Acta Biochimica Polonica

Vol. 53 No. 3/2006, 553–561

on-line at: www.actabp.pl

Regular paper

Molecular cloning and functional expression of human cytosolic acetyl-CoA hydrolase

Naoya Suematsu and Fumihide Isohashi✉

Department of Biochemistry, St. Marianna University School of Medicine, Kanagawa, Japan;
✉e-mail: n2sue@marianna-u.ac.jp

Received: 09 March, 2006; revised: 09 June, 2006; accepted: 27 June, 2006
available on-line: 02 September, 2006

A cDNA encoding human cytosolic acetyl-CoA hydrolase (CACH) was isolated from a human liver cDNA library, sequenced and functionally expressed in insect cells. The human CACH cDNA encodes a 555-amino-acid sequence that is 81.4%/78.7% identical to those of the mouse/rat homologue, suggesting a conserved role for this enzyme in the human and rodent livers. Bioinformatical study further reveals a high degree of similarity among the human and rodent CACHs as follows: First, the gene is composed of 15 exons ranging in size from 56 to 157 bp. Second, the protein consists of two thioesterase regions and a C-terminal steroidogenic acute regulatory protein-related lipid transfer (START) domain. Third, the promoter region is GC-rich and contains GC boxes, but lacks both TATA and CCAAT boxes, the typical criteria of housekeeping genes. A consensus peroxisome proliferator responsive element (PPRE) present in the rodent CACH promoter regions supports marked CACH induction in rat liver by peroxisome proliferator (PP).

Keywords: acetyl-CoA hydrolase, PCR, cDNA sequence, *Spodoptera frugiperda*, functional expression, housekeeping-type promoter

INTRODUCTION

The cytosolic or extramitochondrial acetyl-CoA hydrolase (CACH) hydrolyzes acetyl-CoA to acetate and CoA. It seems superficially wasteful to hydrolyze the most common energy-rich metabolite acetyl-CoA without recovering any energy. So it is exciting to understand the physiological role of the enzyme comprehensibly.

The enzyme has been detected in rat liver (Prass *et al*., 1980) and kidney (its cytosolic CACH specific activity was 5% of that of the liver enzyme) (Matsunaga *et al*., 1985). Its activity in the liver increases notably in the opposite metabolic states: during enhanced fatty acid oxidation and in increased fatty acid synthesis (Matsunaga *et al*., 1985). Further, a marked induction in the liver was observed by thyroid hormones (Matsunaga *et al*., 1985) and also by 2-(*p*-chlorophenoxy) isobutyric acid (Ebisuno *et al*., 1988), a hypolipidemic drug or peroxisome proliferator (PP), which enhances β-oxidation in rat liver mitochondria and peroxisomes (Mannaerts *et al*., 1979) and increases cytosolic CoA level (Berge *et al*., 1983; Horie *et al*., 1986). These findings reveal that the enzyme plays a vital role specifically in fat metabolism by supplying cytosolic free CoA necessary for both fatty acid synthesis and oxidation (Matsunaga *et al*., 1985).

The enzyme had rejected earlier an enough of purification due to its extreme cold lability (Isohashi *et al*., 1983a; Suematsu *et al*., 1996) and little activity in the absence of ATP (Söling & Rescher,

***Note**: Nucleotide sequence data are available in the DDBJ/EMBL/GenBank databases under the accession number AB078619. **Enzymes**: acetyl-CoA hydrolase (EC 3.1.2.1); acyl-CoA thioesterase (EC 3.1.2.2); 4-hydroxybenzoyl-CoA thioesterase (EC 3.1.2.23).

Abbreviations: CACH, cytosolic acetyl-CoA hydrolase; ESTs, expressed sequence tags; NCBI, National Center for Biotechnology Information; TESS, Transcription Element Search Software; ORF, open reading frame; START, steroidogenic acute regulatory protein-related lipid transfer; 4HBT, 4-hydroxybenzoyl-CoA thioesterase; PP, peroxisome proliferator; PPRE, peroxisome proliferator responsive element; bHLH, basic helix-loop-helix.

1985; Nakanishi *et al*., 1988). At long last, we successfully purified the enzyme in the presence of a protease inhibitor at room temperature (Ebisuno *et al*., 1989; Nakanishi *et al*., 1993). Characterization of the purified enzyme demonstrated that at room temperature it was usually present as active forms: homodimer (135 kDa) and homotetramer (240 kDa) whose K_m values for acetyl-CoA were 170 μM and 60 μM, respectively (Isohashi *et al*., 1983a). But under cold conditions, they dissociate into an inactive monomer (63 kDa) showing no hydrolyzing activity any more even at 25°C (Isohashi *et al*., 1984). Our *in vitro* study further revealed that CACH is an allosteric enzyme regulated by ATP (activator) and ADP (inhibitor) (Isohashi *et al*., 1983b; Nakanishi *et al*., 1994), suggesting it is presumably a key enzyme involved in energy metabolism. It should be noted here that ATP is not a substrate but an allosteric effector because ATP works well without production of either ADP or inorganic phosphate in the absence of Mg^{2+} (Prass *et al*., 1980). Recently, we found that low concentration of Triton X-100 prevents cold inactivation of CACH and further partially reactivates the cold-inactivated enzyme at 37°C (Suematsu *et al*., 2003), allowing us to purify the enzyme efficiently under usual cold conditions at 4°C.

We previously reported molecular cloning of rat and mouse CACH cDNAs, demonstrating that the enzyme is a novel thioesterase (Suematsu *et al*., 2001; 2002). Here we report molecular cloning and recombinant expression of a human homologue cDNA, as the third example of mammalian species. We have further analyzed the corresponding gene in the established databases and describe its exon-intron structure. We also present an initial search for its *cis*-regulatory elements. Molecular analysis of the newly discovered human CACH gene might provide clues to its expression and physiological functions of the enzyme and further our understanding of the implications of peroxisome proliferator-induced pleiotropic responses to human health.

MATERIALS AND METHODS

Chemicals. Acetyl-CoA was synthesized as described previously (Simon & Shemin, 1953). All other chemicals and reagents used were of analytical grade or better.

