Geacintov, N. E., Ibanez, V., Gagliano, A. G., Jacobs, S. A., & Harvey, R. G. (1984b) *J. Biomol. Struct. Dyn. 1*, 1473-1484.
Gelboin, H. V. (1980) *Physiol. Rev. 60*, 1107-1166.
Harvey, R. G. (1981) *Acc. Chem. Res. 14*, 218-226.
Hogan, M. E., Dattagupta, N., & Whitlock, J. P. (1981) *J. Biol. Chem. 265*, 4504-4513.
Jeffrey, A. M., Weinstein, I. B., Jennette, K. W., Grezeskowiak, K., Harvey, R. G., Autrup, H., & Harris, C. (1977) *Nature (London) 269*, 348-350.
Jernstrom, B., Lycksell, P. O., Graslund, A., & Norden, B. (1984) *Carcinogenesis (London) 5*, 1129-1135.
Kakefuda, T., & Yamamoto, H. A. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 415-419.
Koreeda, M., Moore, P. D., Wislocki, P. G., Levine, W., Cooney, A. H., Yagi, H., & Jerina, D. M. (1978) *Science (Washington, D.C.) 199*, 778-781.
King, H. W. S., Osborne, M. R. M., Beland, F. A., Harvey, R. G., & Brookes, P. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 2679-2681.
MacLeod, M. C., & Selkirk, J. K. (1982) *Carcinogenesis (London) 3*, 287-292.
MacLeod, M. C., & Tang, M.-S. (1985) *Cancer Res. 45*, 51-56.
Meehan, T., & Straub, K. (1979) *Nature (London) 277*, 410-442.
Meehan, T., Straub, K., & Calvin, M. (1977) *Nature (London) 269*, 725-727.
Meehan, T., Gamper, H., & Becker, J. F. (1982) *J. Biol. Chem. 257*, 10479-10485.
Osborne, M. R., Beland, F. A., Harvey, R. G., & Brookes, P. (1976) *Int. J. Cancer 18*, 362-368.
Shahbaz, M., Geacintov, N. E., & Harvey, R. G. (1986) *Biochemistry 25*, 3290-3296.
Singer, D., & Grunberger, D. (1983) *Molecular Biology of Mutagens and Carcinogens*, Plenum Press, New York.
Taylor, E. R., Miller, K. J., & Bleyer, A. J. (1983) *J. Biomol. Struct. Dyn. 1*, 883-904.
Undeman, O., Lycksell, P. O., Graslund, A., Astlind, T., Ehrenberg, A., Jernstrom, B., Tjerneld, F., & Norden, B. (1983) *Cancer Res. 43*, 1851-1860.
Waring, M. (1970) *J. Mol. Biol. 54*, 247-279.
Weinstein, I. B., Jeffrey, A. M., Jennette, K. W., Blobstein, S. H., Harvey, R. G., Harris, C., Autrup, H., Kasai, H., & Nakanishi, K. (1976) *Science (Washington, D.C.) 193*, 592-595.
Yoshida, H. (1984), Ph.D. Thesis, New York University, New York, NY.

Molecular Cloning and Sequence Analysis of Complementary DNA Encoding Rat Mammary Gland Medium-Chain *S*-Acyl Fatty Acid Synthetase Thio Ester Hydrolase

Richard Safford,* Jacquie de Silva, Clare Lucas, John H. C. Windust, James Shedden, Christopher M. James, Christopher M. Sidebottom, Antoni R. Slabas, Michael P. Tombs, and Stephen G. Hughes

Biosciences Division, Unilever Research, Bedford MK44 1LQ, U.K.

Received July 29, 1986; Revised Manuscript Received September 26, 1986

ABSTRACT: Poly(A)+ RNA from pregnant rat mammary glands was size-fractionated by sucrose gradient centrifugation, and fractions enriched in medium-chain *S*-acyl fatty acid synthetase thio ester hydrolase (MCH) were identified by in vitro translation and immunoprecipitation. A cDNA library was constructed, in pBR322, from enriched poly(A)+ RNA and screened with two oligonucleotide probes deduced from rat MCH amino acid sequence data. Cross-hybridizing clones were isolated and found to contain cDNA inserts ranging from ~1100 to 1550 base pairs (bp). A 1550-bp cDNA insert, from clone 43H09, was confirmed to encode MCH by hybrid-select translation/immunoprecipitation studies and by comparison of the amino acid sequence deduced from the DNA sequence of the clone to the amino acid sequence of the MCH peptides. Northern blot analysis revealed the size of the MCH mRNA to be 1500 nucleotides, and it is therefore concluded that the 1550-bp insert (including G·C tails) of clone 43H09 represents a full- or near-full-length copy of the MCH gene. The rat MCH sequence is the first reported sequence of a thioesterase from a mammalian source, but comparison of the deduced amino acid sequences of MCH and the recently published mallard duck medium-chain *S*-acyl fatty acid synthetase thioesterase reveals significant homology. In particular, a seven amino acid sequence containing the proposed active serine of the duck thioesterase is found to be perfectly conserved in rat MCH.

In animals, de novo synthesis of fatty acids is catalyzed by a multienzyme complex, fatty acid synthetase, which elongates acetyl coenzyme A (acetyl-CoA) by successive additions of two-carbon units derived from malonyl-CoA. The acyl chain is covalently bound, via a thio ester linkage, to the 4′-phosphopantetheine moiety of the fatty acid synthetase (Phillips et al., 1970). Chain termination and release of the free fatty acid product is achieved by hydrolysis of the thio ester by a component of the synthetase complex, thioesterase I. The chain length of the released fatty acid is usually C16. However, in some specialized tissues, such as the mammary glands of nonruminant mammals (Libertini & Smith, 1978; Knudsen

* Author to whom correspondence should be addressed.

