1994

Cloning, Characterization And Targeted Disruption Of The Mouse Medium-Chain Acyl-Coa Dehydrogenase Gene.

Ravi Jiwat Tolwani
University of Alabama at Birmingham

Order Number 9520557

Cloning, characterization and targeted disruption of the mouse medium-chain acyl-CoA dehydrogenase gene

Tolwani, Ravi Jiwat, Ph.D.

University of Alabama at Birmingham, 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

CLONING, CHARACTERIZATION AND TARGETED DISRUPTION OF THE MOUSE MEDIUM-CHAIN ACYL-COA DEHYDROGENASE GENE

by

RAVI J. TOLWANI

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Molecular and Cellular Pathology in the
Department of Comparative Medicine, The Graduate School,
The University of Alabama at Birmingham

BIRMINGHAM, ALABAMA

1994

ABSTRACT OF DISSERTATION
GRADUATE SCHOOL, UNIVERSITY OF ALABAMA AT BIRMINGHAM

Degree Ph.D. Major Subject Molecular and Cellular Pathology

Name of Candidate Ravi J. Tolwani

Title Cloning, Characterization and Targeted Disruption of the Mouse Medium-Chain Acyl-CoA Dehydrogenase Gene

Medium-chain acyl-CoA dehydrogenase (enzyme symbol MCAD, gene symbol *Acadm*) is one of four dehydrogenases involved in the first step of fatty acid β-oxidation. Inherited MCAD deficiency exists in humans as an autosomally recessive disorder. MCAD deficient patients experience a wide range of metabolic problems whose pathogenesis is complex and poorly understood. In order to better understand MCAD deficiency, we have cloned and characterized the mouse MCAD (*Acadm*) gene and are developing a mouse model with MCAD deficiency via gene targeting in embryonic stem (ES) cell.

The *Acadm* cDNA was cloned by library screening and reverse transcription-polymerase chain reaction. The *Acadm* cDNA coding region is 1263 basepairs (bp) long with a 3' untranslated region of 576 bp and encodes a 421 amino acid precursor protein. Similarity at the nucleotide and deduced amino acid level between mouse compared to rat and human MCAD cDNAs implies similar structure of the MCAD monomer.

The *Acadm* gene, cloned by library screening, spans 25 kilobases and consists of 12 exons. The promoter region does not have a TATA or CCAAT box and is GC rich with several putative Sp1 binding sites. Also present are several consensus elements to

bind various steroid nuclear receptors such as the retinoic acid receptors. Two processed *Acadm* pseudogenes were also cloned via genomic library screening.

Genomic and cDNA fragments have been used to map the chromosomal location of the *Acadm* gene via restriction fragment length polymorphism (RFLP) linkage analysis. The structural *Acadm* gene was mapped to the distal end of chromosome 3, while the *Acadm* processed pseudogenes were mapped to mouse chromosome 8 and 11.

Acadm genomic fragments were used to design insertion vectors to disrupt the *Acadm* gene in ES cells. These vectors were designed to produce a mutation that would inhibit MCAD monomers from forming functional homotetramers. ES cells with the desired recombination were detected by PCR dependent on gap repair and confirmed with RFLP analysis. Subsequently, disrupted ES cells were used to generate chimeric mice.

Abstract Approved by: Committee Chairman Philip A. Wood

Program Director

Date 12/15/94 Dean of Graduate School Joan F. Logan

PS-5744

DEDICATION

I thank my wife, Lovey, for her love, support and enormous sacrifices during my training. It is to her that I dedicate this dissertation.

ACKNOWLEDGEMENTS

I wish to thank my mentor, Phil Wood, for his support and guidance during my tenure in his lab. Before joining his lab I was used to measurements in liter quantities, i.e., how many liters of fluids a dehydrated German shepherd dog would receive. Now I deal with volumes in microliter quantities. Thanks to Phil, I have learned a great deal about science and molecular biology. I thank him for making a significant contribution to my education.