Enzyme assay. Enzyme activity was routinely assayed at 25°C as previously described (Prass *et al*., 1980). One unit of activity is that required to hydrolyze 1 μmol of acetyl-CoA × min^{-1} under the conditions of the assay. Acetyl-CoA hydrolase activity was estimated by subtracting the nonenzymatic rate measured in the presence of 2 mM ADP, which inhibits the enzymatic activity, from that observed in the 2 mM ATP. All determinations were carried out in triplicate.

cDNA cloning from human liver cDNA library. Oligonucleotide PCR primers were designed (Table 1) based on the two human expressed sequence tags (ESTs) AV693695 and AV685167 in the public database: sense primers S1 and S2 correspond to nucleotide sequences at 1/21 and 7/27 of the 5′-terminus region of AV693695, respectively, and antisense primers A1 and A2 correspond to those at 363/386 and 329/352 of the 3′-terminus region of AV685167, respectively. A human liver cDNA library (QUICK-Clone cDNA, CLONTECH Laboratories, Inc.), derived from whole liver mRNA from a 41-yr-old Caucasian, was used as the template for nested PCR with *Ex Taq* DNA Polymerase (TaKaRa), using the primer set of S1/A1 for the first step and then the nested set of S2/A2 for the second. Both strands of the amplified cDNA were directly sequenced. Then for the cDNA cloning, the nested PCR step was performed with Platinum *Pfx* DNA Polymerase (Life Technologies, Inc.), using a primer set of S5/A2 (Table 1). The anchor primer S5 includes a restriction site to facilitate subsequent cloning. The amplified cDNA was then inserted into the *Xma*I/*Psh*AI site of the baculovirus transfer vector pTriEx-4 for expressing His-tag fusion protein, allowing a His-tag pull

Table 1. PCR primers used for cloning human acetyl-CoA hydrolase cDNA

Nucleotide positions are numbered as in Fig. 2. The restriction site used for the cDNA cloning is highlighted in bold type. S, sense; A, antisense; CDS, coding sequence.

Primer	Sequence (5′-3′)	Position	Location
S0	AAATTTGAACGCCAGCACATGGAC	(specific to pTriEx-4 vector)	
S1	GCCGCAGGCTTAGCGCTCTCG	–29/–9	5′-UTR
S2	GGCTTAGCGCTCTCGCCCTGG	–23/–3	5′-UTR
S3	TTCAGAGCATTGAACTGGTCCTCC	548/571	CDS
S4	AAAAGGGGTTGGGAGGTTACCAGC	1093/1116	CDS
S5	TCTCG**CCCGGG**CCATGGAGC	–13/7	anchor primer
A0	TCGATCTCAGTGGTATTTGTGAGC	(specific to pTriEx-4 vector)	
A1	CCAAATCACTGGTAGACAAGATGC	1809/1832	3′-UTR
A2	CTCCGTCACTGCATTTTGCTTAGG	1774/1797	3′-UTR
A3	TTTTCCACAGTGCTGGTAACCTCC	1104/1127	CDS
A4	AAATGTGTTTCCGTGGTGGTTTGC	580/603	CDS

down approach. The resulting recombinant transfer vector was designated as pTriEx-4/*hCACH* (Fig. 1A).

Recombinant expression and affinity purification. The procedures for amplification of the recombinant virus and the assessment of recombinant expression were previously described in detail (Suematsu *et al*., 2002). The recombinant virus was isolated and then used to transform monolayer *Spodoptera frugiperda* (*Sf9*) insect cells at a multiplicity of infection of 10 for recombinant expression of the His-tag fusion protein (Fig. 1B). Affinity purification of the expressed human His-tag CACH was performed on Ni^{2+}-charged resin using HISTAGcatcher (Cytosignal) according to the procedures provided with the purification system. The purified His-tag fusion protein was then relieved of the fused peptide by thrombin cleavage using Thrombin Cleavage Capture Kit (Novagen) according to the manufacturer's instructions.

Bioinformatics. The nucleotide sequences of the non-coding region of mammalian CACH genes were from established databases. Nucleotide and amino-acid sequence homology was searched with a computer using BLAST 2.0, and sequence alignments were made using the DiAlign computer algorithm (Genomatix Software GmbH, München). Subcellular localization was predicted by PSORT WWW Server (http://psort.nibb.ac.jp). Functional domains were predicted using the National Center for Biotechnology Information (NCBI) Conserved Domain Search website (www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Consensus *cis*-elements were searched using the Transcription Element Search Software (TESS) website (Schug & Overton, 1997) browsing the TRANSFAC database.

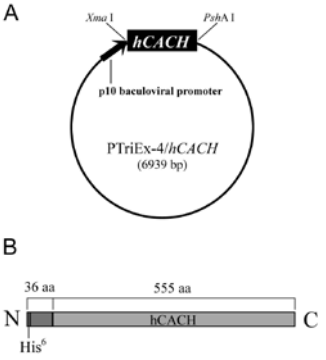

Figure 1. Schematic depiction of the strategy for recombinant expression of a full-length ORF of human cytosolic acetyl-CoA hydrolase.
A. Construction of the recombinant baculovirus transfer vector containing the complete coding region of human cytosolic acetyl-CoA hydrolase (hCACH). The full-length ORF encoding hCACH was inserted into the multiple cloning site of the baculovirus transfer vector pTriEx-4 at the *Xma*I/*Psh*AI site. The resultant vector was 6939 bp long and designated as pTriEx-4/*hCACH*. **B.** A schematic representation of the recombinant CACH containing a polyhistidine (His^6) tag.

RESULTS

Molecular cloning and sequence analysis of a human CACH cDNA covering a full-length ORF

