C	A C T T
ТСТТЕССАТТ СЕЛАТСТАЕТ ССААБЛАТСЯ АССАЕВААСТ АСТОВСАТТСС АСАТАСЛАА СААСААТАСА СТТОАТАААС САСТСААССТ ТТАААССССС ТСТАСССТТС СТТССАТСА ААСТААСТТА ССААТТАААС САСТСААССТ ТТАААССССС ТСТАСССТТС СТТССАТСА ТААСАСТТА ССААТТАААС САСТСААТТАТА ААСТТСАСАТ СААТТАААТС ТТ <u>ААТААА</u> С ТААСАСТТТА ТССТТАСТА ТТГССССТТС СТТТССТСС ТТГСТССА САТТАТАТ	
ТСТТВССАТТ GGAATGTAGT ССААGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAACCCCC TCTACGGTTG GTTGCATCA AACTAAGTTA CCAATTAAAG GAGATATATA AAGTTGAGAT CAATTAAATC TT <u>AATAAAG</u> TAACAATTTA TGGTTTAGTA TTICCCCTTC CTTTTCCTCC TTTGCTCCAA GATTATATT TAATAGTTTT (TCCCTAGATA GGCTTTTCAA CTGAGCCTGT TATCATCAGC TGATACAGA	
ТСТТВССАТТ СБААТБТАБТ ССААБААТБА АССАББААСТ АБТБАСТТББ АБАТАБААА БААБААТАБА СТТБАТАААБ САСТБААССТ ТТАААССССС ТСТАССБТТБ СТТБСАТСА ААСТААБТТА ССААТТАААБ САСТБААССТ ТТАААССССС ТСТАССБТТБ СТТБСАТСА ТААСАБТТТА ТБСТТТААБТА ТТТСССТТС СТТТТССТБС ТТГБСТСТСА БАТТАТАТ ТААТАБТТТТ (ТСССТАБАТА БЕСТТТТСА СТБАССТТ ТАТСАТСАБС ТОАТАСАБА ТТТТААСТАС АБАББДААААА АААТБАТАСТ БТААТТТТАТ ТАТСАБАСТТ ТСАСААТТА	
TCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAACCCCC TCTACGGTTG GTTGCATCA AACTAAGTTA CCAATTAAAG GAGATATATA AAGTTGAGAT CAATTAAATC TT <u>AATAAAC</u> TAACAGTTTA TGGTTTAGTA TTTCCCCTTC CTTTTCCTGC TTTGCTCCAA GATTATATT TAATAGTTTT (TCCCTAGATA GGCTTTTCAA CTGAGCCTG TATCATCAGCAG TITTAACTAC AGAGGAAAAA AAATGATACT GTAATTTTAT TATCAGACTT TCACAGACT ATTCATTGAA GTCCTTTAAT GTGATAGTTA CTGTTCATTT TAGGCTTATT TCAGCAAT	
TCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAACCCCC TCTACGGTTG GTTGCATCA AACTAAGTTA CCAATTAAAG GAGATATATA AAGTTGAGAT CAATTAAATC TT <u>AATAAAC</u> TAACAGTTTA TGGTTTAGTA TTTCCCCTTC CTTTTCCTGC TTTGTCTCAA GATTATATT TAATAGTTTT (TCCCTAGATA GGCTTTTCCAA CTGAGCCTGT TATCATCAGC TGATACAGA TTTTAACTAC AGAGGAAAAA AAATGATACT GTAATTTTAT TATCAGACTT TCCACATTA ATTCATTGAA GTCCTTTAAT GTGATAGTTA CTGTTCATTT TATCAGACAT TTCAGTCAGA CTTTGAGCAT GTCCTTTCC CCCTAACTTT GATAGAAACA CGTATTCAC	
TCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAACCCCC TCTACGGTTG GTTGCATCA AACTAAGTTA CCAATTAAAG GAGATATATA AGTTGAGAT CAATTAAATC TT <u>AATAAAC</u> TAACAGTTA TGGTTTAGTA TTTCCCCTTC CTTTTCCTGC TTTGTCTCAA GATTATATT TAATAGTTTI (TCCCTAGATA GGCTTTTCAA CTGAGCCTGT TATCATCAGC TGATACAGA TTTAACTAC AGAGGAAAAA AAATGATACT GTAATTTTAT TATCAGGACTT TCCAGCCATG TTCATTGAA GTCCTTTAAT GTGATAGTTA CTGTTCATTT TATCAGGACTT TCCAGCCATG TTCAGTCAGA CTTTGAGCAT GTCCTTTCT CCTTACATTT TACGGCTTATT TCCAGCCATG TTCAGTCAGA CTTTGAGCAT GTCCTTTCT CCTTACATTT GATAGAACA CGTATTCAC AQTACTTTGC ATCATGCAAA GTCCTACAAA TTTTAGGCACA GGATTGATTT TACTAGGAA AGAACAGAA AAATTACAT CAGTCATATA GTCACTGCGA GTGCTTAAAA	
TCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAACCCCC TCTACGGTTG GTTGCATCA AACTAAGTTA CCAATTAAAG GAGATATATA AGTTGAGAT CAATTAAATC TT <u>AATAAAC</u> TAACAGTTA TGGTTTAGTA TTTCCCCTTC CTTTTCCTGC TTTGTCTCAA GATTATATT TAATAGTTTI (TCCCTAGATA GGCTTTTCAA CTGAGCCTGT TATCATCAGC TGATACAGA TTTAACTAC AGAGGAAAAA AAATGATACT GTAATTTTAT TATCAGGACTT TCCAGCCATG TTCATTGAA GTCCTTTAAT GTGATAGTTA CTGTTCATTT TATCAGGACTT TCCAGCCATG TTCAGTCAGA CTTTGAGCAT GTCCTTTCT CCTTACATTT TACGGCTTATT TCCAGCCATG TTCAGTCAGA CTTTGAGCAT GTCCTTTCT CCTTACATTT GATAGAACA CGTATTCAC AQTACTTTGC ATCATGCAAA GTCCTACAAA TTTTAGGCACA GGATTGATTT TACTAGGAA AGAACAGAA AAATTACAT CAGTCATATA GTCACTGCGA GTGCTTAAAA	
TCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAACCCCC TCTACGGTTG GTTGCATCA AACTAAGTTA CCAATTAAAG GAGATATATA AAGTTGAGAT CAATTAAATC TT <u>AATAAAC</u> TAACAGTTTA TGGTTTAGTA TTTCCCCTC CTTTTCCTGC TTTGTCTCAA GATTTAATT TAATAGTTTT (TCCCTAGATA GGCTTTTCCA CTGAGCCTG TATCATCAGC TGATACAGA TTTTAACTAC AGAGGAAAAA AAATGATACT GTAATTTAT TATCAGACTT TCACAATTA ATTCATTGA GTCCTTTAAT GTGATAGTTA CTGTTCATTT TATCAGACTT TCACAATTA ATTCATTGAG GTCTTTAAT GTGATAGTTA CTGTTCATTT TAGGCTTATT TCACAGTCAGA CTTTGAGCAT GTCCTTTCC CCCTAACTTT GATAGAAACA CGTATTCAC AQTACTTGC ATCATGCAAA GTCCTACAAA TTTTAGACAA AGATTGTTTT TACTAGGAAC	A C T T T T A C C G A A C
TCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAACCCCC TCTACGGTTG GTTGCATCA AACTAAGTTA CCAATTAAAG GAGATATATA AGTTGAGAT CAATTAAATC TT <u>AATAAAC</u> TAACAGTTA TGGTTTAGTA TTTCCCCTTC CTTTTCCTGC TTGGTCTCAA GATTATATT TAATAGTTTI (TCCCTAGATA GGCTTTTCAA CTGAGGCTGT TATCATCAGC TGATACAGA TTTAACTAC AGAGGAAAAA AAATGATACT GTAATTTTAT TATCAGGACTT TCCAGCCATG TTCATTGAA GTCCTTTAAT GTGATAGTAC TGTACATCAGC TGATACAGA ATTCATTGAA GTCCTTTAAT GTGATAGTAC TGTACATCATT TACGGCTTATT TCCAGCCATG TTCAGTCAGA CTTTGAGCAT GTCCTTTCT CCTACACTT GATAGAAACA CGTATCAAC AGTACTTGC ATCATGGAA GTCCTACAAA TTTTAGGCATA GGATTGTTT TACTAGGAA AGAACAGAA AAGTTACACT CAGTCATATA GTCACTGCA GTGCTTAAAA CCGTTGTGC CAATGAAAAG GATCAATTGC TATATCCACG TGAATGTCA TGAAAGTTGT TTTTAACAA	
TCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAACCCCC TCTACGGTTG GTTGCATCA AACTAAGTTA CCAATTAAAG GAGATATATA AAGTTGAGAT CAATTAAATC TAACAGTTTA TGGTTTAGTA TTTCCCCTTC CTTTTCCTGC TTTGTCTCAA GATTATATT TAATAGTTTI (TCCCTAGATA GGCTTTTCAA CTGAGGCTGT TATGATCAGC TGATAGAGA TTTAACTACTAC AGAGGAAAAA AAATGATACT GTAATTTTAT TATCAGGACTT TCACACAGA ATTCATTGAA GTCCTTTAAT GTGATAGTTA CTGTTCATTT TATCAGCACTT TCACACATA ATTCATTGAA GTCCTTTAAT GTGATAGTTA CTGTTCATTT TAGGAACA COTTTCACCATG TTCAGTCAGA GTCCTTTAAT GTGATAGTTA CTGTTCATTT TAGGAACA COTTTCACCAG AGTACTTGC ATCATGCAAA GTCCTACAAA TTTTAGGACAA AGATTGTTTT TACTAGGAA AGAACAGAAA AAGTACAGT CAGTCATATA GTCACTGGCA GTGCTTAAAA CCGTTGGG CAATGAAAAG GATCACAGT CAGTCATATA GTCACTGGCA TGAGTGTTAAAA CCGTTGTGG CAATGAAAAG GATCACATTGC TATATCCCAG TGAATGTCA TGAAAGTTGTT TTTAACACAG TAGTCAAGAG TGAGTGAACA TTTATTTATA AACTAGACTC CATATTTTCA GAATGCTT	
TCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAACCCCC TCTACGGTTG GTTGCATCA AACTAAGTTA CCAATTAAAG GAGATATATA AAGTTGAGAT CAATTAAATC TAACAGTTTA TGGTTTAGTA TTTCCCCTTC CTTTTCCTCC TTTGCTCAA GATTATATT TAATAGTTTI (TCCCTAGATA GGCTTTTCAA CTGAGCCTG TATCATCAGC TGATACGAT ATTCAATGATA GGCGAAAAA AAATGATACT GTAATTTTA TATCATGCAGCTT TCACGACA ATTCAATGAA GTCCTTTAAT GTGATAGTTA CTGTCCATT TACGACACTT TCACGCATG TTCAGTCAGA CTTTGAGCAT GTCCTTTCC CCCTAACTTT TAGGCTTATT TCCAGCCATG TTCAGTCAGA CTTTGAGCAT GTCCTTTCC CCCTAACTTT TAGGCATATT TCAGGCATG ATCATTGAA GTCCTTTAAT GTGATAGTTA CTGTCCATTT TAGGCTATT TCCAGCCATG TCAGTCAGA