0006-2960/87/0426-1358$01.50/0

et al., 1976) and the uropygial glands of certain waterfowl (De Renobales et al., 1980; Rogers et al., 1982), the synthetase produces predominantly medium-chain length (C8, C10, C12) fatty acids. This results from an interaction of the synthetase with a second thioesterase (thioesterase II or medium-chain *S*-acyl fatty acid synthetase thio ester hydrolase, MCH) that is not part of the fatty acid synthetase but is present as a separate polypeptide. In rats, MCH is exclusively localized in mammary epithelia (Nolin et al., 1982), thereby conferring on this cell type the unique capacity for synthesis of medium-chain fatty acids. In response to hormonal influences, the MCH content of mammary gland tissue increases dramatically during late pregnancy, peaking just prior to parturition and remaining high during lactation (Smith & Ryan, 1979). MCH has been shown to be expressed in virgin mammary epithelia (Smith et al., 1983), and although some increase of MCH content occurs, on a per cell basis, during glandular development it has been proposed that the high levels of MCH in lactating tissue arise primarily from hormone-induced proliferation of epithelial cells during mammogenic development (Nolin et al., 1982).

Further insight into the developmental regulation of MCH expression in mammary glands will be gained by studies at the gene level. As a first step toward this, we report here the molecular cloning of cDNA encoding rat mammary gland MCH. Recently, the cDNA sequence of the medium-chain thioesterase from the uropygial gland of the mallard duck has been reported (Poulose et al., 1985), and comparison of the deduced amino acid sequence of this thioesterase with that of the rat mammary gland thioesterase reveals significant homology. In particular, a region surrounding the proposed active serine of the mallard thioesterase is found to be perfectly conserved in the rat enzyme.

Materials and Methods

MCH Purification. MCH was isolated from mammary glands of 18-day lactating rats (Colworth-Wistar strain) as previously described (Slabas et al., 1983) and further purified by DEAE-5PW HPLC[1] chromatography.

Preparation of Affinity-Purified MCH Antibodies. MCH denatured with 9.5 M urea and cross-linked with 2% glutaraldehyde was used to immunize rabbits in Freund's complete adjuvant. Second-bleed antiserum was affinity purified by using native MCH immobilized on CNBr-activated Sepharose 4B (Jockusch et al., 1978).

MCH Amino Acid Sequencing. After reductive alkylation with [^{14}C]iodoacetamide, MCH was cleaved with either trypsin or CNBr. The resulting peptides were separated by reverse-phase HPLC, and some were sequenced by automated Edman degradation (see Figure 6). Two of these peptides were used for provision of oligonucleotide probes (Figure 3).

Polyacrylamide Gel Electrophoresis and Western Blotting. Proteins were separated on 10–20% polyacrylamide gradient SDS gels (Laemmli, 1970). For Western blots, fractionated proteins were electrophoretically transferred to nitrocellulose (Towbin et al., 1979), and after blocking with 3% hemoglobin, filters were incubated with rabbit a-MCH antibody followed by ^{125}I-protein A.

RNA Isolation. Mammary glands, removed from 20-day pregnant Colworth-Wistar rats, were ground in liquid nitrogen and lyophilized. Lyophilized tissue was homogenized in guanadinium thiocyanate solution (Chirgwin et al., 1979) for 60 s with a Polytron homogenizer at full speed. RNA was isolated by centrifugation through 5.7 M CsCl solution (Glisin et al., 1974), dissolved in guanadinium chloride solution by briefly heating at 65 °C, and ethanol precipitated. After water extraction, RNA was selectively precipitated in 2 M LiCl. Polyadenylated [poly(A)+] RNA was isolated from total RNA on poly(U)–Sephadex (BRL) according to the manufacturer's protocol.

Sucrose Gradient Fractionation of Pregnant Rat Mammary Gland (PRMG) Poly(A)+ RNA. 30 μg of RNA was heated at 70 °C for 1 min in 1 mM EDTA, cooled on ice, layered onto a 15–40% sucrose gradient (6 mL) in 0.1 M NaCl, 10 mM Tris-HCl, pH 7.5, and 10 mM EDTA, and centrifuged at 200000*g* for 14 h at 12 °C. Fractions (0.4 mL) were collected, ethanol precipitated, and assayed for MCH mRNA content by in vitro translation and immunoprecipitation with rabbit a-MCH antibody as described below.

Immunoprecipitation of PRMG in Vitro Translation Products with Rabbit a-MCH Antibody. PRMG poly(A)+ RNA (0.5 μg) was translated in a rabbit reticulocyte lysate system (Amersham) with L-[^{35}S]methionine (1 μCi/μL) for 90 min at 30 °C. Translation products were diluted 10-fold in water, made 2% in SDS, and heated for 3 min at 100 °C. Samples were diluted to 1 mL with immunoprecipitation buffer (1% Triton X-100, 0.3 M NaCl, 50 mM Tris-HCl, pH 8.0) and incubated with rabbit a-MCH antibody for 2 h at 20 °C. Immune complexes were precipitated with protein A–Sepharose beads (Pharmacia). After thorough washing with immunoprecipitation buffer, immune complexes were released from the beads by heating at 100 °C for 3 min in 1.5X Laemmli sample buffer (Laemmli, 1970). Products were run on 10–20% gradient SDS–PAGE and fluorographed (Chamberlain, 1979).

cDNA Synthesis. This method is based on a modification (J.-L. Darlix, unpublished results) of the Gubler and Hoffman (1983) RNase H protocol. MCH-enriched mRNA (3 μg) was annealed with 2.5 μg of oligo(dT$_{12-18}$) in 20 μL of water by heating for 30 s at 100 °C and quenching in ice–water. For first-strand cDNA synthesis, 20 μL of annealed mix was added to 30 μL of 2X reverse transcriptase buffer [0.1 M Tris-HCl, pH 8.3 (adjusted at 43 °C), 10 mM $MgCl_2$, 150 mM KCl, 20 mM dithiothreitol, 1 mM ATP, 1 mM GTP, 1 mM CTP, and 0.2 mM TTP], 3.3 μL of water, and 50 units of Super RT reverse transcriptase (Anglian Biotech), and incubated for 45 min at 43 °C. For second-strand synthesis, 150 μL of buffer (40 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 75 mM KCl, 25 μg/mL bovine serum albumin), 25 μCi of [α-^{32}P]TTP (3000 Ci/mmol specific activity from Amersham), 52 units of DNA polymerase I (Amersham), and 8 units of RNase H (P-L Biochemicals) were added to the first-strand mix and incubated at 11 °C for 60 min. Gap repair was achieved by incubating the second-strand mix with 2 μL of 20 mM ATP and 420 units of DNA ligase (Amersham) at 18 °C for 2 h. The reaction was stopped by addition of 20 μL of 0.5 M EDTA and 2 μL of 10% SDS, and the cDNA was purified on Sephadex G-50.