I thank my committee members, Kevin Dybvig, Lane Rutledge, Jerry Thompson and Tim Townes, for sharing their time and sharing their expertise. Their comments and suggestions have been very helpful over the years.

I also wish to thank Dr. Russell Lindsey for his advice and guidance throughout my fellowship at the UAB. His review of this manuscript is also very appreciated.

All my colleagues in the lab, Susan Farmer, Doug Hamm, Myron Hinsdale, Lisa Kelly and David Kurtz, have significantly helped in many ways. I have been fortunate to have friends like them.

I thank my parents and sister for their love and support. They have always been encouraging to whatever my decisions have been.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (continued)

INTRODUCTION

Literature Review

Biochemical and Molecular Aspects of Fatty Acid Oxidation

A major source of energy during fasting periods is derived from β-oxidation of fatty acids. Fatty acids circulate in the bloodstream primarily bound to albumin. Fatty acid oxidation predominantly occurs in liver, cardiac and skeletal muscle, kidney and adrenal glands (Kelly *et al.*, 1989). The brain oxidizes ketone bodies which are derived from fatty acid oxidation by the liver.

Fatty acids are taken up by cells in the liver and other tissues. These fatty acids are then activated by acyl-CoA synthase, a cytosolic enzyme, to form coenzyme A (CoA) esters. These acyl-CoA esters are directed toward mitochondrial β-oxidation primarily under fasting conditions. Since β-oxidation occurs in mitochondria, these acyl-CoA esters must traverse the mitochondrial membrane into the mitochondrial matrix. Long chain acyl-CoA esters are *trans*-esterified by carnitine palmitoyltransferase I (CPT I), an outer mitochondrial membrane enzyme, to form acylcarnitines. These acylcarnitines cross the mitochondrial membrane mediated by carnitine translocase. Once inside the matrix the acylcarnitines are reesterified by carnitine palmityoltransferase II (CPT II) to reform the acyl-CoA molecule. Medium-chain length and short-chain length acyl-CoA molecules can enter the mitochondrial matrix without being *trans*-esterified to an acylcarnitine intermediate (Roe and Coates, 1989).

Once inside the mitochondria, β-oxidation of the fatty acid proceeds via a repeating circuit of 4 sequential steps (Fig. 1). Completion of each circuit results in the reduction of the acyl-CoA molecule chain length by 2 carbons in size and the release of an acetyl-CoA molecule. This acetyl-CoA molecule can be oxidized to carbon dioxide and water by the tricarboxylic acid (TCA) cycle in the liver and heart. Additionally, this acetyl-CoA molecule can be converted to ketone bodies such as β-hydroxybutyrate and acetoacetate in the liver and kidney. The ketone bodies can then be oxidized for energy by tissues such as brain to spare glucose during fasting periods.

The four sequential steps of mitochondrial β-oxidation are mediated by enzymes which have substrate chain length specificity. The first step is catalyzed by acyl-CoA dehydrogenases which result in the α-β dehydrogenation of the acyl-CoA molecule with the transfer of electrons to electron transfer flavoproteins (ETF). The electrons are then transferred to ETF dehydrogenase and ultimately to the electron transfer chain. The next step of β-oxidation is catalyzed by enoyl-CoA hydratase which results in the addition of a hydroxyl group to the β-carbon of the acyl-CoA molecule. Next, the 3-hydroxyacyl-CoA dehydrogenase, along with nicotinamide adenine dinucleotide (NAD), converts the 3-hydroxyacyl-CoA to 3-ketoacyl-CoA. Finally, the 3-ketoacyl-CoA thiolase cleaves the α-β bond of the 3-ketoacyl-CoA with the presence of reduced CoA. Completion of this circuit results in the release of an acetyl-CoA molecule and the reduction in size of the acyl-CoA molecule by a 2 carbon unit. This acyl-CoA molecule can once again reenter the β-oxidation circuit and continue to generate acetyl-CoA and be further reduced in chain length.