CTTTGAGCAT GTCCTTCTC CCCTAACTTT GATAGAAACA CGTATTCAC AQTACTTTGC ATCATGCAAA GTCCTACAAA TTTTAGCAA AGATTGTTTT TACTAAGTA AGAAACAGAAA AAGTTACAGT CAGTCATATA GTCACTGCA GGTCTTAAAA CGTATGTGCA TAGTCAAAGA GAGTCAATTGC TATATCCCACG TGAATGTCT TTTTAACAA TAGTCAAAGG GTGCAATTGC TATATCCCACG TGAATGTCT TTTTAACAAGT AGATGAAAAG GTGCAATTGC TATATCCCAGG TGAATGTCT CATATTTCC AGAGGTCATT CTATATATTA TTATCAAATG ATAGAGAAGT CAAATATTTC AGAGGCTATT ACTGAAAATG	
TCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAACCCCC TCTACGGTTG GTTGCATCA AACTAAGTTA CCAATTAAAG GAGATATATA AAGTTGAGAT CAATAAAAC TAACAGTTA CGATTAAAG GAGATATATA AAGTTGAGAT CAATAAAAC TAACAGTTA TGGTTAGTA TTTCCCCTTC CTTTTCCTGC TTTGCTCAA GATTATATT TAATAGTTTI (TCCCTAGATA GGCTTTTCAC CTGAGGCTG TATCATCAG TATCATTGAA GACCTTTAAT GTGATAGTA CTGAGGCTG TATCATCAG TAATACAG TTTAACTAC AGAGGAAAAA AAATGATACT GTAATTTTAT TATCAGGACTT TCACGCAG TTCAGTCAGA GTCCTTTAAT GTGATAGTA CTGTTCATTT TAGGCTATT TCAGCCATG TTCAGTCAGA GTCCTTTAAT GTGATAGTA CTGTTCATTT TAGGCAACA CCATTCAC AGTACTTTGC ATCATGCAAA GTCCTACAAA TTTTAGGCAA AGATTGTTTT TACTAAGAA AGAACAGAA AAGTTACAGT CAGTCATATA GTCACTGCA GGAGCTTAAAA CCGTTGTGG CAATGAAAAG GATCAATGC TATATCCACG TGAATGTCA TGAAGATGTTT ACTAAGAA TAGTCAAGAG TGAGTGAACA TTTATTTAT AACTAGAACT CATATTTCA GAAGCTTT CTATATATTA TTATCAAATG ATAGAGAAGT CAATATTTC AGAGGCTATT ACTGAAAAT CAAACTGTAGA TATGTGTGG GTGTCTGAAC ACATATTTCA GGAGGCTATT ACTGAAAAT CAAACTGTC AGGATGGCT CGGAAGGTA ATGATTAACG GGTACATTA ACTGAAAAT CAAACTGTC CAGCATGGC TGCGGAACA CTCTTCAACA CGATCGCAGCACATCA	A C T T T T A C C C A A C T T C C A T
TCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAACCCCC TCTACGGTTG GTTGCATCA AACTAAGTTA CCAATTAAAG GAGATATATA AAGTTGAGAT CAATTAAATC TT <u>AATAAAC</u> TAACAGTTTA TGGTTTAGTA TTTCCCCTTC CTTTTCCTGC TTTGCTCAA GATTATATT TAATAGTTTI (TCCCTAGATA GGCTTTTCAA CTGAGGCTT TATCATCAGC TGATACAGA TTTAATAGTTTI (TCCCTAGATA GGCTTTTCAC GTAATTTTAT AATCATTGAA GACCTTTAAT GTGATAGTTA CTGTGCATTT TATCATGACT TCACAATTA ATTCATTGAA GTCCTTTAAT GTGATAGTTA CTGTGCATTT TACGAGACTT TCACAATTA ATTCATTGAA GTCCTTTAAT GTGATAGTTA CTGTTCATTT