Activation by thioesterification to coenzyme A is a prerequisite for most reactions involving fatty acids. Enzymes catalyzing activation, acyl-CoA synthetases, have been classified by their chain length specificities. The most recently identified family is the very long-chain acyl-CoA synthetases (VLCS). Although several members of this group are capable of activating very long-chain fatty acids (VLCFAs), one is a bile acid-CoA synthetase, and others have been characterized as fatty acid transport proteins. It was reported that the Drosophila melanogaster mutant bubblegum (BGM) had elevated VLCFAs and that the product of the defective gene had sequence homology to acyl-CoA synthetases. Therefore, we cloned full-length cDNA for a human homolog of BGM, and we investigated the properties of its protein product, hsBG, to determine whether it had VLCS activity. Northern blot analysis showed that hsBG is expressed primarily in brain. Compared with vector-transfected cells, COS-1 cells expressing hsBG had increased acyl-CoA synthetase activity with either long-chain fatty acid (2.4-fold) or VLCFA (2.6-fold) substrates. Despite this increased VLCFA activation, hsBG-expressing cells did not have increased rates of VLCA degradation. Confocal microscopy showed that hsBG had a cytoplasmic localization in some COS-1 cells expressing the protein, whereas it appeared to associate with plasma membrane in others. Fractionation of these cells revealed that most of the hsBG-dependent acyl-CoA synthetase activity was soluble and not membrane-bound. Immunoaffinity-purified hsBG from transfected COS-1 cells was enzymatically active. hsBG and VLCS are only 15% identical, and comparison with sequences of two conserved motifs from all known families of acyl-CoA synthetases revealed that hsBG along with the D. melanogaster and murine homologs comprise a new family of acyl-CoA synthetases. Thus, two protein families are now known that contain enzymes capable of activating VLCFAs. Because hsBG is expressed in brain but previously described VLCSs were not highly expressed in this organ, hsBG may play a central role in brain VLCFA metabolism and myelinogenesis.

Free fatty acids must be “activated” to their CoA thioesters before participating in most catabolic and anabolic reactions (1). Processes such as incorporation of fatty acids into phospholipids or triacylglycerols, fatty acid elongation or unsaturation, fatty acylation of proteins, and degradation of fatty acids by α-, β-, or ω-oxidation all require activated fatty acid substrates. Activation is catalyzed by acyl-CoA synthetases, also known as acyl-CoA ligases (AMP-forming) (EC 6.3.1.x). The acyl-CoA synthetases have been classified by their specificities for fatty acid substrates of different chain lengths. Biochemically, the most well characterized acyl-CoA synthetases are those that activate 16- and 18-carbon long-chain fatty acids (LCFAs),1 compounds that are abundant both endogenously and in the normal human diet. However, synthetases that activate short- (C2–C4), medium- (C6–C10), and very long-chain fatty acids (VLCFAs, ≥C22) have also been described.

Over the last decade, both biochemical and molecular approaches have revealed that more than one enzyme may exist in each chain length specificity class. For example, five rat long-chain acyl-CoA synthetases (LCS) have been purified and/or cDNA-cloned (2–6). We found that the family of human proteins that contains very long-chain acyl-CoA synthetase (VLCS) includes six members (7). A unifying feature of all acyl-CoA synthetases is the presence of an “AMP-binding domain signature” (PROSITE PS00455). This highly conserved motif consists of 10–12 amino acids and was first described nearly a decade ago (8). Black et al. (9) proposed that a second highly conserved region present in the LCS family was a “signature motif” for this group of enzymes. We reported that a related motif containing 22 residues is characteristic of the VLCS family (7, 10). These observations suggest that other families of acyl-CoA synthetases might be distinguished by the presence of similar domains.

Our understanding of the physiologic role(s) of acyl-CoA synthetases has been facilitated by gene disruption and mutational studies in non-mammalian species (9–12). The Drosophila melanogaster mutant bubblegum is characterized by neurodegeneration and elevated levels of very long-chain fatty acids (12). These findings are similar to those observed in the human disease X-linked adrenoleukodystrophy, in which there also is reduced peroxisomal VLCS activity (13). The gene defective in the bubblegum mutant encodes a protein whose amino acid sequence had similarities to sequences of acyl-CoA synthetases. However, the Drosophila bubblegum protein

1 The abbreviations used are: LCFA, long-chain fatty acids; VLCFAs, very long-chain fatty acids; LCS, long-chain acyl-CoA synthetases; VLCS, very long-chain acyl-CoA synthetases; PCR, polymerase chain reaction; bp, base pair; kb, kilobase pair; 5′-RACE, rapid amplification of cDNA ends.

* This work was supported by National Institutes of Health Grants NS10533, NS 37355, and HD10981. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Kennedy Krieger Institute, 707 N. Broadway, Baltimore, MD 21205. Tel.: 410-502-9493; Fax: 410-502-8279; E-mail: watkins@kennedykrieger.org.

‡ From the Kennedy Krieger Institute and Departments of §Neurology and ¶Pediatrics and the ¶Institute of Genetic Medicine, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205.
Very Long-chain Acyl-CoA Synthetase

35163

(dmBG) lacked both the motif characteristic of the VLCS family and peroxisome targeting signals. Therefore, we cloned cDNA for the human ortholog of bubblegum (hsBG). We investigated the properties of the human bubblegum protein (hsBG). We report here that hsBG is a VLCS that represents a previously undescribed acyl-CoA synthetase family.