Database searching revealed that several human expressed sequence tags (ESTs) represent high homology (around 75%) to the rodent CACH cDNAs (Suematsu *et al*., 2001; 2002), implicating the existence of a human homologue cDNA. One human EST clone (586 bp long, AV693695 in GenBank) shows 82% identity to the 5′-terminus of the rat cDNA in a 114 amino acid overlap, and another (394 bp long, AV685167 in GenBank) shows 70% identity to the 3′-terminus in a 73 amino acid overlap. Based on the nucleotide sequences of the two ESTs, oligonucleotide PCR primers were designed (Table 1) and used to clone a human homologue cDNA from a human liver cDNA library. As outlined in Fig. 2, the amplified human homologue

```
ggcttagcgctctcgccctggccatggagcggccggcgcccggcgaggtggtcatgagccaagccatccag   48
                        M  E  R  P  A  P  G  E  V  V  M  S  Q  A  I  Q    16
ccggcgcacgccactgcgccgcggcgagctgagcgcggggcagctgctcaagtggatcgacaccaccgcctgc  120
P  A  H  A  T  A  R  G  E  L  S  A  G  Q  L  L  K  W  I  D  T  T  A  C    40
ctggcggctgagaaacatgctggagtttcctgcgttacagcctcagtggatgacatacagtttgaggagaca  192
L  A  A  E  K  H  A  G  V  S  C  V  T  A  S  V  D  D  I  Q  F  E  E  T    64
gctagagttggacaagttataaccatcaaagcaaaagttactagagcattcagcacaagcatggagatcagt  264
A  R  V  G  Q  V  I  T  I  K  A  K  V  T  R  A  F  S  T  S  M  E  I  S    88
atcaaggtcatggtacaggatatgctcactggcattgagaagcttgttagtgtggctttctccacatttgta  336
I  K  V  M  V  Q  D  M  L  T  G  I  E  K  L  V  S  V  A  F  S  T  F  V   112
gccaaaccagttggaaaagaaaagattcatttaaaaccagtcacacttctaactgaacaagatcatgtggaa  408
A  K  P  V  G  K  E  K  I  H  L  K  P  V  T  L  L  T  E  Q  D  H  V  E   136
catcatctggctgctgagagaaggaaagttcgattacaacatgaagatacctttaacaatttaatgaaggaa  480
H  N  L  A  A  E  R  R  K  V  R  L  Q  H  E  D  T  F  N  N  L  M  K  E   160
agtagcaaatttgatgatctcatttttgatgaagaggaaggagcggtttccacaaggggcacctccgttcag  552
S  S  K  F  D  D  L  I  F  D  E  E  E  G  A  V  S  T  R  G  T  S  V  Q   184
agcattgaactggtcctcccaccccatgcaaaccatcacggaaatacatttggtggccagattatggcgtgg  624
S  I  E  L  V  L  P  P  H  A  N  H  H  G  N  T  F  G  G  Q  I  M  A  W   208
atggagacagtggctactatttctgcaagccgcctgtgttgggctcatccctttctgaagtccgtagatatg  696
M  E  T  V  A  T  I  S  A  S  R  L  C  W  A  H  P  F  L  K  S  V  D  M   232
tttaagttccgggggaccatctacagttggagatcgtcttgtcttcactgccattgtcaacaatacatttcag  768
F  K  F  R  G  P  S  T  V  G  D  R  L  V  F  T  A  I  V  N  N  T  F  Q   256
acctgtgttgaagttggagttcgcgtggaggcctttgactgtcaggaatgggccgagggccgagggcgtcac  840
T  C  V  E  V  G  V  R  V  E  A  F  D  C  Q  E  W  A  E  G  R  G  R  H   280
atcaacagtgcttttctcatttacaatgctgctgatgataaggaaaatctcatcacgtttcccagaatccaa  912
I  N  S  A  F  L  I  Y  N  A  A  D  D  K  E  N  L  I  T  F  P  R  I  Q   304
cccatttcaaaggatgatttcagacgctatcggggagctattgcacgcaagcgaattcgcctaggcagaaaa  984
P  I  S  K  D  D  F  R  R  Y  R  G  A  I  A  R  K  R  I  R  L  G  R  K   328
tatgttatttcccacaaagaagaggttccactttgcatacactgggatatcagcaagcaggcatccctgagt 1056
Y  V  I  S  H  K  E  E  V  P  L  C  I  H  W  D  I  S  K  Q  A  S  L  S   352
gacagcaatgtggaggccctcaaaaaactggcagccaaaaggggttgggaggttaccagcactgtggaaaag 1128
D  S  N  V  E  A  L  K  K  L  A  A  K  R  G  W  E  V  T  S  T  V  E  K   376
ataaaaatatatactctggaagagcatgatgttttatctgtttgggttgaaaagcacgtgggaagtccagca 1200
I  K  I  Y  T  L  E  E  H  D  V  L  S  V  W  V  E  K  H  V  G  S  P  A   400
catttggcttatcgtctcttgtctgactttacaaagcgacctttgtgggaccccccattttgtgtcctgtgaa 1272
H  L  A  Y  R  L  L  S  D  F  T  K  R  P  L  W  D  P  H  F  V  S  C  E   424
gtcatagactgggtgagtgaagatgatcagctgtatcacatcacctgtcctatactgaatgatgacaaaccc 1344
V  I  D  W  V  S  E  D  D  Q  L  Y  H  I  T  C  P  I  L  N  D  D  K  P   448
aaagacttggtagtactcgtatcacgaagaaaaccccctcaaagatggtaacacttacacagtggcagtgaag 1416
K  D  L  V  V  L  V  S  R  R  K  P  L  K  D  G  N  T  Y  T  V  A  V  K   472
tcggtcattttgccatcggtccccccgtctccacagtacatcagaagtgaaatcatatgtgccggatttctc 1488
S  V  I  L  P  S  V  P  P  S  P  Q  Y  I  R  S  E  I  I  C  A  G  F  L   496
atccatgctattgacagcaattcatgcatcgtatcttacttaaccatatgtctgctagcatcctcttctac 1560
I  H  A  I  D  S  N  S  C  I  V  S  Y  F  N  H  M  S  A  S  I  L  P  Y   520
tttgctggaaatcttggtggctggtcaaaatccattgaagaaacagcagcctcttgtatacagttcttagag 1632
F  A  G  N  L  G  G  W  S  K  S  I  E  E  T  A  A  S  C  I  Q  F  L  E   544
aatcctcctgatgatgggtttgtaagcacattttaaaggtcaacttccaattactggtaatttaatttccca 1704
N  P  P  D  D  G  F  V  S  T  F  *                                       555
cttttaattccaagcacccttagccctgacatctgtcaagctttggggccacaaaataatttaatataacct 1776
aagcaaaatgcagtgacggag                                                   1797
```

Figure 2. Nucleotide and deduced amino-acid sequences of the cDNA encoding full-length ORF of human cytosolic acetyl-CoA hydrolase.
The nucleotide and predicted amino-acid residues are numbered on the right from the first base of the ATG start codon (shown in bold face). The underlined nucleotide sequences correspond to the gene-specific primers used for sequencing (Table 1). The asterisk denotes the TAA stop codon. Sequence data have been submitted to DDBJ under the accession number AB078619.