TAGGAAACA CATATCAAG AGAACAGAA AAGTTACAGT GTCTTTTCC CCTAACTTT TAGGAACA CATATCAAG AGAACAGAA AAGTTACAGT CAGTCATAAA TTTTAGGCAA AGATTGTTTT TACTAGGAA GGAACAGAA AAGTTACAGT CAGTCATAAA TTTTAGGCA AGATTGTTTT TACTAGGAA AGAACAGAA AAGTTACAGT CAGTCATAAA GTCACAGCA GAATTGTCA TGAAGAACA CAGTTGCA TAGTCAAGAG TGAGTGAACA TTTATTATA AACTAGATCT CATATTTCA GAAGCTCT CTATATATTA TTATCAAATG ATAGAGAAGT CAAATATTCA GGAGGTAAT ACTGAAAAT CAACTGTAA AGTGCTGGG GTGTCTGAAC ACATATTCCA GGAGGCTATT ACTGAAAAT CAACTGTACA ATGGTGTGGG GTGTCTGAAC ACATATTCCA GGAGCTATT ACTGAAAAG GATTATTGC CAGCATGGC CGGAAGGTAA ATGATTAAGA GGAGAAGC CCACATACA GATATTGGT CAGCATGGT GGGAGGTAA ATGATTAAGA GGAGAAGC GGTGTGCGC GTACAAAAGT GAATGAAGC TGTGATGAC ATCATTCAC GTATGAAGCC CCACATACA GATAGAAAATG ATGGATGG TGCCGGAACA CCATATTCCA TGAGAATGC GGTGTGCC GTAGAAAAGT GAATGAATGC TGTGATTGAC ATCATTCAC GTATGAAGAC GGTGTGCCC	A C T T T T A C C C A A C T T C C A T
TCTTGCCATT GGAATGTAGT CCAAGAATGA ACCAGGAACT AGTGACTTGG AGATAGAAA GAAGAATAGA GTTGATAAAG CACTGAACCT TTAAACCCCC TCTACGGTTG GTTGCATCA AACTAAGTTA CCAATTAAAG GAGATATATA AAGTTGAGAT CAATAAAAC TAACAGTTA CGATTAAAG GAGATATATA AAGTTGAGAT CAATAAAAC TAACAGTTA TGGTTAGTA TTTCCCCTTC CTTTTCCTGC TTTGCTCAA GATTATATT TAATAGTTTI (TCCCTAGATA GGCTTTTCAC CTGAGGCTG TATCATCAG TATCATTGAA GACCTTTAAT GTGATAGTA CTGAGGCTG TATCATCAG TAATACAG TTTAACTAC AGAGGAAAAA AAATGATACT GTAATTTTAT TATCAGGACTT TCACGCAG TTCAGTCAGA GTCCTTTAAT GTGATAGTA CTGTTCATTT TAGGCTATT TCAGCCATG TTCAGTCAGA GTCCTTTAAT GTGATAGTA CTGTTCATTT TAGGCAACA CCATTCAC AGTACTTTGC ATCATGCAAA GTCCTACAAA TTTTAGGCAA AGATTGTTTT TACTAAGAA AGAACAGAA AAGTTACAGT CAGTCATATA GTCACTGCA GGAGCTTAAAA CCGTTGTGG CAATGAAAAG GATCAATGC TATATCCACG TGAATGTCA TGAAGATGTTT ACTAAGAA TAGTCAAGAG TGAGTGAACA TTTATTTAT AACTAGAACT CATATTTCA GAAGCTTT CTATATATTA TTATCAAATG ATAGAGAAGT CAATATTTC AGAGGCTATT ACTGAAAAT CAAACTGTAGA TATGTGTGG GTGTCTGAAC ACATATTTCA GGAGGCTATT ACTGAAAAT CAAACTGTC AGGATGGCT CGGAAGGTA ATGATTAACG GGTACATTA ACTGAAAAT CAAACTGTC CAGCATGGC TGCGGAACA CTCTTCAACA CGATCGCAGCACATCA	A C T T T T A C C C A A C T T C C A T