EXPERIMENTAL PROCEDURES

Materials and General Methods—Expressed sequence tag clones were obtained from the American Type Culture Collection (ATCC) and are identified by their ATCC number. COS-1 cells were a gift from Catherine Thompson (Kennedy Krieger Institute), 1-14C]lignoceric acid was synthesized from tricosanol and K14CN (American Radiolabeled Chemicals). [1-14C]Ligadore acid was synthesized from tricosanol and K14CN (American Radiolabeled Chemicals) by the method of Muralidharan and Kimishita (14). DNA sequencing was performed at The Johns Hopkins University Department of Biological Chemistry Bioinformatics and Sequencing Facility using the fluorescent dye terminator method of cycle sequencing on an Applied Biosystems Inc. 377 automated DNA sequencer, following Applied Biosystems protocols. Nucleotide and amino acid sequence alignment, and calculation of sequence identity and similarity, was done using the ClustalW program (15). PCR conditions were as described previously (16). Protein was determined by the method of Lowry et al. (17).

Cloning of hsBG Full-length cDNA—Homology probing of the GenBankTM nonredundant protein data base identified the 634 amino acid "KIAA0631 protein" (accession number AB014531) as a candidate human homolog of dmBG. Amino acid sequence alignment with that of dmBG demonstrated that the KIA0631 sequence was incomplete at the amino terminus. We devised a two-step strategy for obtaining full-length hsBG cDNA. First, the KIA0631 sequence was obtained by PCR amplification; second, 5'-RACE was used to obtain the remaining 5'-sequence. Two overlapping fragments of hsBG were amplified by PCR from a human brain cDNA library (CLONTECH) based on the GenBankTM KIAA0631 sequence. PCR using oligonucleotide primers P12-1 (-5'TCGGAGGCCTCCCTTACTCATGC-3', which incorporates an XhoI site) plus P12-2 (-5'GGAAATGGCAGGCAGTGCAG-3') for fragment 1 and P12-3 (-5'GGTTGGGAGCACATCACCA-3') plus P12-4 (-5'TCTTACACCGGACGCCTGCCCTTATT-3') for fragment 2 yielded amplicons of the expected size (891 and 1126 bp, respectively). The region of overlap contained a Sau3A restriction site. Fragments 1 and 2 were cloned into a TA cloning vector (pGEM-T Easy, Promega). A ligation mixture containing restriction enzyme-restricted Sau3A fragments 1 and 2 (Smal/XhoI), and N-pec-pcDNA3 (Xhol/XbaI), this vector is a modified pcDNA3 (Invitrogen) vector that places an in-frame c-Myc epitope at the amino terminus; gift of Stephen Gould, Johns Hopkins University) was homogenized using a precision ball-bearing homogenizer (18) and a postnuclear supernatant was prepared by centrifugation at 1500 g for 30 s. A fraction enriched in mitochondria and peroxisomes (ML fraction) was prepared by centrifugation of the postnuclear supernatant for 15 min at 18,000 g in a Beckman JA-20 high speed centrifuge. The organelle-free supernatant was then centrifuged at 435,000 g for 15 min in a Beckman Optima TLX ultracentrifuge (TLA 120.2 rotor) to separate the cytoplasm (supernatant) from mitochondria and peroxisomes (pellet). To prepare a true cytosolic fraction, transfected cells were homogenized using a precision ball-bearing homogenizer (18) and a postnuclear supernatant was prepared by centrifugation at 1500 g for 30 min at 4 °C. A fraction enriched in mitochondria and peroxisomes (ML fraction) was prepared by centrifugation of the postnuclear supernatant for 15 min at 18,000 g in a Beckman JA-20 high speed centrifuge. The organelle-free supernatant was then centrifuged at 435,000 g for 15 min in a Beckman Optima TLX ultracentrifuge (TLA 120.2 rotor) to separate microsomes and cytosol. Assay of activation of 1-14C-labeled fatty acids was performed as described previously (19). Fatty acid specific activity was measured in freshly isolated COS-1 cells as described previously (20) 3 days post-transfection.

Immunofluorescence—Following transfection with N-myc hsBG, COS-1 cells were plated on glass coverslips. After 3 days, cells were fixed, permeabilized with Triton X-100, and incubated first with mouse anti-Myc (gift of Gerald Johnson, San Diego State University) and then with donkey anti-mouse IgG fluorescein isothiocyanate conjugate (Jackson ImmunoResearch) as described previously (21). For double labeling, cells were transfected with both N-myc hsBG and human adrenoleukodystrophy protein (in pcDNA). After fluorescein labeling of hsBG as above, coverslips were incubated with anti-adenosine 3',5'-cyclic monophosphate (cAMP) antibody (21) and then with donkey anti-rabbit IgG isotype rhodamine conjugate (Jackson ImmunoResearch). Fluorescence was visualized using a Leitz Aristoplan fluorescent microscope. Confocal microscopy was performed at The Johns Hopkins University School of Medicine Microscope Facility. Images were acquired with a Nikon Oz confocal laser scanning system using Intervision (version 6.5) software on a Silicon Graphics O2 platform.