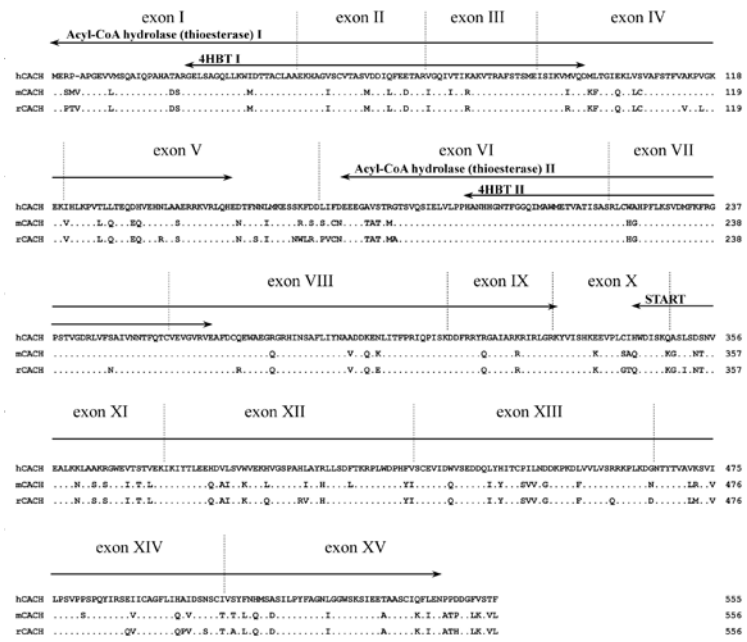

Figure 3. Comparison of the deduced amino-acid sequence of cytosolic acetyl-CoA hydrolase among human and the rodent species.

The amino-acid residues are numbered on the right. Residues identical with those in the first sequence were denoted by dots. A dash has been introduced to maximize alignment. Exons I–XV are presented with their boundaries indicated by broken line. Arrows represent the extent of functional domains revealed by a publicly available database at the NCBI Conserved Domain Search: Acyl-CoA hydrolase (thioesterase); 4HBT (4-hydroxybenzoyl-CoA thioesterase); START (steroidgenic acute regulatory (StAR)-related lipid-transfer) domain. The compared sequences are as follows: hCACH, human cytosolic acetyl-CoA hydrolase (GenBank accession AB078619); mCACH, mouse cytosolic acetyl-CoA hydrolase (GenBank accession AB078618); rCACH, rat cytosolic acetyl-CoA hydrolase (GenBank accession AB040609).

cDNA is 1820 bp long and comprises the entire coding region of 1668 bp. The open reading frame (ORF) for a polypeptide of 555 amino-acid residues with a calculated molecular mass of 62033 Da, starts with an ATG initiation codon (Fig. 2, shown in bold face), whose first nucleotide is numbered as +1, and terminates with a TAA stop codon (at positions 1669/1671, shown with an asterisk). The first AUG agrees pretty well with the Kozak's consensus sequence GCCGCC(A/G)CCAUGG for initiation in higher eukaryotes (Kozak, 1987; 1991). As shown in Fig. 3, the deduced amino-acid sequence of the human homologue cDNA exhibited extensive homology throughout the entire ORF to those of the mouse and rat CACH (81.4% and 78.7% identity, respectively, see Table 2). The human cDNA sequence is available from DDBJ/EMBL/GenBank under the accession number AB078619.

Expression and purification of recombinant CACH in *Sf9* insect cells

Recombinant expression of the human CACH cDNA in *Sf9* insect cells resulted in overproduction of a His-tag fusion protein with the expected enzymatic activity. The activity was detected as early as two days after infection in the lysate prepared from infected cells, and was maximal at four days post-infection. The expressed human recombinant enzyme hydrolyzed acetyl-CoA in the presence of 2 mM ATP and the activity was completely inhibited by ADP. Further, it exhibited cold lability and its cold-inactivation could be partially abolished through the incubation at 37°C in the presence of 0.16 mM Triton X-100 just as the rodent enzymes (Suematsu *et al*., 2001; 2002). The expressed His-tag

Table 2. Comparison of nucleotide and deduced amino-acid sequence of cytosolic acetyl-CoA hydrolase ORF among human and the rodents.

Percentage identities of nucleotide and amino-acid residues are shown. The compared sequences are described in the legend to Fig. 3.

	Length	Identical residues with human CACH	
hCACH	555 aa		
mCACH	556 aa	1371/1665 bp (82.3%)	452/555 aa (81.4%)
rCACH	556 aa	1356/1665 bp (81.4%)	437/555 aa (78.7%)

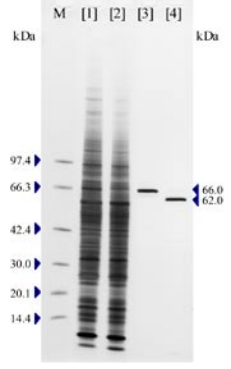

Figure 4. SDS/PAGE analysis of purified recombinant hCACH protein.

Protein samples from purification steps for the recombinant hCACH protein were subjected to SDS/PAGE on a 10–20% polyacrylamide gradient gel followed by silver staining. The samples, prepared as described in Materials and Methods, were the clarified cell lysate from the infected insect cells [1], the flow through from the Ni^{2+}-charged resin [2], the eluted His-tag fusion protein [3] and the hCACH protein free of the fused peptide [4]. Numbers to the left and right denote the molecular mass (in kDa) of the markers (lane M) and the purified proteins (lanes [3] and [4]). The molecular mass markers are phosphorylase *b* (97400), bovine serum albumin (66300), aldolase (42400), carbonic anhydrase (30300), soybean trypsin inhibitor (20100) and lysozyme (14400).

Table 3. Affinity purification and characterization of a recombinant form of human cytosolic acetyl-CoA hydrolase.

(A) Recombinant human cytosolic acetyl-CoA hydrolase protein expressed in infected *Sf9* insect cells (4 μl cell pellet) was purified as described in the text. (B) Constants representing catalytic activity of purified human recombinant enzyme (hCACH) and purified rat counterpart (rCACH, Suematsu *et al*., 2001). Proteins were determined according to the method of Bradford.

(A)

Step	Total activity (U)	Total protein (μg)	Specific activity ($U \cdot mg^{-1}$)	Purification (-fold)	Yield (%)
Lysate	5.2	655	7.9	–	100
Eluate	5.0	1.51	3300	419	97

(B)

Enzyme	Specific activity ($U \cdot mg^{-1}$)	Michaelis constant K_m (M)	Catalytic constant k_{cat} (s^{-1})	Catalytic efficiency k_{cat}/K_m ($M^{-1} \cdot s^{-1}$)
hCACH	3300	1.5×10^{-4}	7.0×10^{3}	4.6×10^{7}
rCACH	3100	1.5×10^{-4}	6.6×10^{3}	4.3×10^{7}

fusion protein was purified to homogeneity with 97% yield by affinity purification using Ni^{2+}-charged resin (Table 3A) and then relieved of the fused peptide by thrombin cleavage (Fig. 4). The purified human CACH represented comparable specific activity (3300 U/mg), Michaelis constant for acetyl-CoA ($K_m = 1.5 \times 10^{-4}$ M), catalytic constant ($k_{cat} = 7.0 \times 10^3$/s) and catalytic efficiency ($k_{cat}/K_m = 4.6 \times 10^7$ /M·s) to those of the rat enzyme (Table 3B).

