Volume 276, number 1,2, 131-134 FEBS 09199 December 1990

Sequence identity between human pancreatic cholesterol esterase and bile salt-stimulated milk lipase

David Y. Hui and James A. Kissel

Department of Pathology and Laboratory Medicine, University of Cincinnati, College of Medicine, Cincinnati, OH 45267-0529, USA

Received 6 August 1990; revised version received 8 October 1990

Three overlapping cDNA clones covering the entire primary sequence of the bile salt stimulated lipase in human milk were isolated from a human breast lambda gt10 cDNA library by screening with the rat pancreatic cholesterol esterase cDNA. Nucleotide sequencing of the cDNA showed that the human milk lipase mRNA encodes a 748-residue protein, including a 23-residue signal peptide. The human milk lipase cDNA is highly homologous to rat pancreatic cholesterol esterase, suggesting that the milk lipase may be identical to the cholesterol esterase in human pancreas. This conclusion was confirmed by isolation and sequencing of the cDNA for human pancreatic cholesterol esterase. Analysis of the sequence for the human cholesterol esterase/milk lipase revealed similarities to other serine esterases in three distinct regions of the protein. These domains may represent the active site triads of these proteins.

Cholesterol esterase; Lipase; cDNA cloning; Serine esterase

1. INTRODUCTION

Pancreatic cholesterol esterase catalyzes the hydrolysis of cholesteryl esters to free cholesterol and is important for catalyzing the absorption of dietary cholesterol and fat-soluble vitamins (reviewed in [1]). Recently, Cox et al. [2] have shown that cholesterol esterase expression in the pancreas is low during the juvenile period and increases 1300-fold during maturation. This observation suggests another protein may be involved in mediating cholesterol and vitamin absorption during the developmental period.

The human milk contains a lipase with properties almost identical to the cholesterol esterase of the pancreas [3-6]. Like the pancreatic cholesterol esterase, the catalytic activity of the human milk lipase is activated by the presence of bile salt [3]. The activity of the milk lipase is also inhibited by reagents that inhibit the pancreatic cholesterol esterase [6]. Moreover, antibodies prepared against pancreatic cholesterol esterase cross-reacted with the bile salt stimulated lipase in human milk [7]. Thus, these data suggest that human milk lipase is very similar to pancreatic cholesterol esterase and may be responsible for catalyzing fat absorption during prenatal and neonatal periods.

The cDNA clones and the sequences for rat and bovine pancreatic cholesterol esterase have been reported recently [8,9]. The N-terminal sequences of both cholesterol esterases are strikingly similar to that of the human milk lipase [5], suggesting that the human milk enzyme may be identical to the cholesterol esterase in mature human pancreas. In the current study, we report the cloning and sequencing of the cDNA for both the human milk lipase and pancreatic cholesterol esterase. The results of this study provided definitive proof of identity between the two proteins.

Correspondence address: D.Y. Hui, Department of Pathology and Laboratory Medicine, University of Cincinnati, College of Medicine, 231 Bethesda Ave, Cincinnati, OH 45267-0529, USA

The nucleotide sequence reported in this paper has been deposited into the GenBank database and has been assigned the accession number M37044

2. MATERIALS AND METHODS

2.1. *Screening of the cDNA library*

Full-length rat pancreatic cholesterol esterase cDNA [8] was labeled to high specific activity with [^{32}P]dATP and used to screen a human breast cDNA library in lambda gt10 vector (Clontech Laboratories, Inc.). According to the manufacturer, the mRNA used for the synthesis of this cDNA library was obtained from human breast tissue excised during the eighth month of pregnancy. The tissue was well-differentiated and was lactational competent. The ^{32}P-labeled rat cholesterol esterase cDNA was also used to screen two human pancreatic cDNA libraries (Clontech Laboratories, Inc.) to isolate the cDNA for human pancreatic cholesterol esterase. The hybridization and washing conditions for library screening were exactly the same as those described previously [8].