RESULTS

Cloning and Characteristics of hsBG cDNA—Comparison of the amino acid sequence of dmBG to human proteins in the GenBankTM nonredundant protein data base identified only one protein with significant homology. The "KIAA0631 protein" (an incomplete open reading frame consisting of 634 amino acids) and dmBG (666 amino acids) were 41% identical and 64% similar. Sequence alignment using the ClustalW program revealed that the amino-terminal region of KIAA0631 was incomplete. By using PCR, cdNA encoding the KIAA0631 protein
was amplified from a human brain cDNA library, and additional 5'-sequence was obtained using a RACE protocol as described under “Experimental Procedures.” The complete nucleotide sequence and predicted amino acid sequence of hsBGM cDNA are shown in Fig. 1. Two upstream stop codons (indicated by *) were found in the 5'-untranslated region at bp -30 and -75 relative to the putative start codon. An additional stop codon was also found in the 3'-untranslated region at bp 2196. Motif 1 (AMP-binding domain; amino acids 281–290) and motif 2 (amino acids 561–585) are shown in bold.

![Fig. 1. Nucleotide sequence and deduced amino acid sequence of hsBGM.](image)

Full-length cDNA for hsBGM was cloned into the pcDNA3 vector as described under “Experimental Procedures” and was sequenced in its entirety. Two upstream stop codons (indicated by *) were found in the 5'-untranslated region at bp -30 and -75. Although this clone contains only 48 bp of the 3'-untranslated region, the 2175-bp open reading frame is compatible with this size transcript. The high level expression in brain is consistent with the observation that mutation of this protein in Drosophila results in neurodegeneration (12). In addition to the 3.0-kb transcript, a larger but less abundant transcript of 4.4 kb was seen in brain. An even longer transcript is also seen in brain, but its sequence has not been determined. The predicted amino acid sequence of hsBG contains 724 residues. Although this sequence is longer than that of dmBG (647 amino acids), we noted that the hsBG lacked a methionine residue at or near the proposed initiator methionine position of dmBG (12). A molecular mass of 81,290 daltons for hsBG was calculated from the predicted amino acid sequence. The pI of hsBG was calculated to be 5.73. Amino acids 281–290 (YTSGTTGNPK) comprise the putative AMP-binding domain characteristic of acyl-CoA synthetases and related enzymes (8).

Tissue Expression of hsBG—To determine which tissues normally express hsBG, a human multiple tissue Northern blot was probed with a radiolabeled fragment of hsBGM cDNA as described under “Experimental Procedures.” A strong signal corresponding to a 3.0-kb mRNA species was detected in brain (Fig. 2). Whereas the exact sizes of the 3'- and 5'-untranslated regions of hsBGM are not known, the 2175-bp open reading frame is compatible with this size transcript. The high level expression in brain is consistent with the observation that mutation of this protein in Drosophila results in neurodegeneration (12). In addition to the 3.0-kb transcript, a larger but less abundant transcript of 4.4 kb was seen in brain. An even longer transcript is also seen in brain, but its sequence has not been determined.
larger transcript of 6.5 kb was observed in heart and skeletal muscle. The significance of these larger transcripts is currently under investigation.

**Acyl-CoA Synthetase Activity of COS-1 Cells Expressing hsBG**—The presence of an AMP-binding domain and the homology of hsBG and dmBG to other acyl-CoA synthetases suggested that these proteins were also acyl-CoA synthetases. To investigate this, we expressed the full-length cDNA encoding hsBG in COS-1 cells. Three days post-transfection, cell suspensions were assayed for their ability to activate fatty acids to their CoA derivatives. As shown in Table I, cells transiently expressing hsBG activated both palmitic acid (C16:0), a representative LCFA, and lignoceric acid (C24:0), a representative VLCFA at a greater rate than did vector-transfected cells. Activation of the LCFA was increased 2.4-fold, and activation of the VLCFA 2.6-fold compared with vector-transfected cells. Activities of non-transfected cells and vector-transfected cells were similar (data not shown). No differences in the acyl-CoA synthetase activity of cells expressing hsBG or N-myc-hsBG were observed (data not shown).

**Effect of hsBG Overexpression on Fatty Acid β-Oxidation**—The observation of increased concentrations of VLCFA in the *Drosophila* mutant *bubblegum* and the ability of hsBG expressed in COS-1 cells to activate VLCFA suggested that this enzyme may be relevant to cellular VLCFA homeostasis. Excess VLCFA are normally degraded by the peroxisomal β-oxidation pathway (22). Therefore, we investigated the possibility that overexpression of hsBG in COS-1 cells might increase the rate of degradation of lignoceric acid, a 24-carbon VLCFA. Despite the increased rate of activation of VLCFA in total cellular extracts of cells overexpressing hsBG (Table I), no increase in the rate of VLCFA β-oxidation was found. The potential role(s) of hsBG in other facets of VLCFA metabolism are currently under investigation.