Highly conserved genomic organization of mammalian CACHs

We here describe a comparative analysis of the genomic organization of human, mouse and rat CACHs. It was verified by mRNA-genomic alignments and published in established databases that the CACH gene is mapped at human chromosome 5q14.1, mouse chromosome 13C3 and rat chromosome 2q12 and covers about 64-kb, 45-kb and 50-kb genomic regions, respectively (Fig. 5). The mammalian CACH genes are similarly composed of 15 exons, the size of each of which is identical among the three mammalian species with an exception that exon I codes for 43 amino-acid residues in human and 44 residues in the rodents (Fig. 3). Introns are generally longer in human than in the rodents with the exceptions of introns I, IX, and X (Fig. 5).

Putative domain structure of mammalian CACH proteins

Bioinformatical work using the NCBI Conserved Domain Search website revealed that human, mouse and rat CACH proteins commonly consist of uniquely aligned functional domains as follows: two acyl-CoA hydrolase (thioesterase) regions (accession: COG1607 (Dillon & Bateman, 2004)) encoded by exons I–V and VI–IX, respectively, and a C-terminal Steroidgenic acute regulatory (StAR) related lipid transfer (START) domain (accession: pfam01852; (Tsujishita & Hurley, 2000)) encoded by exons X–XV (Fig. 3). Because the rest of the domains might be too small, ranging in length from only 10 to 23 amino-acid residues, to contribute to the enzymatic function, all or some of the three functional domains presumably contribute to the CACH activity. As shown in Fig. 3, 4-hydroxybenzoyl-CoA thioesterase (4HBT) domains (accession: pfam03061) I, II, encoded by exons I–IV and VI–VIII, respectively, were also demarcated in the midst of the thioesterase regions I and II, respectively. The 4HBT domains I and II show especially high amino-acid sequence homology between human and mouse/rat (I: 87.5%/86.1% identity; II: 93.2%/93.2% identity). On the other hand, the C-terminal START domain shows a relatively low amino-acid sequence homology between human and mouse/rat (77.1%/74.6% identity).

Characterization of the putative promoter region of mammalian CACH genes

The putative promoter regions of human, mouse and rodent CACH genes are unexpectedly found to be of the housekeeping-type, namely GC-rich with GC box(es), but lacking both TATA and CCAAT boxes, as described below. The 5′-flanking nucleotide sequences 100 bp long show high GC contents (human, 83%; mouse, 74%; rat, 70%). A GC box (GGGCGG) is commonly found in the upstream regions of the three mammalian species, especially in human, where six tandemly arranged GC boxes are found in the region –107/–54 (Fig. 6). On the other hand, although we searched the 5′-

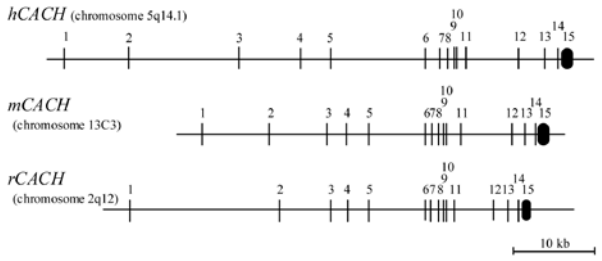

Figure 5. Comparison of exon-intron structure of cytosolic acetyl-CoA hydrolase genes among human and the rodents.

Exons are numbered from 1 to 15. The data here taken from publicly available RefSeq genes as follows: *hCACH*, human *cytosolic acetyl-CoA hydrolase* (GenBank accession NM_130767); *mCACH*, mouse *cytosolic acetyl-CoA hydrolase* (GenBank accession NM_028790); *rCACH*, rat *cytosolic acetyl-CoA hydrolase* (GenBank accession NM_130747).

Figure 6. Comparison of the 5′-upstream region of cytosolic acetyl-CoA hydrolase genes among human and the rodents.
The nucleotide residues are numbered on the right, relative to the adenine residue of the AUG translational start triplet marked as +1. Asterisks denote initiation sites of transcripts (*3, AK034622; *4, AK004905, AK033663; *6, AB040609) or ESTs *1, AV693695, AV693776; *2, BB843576; *4, BB626692, BY705110; *5, BB625271) from GenBank. Consensus binding sites are indicated with bold letters: 'GC boxes' **gggcgg** for Sp1; 'E box' **CACCTG** for bHLH is underlined; **cggggt/tggggt** for Adr1p; **CCTGC** for LVc (leukemia virus factor c); PPRE **TGACCTcTCACCTccta** is shaded. The compared sequences are described in the legend to Fig. 5.

flanking region of the CACH genes up to the position –1200, neither a canonical core promoter motif TATA box (TATAAA) nor a CCAAT box was found in mouse and rat. These *cis*-elements were also absent in the proximal promoter region of the human gene, whereas in a farther upstream region, a sequence TTTATA at –824/–819, corresponding to the inverted sequence of TATAAA, a typical CCAAT box at –394/–390, and its inverted sequence ATTGG at –767/–763, were found.

A peroxisome proliferator responsive element (PPRE) motif (TGACCTcTGACCT) (Kliewer *et al*., 1992) was, as expected, found in the rodent CACH promoter regions (Fig. 6), consistent with the marked CACH induction by peroxisome proliferator (PP) observed in rat liver (Prass *et al*., 1980). In human the PPRE was found not in the putative promoter region but in introns II and III and so forth of the CACH gene (Table 4). Further, a computer analysis using TESS revealed that possible well-characterized *cis*-elements are commonly found in the upstream regions of the mammalian CACH genes (Fig. 6). These features observed are discussed later in relation to the CACH expression.