Insertion of cDNA into pBR322 and Transformation of Escherichia coli. Double-stranded cDNA was dC-tailed by using terminal transferase (Amersham) and, after phenol extraction, annealed to an equimolar amount of dG-tailed *Pst*I-cleaved pBR322 (BRL) in 10 mM Tris-HCl, pH 7.5, and 0.1 M NaCl by heating at 65 °C for 2 h and then cooling to 40 °C over 2 h. The annealed mix was used to transform *E. coli* HB101, essentially as described by Mandel and Higa

[1] Abbreviations: SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; CNBr, cyanogen bromide; HPLC, high-performance liquid chromatography; CsCl, cesium chloride; bp, base pair(s); EDTA, ethylenediaminetetraacetic acid; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride.

(1970). A total of 6000 transformants were selected by tetracycline resistance and stored in microtiter dishes at −70 °C in 30% glycerol.

Oligonucleotide Synthesis. Mixed oligonucleotides were synthesized by the phosphoamidate method (Adams et al., 1983) at Unilever Research, Vlaardingen, Holland. Following purification on 20% polyacrylamide gels, oligonucleotides were 5′ end labeled by using [γ-^{32}P]ATP and T_4 polynucleotide kinase.

Oligonucleotide Screening of cDNA Library. Colony blots of chloramphenicol-amplified recombinant clones were prepared on Biodyne A membrane (Pall), according to the manufacturer's protocol. Blots were prewashed in 50 mM Tris-HCl, pH 8.0, 1 M NaCl, 1 mM EDTA, and 0.1% SDS for 2 h at 42 °C and prehybridized in 5X standard saline citrate (SSC), 5X Denhardt's solution, 0.2% SDS, and 250 μg/mL denatured salmon sperm DNA at 65 °C for 4–16 h. Hybridizations were carried out with $\sim 5 \times 10^5$ cpm/mL ^{32}P-labeled oligonucleotides for 16 h at $\sim$5 °C below T_d range for the mixed probes [T_d = 4(G·C bp) + 2(A·T bp) °C according to Suggs et al. (1981)]. Washing stringency was 5 X SSC and 1% SDS at the hybridization temperature (2 × 30 min).

Southern Blot Analysis. Plasmid DNA was isolated by the alkaline lysis method (Birnboim & Doly, 1979) and further purified by CsCl–ethidium bromide gradient centrifugation. cDNA inserts excised by restriction with *Pst*I were electrophoresed in 1% agarose gels and blotted onto nitrocellulose (Southern, 1975).

Northern Blot Analysis. Mammary gland poly(A)+ RNA was glyoxalated (McMaster & Carmichael, 1977), excluding dimethyl sulfoxide, electrophoresed in 1% agarose gel in 10 mM sodium phosphate, pH 6.5, and transferred to Genescreen membrane (NEN) according to the manufacturer's protocol. Hybridization with nick-translated clone 43H09 cDNA insert was under conditions similar to those described for Southern blot analysis.

Hybrid Selection of MCH mRNA. Clone 43H09 cDNA insert was excised by *Pst*I digestion and purified by gel electrophoresis and electroelution onto DEAE membrane (Dretzen et al., 1978). The cDNA, after denaturation at 100 °C for 5 min in water, was applied onto sterile nitrocellulose squares and baked for 2 h at 80 °C under vacuum. Hybrid selection was carried out, using 25 μg of PRMG poly(A)+ RNA, as previously described (Parnes et al., 1981). Selected mRNAs were translated in vitro and the products immunoprecipitated with rabbit a-MCH antibody and analyzed by 10–20% gradient SDS-PAGE/fluorography as described above.

DNA Sequencing. Clone 43H09 cDNA insert was subcloned into M13 and sequenced by the dideoxy method (Sanger et al., 1977).

Results

Our strategy for cloning rat MCH cDNA was to construct a cDNA library from size-fractionated mammary gland mRNA, screen it with oligonucleotide probes derived from rat MCH amino acid sequence data, and confirm putative MCH clones by hybrid-select translation and immunoprecipitation.

MCH Purification and Characterization of MCH Antibodies. MCH was isolated from lactating rat mammary glands as previously described (Slabas et al., 1983) and purified to homogeneity by DEAE-5PW HPLC chromatography. This preparation was used to raise antibodies and to generate amino acid sequence data suitable for the provision of oligonucleotide probes. Antibodies raised in rabbits were affinity purified against immobilized native MCH (Jockusch et al., 1978) and characterized by immunoblotting (Towbin et al., 1978) (Figure 1). Under the denaturing conditions employed in this technique the antibodies both specifically recognized (Figure 1, lane 3) and quantitatively immunoprecipitated (Figure 1, lane 2) the 29-kDa MCH polypeptide from mammary gland cytosol preparations.

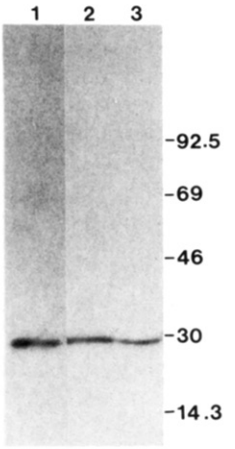

FIGURE 1: Characterization of rabbit a-MCH antibody by Western blot analysis. Lactating rat mammary gland (LRMG) cytosol was prepared as previously described (Slabas et al., 1983) and reacted with MCH antibody. Both the immunoprecipitate and untreated LRMG cytosol were run on 10–20% gradient SDS–PAGE, transferred to nitrocellulose, and reacted with MCH antibody followed by ^{125}I-protein A: lane 1, purified MCH; lane 2, MCH immunoprecipitate from 50 μL of LRMG cytosol; lane 3, 50 μL of LRMG cytosol. ^{14}C-Labeled molecular weight markers, in K, are shown on the right.