2.2. *DNA sequencing*

The cDNA from positive clones were isolated by plaque purification. The recombinant DNA was digested with *Eco*RI and then subcloned into pUC-13 plasmid vector for propagation. The resulting plasmids were analyzed by restriction mapping analysis and various restriction fragments of the cDNA were subcloned into M13mp18 or M13mp19 vectors. Single-stranded M13 DNA was prepared and used for sequencing [10] with T7 DNA polymerase. Sequence information was obtained from both strands to cover both orientations and overlapping sequences.

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

3. RESULTS AND DISCUSSION

Three overlapping cDNA clones encompassing the entire coding region of the bile salt-stimulated lipase in human milk have been isolated by screening of a human breast cDNA library. Nucleotide sequencing of the cDNA showed that the mRNA for bile salt stimulated lipase is at least 2340 bases in length. The translational initiation codon has been identified at residue 9 of the cDNA. The polyadenylation signal, AATAAA is located at 14 nucleotides upstream from the poly(A) tail (Fig. 1).

The primary sequence of the milk lipase was deduced from the nucleotide sequence of the cDNA. Beginning with the first ATG, at position 9, a long open reading frame coding for a protein of 745 amino acid residues was observed (Fig. 1). Comparison of this deduced sequence with the sequence of the N-terminal peptide [5] demonstrated that the human milk lipase is synthesized as a precursor protein with a 23-residue signal peptide. The sequence between the bile salt-stimulated lipase in human milk and the rat pancreatic cholesterol esterase is highly conserved with over 78% identity between the two proteins. The differences observed are mostly conserved substitutions and may reflect differences between the rat and human enzymes.