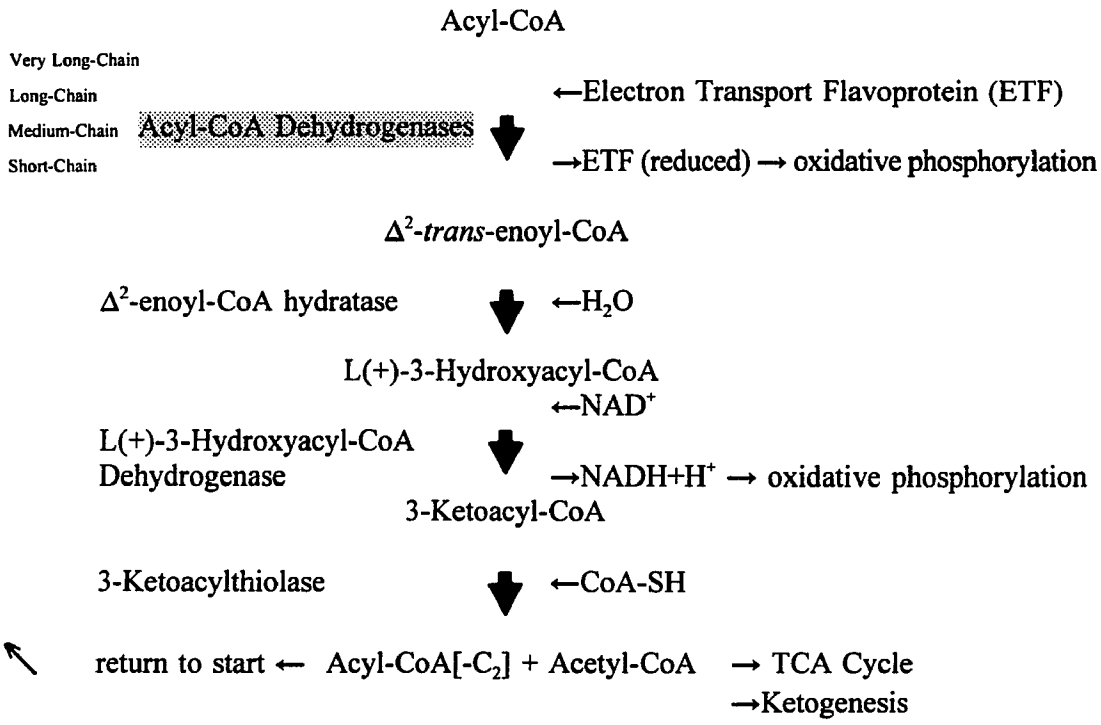
Specifically, there are four straight chain acyl-CoA dehydrogenase enzymes involved in the initial step of β-oxidation. The very long-chain (VLCAD), long-chain (LCAD), medium-chain (MCAD) and short-chain (SCAD) acyl-CoA dehydrogenases act on acyl-CoA molecules of 14-20 carbons in length, 12-16 carbons in length, 4-14 carbons in length and 4-6 carbons in length, respectively (Roe and Coates, 1989). These proteins are encoded for by nuclear genes and translated in the cytosol and translocated into the mitochondria.

Medium-chain acyl-CoA dehydrogenase (MCAD), a flavoenzyme, is functionally active as a homotetramer in the mitochondrial matrix. The 46,600 molecular weight precursor protein is synthesized in the cytosol and translocated into the mitochondrial matrix (Matsubara *et al.*, 1987; Kelly *et al.*, 1987; Ikeda *et al.*, 1987). This mitochondrial translocation results in the proteolytic cleavage of the leader peptide sequence producing the mature MCAD monomer with a molecular weight of 43,700 (Ikeda *et al.*, 1987; Ikeda *et al.*, 1986). To gain enzymatic function these monomers are assembled into homotetramers with each monomer containing one molecule of flavin adenine dinucleotide (FAD) (Thorpe *et al.*, 1979).