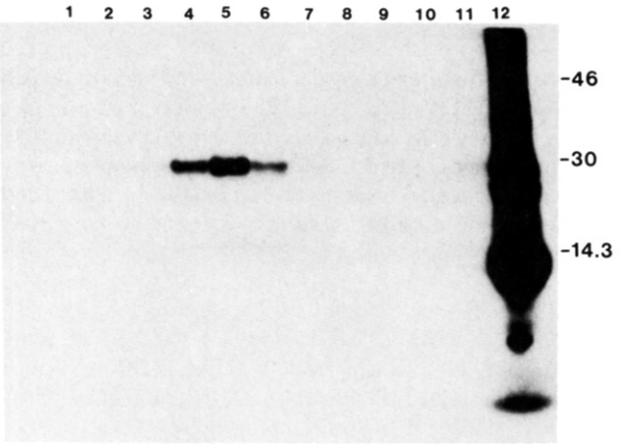

FIGURE 2: Fluorograph of PRMG mRNA sucrose gradient fraction in vitro translation products immunoprecipitated with rabbit a-MCH antibody: lanes 1–10, PRMG mRNA sucrose gradient fraction translation products immunoprecipitated with MCH antibody; lane 11, MCH immunoprecipitate of unfractionated PRMG mRNA translation products; lane 12, unfractionated PRMG mRNA translation products. ^{14}C-Labeled molecular weight markers, in K, are shown on the right.

MCH Amino Acid Sequencing. MCH peptides generated by cleavage with trypsin, following reductive alkylation with [^{14}C]iodoacetamide, and by cyanogen bromide treatment were separated by reverse-phase HPLC chromatography and sequenced (see Figure 6). Two of these peptides, namely, FIFDKP and MEPLH, were selected for provision of oligonucleotide probes (see Figure 3).

Enrichment of MCH mRNA and cDNA Cloning. mRNA isolated from pregnant mammary glands was translated in vitro and immunoprecipitated with MCH antibodies. Analysis of the products by SDS gel electrophoresis/fluorography (Figure 2) showed specific immunoprecipitation of an MCH polypeptide with the same molecular weight (29K) as that

```
                                                                    Arg
MCH peptide          Phe Ile Phe Asp Lys Pro      Met Glu Pro Leu His

Complementary          A   A   A   A   T                T     A   TC
probe          3' AAG TAG AAG CTG TTC GG 5'  3' TAC CTC GGN GAN GT 5'
                         T

                        17mer                        14mer
```

FIGURE 3: Synthetic oligonucleotide probes for MCH cDNA screening. Probes were of the mixed-synthesis type, complementary to the coding sequence. The 17-mer was synthesized as two mixes, one using TTT and the other TTC for the Lys codon. As assignment of the terminal amino acid of the peptide used for the 14-mer was equivocal (Arg or His), a mix encoding both possibilities was used. The 14-mer was made in two syntheses, one containing GAN and the other AA^{T}_{C} for the Leu codon. N = A, C, G, or T.

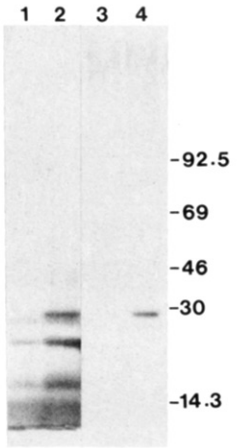

FIGURE 4: In vitro translation of mRNA hybrid selected with clone 43H09. Purified clone 43H09 cDNA insert and control pBR322 DNA were hybridized to PRMG mRNA, selected mRNAs translated in vitro and immunoprecipitated with MCH antibody, and the products analyzed by SDS–PAGE/fluorography: lane 1, pBR322-selected mRNA translation products; lane 2, clone 43H09 selected mRNA translation products; lanes 3 and 4, as for lanes 1 and 2 after immunoprecipitation. ^{14}C-Labeled molecular weight markers, in K, are shown on the right.

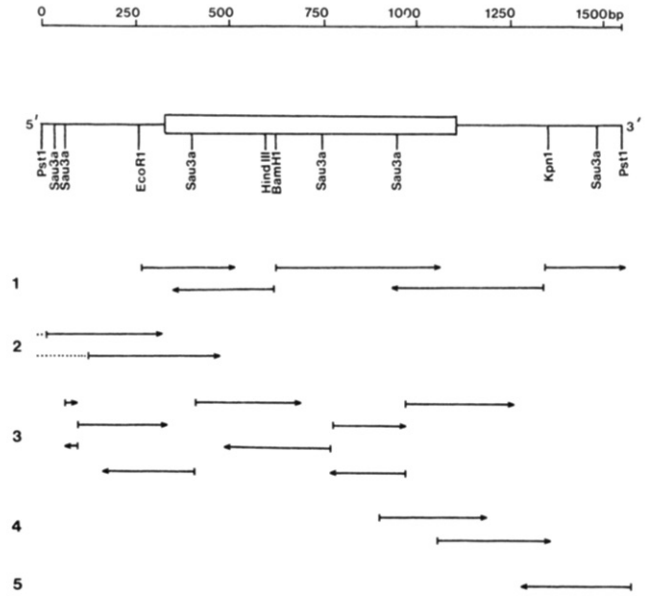

FIGURE 5: Restriction map and sequencing strategy of clone 43H09. Arrows represent the direction and extent of dideoxy sequences derived in M13: 1, *Bam*HI–*Kpn*I/*Bam*HI–*Eco*RI fragments; 2, *Bal*31 exonuclease deleted *Pst*I–*Bam*HI fragments; 3, *Sau*3A fragments; 4, *Bam*HI–*Kpn*I fragments primed with MCH-specific oligonucleotides; 5, intact cDNA insert. The boxed area represents the MCH coding region.

observed for the mature protein (Figure 1), demonstrating that MCH is not made in a precursor form. Although the MCH antibody preparation was monospecific, contamination of the immunoprecipitate with caseins could occur, as a result of binding of these hydrophobic, micellar forming proteins to Protein A–Sepharose. This problem was overcome by incorporating a heating step in the presence of 2% SDS prior to immunoprecipitation (Suard et al., 1982). While MCH protein was routinely purified from 18-day lactating tissue, initial in vitro translation studies showed this tissue to be a poor source of MCH mRNA. A considerably richer source was found to be pregnant, preparturition glands, and these were subsequently used for RNA isolation. Even so, the abundance of MCH mRNA from this source was still low (~0.05%), and therefore, prior to cloning, PRMG poly(A)+ RNA was size-fractionated by sucrose gradient centrifugation. Fractions were assayed for MCH mRNA content by in vitro translation and immunoprecipitation with MCH antibodies (Figure 2). Fraction 5 is seen to be considerably enriched for MCH mRNA and was subsequently used to construct a cDNA library in pBR322 by using a modification (J.-L. Darlix, unpublished results) of the RNase H protocol of Gubler and Hoffman (1983).