To definitely show that human milk lipase is identical to pancreatic cholesterol esterase, cDNAs were isolated from two different human pancreatic cDNA libraries for comparison. Although the libraries failed to produce a full-length cDNA for pancreatic cholesterol esterase, the complete primary sequence of the protein could be deduced based on the sequences of overlapping partial clones and from previously published data

```
A
  1 AGAGGCTG ATG CTC ACC ATG GGG CGC CTG CAA CTG GTT GTG TTG GGC CTC ACC TGC TGC TGG    62
-23          Met Leu Thr Met Gly Arg Leu Gln Leu Val Val Leu Gly Leu Thr Cys Cys Trp    -6

 63 GCA GTG GCG AGT GCC GCG AAG CTG GGC GCC GTG TAC ACA GAA GGT GGG TTC GTG GAA GGC   122
 -5 Ala Val Ala Ser Ala Ala Lys Leu Gly Ala Val Tyr Thr Glu Gly Gly Phe Val Glu Gly    15

123 GTC AAT AAG AAG CTC GGC CTC CTG GGT GAC TCT GTG GAC ATC TTC AAG GGC ATC CCC TTC   182
 16 Val Asn Lys Lys Leu Gly Leu Leu Gly Asp Ser Val Asp Ile Phe Lys Gly Ile Pro Phe    35

183 GCA GCT CCC ACC AAG GCC CTG GAA AAT CCT CAG CCA CAT CCT GGC TGG CAA GGG ACC CTG   242
 36 Ala Ala Pro Thr Lys Ala Leu Glu Asn Pro Gln Pro His Pro Gly Trp Gln Gly Thr Leu    55

243 AAG GCC AAG AAC TTC AAG AAG AGA TGC CTG CAG GCC ACC ATC ACC CAG GAC AGC ACC TAC   302
 56 Lys Ala Lys Asn Phe Lys Lys Arg Cys Leu Gln Ala Thr Ile Thr Gln Asp Ser Thr Tyr    75

303 GGG GAT GAA GAC TGC CTG TAC CTC AAC ATT TGG GTG CCC CAG GGC AGG AAG CAA GTC TCC   362
 76 Gly Asp Glu Asp Cys Leu Tyr Leu Asn Ile Trp Val Pro Gln Gly Arg Lys Gln Val Ser    95

363 CGG GAC CTG CCC GTT ATG ATC TGG ATC TAT GGA GGC GCC TTC CTC ATG GGG TCC GGC CAT   422
 96 Arg Asp Leu Pro Val Met Ile Trp Ile Tyr Gly Gly Ala Phe Leu Met Gly Ser Gly His   115

423 GGG GCC AAC TTC CTC AAC AAC TAC CTG TAT GAC GGC GAG GAG ATC GCC ACA CGC GGA AAC   482
116 Gly Ala Asn Phe Leu Asn Asn Tyr Leu Tyr Asp Gly Glu Glu Ile Ala Thr Arg Gly Asn   135

483 GTC ATC GTG GTC ACC TTC AAC TAC CGT GTC GGC CCC CTT GGG TTC CTC AGC ACT GGG GAC   542
136 Val Ile Val Val Thr Phe Asn Tyr Arg Val Gly Pro Leu Gly Phe Leu Ser Thr Gly Asp   155

543 GCC AAT CTG CCA GGT AAC TAT GGT CTT CGG GAT CAG CAC ATG GCC ATT GCT TGG GTG AAG   602
156 Ala Asn Leu Pro Gly Asn Tyr Gly Leu Arg Asp Gln His Met Ala Ile Ala Trp Val Lys   175

603 AGG AAT ATC GCG GCC TTC GGG GGG GAC CCC AAC AAC ATC ACG CTC TTC GGG GAG TCT GCT   662
176 Arg Asn Ile Ala Ala Phe Gly Gly Asp Pro Asn Asn Ile Thr Leu Phe Gly Glu Ser Ala   195

663 GGA GGT GCC AGC GTC TCT CTG CAG ACC CTC TCC CCC TAC AAC AAG GGC CTC ATC CGG CGA   722
196 Gly Gly Ala Ser Val Ser Leu Gln Thr Leu Ser Pro Tyr Asn Lys Gly Leu Ile Arg Arg   215

723 GCC ATC AGC CAG AGC GGC GTG GCC CTG AGT CCC TGG GTC ATC CAG AAA AAC CCA CTC TTC   782
216 Ala Ile Ser Gln Ser Gly Val Ala Leu Ser Pro Trp Val Ile Gln Lys Asn Pro Leu Phe   235

783 TGG GCC AAA AAG GTG GCT GAG AAG GTG GGT TGC CCT GTG GGT GAT GCC GCC AGG ATG GCC   842
236 Trp Ala Lys Lys Val Ala Glu Lys Val Gly Cys Pro Val Gly Asp Ala Ala Arg Met Ala   255

843 CAG TGT CTG AAG GTT ACT GAT CCC CGA GCC CTG ACG CTG GCC TAT AAG GTG CCG CTG GCA   902
256 Gln Cys Leu Lys Val Thr Asp Pro Arg Ala Leu Thr Leu Ala Tyr Lys Val Pro Leu Ala   275

903 GGC CTG GAG TAC CCC ATG CTG CAC TAT GTG GGC TTC GTC CCT GTC ATT GAT GGA GAC TTC   962
276 Gly Leu Glu Tyr Pro Met Leu His Tyr Val Gly Phe Val Pro Val Ile Asp Gly Asp Phe   295

963 ATC CCC GCT GAC CCG ATC AAC CTG TAC GCC AAC GCC GCC GAC ATC GAC TAT ATA GCA GGC  1022
296 Ile Pro Ala Asp Pro Ile Asn Leu Tyr Ala Asn Ala Ala Asp Ile Asp Tyr Ile Ala Gly   315

1023 ACC AAC AAC ATG GAC GCC CAC ATC TTC GCC AGC ATC GAC ATG CCT GCC ATC AAC AAG GGC  1082
316 Thr Asn Asn Met Asp Gly His Ile Phe Ala Ser Ile Asp Met Pro Ala Ile Asn Lys Gly   335

1083 AAC AAG AAA GTC ACG GAG GAG GAC TTC TAC AAG CTG GTC AGT GAG TTC ACA ATC ACC AAG  1142
336 Asn Lys Lys Val Thr Glu Glu Asp Phe Tyr Lys Leu Val Ser Glu Phe Thr Ile Thr Lys   355
```

Fig. 1. Nucleotide sequence of cDNA encoding human bile salt-stimulated lipase and its deduced amino acid sequence. The N-terminal peptide sequence reported in ref. [5] is underlined. The N-terminal residue in the mature protein is numbered as residue 1.