DISCUSSION

A human cytosolic acetyl-CoA hydrolase (CACH) cDNA was cloned from a liver cDNA library and characterized for the first time. Cytoplasmic localization of the human CACH was predicted by PSORT WWW Server from the cDNA nucleotide sequence not containing any targeting signals for organella including peroxisome, mitochondria or nucleus, as well as the rodent enzymes (Suematsu *et al*., 2001; 2002). The extensive homology of its deduced amino-acid sequence with the rodent enzymes (Table 2) implies that the obtained human homologue cDNA encodes no other acyl-CoA thioesterase (Broustas *et al*., 1996; Hunt *et al*., 1999) but CACH. The authenticity of the human cDNA was further confirmed by its functional expression. The expressed recombinant form of human CACH was found to be a cold labile allosteric enzyme activated by ATP and completely inhibited by ADP, and shared comparable enzymatic characteristics with the rat enzyme (Table 3B). Thus, our previous and current studies clearly demonstrate that the mammalian CACH proteins have been highly conserved.

Here we proposed putative functional domains of CACH (Fig. 3). Among them, the highly conserved 4HBT thioesterase domains (Benning *et al*., 1998) presumably contribute greatly to the enzymatic activity of active CACH homooligomers. Although the actual domain functions remain to be established, the highly conserved unique alignment of the putative functional domains (Fig. 3) together with the overall high amino-acid sequence similar-

Table 4. Putative peroxisome proliferator responsive elements (PPRE) in human, mouse and rat cytosolic acetyl-CoA hydrolase genes.

The nucleotide sequences are described in the legend to Fig. 5. PPRE was searched using TESS. Nucleotide positions are numbered as in Fig. 2. Boldface letters represent highly conserved sequences in agreement with the consensus sequences for PPRE (Kliewer *et al*., 1992).

Gene	Sequence (5′-3′)	Position in gene	Location
hCACH	**g**gta**AGGTCAtAG**A**TCA**	–39,165/–39,149	5′-upstream region
	ag**gt**T**TGTCAaAGGTCA**	15,930/15,946	intron II
	agt**cAGGTCAcAGGT**A**A**	22,527/22,543	intron III
mCACH	**TGACCAtTGTCCTgc**at	–266/–249	5′-flanking region
	TGACCTcTCACCTcct**a**	–101/–85	5′-flanking region
rCACH	**TGACCTcTCACCTcc**t**a**	–105/–89	5′-flanking region
	attcAGGTCAaAGT**TCA**	33,614/33,630	intron V

ity among the three mammalian species (Table 2) implies a crucial role of CACH in mammalian metabolism.

In the mammalian CACH genes, all the introns are found to be bordered by consensus GT-AG splice sites, which are usually excised by the major spliceosome utilizing U2 snRNA (Sharp & Burge, 1997), and further their splicing donor and acceptor sites share common pyrimidine stretches (Breathnach *et al*., 1978; Breathnach & Chambon, 1981). In the human CACH gene, as shown in Table 5, exons range in length from 56 bp (exon IX) to 157 bp (exon VI), while the introns from 270 bp (intron XIV) to 13946 kb (intron II). As commonly found in the mammalian CACH genes, introns I–V are generally larger, whereas introns VI–X are distinctly compact (Fig. 5), implying separate origins of the two thioesterase domains, composed of exons I–V and VI–IX, respectively, and their conjugation by domain shuffling during evolution of the genomes. This interesting inference is further supported by the finding that intron V located at the domain border is of phase I, which is classified according to the number of bases of the final codon generated in the previous exon (Cawley *et al*., 2001). It has been proposed that domain shuffling within phase I introns played an important role in the evolution of the human genome (Kaessmann *et al*., 2002).

The three major core promoter motifs containing binding sites for RNA polymerase II have been reported to be the TATA box, the initiator and the downstream promoter element (Burke & Kadonaga, 1997). In the present paper, we analyzed mammalian CACH putative promoter regions for the first time. To estimate the transcription start sites of CACH, each 5′-end of mammalian CACH transcripts and EST entries extending into the 5′-direction available from the GenBank was marked on the 5′-upstream region of the mammalian CACH genes (Fig. 6). Unexpectedly, we found that none of them matched the consensus sequence of mammalian initiator Py-Py-A_{+1}-N-T/A-Py-Py, whose A_{+1} is the transcription start site (Burke & Kadonaga, 1997). We here report that the CACH putative promoters do not have any of the three known core promoter motifs, suggesting the presence of other core promoter motifs responsible for accurate positioning of RNA polymerase relative to the initiation site.

It should be noted that a number of promoters lack classical TATA or CCAAT boxes but have an increased GC content with specificity protein 1 (Sp 1) binding site(s), features typical of 'housekeeping' genes (La Thangue & Rigby, 1988; Rundlöf *et al*., 2001). The ubiquitously expressed TATA-less genes have been reported to be controlled by Sp proteins (Dynan *et al*., 1986; Suske, 1999; Müller *et al*., 1999), which have been shown to bind the GC box, the hexanucleotide sequence GGGCGG (Lania *et al*., 1997). Interestingly, the mammalian CACHs have a typical housekeeping-type promoter (Fig. 6), which presumably stimulates constitutive transcription of CACH, implying that the enzyme might be generally essential for mammalian cells. The ubiquitously expressed Sp1 was reported as a member of a transcriptionally active multiprotein–DNA complex (Foti *et al*., 2003), and also as a repressor by recruiting histone deacetylase 1 (HDAC1) to the growth-regulated murine thymidine kinase gene (*TK*) promoter (Doetzlhofer *et al*., 1999). Thus, it is likely that Sp1 acts in a combinatorial manner with other transcription factors, which may have more pronounced temporally or spatially restricted expression patterns. It has been also shown that the Sp1-related BTEB (basic transcription element binding protein) protein is a repressor of a promoter containing a single GC box, however, when the GC box was repeated five times, BTEB turned out to be an activator (Lania *et al*., 1997). Thus, Sp1-related proteins like BTEB could possibly influence the CACH transcription differently between the rodents and human, since the former contains a single GC box and the latter contains six tandemly arranged GC boxes in their putative promoter regions of CACH. How the Sp proteins actually control the CACH gene expression remains to be established.