Screening for MCH Clones. The cDNA library was screened for MCH clones by hybridization to synthetic oligonucleotides derived from MCH amino acid sequence data. To reduce the number of false positive clones, screening was carried out with two sets of oligonucleotides derived from two different peptides, MCH clones being identified by hybridization to both probes. The two oligonucleotide probes were of the mixed synthesis type, a 17-mer and a 14-mer (Figure 3).

Colony blot screening of the 6000 recombinant clones with the ^{32}P-labeled 17-mer mix resulted in identification of 18 positive clones. Rescreening the library with the 14-mer mix gave 16 positive clones, 6 of which had previously hybridized to the 17-mer probe. Plasmid DNA was isolated from the six cross-hybridizing clones, and gel electrophoresis of the cDNA inserts excised by *Pst*I digestion showed a range from 1100–1550 bp. Restriction analysis of the six inserts showed them to be structurally related. This was further confirmed by Southern blotting, when all six inserts gave positive signals when hybridized, under stringent conditions, to nick-translated insert from the longest clone (43H09). As a result of this, clone 43H09 was subjected to further characterization by hybrid-select translation and M13 dideoxy sequencing.

Hybrid-Select Translation of Clone 43H09. When hybridized to PRMG poly(A)+ RNA, immobilized clone 43H09 cDNA insert was found to select a mRNA coding for a 29-kDa polypeptide (Figure 4, lane 2) which was immunoprecipitable with MCH antibodies (Figure 4, lane 4). The 29-kDa polypeptide was not detected in control experiments with immobilized pBR322 DNA (Figure 4, lanes 1 and 3).

```
GGGGGGGTTGCTGTGCCTCTCACTGTGGTTCGATCTCTCCTCTGGAAGGAGCATAGCTGAAAGCTGATCTTCTGGTCTCCAGAACGCTTCTGAGATTTTGCCATCGAGTCCATTTCCTTC
        10        20        30        40        50        60        70        80        90       100       110       120

TCTTCTTCCCTATCTCTTTTTCTTCACTGCCCTTTAAAAGCAGGATTCTGAGTGACAGCCATAGTCAACAGGTTAAAAAGAAAGCTGGTCCCTTGGTATACTGAAGGCTTACTGGAGATT
       130       140       150       160       170       180       190       200       210       220       230       240

                                                                                  M  E  T  A  V  N  A  K  S  P  R  N  E  K  V
TTTGTGCCCAGAGGGAATTGAATTCAGCGTGGACACAGAAGGTTGAAGGAGCAACAACCCCTAATCAACTCACAGAATGGAGACAGCAGTCAATGCTAAGAGTCCCAGGAATGAAAAGGT
       250       260       270       280       290       300       310       320       330       340       350       360

  L  N  C  L  Y  Q  N  P  D  A  V  F  K  L  I  C  F  P  W  A  G  G  G  S  I  H  F  A  K  W  G  Q  K  I  N  D  S  L  E  V
TTTGAACTGTTTGTATCAAAATCCTGATGCAGTTTTCAAGCTGATCTGCTTCCCTTGGGCAGGAGGCGGCTCCATCCATTTTGCCAAGTGGGGCCAAAAGATTAACGACTCTCTGGAAGT
       370       380       390       400       410       420       430       440       450       460       470       480

  H  A  V  R  L  A  G  R  E  T  R  L  G  E  P  F  A  N  D  I  Y  Q  I  A  D  E  I  V  T  A  L  L  P  I  I  Q  D  K  A  F
GCATGCTGTAAGACTGGCTGGAAGAGAAACCCGACTTGGAGAACCTTTCGCAAATGACATCTACCAGATAGCTGATGAAATCGTGACCGCCCTGTTGCCCATCATTCAGGATAAAGCTTT
       490       500       510       520       530       540       550       560       570       580       590       600

  A  F  F  G  H  S  F  G  S  Y  I  A  L  I  T  A  L  L  L  K  E  K  Y  K  M  E  P  L  H  I  F  V  S  G  A  S  A  P  H  S
TGCGTTTTTTGGCCACAGTTTTGGATCCTACATTGCTCTTATTACTGCTCTGCTCCTAAAGGAGAAATACAAAATGGAGCCGCTGCATATTTTTGTATCCGGTGCATCCGCCCCTCACTC
       610       620       630       640       650       660       670       680       690       700       710       720

  T  S  R  P  Q  V  P  D  L  N  E  L  T  E  E  Q  V  R  H  H  L  L  D  F  G  G  T  P  K  H  L  I  E  D  Q  D  V  L  R  M
AACATCCCGGCCTCAAGTTCCTGATCTTAACGAATTGACAGAAGAACAAGTCAGACATCACCTTCTGGATTTCGGAGGCACGCCCAAGCATCTCATAGAAGACCAGGATGTTCTGAGGAT
       730       740       750       760       770       780       790       800       810       820       830       840

  F  I  P  L  L  K  A  D  A  G  V  V  K  K  F  I  F  D  K  P  S  K  A  L  L  S  L  D  I  T  G  F  L  G  S  E  D  T  I  K
GTTCATTCCTTTGCTGAAGGCAGATGCTGGCGTTGTGAAAAAATTCATCTTTGACAAGCCCTCCAAAGCTCTTCTCTCTCTGGACATAACGGGCTTCCTTGGATCTGAAGATACAATAAA
       850       860       870       880       890       900       910       920       930       940       950       960

  D  I  E  G  W  Q  D  L  T  S  G  K  F  D  V  H  M  L  P  G  D  H  F  Y  L  M  K  P  D  N  E  N  F  I  K  N  Y  I  A  K
GGACATAGAAGGCTGGCAAGACCTAACCAGTGGGAAGTTTGATGTCCACATGCTGCCAGGCGACCACTTTTATCTGATGAAGCCCGACAACGAGAACTTTATCAAGAACTACATAGCCAA
       970       980       990      1000      1010      1020      1030      1040      1050      1060      1070      1080

  C  L  E  L  S  S  L  T  *
GTGCTTGGAACTCTCGTCACTCACTTGACTACTTTTAGATGAGCTTTCTTTGGGGCTGTGGATATGCAGACGGTTCAAAAGCTGCTCCTCTGGGTCCAGCTTTTAGTTCTTAGAGGTTGG
      1090      1100      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200

AAAGGTGCATATTAATGAATTTCTTCGTGGAGCATTTATCTCTCTGGGGAAAGGACACGGCTGAAGAGGCGCACAGGCACTATTGGGAGCTAACAGAGGGGAGAGGGGGGTGGTTTTCTT
      1210      1220      1230      1240      1250      1260      1270      1280      1290      1300      1310      1320

TGGGGGGATGTTGCCCTCAGTAGGTACCCATGCTCCAGTAAATGACACTACACCCTGTGGCACATGCAGGCAGCTCTAATGCAGACTCAGTGGGAAGTTGGGAGGCATACATGTTGGGGG
      1330      1340      1350      1360      1370      1380      1390      1400      1410      1420      1430      1440

ACATCTGTCACGGGGCTTGGAAAGACAAGTGAGGGTCTGATCAAAACCTGCTGTTTACATTTTCAAAGAATAAATAAAAAAATTATGCTAAAAAAACCCCCCCCCCCCCCC
      1450      1460      1470      1480      1490      1500      1510      1520      1530      1540
```