The 3-dimensional structure of the MCAD monomer has been determined by X-ray defraction (Kim and Wu, 1988; Kim *et al.*, 1993). The protein is folded into three domains. The N-terminal domain consists of α-helixes A to F and the C-terminal domain consists of α-helixes G to L. The middle domain consists of orthogonal β-sheets 1-6. The flavin ring lies between the C-terminal and β-sheet domains. Helices H and I in the C-terminal domain make the intersubunit contacts with the formation of the homotetramer.

FIG. 1. β-oxidation pathway. Long-chain acyl-CoA esters enter the mitochondria via carnitine mediated transport. Medium- and short-chain acyl-CoA esters do not require carnitine mediated transport. Fatty acid oxidation proceeds in a repeating four step circuit within the mitochondria. The completion of each cycle results in removal of two carbons from the chain and production of an acetyl-CoA molecule.

Overview of Mitochondrial β-Oxidation of Fatty Acids



The rat and human MCAD cDNAs have been cloned and consist of coding sequence of 1263 nucleotides each (Kelly *et al.*, 1987; Matsubara *et al.*, 1987). They encode for a 421 amino acid precursor protein, including a 25 amino acid leader peptide which is removed upon mitochondrial translocation. The mature MCAD monomer, therefore, is 396 amino acids in length.

The human MCAD gene spans 44 Kb and consists of 12 exons (Zhang *et al.*, 1992). The human MCAD promoter lacks TATA and CAAT boxes and is GC rich. The human MCAD 5' regulatory region contains a series of hexamer sequences which bind to several transcriptional factors of the nuclear receptor superfamily. The region at -341 to -308 to the transcriptional start site contains 4 potential nuclear receptor hexamer sites and has been labeled as the nuclear receptor response element-1 (NRRE-1) (Carter *et al.*, 1994). These elements have been shown to bind to all-*trans* retinoic acid receptors (RARs) and 9-*cis* retinoic acid receptors (RXRs), resulting in transcriptional activation of the MCAD gene (Raisher *et al.*, 1992). This region also binds to hepatocyte nuclear factor-4 (HNF-4), an orphan receptor, also resulting in transcriptional activation of the human MCAD gene (Carter *et al.*, 1993). Most recently, the NRRE-1 has been shown to interact with chicken ovalbumin upstream promoter transcriptional factor (COUP-TF) to repress transcription of the human MCAD gene (Carter *et al.*, 1994).

Interestingly, the human MCAD precursor mRNA has been shown to undergo abnormal missplicing (Zhang *et al.*, 1992). Eight to 10 percent of PCR derived cDNA MCAD clones from normal individuals were misspliced. These missplicing events resulted in the deletion of either exon 2 or exon 5. This is the only gene reported to undergo aberrant missplicing from the normal gene.

The human MCAD gene has been mapped to chromosome 1p31 by *in situ* hybridization using the human MCAD cDNA as a probe (Kidd *et al.*, 1990). Additionally, the mouse MCAD gene has been mapped to mouse chromosome 8 (Bahary *et al.*, 1991). A 1000 bp fragment produced by PCR amplification from a rat liver cDNA library was used as a probe. Interestingly, the region of mouse chromosome 8 containing the MCAD gene contains homology to human chromosome 16q with no homology to human chromosome 1.