(continued on p. 133)

on the amino terminal sequence of the protein [7]. The comparison of the deduced amino acid sequences for the pancreatic cholesterol esterase with milk lipase revealed complete identity between the two proteins.

The human cholesterol esterase/lipase is structurally similar to other serine esterases. As noted previously for the rat pancreatic cholesterol esterase [8], the human enzyme also shares homology with a 63-residue domain in acetylcholinesterase (Table I). This region, encompassing the active site serine of acetylcholinesterase [11], has also been demonstrated recently to contain the active site serine (Ser^{194}) of rat pancreatic cholesterol esterase [12]. Thus, Ser^{194} in human cholesterol esterase/lipase is most likely the active site serine of this protein.

The cholesterol esterase/lipase and cholesterol esterase are similar to acetylcholinesterase and cholinesterase at two additional domains. One of these domains is located at residues 78 to 88 where 100% identity was observed between the lipases and acetylcholinesterase (Table II). The sequence for cholinesterase [13] is similar with two conserved substitutions (Table II). The other domain with similarities is within residues 430–446. The human lipase and rat pancreatic cholesterol esterase are identical in this region and are 47% similar to the cholinesterases. The similarities increase to 88% if conserved substitutions are considered.

Although the functional significance of these homologous domains remains unknown at the present time, it is noteworthy that the activity of these proteins depends on a catalytic triad involving serine, histidine, and an acidic amino acid [6, 14–15]. The key histidine and the acidic residue have not been identified in any of these proteins. In view of the observation that His^{435} in cholesterol esterase/lipase is the only histidine conserved with the cholinesterases, this residue may be the key histidine involved in the catalytic activity of the protein.