**Subcellular Localization of hsBG**—Because of the similarities between the phenotype of the fruit fly *bubblegum* mutant and the human peroxisomal disease X-linked adrenoleukodystrophy, it was hypothesized that hsBG might be a peroxisomal protein. Although no peroxisome targeting signal I (PTS1) was present, a potential PTS2 (RIDDPSCPQKL) was found at amino acid residue 93. When analyzed by SignalP version 1.1 (23) and PSORT (24) prediction programs, no signal sequence or endoplasmic reticulum membrane retention signal was found. No mitochondrial or nuclear targeting signals were detected. The TopPred 2 program (25) identified a single candidate membrane-spanning segment. In contrast, the Predict Protein program (26) indicated that hsBG was a compact globular protein and PSORT predicted a cytoplasmic location. To investigate the subcellular location of hsBG, we used indirect immunofluorescence and confocal microscopy of cells transiently expressing the protein. Because antibody to hsBG is not yet available, we transfected COS-1 cells with *hsBGM* cDNA containing the c-Myc epitope in frame at the amino terminus. Fluorescence appeared to be associated with the plasma membrane, but fluorescence of the cytoplasm and intracellular organelles could not be ruled out (data not shown). Therefore, confocal microscopy was used to examine COS-1 cells co-transfected with hsBG and hsALDP (a peroxisomal membrane protein) (Fig. 3). This method did not clearly distinguish between a cytoplasmic and plasma membrane location for hsBG (Fig. 3A). Whereas hsALDP had a punctate immunostaining pattern typical of peroxisomes (Fig. 3B), there was no evidence for localization of hsBG with peroxisomes (Fig. 3C).

Further analysis of cells expressing hsBG was performed by obtaining confocal images of serial sections and using these to reconstruct a three-dimensional image of the cell (Fig. 4). Two populations of COS-1 cells expressing N-myc-hsBG were observed. In some cells, hsBG was predominantly cytoplasmic (Fig. 4, A and B), and in others, the majority of the hsBG immunofluorescence appeared to be close to or associated with the plasma membrane (Fig. 4, C and D). There was no evidence of fluorescent labeling of intracellular organelles. Similar results were obtained when cells were transfected with C-myc-hsBG (data not shown).

The association of some hsBG with the plasma membrane and the possibility that hsBG contains one membrane-spanning region prompted us to conduct additional immunofluorescence studies. COS-1 cells were transfected with either N-myc-hsBG or C-myc-hsBG. When cells were permeabilized with 1% Triton X-100 prior to incubation with anti-Myc antibody (standard procedure), results were identical to those in Fig. 4 (data not shown). When the Triton X-100 step was eliminated (no permeabilization), no immunofluorescent signal was detected in cells expressing either construct (data not shown). If either the carboxyl or amino terminus of hsBG protruded through the plasma membrane, fluorescence would have been detected in the absence of detergent. Thus, we found no evidence that hsBG is an integral plasma membrane protein.

**hsBG-dependent Acyl-CoA Synthetase Activity Is Primarily Cytosolic**—To investigate the subcellular distribution of hsBG by biochemical methods, digitonin was initially used to fractionate transfected cells into crude cytoplasmic and crude membrane/organelle fractions. Most of the hsBG-dependent
acyl-CoA synthetase activity was in the cytoplasmic fraction (Table II). To determine whether this activity was truly soluble, COS-1 cells transiently expressing hsBG were subjected to a more rigorous subfractionation procedure as described under “Experimental Procedures.” A peroxisome- and mitochondria-enriched ML fraction, a microsomal fraction, and cytosol were prepared and assayed for acyl-CoA synthetase activity. As expected, nearly all of the acyl-CoA synthetase activity of vector-transfected COS-1 cells was associated with the ML fraction and microsomes (Table II). In contrast, the hsBG-dependent acyl-CoA synthetase activity was mainly in the soluble (cytosolic) fraction (Table II). Small increases in activity of the ML and microsomal fractions in hsBG-expressing cells were also observed. It is not clear whether the plasma membrane-associated hsBG is responsible for these small increases.

hsBG Isolated from COS-1 Cells Has Enzyme Activity—Although amino acid sequence homology suggests that hsBG is an acyl-CoA synthetase, it is possible that the increased enzyme activity of COS-1 cells expressing hsBG is a secondary phenomenon. To rule out this possibility, we purified the expressed N-myc-hsBG from the cytosolic fraction using anti-Myc antibody covalently linked to agarose beads. No acyl-CoA synthetase activity was bound to beads incubated with cytosol from cells transfected with the vector (N-myc-pcDNA3) alone (Table III). In contrast, 26–30% of the cytosolic activity in hBG-transfected cells was found associated with the anti-Myc beads (Table III). These data confirm that the hBG protein itself has acyl-CoA synthetase activity.

hsBG Represents a New Family of Acyl-CoA Synthetases—In our investigations of the VLCS protein family, we noted two highly conserved motifs present in their amino acid sequences. One motif is the previously mentioned AMP-binding domain. The second motif is 240–250 amino acids downstream and overlaps a highly conserved region in the long-chain acyl-CoA synthetase family described by Black et al. (9) as the signature motif...
motif for this family. Inspection of the comparable regions in both dmBG and hsBG amino acid sequences revealed that these two proteins lack this motif and thus were not members of either the VLCS family or the LCS family. To determine whether hsBG and dmBG were similar to any other known acyl-CoA synthetases, we screened the GenBank™ nonredundant protein data base. Amino acid sequences of all identifiable vertebrate and yeast acetyl-CoA synthetases, medium- and long-chain acyl-CoA synthetases, and unknown or putative acyl-CoA synthetases were aligned and regions corresponding to VLCS motifs 1 and 2 were identified. In some cases Caenorhabditis elegans, D. melanogaster, Arabidopsis thaliana, and/or bacterial sequences were included. To facilitate assignment of sequences to protein families, we extended motif 1 by 5 amino acids in each direction and motif 2 by 10 amino acids. Alignments of these sequences are shown in Table IV.