FIGURE 6: Nucleotide sequence and deduced amino acid sequence of rat MCH determined from clone 43H09 cDNA. The boxed regions indicate those residues identified by amino acid sequencing of MCH peptides. The consensus polyadenylation sequence AATAAA, located 14 bases upstream from the poly(A) tail, is also boxed.

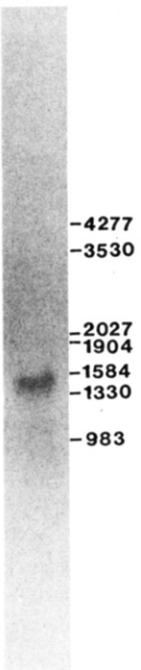

FIGURE 7: Northern blot sizing of MCH mRNA. Glyoxalated PRMG mRNA was electrophoresed on 1% agarose, transferred to Genescreen membrane, and hybridized to nick-translated clone 43H09 cDNA insert in 5 × SSC at 65 °C. Glyoxalated *Hin*dIII/*Eco*RI λ DNA fragments were used as size markers.

DNA Sequencing of Clone 43H09 cDNA Insert. The 1550-bp cDNA insert of clone 43H09 was subcloned into M13 vectors according to the strategy shown in Figure 5 and sequenced by using the dideoxy chain-termination method (Sanger, 1977). The complete nucleotide sequence of the cDNA and the derived amino acid sequence are shown in Figure 6. Confirmation that the cDNA codes for MCH was provided when matches (boxed sequences in Figure 6) were found in the derived amino acid sequence for the MCH peptides FIFDKPSK, WGQK, NYIAK, DIEG, MEPLH, INDSL, and VLN(?)LYQN. In addition, an 11 amino acid mixed peptide was resolved by DNA sequencing, to give the sequences DIEGWQDLTSG and LMKPDNENFIK. All these sequences were found in a single open reading frame which was defined by a translational stop codon (TGA) at nucleotide 1106, together with an ATG at nucleotide 317, with no alternative upstream translational start codons. This gives rise to a translation product of 263 amino acids with a molecular mass of 29 300 Da. This is in agreement with the apparent M_r of 29 000 determined for MCH from SDS–PAGE and confirms also that MCH is not made in a precursor form. Correct assignment of the coding sequence was further confirmed by the close match of the derived amino acid composition with a previously published MCH composition (Libertini & Smith, 1978) (data not shown). Assignment of the coding region determined that clone 43H09 had a 310-nucleotide 5′-untranslated region and a 3′-untranslated sequence of 423 nucleotides, in which the transcription termination signal AATAAA (Proudfoot & Brownlee, 1976) was located 14 bases upstream of the poly(A) tail.

Northern Blot Analysis. To determine the size of MCH mRNA, a Northern blot of PRMG poly(A)+ RNA was