MCAD Deficiency

Defects resulting in deficiencies of each of the 4 straight chain acyl-CoA dehydrogenases exist with MCAD deficiency being by far the most common (Roe and Coates, 1989). MCAD deficiency was first described by Kolvraa and co-workers in 1982. MCAD deficiency is an autosomal recessive disorder and has been described in hundreds of patients worldwide (Roe and Coates, 1989; Coates, 1992; Matsubara *et al.*, 1992; Ding *et al.*, 1992; Workshop on Molecular Aspects of MCAD Deficiency, 1992). The clinical signs of MCAD deficiency are varied, indicating considerable phenotypic heterogeneity. Clinical signs usually occur at a young age with the earliest onset of symptoms at one day of age (Catzeflis *et al.*, 1990). The greatest frequency of initial presentation is between 5 and 24 months of age (Roe and Coates, 1989; Coates, 1992). Patients can present with a range of signs from clinically asymptomatic to sudden death. MCAD patients usually present within the first 2 years of life with episodes of lethargy, vomiting and coma precipitated by fasting. Patient history may indicate a recent upper respiratory or gastrointestinal virus and may include similar previous episodes. Generally, recurrent illness decreases in frequency after 4 years of age. Patients experiencing clinical episodes

develop hypoglycemia, impaired ketogenesis, hyperammonemia and increased levels of liver transaminases, acidosis and aciduria predominantly consisting of medium-chain dicarboxylic acids. Patients also develop a secondary carnitine deficiency. Patients can develop micro- and macro-vesicular fatty liver degeneration and encephalopathy (Roe and Coates, 1989; Coates, 1992). Additionally, sudden infant death, recurrent Reye syndrome and episodic hypoglycemic coma have all been described in MCAD deficiencies (Strauss *et al.*, 1990).

The mortality among MCAD deficient children is significant. Each episodic illness, whether the first or a repeated episode, can lead to death. In one study, 4 of 30 children (13%) died during their first episode in the first year of life. Fifteen of 35 (43%) of children died after 12 months of age. Mortality is highest (59%) between the ages of 15 and 26 months (Roe and Coates, 1989).

The incidence of MCAD deficiency has been estimated at 1 in 10,000 to 1 in 25,000 in the Caucasian population (Bennett *et al.*, 1987; Roe and Coates, 1989). Another study reported incidence of the disease to be 1 in 6,400 in England, 1 in 20,000 in Australia and 1 in 46,000 in North America (Matsubara *et al.*, 1992). The carrier frequency, therefore, has been estimated to be between 1 in 50-80 in the Caucasian population (Matsubara *et al.*, 1992). MCAD deficiency, therefore, is as common as phenylketonuria, a disorder for which a newborn mass screening program exists in several countries (Scriver *et al.*, 1989).

The most prevalent mutation in MCAD deficient patients is a point mutation at nucleotide 985 of the MCAD cDNA (Roe and Coates, 1989; Ding *et al.*, 1992; Matsubara *et al.*, 1992; Coates, 1992; Kelly *et al.*, 1990; Strauss, 1990; Workshop on Molecular

Aspects of MCAD Deficiency, 1992). This mutation is a deoxyadenine (A) to deoxyguanine (G) replacement which results in the substitution of lysine (K) for a glutamic acid (E) at amino acid position 329 of the precursor MCAD protein or at amino acid 304 of the mature subunit. This K329E mutation is present in the majority of the MCAD deficient patients. It has been estimated that approximately 90% of affected patients are homozygous for this mutation and approximately 7% are compound heterozygotes for this mutation (Workshop on Molecular Aspects of MCAD Deficiency, 1992). No other mutation has been described with a frequency greater than 1%.

Other mutations have been identified in MCAD deficient patients. One patient had a mutant allele which contained a 13 basepair tandem repeat at position 999 of the MCAD cDNA, resulting in a premature stop codon (Yokota *et al.*, 1990). Another patient had a 12 basepair insert in the region of the MCAD gene encoding for the mitochondrial transit peptide (Strauss, 1990). This probably disturbed the secondary structure of the leader transit peptide, causing improper or inefficient translocation of the MCAD precursor into mitochondria. Also, a patient has been identified with a 4 base pair deletion at nucleotide 1100 (Workshop on Molecular Aspects of MCAD Deficiency, 1992). Additionally, many other point mutations have been identified: substitutions of G to A at nucleotide 799, C to T at nucleotide 157, G to A at nucleotide 447, T to C at nucleotide 730, T to C at nucleotide 1124. These result in amino acid substitutions of glycine to arginine at amino acid 267, arginine to cysteine at amino acid 53, methionine to isoleucine at amino acid 149, cysteine to arginine at amino acid 244 and isoleucine to threonine at amino acid 375, respectively (Workshop on Molecular Aspects of MCAD Deficiency, 1992).