Visual inspection of motif 2 in most cases confirmed that proteins previously designated as acetyl-CoA synthetases, LCSS and VLCSs, were indeed closely related to each other but proteins previously designated as acetyl-CoA synthetases, were assayed for acyl-CoA synthetase activity using C16:0 and C24:0 substrates. Results of two experiments are shown, each using cytosol from an independent transfection. To calculate the percent activity bound to beads, endogenous acyl-CoA synthetase activity (vector-transfected cytosol, A) was subtracted from that of the hsBG-transfected cytosol (B).

### Table III

| N-mylocDNA3  | Cytosol (A) | Bound | C16:0 | C24:0 |
|--------------|-------------|-------|-------|-------|
| hsBG         | 553         | 600   | 106   |
| N-myloc-hsBG | 0           | 0     | 0     |
| Cytosol (B)  | 9164        | 8750  | 1219  |
| Cytosol, hBG-specific (B-A) | 8611 | 8150 | 1113 |
| Bound        | 2252        | 3173  | 442   |
| (% bound)    | (26%)       | (39%) | (40%) |

**Discussion**

Our previous investigations of enzymes that activate VLCFA to their CoA derivatives suggested that although hsVLCS, hs-VLCS-H2, and mmFATP1 have VLCS activity, additional enzymes with VLCS activity must exist (16, 29–31). The elevation of VLCSA levels in the bubblegum mutant and the similarity of dmBG to other acyl-CoA synthetases (12) suggested that BG might have VLCS activity. As we report here, hsBG is indeed capable of activating VLCFA and, like hsVLCS, will also activate long-chain fatty acid substrates.

At present, the physiologic role of hsBG in fatty acid metabolism is not known. Because patients with XALD and fruit flies with the bubblegum mutation have similarly elevated VLCFA levels, it was reasonable to investigate the hypothesis that hsBG is a peroxisomal protein that activates VLCFA for degradation by peroxisomal β-oxidation. However, neither aspect of this hypothesis was validated in COS-1 cells expressing hsBG. In these cells, hsBG could not be detected in peroxisomes, and a role for this protein in VLCFA β-oxidation was not demonstrated. However, if concentrations of other enzymes or proteins required for normal peroxisomal VLCFA β-oxidation are rate-limiting, increasing the concentration of hsBG would likely be without effect. Nonetheless, studies to determine whether hsBG is involved in XALD biochemical pathology are currently in progress.

We also considered the possibility that hsBG might play a role in complex lipid synthesis. Although we did not observe increased incorporation of either C16:0 or C24:0 into neutral or polar lipid classes in COS-1 cells expressing hsBG (data not shown), these experiments are subject to the same concerns as the β-oxidation experiments, i.e. VLCFA activation may not be rate-limiting. Thus, COS-1 cells overexpressing hsBG may not be an appropriate system for demonstrating a role for this enzyme in complex lipid synthesis. Studies to evaluate the effect on VLCFA metabolism of decreasing the level of hsBG in cells endogenously expressing the protein are currently in progress.

The observation that hsBG expressed in COS-1 cells is primarily soluble was surprising because most endogenous acyl-CoA synthetase activity has been reported to be in either organellar membranes (microsomal, peroxisomal, or mitochondrial) or the plasma membrane (1). Confocal microscopy showed hsBG localizing near the plasma membrane in some transfected COS-1 cells, whereas the protein was clearly cytoplasmic in others. These observations suggest a potential role of hsBG in “trapping” fatty acids inside cells as their CoA derivatives as they enter the cell from the extracellular medium. Furthermore, this raises the question of whether alterations in the physiologic status of cells or the composition of the culture medium could influence the distribution of hsBG within the cell.

Prior to this report, the only enzymes capable of activating VLCFA were members of the VLCS/FATP protein family (16, 29–31). It is noteworthy that hsBG belongs to a separate protein family. The bubblegum family, established by homology within motifs 1 and 2 of Table IV, includes hsBG (this report), mmBG,² dmBG (12), and a homolog of dmBG that derived from the Dro sophila genome sequencing project (GenBank™ accession number AAF44850). In addition to these proteins, we have postulated the existence of homologs of hsBG and mmBG based on sequences of expressed sequence tags.² These additional proteins are most likely acyl-CoA synthetases and potentially have VLCS activity.