Although the mammalian CACHs have a typical housekeeping-type promoter, the enzyme activity is detectable only in liver and kidney of the rat tissues tested (Matsunaga *et al*., 1985). The tissue-specific expression must arise from *cis*-acting elements other than a GC box. Web-based search using TESS revealed several other possible *cis*-elements commonly found in the mammalian CACH putative promoter regions (Fig. 6) as described below. First, consensus (C/T)GGGGT sequence, known as a binding site of yeast Adr1p, locates at –42/–37 (human), –22/–17 (mouse) and –145/–140 (rat). Adr1p has been reported to govern fatty acid degradation and peroxisome proliferation (Gurvitz *et al*., 2001). Second, consensus CCTGC sequence, known as a binding site of LVc (leukemia virus factor c), lies at –26/–22 (human), –37/–33 (mouse) and –43/–39, –92/–88 (rat). LVc interacts with leukemia virus enhancer (Speck & Baltimore, 1987). Third, an E box motif (consensus CANNTG) exists at –134/–129 (human), –101/–96 (mouse) and –106/–101 (rat). It is usually found in promoter or enhancer regions and is known as a binding site of ubiquitous basic helix-loop-helix (bHLH) transcription factors, which induce the expression of a number of genes in the appropriate cell type, resulting in a tissue-specific phenotype (Murre *et al*., 1994). The bHLH proteins can be categorized into three classes: Class A and B function as transcriptional activators, while Class C as transcrip-

tional repressors (Azmi *et al*., 2003). Since E box sites appear in a wide variety of promoters and enhancer regions (Murre *et al*., 1994), other flanking factors would specify their functions. In the rodent CACH promoter region, an E box is located close to a PPRE element, and in rat it is additionally found directly adjacent to a GC box (Fig. 6), suggesting possible synergistic transactivation of the CACH gene by their ligands. A PPRE element found in the rodent CACHs putative promoter region (Fig. 6) must be responsible for the marked enzyme induction by a peroxisome proliferator (PP) (Prass *et al*., 1980). On the other hand, the human PPRE elements found outside of the putative promoter region (Table 4) remain to be determined whether active or not. A significant number of studies have suggested that human hepatocytes are non-responsive to PPs — this includes epidemiological studies using hypolipidemic drugs, and *in vitro* experiments with human hepatocytes (Tugwood *et al*., 1996). To address the question whether hCACH is induced by PPs, further *in vitro* experiments with human hepatocytes should be carried out. Precise demarcation of the promoter and identification of all the elements involved in the gene transcription awaits further experiments including *in vivo* transfection.

In conclusion, this paper for the first time describes the entire cDNA sequence of human cytosolic acetyl-CoA hydrolase and its overproduction, allowing future studies on its physiological functions and physicochemical characteristics. Although a detailed characterization of its promoter remains to be established, the bioinformatical promoter analysis presented here has provided insights into the regulation of the gene expression. Transgenic approach is currently underway to explore physiological roles of the enzyme. The *in vivo* study would provide direct evidence on the enzyme functions and facilitate conclusive establishment of its significance.

Acknowledgements

This research was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan and by a grant from the Vitamin Society of Japan.

REFERENCES

Azmi S, Sun H, Ozog A, Taneja R (2003) mSharp-1/DEC2, a basic helix-loop-helix protein functions as a transcriptional repressor of E box activity and *Stra13* expression. *J Biol Chem* **278**: 20098–20109.

Benning MM, Wesenberg G, Liu R, Taylor KL, Dunaway-Mariano D, Holden HM (1998) The three-dimensional structure of 4-hydroxybenzoyl-CoA thioesterase from *Pseudomonas* sp. strain CBS-3. *J Biol Chem* **273**: 33572–33579.

Berge RK, Aarsland A, Bakke OM, Farstad M (1983) Hepatic enzymes, CoASH and long-chain acyl-CoA in subcellular fractions as affected by drugs inducing peroxisomes and smooth endoplasmic reticulum. *Int J Biochem* **15**: 191–204.

Breathnach R, Chambon P (1981) Organization and expression of eucaryotic split genes coding for proteins. *Annu Rev Biochem* **50**: 349–383.

Breathnach R, Benoist C, O'Hare K, Gannon F, Chambon P (1978) Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Proc Natl Acad Sci USA* **75**: 4853–4857.

Broustas CG, Larkins LK, Uhler MD, Hajra AK (1996) Molecular cloning and expression of cDNA encoding rat brain cytosolic acyl-coenzyme A thioester hydrolase. *J Biol Chem* **271**: 10470–10476.

Burke TW, Kadonaga JT (1997) The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by $TAF_{II}60$ of *Drosophila*. *Genes Dev* **11**: 3020–3031.

Cawley SE, Wirth AI, Speed TP (2001) Phat—a gene finding program for *Plasmodium falciparum*. *Mol Biochem Parasitol* **118**: 167–174.

Dillon SC, Bateman A (2004) The Hotdog fold: wrapping up a superfamily of thioesterases and dehydratases. *BMC Bioinformatics* **5**: 109–122.

Doetzlhofer A, Rotheneder H, Lagger G, Koranda M, Kurtev V, Brosch G, Wintersberger E, Seiser C (1999) Histone deacetylase 1 can repress transcription by binding to Sp1. *Mol Cell Biol* **19**: 5504–5511.

Dynan WS, Sazer S, Tjian R, Schimke RT (1986) Transcription factor Sp1 recognizes a DNA sequence in the mouse dihydrofolate reductase promoter. *Nature* **319**: 246–248.

Ebisuno S, Isohashi F, Nakanishi Y, Sakamoto Y (1988) Acetyl-CoA hydrolase: relation between activity and cholesterol metabolism in rat. *Am J Physiol* **255**: R724–R730.

Ebisuno S, Isohashi F, Nakanishi Y, Higashi T, Sakamoto Y (1989) The biphasic change of cytosolic acetyl-CoA hydrolase in rat liver during 3'-methyl-4-dimethylaminoazobenzene hepatocarcinogenesis. *Jpn J Cancer Res* **80**: 132–135.

Foti D, Iuliano R, Chiefari E, Brunetti A (2003) A nucleoprotein complex containing Sp1, C/EBPβ, and HMGI-Y controls human insulin receptor gene transcription. *Mol Cell Biol* **23**: 2720–2732.

Gurvitz A, Hiltunen JK, Erdmann R, Hamilton B, Hartig A, Ruis H, Rottensteiner H (2001) *Saccharomyces cerevisiae* Adr1p governs fatty acid β-oxidation and peroxisome proliferation by regulating *POX1* and *PEX11*. *J Biol Chem* **276**: 31825–31830.

Horie S, Isobe M, Suga T (1986) Changes in CoA pools in hepatic peroxisomes of the rat under various conditions. *J Biochem* **99**: 1345–1352.

Hunt MC, Nousiainen SEB, Huttunen MK, Orii KE, Svensson LT, Alexson SEH (1999) Peroxisome proliferator-induced long chain acyl-CoA thioesterases comprise a highly conserved novel multi-gene family involved in lipid metabolism. *J Biol Chem* **274**: 34317–34326.