FIG. 3. Restriction endonuclease maps of the cDNA inserts 49R2, 49R3, and 54R1 are shown at the top of the left-hand column: P, S, BgII, N, B, and E indicate *Pst* I, *Sma* I, *Bgl* II, *Nco* I, *Bam*HI, and *Eco*RI restriction sites, respectively; the bold line in 49R2 indicates the region encoding the LPL amino acid sequence. The remainder of the figure shows the complete nucleotide sequence of the 49R2 cDNA and the predicted amino acid sequence of the bovine LPL (standard one-letter amino acid symbols, given below the nucleotide sequence). The sequence of the NH<sub>2</sub>-terminal peptide (see Fig. 1*B*, peak N) is underlined. Also underlined are the regions corresponding to the peptides (I, II, and III) on which the synthetic hybridization probes A (23-mer), B (21-mer), and C (20-mer) were based. Broken underlining shows the two potential N-glycosylation sites. The two potential cleavage/polyadenylylation signals (AATAAA) at bases 2033–2038 and 2882–2887 are underlined. The nucleotide sequence was determined by reading both strands except for the region (bases 2111–2342) in parentheses.

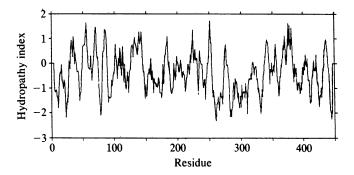


FIG. 4. Hydrophobic and hydrophilic regions of bovine milk LPL. The predicted primary amino acid sequence of LPL was computer-scanned for hydrophobicity, using the program of Kyte and Doolittle (23). The search length was nine amino acids. Hydrophobicity (hydropathy) increases with distance above the abscissa.

nucleotide sequencing. Restriction mapping of the other two inserts (49R3 and 54R1) showed that they lacked the 3'-terminal regions of 49R2 (Fig. 3).

Nucleotide Sequence of Cloned cDNA. The major emphasis in sequence analysis was to locate the coding region for the NH<sub>2</sub>-terminal and three internal peptides of LPL whose amino acid sequence had been established by microsequencing. As shown in Fig. 3, the complete nucleotide sequence of 49R2 (2903 bases) was determined. The region coding for the NH<sub>2</sub>-terminal peptide Asp-Arg-Ile-Thr-Gly-Gly-Lys-Asp-Phe-Arg-Asp-Ile-Glu-Ser-Lys-Phe-Ala-Leu-Arg-Thr-Pro-Glu was found at the 5' end of the cDNA insert (bases 13-78). The upstream region that should code for the signal peptide was absent in this clone, except for the sequence TCGC-CGCA that encodes three amino acids (Xaa-Ala-Ala, where Xaa is Phe, Leu, Ile, or Val) absent from the mature LPL. Regions that code for peptides I, II, and III (Fig. 1) were found at bases 472-516, 874-921, and 361-387 (Fig. 3), respectively. All four peptides were in the same open reading frame, which was interrupted by a terminator TGA at bases 1363-1365. There were many stop codons thereafter. This reading frame encompasses 1350 bases capable of encoding a polypeptide consisting of 450 amino acid residues. These sequence data predict that the mature bovine LPL has a molecular weight of 50,548, in its unglycosylated form.