In our analysis of the amino acid sequences of the known

---

² A. K. Heinzer, K. D. Smith, and P. A. Watkins, unpublished data.
| Protein | Motif 1 | Motif 2 | Accession No. |
|---------|---------|---------|---------------|
| Acetyl-CoA synthetase family | | | |
| hsAcCS1 | MFLLMYTGSTGMPGKGVHT | GARYTEGGYQITGRMDVINIS-HGRLTAEIAEDAIHPAVFE | CAB75500 |
| hsAcCS2 | PFLFYMTGSTGMPGKGVHT | GCRDQQGYNTRGIDMNLVNS-HGRLTAEIAEDAIHPAVFE | AAF75064 |
| mmAcCS | PFLFYMTGSTGMPGKGVHT | GARRDEGGYLMGIDMNLVNS-HGRLTAEIAEDAIHPAVFE | CA91274 |
| ceAcCS | PFLFYMTGSTGMPGKGVHT | GARRDEGGYLMGIDMNLVNS-HGRLTAEIAEDAIHPAVFE | S92154 |
| dmAcCS | PFLFYMTGSTGMPGKGVHT | GAADGYYLMGIDMNLVNS-HGRLTAEIAEDAIHPAVFE | S90018 |
| scAcCS2 | PFLFYMTGSTGMPGKGVHT | GARRDGYLMGIDMNLVNS-HGRLTAEIAEDAIHPAVFE | S65002 |
| ncAcCS | PFLFYMTGSTGMPGKGVHT | GARRDGYLMGIDMNLVNS-HGRLTAEIAEDAIHPAVFE | P16929 |
| eeAcCS | PFLFYMTGSTGMPGKGVHT | GARRDGYLMGIDMNLVNS-HGRLTAEIAEDAIHPAVFE | P27550 |
| MCS/SA family | | | |
| hsMC/SA | MFAIFTGSTGMPGKMAHS | RADIEMEGYWFLGRSDVINAS-YGRVPGAPVSLAEPVAE | ADF16145 |
| hsSA | MIAIFTGSTGMPKTAHMT | RGYMDQGDFWYVFARVLS-YGRVPGAPVSLAEPVAE | AAC1667 |
| rnSA | MLAIYMTGSTGMPKVEHS | RAKMDQGDFWYVFARVLS-YGRVPGAPVSLAEPVAE | BAA91273 |
| mmSA | MAIYMTGSTGMPKVEHS | RAKMDQGDFWYVFARVLS-YGRVPGAPVSLAEPVAE | D76385 |
| LCS family | | | |
| hsLCS | LAVICTGSTGMPGKMWT | IGIWKLPNTLIDIRKIFKILALQVEYAPKIEKYNRSVPQV | L09229 |
| hsLCS2 | LAVICTGSTGMPGKMWT | IGIWKLPNTLIDIRKIFKILALQVEYAPKIEKYNRSVPQV | P31212 |
| hsLCS3 | IAVMYMTGSTGMPGVMIS | IGEFHPDGKIDRIRKDKVAGKQLGGVEAGLKLQVNL | D10910 |
| hsLCS4 | LFSICGSTGMPGKMIT | IGIWKLPNTLIDIRKIFKILALQVEYAPKIEKYNRSVPQV | D10041 |
| hsLCS5 | LFSICGSTGMPGKMIT | IGIWKLPNTLIDIRKIFKILALQVEYAPKIEKYNRSVPQV | D10041 |
| hsKIAA0837 | LSIVGSTGMPGKMPLT | IGIWKLPNTLIDIRKIFKILALQVEYAPKIEKYNRSVPQV | D90866 |
| mmLCS | LAVICTGSTGMPGKMWT | IGIWKLPNTLIDIRKIFKILALQVEYAPKIEKYNRSVPQV | D38519 |
| mmAcCS2 | LAVICTGSTGMPGKMWT | IGIWKLPNTLIDIRKIFKILALQVEYAPKIEKYNRSVPQV | BAA12933 |
| mmAcCS4 | LAVICTGSTGMPGKMWT | IGIWKLPNTLIDIRKIFKILALQVEYAPKIEKYNRSVPQV | N0_032007 |
| scFaa1p | LAVICTGSTGMPGKMWT | IGIWKLPNTLIDIRKIFKILALQVEYAPKIEKYNRSVPQV | BAA59929 |
| scFaa2p | LAVICTGSTGMPGKMWT | IGIWKLPNTLIDIRKIFKILALQVEYAPKIEKYNRSVPQV | BAA59929 |
| scFaa3p | LAVICTGSTGMPGKMWT | IGIWKLPNTLIDIRKIFKILALQVEYAPKIEKYNRSVPQV | A59910 |
| scFaa4p | LAVICTGSTGMPGKMWT | IGIWKLPNTLIDIRKIFKILALQVEYAPKIEKYNRSVPQV | S65060 |
| VFLCS/FATP family | | | |
| hsVFLCS | PALYMTGSTGMPKAMT | LLWDHNNFIFYFRDVRGTDVRFFK-GENAVTEAVTDVGLDFQVE | NP_003636 |
| hsVFLCS-H1 | TCLYMTGSTGMPKAALS | LTVQDENFIFYFRDVRGTDVRFFK-GENAVTEAVTDVGLDFQVE | AAD29443 |
| hsVFLCS-H2 | PFLYMTGSTGMPKAALS | LTVQDENFIFYFRDVRGTDVRFFK-GENAVTEAVTDVGLDFQVE | AAD29443 |
| hsFATP4 | KLFLYMTGSTGMPKAALS | LNGMDQFRGDFVRGTDVRFFK-GENAVTEAVTDVGLDFQVE | NP_005085 |
| mmVFLCSR | PAKFMTGSTGMPKAALS | LNGMDQFRGDFVRGTDVRFFK-GENAVTEAVTDVGLDFQVE | O35488 |
| mmVFLCSR | PAFYMTGSTGMPKAALS | LNGMDQFRGDFVRGTDVRFFK-GENAVTEAVTDVGLDFQVE | NP_030617 |
| mmFATP1 | KLFLYMTGSTGMPKAALS | LNGMDQFRGDFVRGTDVRFFK-GENAVTEAVTDVGLDFQVE | AAC40187 |
| mmFATP3 | TLYCMTGSTGMPKAALS | LNGMDQFRGDFVRGTDVRFFK-GENAVTEAVTDVGLDFQVE | AAC40187 |
| mmFATP4 | TLYCMTGSTGMPKAALS | LNGMDQFRGDFVRGTDVRFFK-GENAVTEAVTDVGLDFQVE | AAC40187 |
| mmNCS | TLYCMTGSTGMPKAALS | LNGMDQFRGDFVRGTDVRFFK-GENAVTEAVTDVGLDFQVE | AAC40187 |
| scFAT1p | TLYCMTGSTGMPKAALS | LNGMDQFRGDFVRGTDVRFFK-GENAVTEAVTDVGLDFQVE | AAC40187 |
| Bumblegum family | | | |
| hsBGS | CCLYMTGSTGMPGKMA | AGRDADGFLYIFTRGKLELITAGENMPVFVFVEAAKRMELPII | A111606 |
| mmBGS | CCLYMTGSTGMPGKMA | AGRDADGFLYIFTRGKLELITAGENMPVFVFVEAAKRMELPII | A111606 |
| dmBGS | CCLYMTGSTGMPGKMA | AGRDADGFLYIFTRGKLELITAGENMPVFVFVEAAKRMELPII | A111606 |
| dmBGS-H1 | CCLYMTGSTGMPGKMA | AGRDADGFLYIFTRGKLELITAGENMPVFVFVEAAKRMELPII | A111606 |
| scFat2p family | | | |
| scFat2p | VALIYMTGSTGMPGKMA | QSFYDPFFGFLRTGKLELITAGENMPVFVFVEAAKRMELPII | A111606 |
| scFat2p | VALIYMTGSTGMPGKMA | QSFYDPFFGFLRTGKLELITAGENMPVFVFVEAAKRMELPII | A111606 |
| scFat2p | VALIYMTGSTGMPGKMA | QSFYDPFFGFLRTGKLELITAGENMPVFVFVEAAKRMELPII | A111606 |
| scFat2p | VALIYMTGSTGMPGKMA | QSFYDPFFGFLRTGKLELITAGENMPVFVFVEAAKRMELPII | A111606 |
| mtFatD7 | DAIMFGMTGSTGMPGKMA | LGLSLASDLSRGLKLELITAGENMPVFVFVEAAKRMELPII | C03808 |