B

```
1143  GGG CTC AGA GGC GCC AAG ACG ACC TTT GAT GTC TAC ACT GAG TCC TGG GCC CAG GAC CCA  1202
 356  Gly Leu Arg Gly Ala Lys Thr Thr Phe Asp Val Tyr Thr Glu Ser Trp Ala Gln Asp Pro   375

1203  TCC CAG GAG AAT AAG AAG AAG ACT GTG GTG GAC TTT GAG ACC GAT GTC CTC TTC CTG GTG  1262
 376  Ser Gln Glu Asn Lys Lys Lys Thr Val Val Asp Phe Glu Thr Asp Val Leu Phe Leu Val   395

1263  CCC ACC GAG ATT GCC CTA GCC CAG CAC AGA GCC AAT GCC AAG AGT GCC AAG ACC TAC GCC  1322
 396  Pro Thr Glu Ile Ala Leu Ala Gln His Arg Ala Asn Ala Lys Ser Ala Lys Thr Tyr Ala   415

1323  TAC CTG TTT TCC CAT CCC TCT CGG ATG CCC GTC TAC CCC AAA TGG GTG GGG GCC GAC CAT  1382
 416  Tyr Leu Phe Ser His Pro Ser Arg Met Pro Val Tyr Pro Lys Trp Val Gly Ala Asp His   435

1383  GCA GAT GAC ATT CAG TAC GTT TTC GGG AAG CCC TTC GCC ACC CCC ACG GGC TAC CGG CCC  1442
 436  Ala Asp Asp Ile Gln Tyr Val Phe Gly Lys Pro Phe Ala Thr Pro Thr Gly Tyr Arg Pro   455

1443  CAA GAC AGG ACA GTC TCT AAG GCC ATG ATC GCC TAC TGG ACC AAC TTT GCC AAA ACA GGG  1502
 456  Gln Asp Arg Thr Val Ser Lys Ala Met Ile Ala Tyr Trp Thr Asn Phe Ala Lys Thr Gly   475

1503  GAC CCC AAC ATG GGC GAC TCG GCT GTG CCC ACA CAC TGG GAA CCC TAC ACT ACG GAA AAC  1562
 476  Asp Pro Asn Met Gly Asp Ser Ala Val Pro Thr His Trp Glu Pro Tyr Thr Thr Glu Asn   495

1563  AGC GGC TAC CTG GAG ATC ACC AAG AAG ATG GGC AGC AGC TCC ATG AAG CGG AGC CTG AGA  1622
 496  Ser Gly Tyr Leu Glu Ile Thr Lys Lys Met Gly Ser Ser Ser Met Lys Arg Ser Leu Arg   515

1623  ACC AAC TTC CTG CGC TAC TGG ACC CTC ACC TAT CTG GCG CTG CCC ACA GTG ACC GAC CAG  1682
 516  Thr Asn Phe Leu Arg Tyr Trp Thr Leu Thr Tyr Leu Ala Leu Pro Thr Val Thr Asp Gln   535

1683  GAG GCC ACC CCT GTG CCC CCC ACA GGG GAC TCC GAG GCC ACT CCC GTG CCC CCC ACG GGT  1742
 536  Glu Ala Thr Pro Val Pro Pro Thr Gly Asp Ser Glu Ala Thr Pro Val Pro Pro Thr Gly   555

1743  GAC TCC GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC  1802
 556  Asp Ser Glu Thr Ala Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro   575

1803  ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG  1862
 576  Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val   595

1863  CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC  1922
 596  Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro   615

1923  CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGC  1982
 616  Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly   635

1983  GCC CCC CCC GTG CCG CCC ACG GGT GAC GCC GGG CCC CCC CCC GTG CCG CCC ACG GGT GAC  2042
 636  Ala Pro Pro Val Pro Pro Thr Gly Asp Ala Gly Pro Pro Pro Val Pro Pro Thr Gly Asp   655

2043  TCC GGC GCC CCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCC GTG ACC CCC ACG  2102
 656  Ser Gly Ala Pro Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Thr Pro Thr   675

2103  GGT GAC TCC GAG ACC GCC CCC GTG CCG CCC ACG GGT GAC TCC GGG GCC CCC CCT GTG CCC  2162
 676  Gly Asp Ser Glu Thr Ala Pro Val Pro Pro Thr Gly Asp Ser Gly Ala Pro Pro Val Pro   695

2163  CCC ACG GGT GAC TCT GAG GCT GCC CCT GTG CCC CCC ACA GAT GAC TCC AAG GAA GCT CAG  2222
 696  Pro Thr Gly Asp Ser Glu Ala Ala Pro Val Pro Pro Thr Asp Asp Ser Lys Glu Ala Gln   715

2223  ATG CCT GCA GTC AAT AGG TTT TAG CGTCCCATGAGCCTTGGTATCAAGAGGCCACAAGAGTGGGACCCCAG  2293
 716  Met Pro Ala Val Ile Arg Phe  *                                                   722

2294  GGGCTCCCCTCCCATCTTGAGCTCTTCCTGAATAAAGCCTCATACCCCTGAAA(n)
```

Table I

Serine active site domains for the esterases

```
AChE          PGNVGLLDQRMALQWVHDNIQFFGGDPKTVTIFGESAGGASVGMHILSPGSRDLFRRAILQSG
                 *  *  *  **  **  **      *** *           ****   **** *    *
HUMAN.CEH     PGNYGLRDQHMAIAWVKRNIAAFGGDPNNITLFGESAGGASVSLQTLSPYNKGLIRRAISQSG
                 *                       *   *       *
RAT.CEH       PGNFGLRDQHMAIAWVKRNIAAFGGDPDNITIFGESAGAASVSLQTLSPYNKGLIRRAISQSG
```

Comparison of the active site domains of acetylcholinesterase (AChE) and rat cholesterol esterase (RAT.CEH) with a region of human cholesterol esterase (HUMAN.CEH). The * indicates differences in amino acid. The active site motifs are underlined.