Interestingly, mRNA from patients with homozygous K329E point mutations have been shown to undergo missplicing (Kelly *et al.*, 1990; Zhang *et al.*, 1992). Analysis of cDNA clones derived from a library made from liver poly(A) RNA of a Dutch MCAD deficient patient revealed that the cDNAs underwent deletions due to loss of exons and insertions from retention of introns. Specifically, cDNA clones contained a deletion of exon 2 (11/85 cDNA clones), a deletion of exon 2, exon 5 and exons 8-10 in one clone and various other exon deletions, most using normal splicing donor and acceptor sites. Additionally, cDNA clones retained intron 3, the 5' part of intron 11 and a 12 bp insertion between exons 1 and 2. Approximately 29% of the mRNA was abnormally misspliced. The rate of abnormal missplicing in this patient was 29% which is much higher than the rate of missplicing in normal MCAD individuals (Zang *et al.*, 1992). The splicing abnormality may, therefore, be an occurrence that is enhanced by the K329E mutation.

The lysine residue at 329 is conserved between rat and human species (Kelly *et al.*, 1987; Matsubara *et al.*, 1987). This lysine is located in the H-helix region of the C-terminal domain of the MCAD monomer. The H and I helices are responsible for subunit contacts of the MCAD monomer to form a homotetramer. The substitution of lysine, a basic residue, with glutamic acid, an acidic residue, would change the charge of this region. It is believed that this point mutation results in interference with normal MCAD homotetramer formation, thus abolishing MCAD activity (Kim and Wu, 1988; Kim *et al.*, 1993).

It has been suggested that the mutant MCAD protein due to the K329E mutation is unstable. Western blot studies on MCAD patient fibroblasts using anti-MCAD antibody

raised in rabbits immunized with rat liver purified MCAD protein revealed that patients had very little or no MCAD antigen (Ikeda and Tanaka, 1987; Coates, 1992). MCAD protein undergoes normal translation and immediate post-translational processing in MCAD deficient patients (Ikeda *et al.*, 1986). Since there is very little or no steady state MCAD protein detectable on immunoblots, the mutant MCAD protein within the mitochondria is thought to be unstable (Coates, 1992). Other reports, however, determined that mutant MCAD in patient fibroblasts were detectable by immunoblot assays (Strauss *et al.*, 1990; Kelly *et al.*, 1990). These latest studies used anti-pig kidney MCAD antibodies and different techniques. It is, therefore, unknown whether MCAD patients have unstable MCAD protein. The instability of MCAD protein in patients, if this exists, may be related to the inability of the MCAD monomers to stably form the homotetramer required for MCAD activity (Coates, 1992).

Pathogenesis

A disruption of β-oxidation results in increased acyl-CoA molecules in the liver which participate in triglyceride synthesis and storage (Treem *et al.*, 1986). This results in an accumulation of lipid in micro- and macrovesicular cytoplasmic droplets in the hepatocytes.

Alternate oxidation pathways lead to the formation of toxic metabolites. Excessive acyl-CoA molecules undergo omega (ω) oxidation in liver smooth endoplasmic reticulum via cytochrome P450-linked mixed function oxygenase (Roe and Coates, 1989). Subsequent NAD+ dependent oxidation converts the ω-hydroxy fatty acid into dicarboxylic acids. These dicarboxylic acids are transported from the mitochondrial matrix to the peroxisomes for perixosomal β-oxidation (Gregersen *et al.*, 1983). These

alternate pathways cannot handle the excessive load of intermediate acyl-CoA compounds, resulting in the accumulation of dicarboxylic acids. These excessive dicarboxylic acids are excreted in the urine. Additionally, ω-1 hydroxy acids also have been seen in MCAD deficient patients (Colle *et al.*, 1983). These acids also lead to metabolic acidosis in the patient.