Amino acid sequences of known or putative acyl-CoA synthetases were obtained from GenBank™. Sequence alignments were performed using the ClustalW program. Regions corresponding to Motif 1 and Motif 2 of the VFLCS/FATP family (7, 10) were identified. These sequences were extended in both the 5' - and 3'-directions by 5 residues (Motif 1) or 10 residues (Motif 2) to improve the accuracy of family assignments. Additional family members were identified by using a representative Motif 2 sequence (from a human protein if available) as query sequence in a BLAST search of the GenBank™ non-redundant protein database. All vertebrate sequences identified are shown below. Representative sequences from yeast and other species are also shown, but these listings are not exhaustive. Species abbreviations: hs, Homo sapiens; bt, B. terrestris; rn, Rattus norvegicus; mm, Mus musculus; ce, C. elegans; dm, D. melanogaster; sc, S. cerevisiae; sp, Schizosaccharomyces pombe; nc, Neurospora crassa; mt, Mycobacterium tuberculosis; at, A. thaliana; ec, Escherichia coli. Gene and protein abbreviations: AcCS, acetyl-CoA synthetase; MCS, medium-chain acyl-CoA synthetase; TN, unnamed protein; RSP, kidney-specific protein; LCS, long-chain acyl-CoA synthetase; FACI, fatty acyl-CoA ligase; ACS, acyl-CoA synthetase; FAA, fatty acid activation; VFLS, very long-chain acyl-CoA synthetase; FATP, fatty acid transport protein; Fat, fatty acid transporter; BG, bubblegum; CCLL, commarate-CoA ligase-like, CAS, coenzyme A synthetase.
acyl-CoA synthetases, we also established that Fat2p, a putative acyl-CoA synthetase from *Saccharomyces cerevisiae*, belonged to yet another protein family (Table IV). Enzyme activity of Fat2p has not been firmly established, and if indeed it has activity it will be of interest to determine its substrate specificity. Although no mammalian members of this family were readily detected in public domain data bases, if such homologs exist they will soon be identified by genome sequencing efforts. It is also likely that additional families of acyl-CoA synthetases may be identified as a result of genome sequencing and that additional candidate VLCSs will emerge.

It is significant that hsBG is expressed primarily in brain, since no member of the hsVLCS/FATP family is highly expressed in this tissue. In general, brain has higher concentrations of VLCFA than other tissues (32). VLCFA present in brain include saturated (e.g. C24:0 and C26:0), monoenoic (e.g. C24:1 and C26:1), and polyunsaturated fatty acids (e.g. C24:4ω6 and C22:6ω3) (32). Concentrations of brain VLCFA increase during development (33). These VLCFA are components of complex lipids such as gangliosides, cerebrosides, sulfatides, sphingomyelin, and other phospholipids. Activation by VLCSs is required for incorporation of VLCFA into these complex lipids. Although their role(s) in brain function are not completely understood, many of these VLCFA-containing lipids are components of myelin membranes and disruption of VLCS activity could be responsible for one or more genetic demyelinating disorders. Thus, investigation of brain-specific VLCSs will significantly enhance our understanding of brain fatty acid homeostasis due to the central importance of VLCFA activation in this process.