Table II

Putative domain for active site acidic residue

```
                                  *  *
Cholinesterase (90-100)  E D C L Y L N V W I P
AChE (92-102)            E D C L Y L N I W V P
HUMAN.CEH (78-88)        E D C L Y L N I W V P
RAT.CEH (78-88)          E D C L Y L N I W V P
```

A conserved sequence between human serum cholinesterase, acetylcholinesterase (AChE), the human- (HUMAN.CEH), and rat- cholesterol esterase (RAT.CEH). The differences are indicated by * and the putative acidic residue(s) important for enzyme activities are underlined.

Table III

Putative histidine active site domain

```
AChE (435-451)           W M G V I H G V E I E F V F G L P
CHOLINESTERASE (433-449) W M G V M H G Y E I E F V F G L P
                           *  *    * * * * # *      #
HUMAN.CEH (430-446)      W V G A D H A D D I Q Y V F G K P
RAT.CEH (430-446)        W M G A D H A D D L Q Y V F G K P
```

Sequence comparison between putative histidine active site domains of acetylcholinesterase (AChE), human serum cholinesterase, human- (HUMAN.CEH), and rat- cholesterol esterase (RAT.CEH). Conserved substitutions between the serine esterases and the cholesterol esterases are indicated by * while the non-conserved substitutions are indicated by #. The putative active histidines are underlined.

The identity of Glu^{78} and Asp^{79} in the third homologous domain suggests that one or both of these residues may also be important for enzyme activity. Thus, we propose that Glu^{78}/Asp^{79}, Ser^{194}, and His^{435} form the catalytic triad for cholesterol esterase/lipase.

Acknowledgement: This research was supported by Grant DK40917 from the National Institutes of Health and by a Grant-in-aid (SW-90-15) from the Ohio Affiliates of the American Heart Association. D.Y.H. was the recipient of an Established Investigator Award from the American Heart Association while this research was in progress.

REFERENCES

[1] Rudd, E.A. and Brockman, H.L. (1984) in: Lipase (B. Borgstrom and H.L. Brockman, eds) pp. 185–204, Elsevier, Amsterdam.
[2] Cox, D.G., Leung, C.K.T., Kyger, E.M., Spilburg, C.A. and Lange, L.G. (1990) Biochemistry 29, 3842–3848.
[3] Hernell, O. and Olivercrona, T. (1974) Biochim. Biophys. Acta 369, 234–244.
[4] Blackberg, L., Blind, P.-J., Ljungberg, B. and Hernell, O. (1985) J. Pediatr. Gastroenterol. and Nutr. 4, 441–445.
[5] Wang, C.-S. and Johnson, K. (1983) Anal. Biochem. 133, 457–461.
[6] Abouakil, N., Rogalska, E. and Lombardo, D. (1989) Biochim. Biophys. Acta 1002, 225–230.
[7] Abouakil, N., Rogalska, E., Bonicel, J. and Lombardo, D. (1988) Biochim. Biophys. Acta 961, 299–308.
[8] Kissel, J.A., Fontaine, R.N., Turck, C.W., Brockman, H.L. and Hui, D.Y. (1989) Biochim. Biophys. Acta 1006, 227–236.
[9] Kyger, E.M., Wiegand, R.C. and Lange, L.G. (1989) Biochem. Biophys. Res. Commun. 164, 1302–1309.
[10] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[11] MacPhee-Quigley, K., Taylor, P. and Taylor, S. (1985) J. Biol. Chem. 260, 12185–12189.
[12] DiPersio, L.P., Fontaine, R.N. and Hui, D.Y. (1990) J. Biol. Chem. 265, 16801–16806.
[13] Lockridge, O., Bartels, C.F., Vaughan, T.A., Wong, C.K., Norton, S.E. and Johnson, L.L. (1987) J. Biol. Chem. 262, 549–557.
[14] Mounter, L.A., Alexander, H.C., Tuck, K.D. and Dien, L.T. (1957) J. Biol. Chem. 226, 867–872.
[15] Lombardo, D. (1982) Biochim. Biophys. Acta 700, 67–74.