MCAD deficiency leads to decreased production of acetyl-CoA molecules. Decreased production of acetyl-CoA compromises the TCA cycle, gluconeogenesis and ketone formation. The deficiency of acetyl-CoA also leads to the decreased production of ketone bodies. Under fasting conditions the liver is responsible for the production of ketone bodies to be used as a substrate for brain oxidative metabolism. A block in β-oxidation, however, results in decreased levels of acetyl-CoA and acetoacetyl-CoA molecules, which in turn reduces ketogenesis (Roe and Coates, 1989).

MCAD patients also develop a secondary carnitine deficiency. Carnitine is required for several functions, including the transport of long-chain acyl-CoA molecules into the mitochondria and transport of acetyl-CoA molecules between the mitochondrial matrix and cytoplasm. Additionally, carnitine may be involved in the elimination of excessive acyl-CoA compounds by urinary excretion of acylcarnitine molecules (Bremer, 1983, Rebouche and Paulson, 1986, Stumpf *et al.*, 1985). Specifically, these functions are mediated by carnitine palmitoyltransferase I and II and carnitine acetyltransferase.

MCAD deficiency results in significant decreases in total body carnitine. Specifically, patients have total carnitine levels decreased to 25% of normal in muscle, plasma and liver (Stanley *et al.*, 1983). Other acyl-CoA oxidation defects such as LCAD

deficiency and isovaleryl-CoA dehydrogenase deficiency also result in secondary carnitine deficiency (Stanley, 1987).

MCAD deficiency, therefore, is an important disease with a complex pathogenesis. Understanding the molecular genetic mechanisms involved in normal and disease biology of this enzyme and its related pathways will require a mammalian animal model. Despite intensive effort to identify spontaneous animal models (Wood, unpublished), no animal model of MCAD deficiency has been found. Recent advances in the use of embryonic stem cells to induce mutations make developing an MCAD deficient mouse model feasible.

Mouse Embryonic Stem Cells

Site directed mutagenesis in embryonic stem (ES) cells provides a means of generating mice with specific alterations in their genome. This technology provides a means of developing mouse models of human disease. To date, hundreds of different genes have been targeted in ES cells and successfully introduced into blastocysts to produce mutant mice having specific altered phenotypes. Two different advancements led to the development of this new technology: (i) the establishment of ES cells in culture with subsequent introduction of these cells into the germline of mice and (ii) the establishment of a successful means for introducing homologous recombination in ES cells, along with methods to identify these cells. Both of these innovations led to the establishment of the ES cell system to generate mice with specific genetic alterations.

Mouse ES cells can be grown in culture and reintroduced into blastocysts to generate mice capable of germline transmission of the genotype derived from ES cells (Bradley *et al.*, 1984; Evans and Kaufman, 1981). Mouse ES cells are established from

the inner cell mass of mouse blastocysts. Pluripotential murine ES cells can be maintained in culture in their undifferentiated state. ES cells are grown on a mitotically inactivated feeder cell layer as described by Robertson (1987). Most commonly, these ES cells are grown on a feeder cell layer consisting of either embryonic fibroblasts or SNL 76/7 cells that produce a soluble factor called leukemia inhibitory factor (LIF) and possibly other factors that are necessary for maintaining the ES cells in an undifferentiated state (Nichols *et al.*, 1990). The SNL 76/7 feeder cell layer was originally derived from the STO cell line which had been stably transfected with RV4.0, a neomycin expression vector, and an expression vector producing the LIF growth factor which is responsible for maintaining the ES cells in an undifferentiated state (McMahon and Bradley, 1990; Nichols *et al.*, 1990; Thomas and Capecchi, 1987). ES cells also can be maintained in an undifferentiated state in media supplemented with LIF without the feeder cell layer (Nichols *et al.*, 1990). These pluripotent ES cells are microinjected into mouse blastocysts which are then implanted into a foster mouse. Mice produced from this foster female are chimeric if the ES cells have contributed somatic or germline cells of the pup. If the ES cells have contributed to formation of gonadal tissue in the chimeric mice, they are considered germline chimeras. These germline chimeric mice, when bred to wild type mice, result in offspring with heterozygous genotype derived from ES cells. These heterozygous mice, when bred to each other, can produce mice with the desired genotype derived from the original ES cell line.