Acknowledgments—We thank Joseph Park and Zhenzhen Jia for expert technical assistance and Michael Delannoy for assistance with the confocal microscopy.

REFERENCES
1. Watkins, P. A. (1997) *Prog. Lipid Res.* **36**, 55–83
2. Suzuki, H., Kawarabayasi, Y., Kondo, J., Abe, T., Nishikawa, K., Kimura, S., Hashimoto, T., and Yamamoto, T. (1999) *J. Biol. Chem.* **264**, 8681–8685
3. Fujino, T., and Yamamoto, T. (1992) *J. Biochem. (Tokyo)* **111**, 197–203
4. Fujino, T., Kang, M. J., Suzuki, H., Iijima, H., and Yamamoto, T. (1996) *J. Biol. Chem.* **271**, 16748–16752
5. Kang, M. J., Fujino, T., Sasano, H., Minekura, Y., Yabuki, N., Nagura, H., Iijima, H., and Yamamoto, T. T. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94**, 2880–2884
6. Oikawa, E., Iijima, H., Suzuki, T., Sasano, H., Sato, H., Kamataki, A., Nagura, H., Kang, M. J., Fujino, T., Suzuki, H., and Yamamoto, T. T. (1998) *J. Biochem. (Tokyo)* **124**, 679–675
7. Watkins, P. A., Pevsner, J., and Steinberg, S. J. (1999) *Prostaglandins Leukot. Essent. Fatty Acids* **60**, 223–228
8. Babitt, P. C., Kenyon, G. L., Martin, B. M., Charest, H., Slyvestre, M., Scholten, J. D., Chang, K. H., Liang, P. H., and Dunaway-Mariano, D. (1992) *Biochemistry* **31**, 5594–5604
9. Black, P. N., Zhang, Q., Weimar, J. D., and DiRusso, C. C. (1997) *J. Biol. Chem.* **272**, 4986–4993
10. Watkins, P. A., Lu, J. F., Steinberg, S. J., Gould, S. J., Smith, K. D., and Braiterman, L. T. (1998) *J. Biol. Chem.* **273**, 18210–18219
11. Choi, J. Y., and Martin, C. E. (1999) *Science* **288**, 8681–8685
12. Min, K. T., and Benzer, S. (1999) *Science* **284**, 1985–1988
13. Moser, H. W., Smith, K. D., and Moser, A. B. (1995) in *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds) 7th Ed., pp. 2325–2349, McGraw-Hill Inc., New York
14. Muralidharan, V. B., and Kishimoto, Y. (1984) *J. Biol. Chem.* **259**, 13021–13026
15. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4673–4680
16. Steinberg, S. J., Wang, S. J., Kim, D. G., Mihalik, S. J., and Watkins, P. A. (1999) *Biochem. Cell Biol.* **77**, 615–621
17. Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951) *J. Biol. Chem.* **193**, 265–275
18. Balch, W. E., and Rothman, J. E. (1985) *Arch. Biochem. Biophys.* **240**, 413–425
19. Watkins, P. A., Howard, A. E., and Mihalik, S. J. (1994) *Biochim. Biophys. Acta* **1214**, 288–294
20. Watkins, P. A., Ferrer, E. V., Jr., Pedersen, J. I., and Hoefler, G. (1991) *Arch. Biochem. Biophys.* **289**, 329–336
21. Watkins, P. A., Gould, S. J., Smith, M. A., Braiterman, L. T., Wei, H.-M., Kok, F., Moser, A. B., Moser, H. W., and Smith, K. D. (1995) *Am. J. Hum. Genet.* **57**, 292–301
22. Singh, I., Moser, A. B., Goldfischer, S., and Moser, H. W. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81**, 4203–4207
23. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) *Protein Eng.* **10**, 1–6
24. Nakai, K., and Horton, P. (1999) *Trends Biochem. Sci.* **24**, 34–36
25. von Heijne, G. (1992) *J. Mol. Biol.* **225**, 487–494
26. Rost, B., and Sander, C. (1994) *Proteins* **19**, 55–72
27. Iwai, N., Ohmichi, N., Hanai, K., Nakamura, Y., and Kinoshita, M. (1994) *Hypertension* **23**, 375–380
28. Samani, N. J., and Lodwick, D. (1995) *J. Hum. Hypertens.* **9**, 501–503
29. Steinberg, S. J., Wang, S. J., McGuinness, M. C., and Watkins, P. A. (1999) *Mol. Genet. Metab.* **68**, 32–42
30. Coo, N. R., Smith, A. J., Frohnert, B. I., Watkins, P. A., and Bernlohr, D. A. (1999) *J. Biol. Chem.* **274**, 36300–36304
31. Smith, K. D., Kemp, S., Braiterman, L. T., Lu, J. F., Wei, H. M., Geraghty, M., Stetten, G., Bergin, J. S., Pevsner, J., and Watkins, P. A. (1999) *Neurochem. Res.* **24**, 521–535
32. Poulos, A. (1995) *Lipids* **30**, 1–14
33. Svennerholm, L., and Stallberg-Stenhagen, S. (1968) *J. Lipid Res.* **9**, 215–225