The Bgl II-BamHI fragment located at the 3' end of 49R2 cDNA contains a poly(A) sequence and an AATAAA sequence 15 bases upstream of the poly(A), characteristics of eukaryotic mRNA. There is another AATAAA sequence in the middle of the 3' noncoding region (residues 2033–2038), which is preceded by several AATAAA analogue sequences and followed by a "G/T cluster," indicative of a strong mRNA cleavage/polyadenylylation site (22).

mRNA Analysis by Blot Hybridization. To determine the original size of LPL mRNA, total  $poly(A)^+$  RNA extracted from lactating bovine mammary glands was analyzed by blot hybridization (Fig. 2B) using as probe a <sup>32</sup>P-labeled Pst I-Pst I fragment (541 base pairs) present in all cloned cDNAs. The major hybridizable RNA band migrated at a position corresponding to 2.5 kb. Two other minor, but distinct, RNA bands of 3.2 and 1.7 kb were also detected (Fig. 2B).

Judging from their size and restriction maps, cDNA 49R2, and the two medium-sized cDNAs, 49R3 and 54R1, probably derive from the 3.2-kb and 2.5-kb mRNAs, respectively. While these three cDNAs share the same coding region (from restriction maps) and all contain poly(A) at the 3' ends (data not shown), both 49R3 and 59R1 cDNA have a shorter 3' noncoding region.

Structural Features of Bovine LPL. LPL is known to be a glycosylated protein associated with the luminal surface of capillary endothelium via membrane-bound heparan sulfate (1-3). The predicted amino acid sequence of bovine LPL contains two potential sites [Asn-Xaa-(Thr/Ser)] for N-linked glycosylation at amino acid residues 45-47 and 361-363. Hydropathicity plots (Fig. 4), constructed according to Kyte and Doolittle (23), indicate that bovine LPL is very hydrophilic, with the two potential glycosylation sites occurring at  $\beta$ -turns in hydrophilic domains. There are two regions rich in basic amino acids in the COOH-terminal region of LPL (amino acid residues 262-308 and 405-439) that could associate with the acidic domain of heparan sulfate. Both of these regions have an  $\alpha$ -helix configuration [based on analysis as described by Chou and Fasman (24)].

## DISCUSSION

In this paper we describe the cloning of mRNAs that encode bovine LPL. One of the cloned cDNAs (49R2) apparently contains the complete nucleotide sequence coding for mature LPL, although it lacks the 5' region encoding the signal peptide required for translocation into the lumen of the rough endoplasmic reticulum. Curiously, the 49R2 cDNA was found to contain a very long 3' untranslated region. Two other cDNA clones were apparently identical to 49R2 with respect to the LPL coding region but contained shorter 3' untranslated regions that include terminal poly(A). These results imply that there are multiple species of LPL mRNA in bovine mammary gland. In fact, blot analysis of  $poly(A)^+$ RNA from bovine mammary gland showed that there are at least three sizes of LPL mRNAs. Judging from the intensity of the hybridized bands (Fig. 2B), the medium-sized mRNA is the most abundant species in bovine mammary gland, whereas cDNA insert 49R2 seems to correspond to the longest (3.2-kb) mRNA species.

The biological significance and biogenesis of these three mRNA species remain to be elucidated. However, these mRNAs most likely derive from differential transcription termination, as shown in the adenovirus major late transcription unit (25), since the cDNA clones obtained in this study appear to be different only in the 3' noncoding region. Partial sequence analysis (data not shown) of the medium-size cDNA, 49R3, shows that the upstream AATAAA (at bases 2033–2038) apparently has been used for the cleavage/polyadenylylation of the template mRNA. A similar observation concerning various sizes of mRNA from bovine mammary gland was reported (26) for mRNAs encoding *N*-acetylglucosamine ( $\beta$ 1-4)galactosyltransferase.

Recently, Reddy *et al.* (27) reported the amino acid sequence of a peptide that contains a serine residue reactive to diisopropylphosphorofluoridate, a potential active site of bovine LPL. However, we did not locate the corresponding peptide in our predicted LPL sequence. In fact, no such sequence could be found even in the other two reading frames of the cloned LPL cDNA 49R2. Further analyses of genomic clones and assignment of individual cDNA clones to LPL polypeptides should open the way to understanding of regulation of LPL gene expression and clinical application for either diagnostics or as a remedy for genetic disorders.