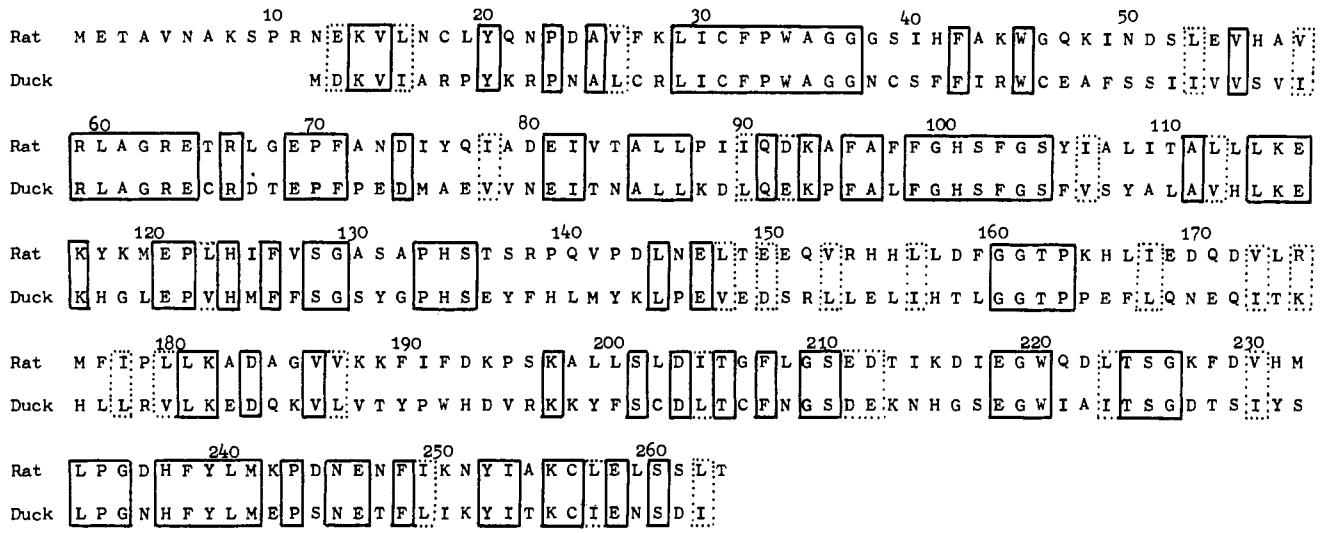

FIGURE 8: Comparison of the amino acid sequences of rat MCH and duck medium-chain thioesterase. The two sequences have been aligned via the proposed active site region, FGHSFGS (residues 98–104). Homologous residues are enclosed in boxes; conservative substitutions are enclosed by broken lines.

probed with nick-translated clone 43H09 cDNA insert. Figure 7 shows specific hybridization to a mRNA of 1500 nucleotides, demonstrating that, within the accuracy of this technique, the 1550-bp cDNA insert of clone 43H09 represents a full- or near-full-length copy of the MCH gene.

Discussion

MCH cDNA clones were identified from a pregnant rat mammary gland library by screening with oligonucleotide probes derived from MCH amino acid sequence data. By screening with two oligonucleotide probes based on two distinct peptides, irrelevant clones, which fortuitously hybridize to either one of the oligonucleotides, were eliminated. Thus, the six putative MCH clones, identified initially by hybridization to both the 17-mer and 14-mer oligonucleotide probes, were all subsequently found, by restriction analysis and Southern blotting, to be derived from a common mRNA. Confirmation of identity was obtained, for the longest cDNA clone (43H09), by hybrid-select translation and matching of translated nucleotide sequences with amino acid sequence data obtained from MCH peptides. Northern blot analysis revealed the size of the MCH mRNA to be ~1500 nucleotides, and therefore the 1550-bp cDNA of clone 43H09 represents a full- or near-full-length clone of MCH. The 263 amino acid open reading frame of clone 43H09 gives rise to a polypeptide of 29 300 Da. This is in agreement with the apparent M_r of 29 000 determined for the protein on SDS–PAGE and is also consistent with the finding, from in vitro translation studies, that MCH is not synthesized in a precursor form. The coding region is flanked by a 310-nucleotide 5′-untranslated sequence and a 423-nucleotide 3′-untranslated region, the latter containing the transcription termination signal AATAAA (Proudfoot & Brownlee, 1976) 14 bases upstream from the poly(A)+ tail.

The first AUG triplet in MCH mRNA is the initiator codon, in agreement with the scanning model proposed for initiation of translation in eucaryotes (Kozak, 1984). In addition, the highly conserved purine A is observed at position −3 from the initiator AUG, although the −5 to +1 sequence of ACAGA(AUG)G has only three nucleotides in common with the proposed consensus sequence $CC^{A}_{G}CC(AUG)G$ (Kozak, 1984).

Recently, the cDNA sequence encoding the medium-chain *S*-acyl fatty acid synthetase thioesterase from the uropygial gland of the mallard duck has been reported (Poulose et al., 1985). The size of the duck thioesterase mRNA was estimated at 1350 nucleotides, and an ~1150-bp cDNA, lacking some of the 5′-untranslated region but containing all the coding region, was cloned and sequenced.

Comparison of the deduced amino acid sequences of the rat and duck medium-chain thioesterases reveals significant homology (Figure 8). In particular, the sequence surrounding the proposed active serine of the duck thioesterase, namely, FGHSFGS, is found to be perfectly conserved in the rat thioesterase (residues 98–104). Alignment of the two sequences via the proposed FGHSFGS active site results in the rat enzyme having an additional 11 amino acids at the N-terminus and 1 extra C-terminal residue. The overall level of homology is 40%, which is increased to 51% if allowance for conservative substitutions (E/D, R/K, I/L/V) is made. Other notable regions of homology include the perfectly conserved LICFPWAGG sequence (residues 29–37) and an eight out of nine match, LPG-HFYLM (residues 233–241). Overall, the two enzymes have broadly similar amino acid compositions, both being highly hydrophobic. One interesting difference, however, is in the cysteine content. Thus, the rat thioesterase contains three C residues, two of which are conserved in the duck enzyme, but the latter contains an additional five C residues, four of which are located in close proximity at the N-terminus of the molecule. It has been shown (Witowski & Smith, 1985) that modification of a single cysteine thiol of the rat thioesterase inhibits functional interaction with rat fatty acid synthetase. It has been inferred (Witowski & Smith, 1985), from affinity chromatographic studies, that this interaction is not as strong as the thioesterase–fatty acid synthetase interaction in the mallard duck, and one could speculate that the additional cysteines of the duck thioesterase are involved in promoting the apparently stronger binding in this system.

Comparison of the rat MCH active center sequence with the active serine sequences of other thioesterases and acyltransferases reveals a high degree of homology (Table I). Another intriguing comparison is with the active serine sequences of lipases, enzymes that are known to hydrolyze fatty acid thio esters. Table I shows that considerable homology exists between rat MCH and rat (Docherty et al., 1985) and porcine (Guidoni et al., 1981) lipase active site sequences.