Isohashi F, Nakanishi Y, Sakamoto Y (1983a) Factors affecting the cold inactivation of an acetyl-Coenzyme-A hydrolase purified from the supernatant fraction of rat liver. *Eur J Biochem* **134**: 447–452.

Isohashi F, Nakanishi Y, Sakamoto Y (1983b) Effects of nucleotides on a cold labile acetyl-CoA hydrolase from the supernatant fraction of rat liver. *Biochemistry* **22**: 584–590.

Isohashi F, Nakanishi Y, Matsunaga T, Sakamoto Y (1984) A cold-labile acetyl-coenzyme-A hydrolase from the supernatant fraction of rat liver: reactivation and reconstitution of the active species from the inactive monomer. *Eur J Biochem* **142**: 177–181.

Kaessmann H, Zöllner S, Nekrutenko A, Li W-H (2002) Signatures of domain shuffling in the human genome. *Genome Res* **12**: 1642–1650.

Kliewer SA, Umesono K, Heyman RA, Mangelsdorf DJ, Dyck JA, Evans RM (1992) Retinoid X receptor–COUP-TF interactions modulate retinoic acid signaling. *Proc Natl Acad Sci USA* **89**: 1448–1452.

Kozak M (1987) At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J Mol Biol* **196**: 947–950.

Kozak M (1991) Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J Biol Chem* **266**: 19867–19870.

La Thangue NB, Rigby PWJ (1988) Trans-acting protein factors and the regulation of eucaryotic transcription. In *Transcription and Splicing* (Hames BD, Glover DM, eds) pp 1–42, IRL Press, Washington, DC.

Lania L, Majello B, De Luca P (1997) Transcriptional regulation by the Sp family proteins. *Int J Biochem Cell Biol* **29**: 1313–1323.

Mannaerts GP, Debeer LJ, Thomas J, De Schepper PJ (1979) Mitochondrial and peroxisomal fatty acid oxidation in liver homogenates and isolated hepatocytes from control and clofibrate-treated rats. *J Biol Chem* **254**: 4585–4595.

Matsunaga T, Isohashi F, Nakanishi Y, Sakamoto Y (1985) Physiological changes in the activities of extramitochondrial acetyl-CoA hydrolase in the liver of rats under various metabolic conditions. *Eur J Biochem* **152**: 331–336.

Müller C, Yang R, Beck-von-Peccoz L, Idos G, Verbeek W, Koeffler HP (1999) Cloning of the *cyclin A1* genomic structure and characterization of the promoter region. *J Biol Chem* **274**: 11220–11228.

Murre C, Bain G, van Dijk MA, Engel I, Furnari BA, Massari ME, Matthews JR, Quong MW, Rivera RR, Stuiver MH (1994) Structure and function of helix-loop-helix proteins. *Biochim Biophys Acta* **1218**: 129–135.

Nakanishi Y, Isohashi F, Ebisuno S, Sakamoto Y (1988) Binding of nucleotides to an extramitochondrial acetyl-CoA hydrolase from rat liver. *Biochemistry* **27**: 4822–4826.

Nakanishi Y, Okamoto K, Isohashi F (1993) Effects of chronic administration of the peroxisome proliferator, clofibrate, on cytosolic acetyl-CoA hydrolase in rat liver. *Biochem Pharmacol* **45**: 1403–1407.

Nakanishi Y, Okamoto K, Isohashi F (1994) Subcellular distribution of ATP-stimulated and ADP-inhibited acetyl-CoA hydrolase in livers from control and clofibrate-treated rats: comparison of the cytosolic and peroxisomal enzyme. *J Biochem* **115**: 328–332.

Prass RL, Isohashi F, Utter MF (1980) Purification and characterization of an extramitochondrial acetyl coenzyme A hydrolase from rat liver. *J Biol Chem* **255**: 5215–5223.

Rundlöf A-K, Carlsten M, Arnér ESJ (2001) The core promoter of human thioredoxin reductase 1. *J Biol Chem* **276**: 30542–30551.

Schug J, Overton GC (1997) *TESS: Transcription Element Search Software on the WWW*. Website (http://www.cbil.upenn.edu/tess/) Computational Biology and Informatics Laboratory School of Medicine University of Pennsylvania. Technical Report CBIL-TR-1997-1001-v0.0.

Sharp PA, Burge CB (1997) Classification of introns: U2-type or U12-type. *Cell* **91**: 875–879.

Simon EJ, Shemin D (1953) The preparation of *S*-succinyl Coenzyme A. *J Am Chem Soc* **75**: 2520.

Söling H-D, Rescher C (1985) On the regulation of cold-labile cytosolic and of mitochondrial acetyl-CoA hydrolase in rat liver. *Eur J Biochem* **147**: 111–117.

Speck NA, Baltimore D (1987) Six distinct nuclear factors interact with the 75-base-pair repeat of the Moloney murine leukemia virus enhancer. *Mol Cell Biol* **7**: 1101–1110.

Suematsu N, Okamoto K, Isohashi F (1996) Effects of various proteins or peptides on reactivation of cold-inactivated acetyl-CoA hydrolase from rat liver. *St Marianna Med J* **24**: 691–697.

Suematsu N, Okamoto K, Shibata K, Nakanishi Y, Isohashi F (2001) Molecular cloning and functional expression of rat liver cytosolic acetyl-CoA hydrolase. *Eur J Biochem* **268**: 2700–2709.

Suematsu N, Okamoto K, Isohashi F (2002) Mouse cytosolic acetyl-CoA hydrolase, a novel candidate for a key enzyme involved in fat metabolism: cDNA cloning, sequencing and functional expression. *Acta Biochim Polon* **49**: 937–945.

Suematsu N, Okamoto K, Isohashi F (2003) Simple and unique purification by size-exclusion chromatography for an oligomeric enzyme, rat liver cytosolic acetyl-coenzyme A hydrolase. *J Chromatogr B, Analyt Technol Biomed Life Sci* **790**: 239–244.

Suske G (1999) The Sp-family of transcription factors. *Gene* **238**: 291–300.

Tsujishita Y, Hurley JH (2000) Structure and lipid transport mechanism of a StAR-related domain. *Nat Struct Biol* **7**: 408–414.

Tugwood JD, Aldridge TC, Lambe KG, Macdonald N, Woodyatt NJ (1996) Peroxisome proliferator-activated receptors: structures and function. *Ann N Y Acad Sci* **804**: 252–265.