We are grateful to Dr. J. Paterniti for encouragement, Ms. L. A. Houghton for technical assistance, and Ms. M. Shida for preparation of the manuscript. This work was supported in part by National Institutes of Health Grant HL32144.

- Nilsson-Ehle, P., Garfinkel, A. S. & Schotz, M. C. (1980) Annu. Rev. Biochem. 49, 667–693.
- 2. Cryer, A. (1981) Int. J. Biochem. 13, 525-541.
- Chimada, K., Gill, P. J., Silbert, J. E., Douglas, W. H. & Fanburg, B. L. (1981) J. Clin. Invest. 68, 995-1002.
- La Rosa, J. C., Levy, R. I., Herbert, P., Lux, S. E. & Fredrickson, D. S. (1970) Biochem. Biophys. Res. Commun. 41, 57-62.

- Mahoney, J. R., Jr., Beutler, B. A., LeJrang, N., Vine, W., Ikeda, Y., Kawakami, M. & Cerami, A. (1985) J. Immunol. 134, 1673-1675.
- Oka, K. & Paterniti, J. R., Jr. (1986) Drug Dev. Res. 7, 377– 392.
- 7. Hahn, P. F. (1943) Science 98, 19-21.
- Beaumont, J.-L., Carlson, L. A., Cooper, G. R., Feifar, Z., Fredrickson, D. S. & Strasser, T. (1970) Bull. W.H.O. 43, 891-915.
- Ben-Avram, C. M., Ben-Zeev, O., Lee, T. D., Haaga, K., Shively, J. E., Goers, J., Pedersen, M. E., Reeve, J. R. & Schotz, M. C. (1986) Proc. Natl. Acad. Sci. USA 83, 4185-4189.
- De Caro, J., Boudouard, M., Guidoni, A., Desnuelle, P. & Rovery, M. (1981) *Biochim. Biophys. Acta* 660, 148–150.
- 11. Bengtsson, G. & Olivecrona, T. (1977) Biochem. J. 167, 109-119.
- 12. Baginsky, M. L. & Brown, W. V. (1979) J. Lipid Res. 20, 548-556.
- 13. Laemmli, U. K. (1970) Nature (London) 227, 680-685.
- 14. Okayama, H. & Berg, P. (1982) Mol. Cell. Biol. 2, 161-170.
- 15. Okayama, H. & Berg, P. (1983) Mol. Cell. Biol. 3, 280-289.

- Grunstein, M. & Hogness, D. (1975) Proc. Natl. Acad. Sci. USA 72, 3961–3965.
- 17. Southern, E. M. (1975) J. Mol. Biol. 98, 503-517.
- Alvine, J. C., Kemp, D. J. & Stark, G. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5350–5354.
- Aviv, H. & Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408-1412.
- 20. Sanger, F. (1981) Science 214, 1205-1210.

1

- Grantham, R., Gautier, C., Gouy, M., Jacobzome, M. & Mercier, R. (1981) Nucleic Acids Res. 9, r43-r74.
- 22. Birnstiel, M. L., Busslinger, M. & Strub, K. (1985) Cell 41, 349-359.
- 23. Kyte, J. & Doolittle, R. F. (1982) J. Mol. Biol. 157, 105-132.
- 24. Chou, P. Y. & Fasman, G. C. (1978) Adv. Enzymol. 47, 45-148.
- 25. Nevins, J. R. & Darnell, J. E., Jr. (1978) Cell 15, 1477-1493.
- Narimatsu, H., Sinha, S., Brew, K., Okayama, H. & Qasba, P. (1986) Proc. Natl. Acad. Sci. USA 83, 4720-4724.
- Reddy, M. N., Maraganore, J. M., Meredith, S. C., Heinrikson, R. L. & Kezdy, F. J. (1986) J. Biol. Chem. 261, 9678–9683.