When the complete amino acid sequence of the rat fatty acid

Table I: Comparison of the Active Center Sequence of Rat MCH with Those from Other Thioesterases, Acyltransferases, and Lipases[a]

	G	H	S	F	G	S
rat MCH (this paper)	G	H	S	F	G	S
duck medium-chain thioesterase (Poulose et al., 1985)	G	H	S	F	G	S
goat FAS malonyltransferase (Mikkelsen et al., 1985)			S	F	G	A
goat FAS malonyltransferase (Mikkelsen et al., 1985)	G	H	S	L	G	E
yeast FAS malonyltransferase (Engeser et al., 1979)	G	H	S	L	G	E
yeast FAS acetyltransferase [Scheweizer quoted in Mikkelsen et al. (1985)]	G	H	S	Q	G	L
rat lingual lipase (Docherty et al., 1985)	G	H	S	Q	G	T
porcine pancreatic lipase (Guidoni et al., 1981)	G	H	S	L	G	S

[a] The sequences are aligned via the active serine (arrow).

synthetase thioesterase domain becomes available, it will be of interest to compare the overall homology with that of the medium-chain thioesterase. The two thioesterase functions are of a very similar size, and one might predict that they arise from common ancestry and would thus show a reasonable level of homology. However, antibodies raised against both intact rat fatty acid synthetase and the proteolytically cleaved thioesterase fragment do not cross-react with rat medium-chain thioesterase (Libertini & Smith, 1978).

MCH is known to be specifically localized in the epithelia of rat mammary gland (Nolin et al., 1982). The availability of cDNA probes should enable investigations into the nature of the regulatory elements involved in this tissue-specific expression. It will also permit studies at the transcriptional level into the origin of the hormonally induced increases in MCH levels that are observed during mammary gland development. Thus, it should be possible to determine if an increase in MCH transcription per se occurs, as part of a differentiation event, or if the increased MCH levels found in the gland result primarily from a hormonally induced proliferation of the epithelial cell type, as has been proposed (Nolin et al., 1982).

ACKNOWLEDGMENTS

We thank our colleagues at Unilever Research, Vlaardingen, Holland, for synthesizing the oligonucleotide probes.

REFERENCES

Adams, S. P., Kauka, K. S., Wykes, E. J., Holder, S. B., & Gallupi, G. R. (1983) *J. Am. Chem. Soc. 105*, 661–663.
Birnboim, H. C., & Doly, J. (1979) *Nucleic Acids Res. 7*, 1513–1523.
Chamberlain, J. P. (1979) *Anal. Biochem. 98*, 132–135.
Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.
De Renobales, M., Rogers, L., & Kolattukudy, P. E. (1980) *Arch. Biochem. Biophys. 205*, 464–477.
Docherty, A. J. P., Bodmer, M. W., Angal, S., Verger, R., Riviere, C., Lowe, P. A., Lyons, A., Emtage, J. S., & Harris, T. J. R. (1985) *Nucleic Acids Res. 13*, 1891–1903.
Dretzen, G., Bellard, M., Sassone-Corsi, P., & Chambon, P. (1978) *Anal. Biochem. 112*, 295–298.
Engeser, H., Hübner, K., Straub, J., & Lynen, F. (1979) *Eur. J. Biochem. 101*, 413–422.
Glisin, V., Crkvenjakov, R., & Byus, C. (1974) *Biochemistry 13*, 2633–2637.
Gubler, U., & Hoffman, B. J. (1983) *Gene 25*, 263–269.
Guidoni, A., Benkouka, F., De Caro, J., & Rovery, M. (1981) *Biochim. Biophys. Acta 660*, 148–150.
Jockusch, B. M., Kelley, K. H., Meyer, R. K., & Burger, M. M. (1978) *Histochemistry 55*, 177–184.
Knudsen, J., Clark, S., & Dils, R. (1976) *Biochem. J. 160*, 683–691.
Kozak, M. (1984) *Nucleic Acids Res. 12*, 857–872.
Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.
Libertini, L. J., & Smith, S. (1978) *J. Biol. Chem. 253*, 1393–1401.
Mandel, M., & Higa, A. (1970) *J. Mol. Biol. 53*, 154.
McMaster, G. K., & Carmichael, G. G. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4835–4838.
Mikkelsen, J., Højrup, P., Rasmussen, M. M., Roepstorff, P., & Knudsen, J. (1985) *Biochem. J. 227*, 21–27.
Nolin, J. M., Thompson, B. J., & Smith, S. (1982) *J. Endocrinol. 94*, 251–256.
Parnes, J. R., Velan, B., Felsenfeld, A., Ramanathan, L., Ferrini, V., Appella, E., & Sigman, J. G. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2253–2257.
Phillips, G. T., Nixon, J. E., Dorsey, J. A., Butterworth, P. H. W., Chesterton, C. J., & Porter, J. W. (1970) *Arch. Biochem. Biophys. 138*, 380–391.
Poulose, A. J., Rogers, L., & Kolattukudy, P. E. (1981) *Biochem. Biophys. Res. Commun. 103*, 377–382.
Poulose, A. J., Rogers, L., Cheesebrough, T. M., & Kolattukudy, P. E. (1985) *J. Biol. Chem. 260*, 15943–15958.
Proudfoot, N. J., & Brownlee, G. G. (1976) *Nature (London) 263*, 211–214.
Rogers, L., Kolattukudy, P. E., & de Renobales, M. (1982) *J. Biol. Chem. 257*, 880–886.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Slabas, A. R., Ormesher, J., Roberts, P. A., Sidebottom, C. M., Tombs, M. P., Jeffcoat, R., & James, A. T. (1983) *Eur. J. Biochem. 134*, 27–32.
Smith, S., & Ryan, P. (1979) *J. Biol. Chem. 254*, 8932–8936.
Smith, S., Pasco, D., & Nandi, S. (1983) *Biochem. J. 212*, 155–159.
Southern, E. M. (1975) *J. Mol. Biol. 98*, 503–517.
Suard, Y. M. L., Tosi, M., & Kraehenbuhl, J.-P. (1982) *Biochem. J. 201*, 81–90.
Suggs, S. V., Hirose, T., Miyake, T., Kawashima, E. H., Johnson, M. J., Itakura, K., & Wallace, R. B. (1981) *ICN-UCLA Symp. Mol. Cell. Biol. 23*, 682–693.
Towbin, H., Staekelin, T., & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4350–4354.
Witowski, A., & Smith, S. (1985) *Arch. Biochem. Biophys. 243*, 420–426.