The means of introducing homologous recombination with exogenous vector sequence in mouse ES cells and subsequent identification of these ES cells was the second advancement leading to the development of ES cell technology (Smithies *et al.*, 1985;

Thomas and Capecchi, 1987; Thomas *et al.*, 1986). Gene transfer of targeting vector DNA into ES cells is usually accomplished by electroporation. Approximately 0.1 % of ES cells stably integrate the targeting vector following electroporation (Frohman and Martin, 1989). Vector constructs also have been introduced via microinjection directly into the ES cell nucleus to obtain a targeted event (Zimmer and Gruss, 1989); however, this is extremely labor intensive. Only a small fraction of the ES cells integrating DNA usually undergoes homologous recombination with the exogenous vector. Homologous recombination is 3- to 40,000-fold less frequent than random integration (DeChiara *et al.*, 1990; Doetschman *et al.*, 1988; Hasty *et al.*, 1991b; Johnson *et al.*, 1989; Joyner *et al.*, 1989; Soriano *et al.*, 1991; Thomas and Capecchi, 1987; Jung *et al.*, 1993). Selection schemes have been designed to enrich for the ES cell colonies that undergo homologous recombination, and schemes will be discussed in detail below.

Generation of Mice with Specific Alterations

After DNA transfection via electroporation, ES cell clones usually are taken through selection in culture to screen for cells that integrate the construct homologously. The vector to be electroporated usually is designed to contain positive and negative selection genes whereby the transfected ES cells lacking stable integrations and random integrations can be eliminated from culture. The positive selection gene is generally a neomycin phosphotransferase gene (NEO), including a promoter, that is inserted within an exon of the cloned gene on the targeting vector. ES cells that integrate the targeting construct after electroporation contain the NEO gene. These ES cells containing the NEO cassette are resistant to G-418, a neomycin analog added to culture. ES cells not integrating the vector (thus not having integrated a NEO gene) are killed during G-418

selection in culture. The herpes simplex virus thymidine kinase (HSV-TK) gene is the most common negative selection marker used in ES cell vectors. The HSV-TK gene is placed usually on one end but can be placed on both ends of the homologous cloned sequence within the vector. ES cells integrating and expressing the HSV-TK gene do not survive when gancyclovir (GANC) is added in the culture medium. ES cells which randomly integrate the vector usually contain both the NEO and HSV-TK genes and therefore die when G-418 and GANC are added to culture medium. ES cells undergoing homologous recombination, on the other hand, should contain the NEO gene but not the HSV-TK gene and therefore survive when challenged with G-418 and GANC in culture. Therefore, when G-418 and GANC are added in culture, ES cells that have randomly integrated the vector construct should die, while ES cells that have undergone homologous recombination should survive. The surviving colonies can then be additionally screened by polymerase chain reaction (PCR). These ES cell colonies can be confirmed for the targeting event via restriction fragment length analysis. Approximately 10 to 15 ES cells which have been confirmed for the targeted event can then be microinjected into 3.5 day old mouse blastocysts. These blastocysts are implanted into foster mothers to potentially develop into a germline chimeric mouse. Successful germline transmission into the next generation leads to a heterozygous mouse with the desired mutation. Eventually, a mouse homozygous for the mutation can be produced by crossing heterozygotes.

One potentially laborious portion of the ES cell system is the identification of the targeted cell line over the background of the ES cells randomly integrating the vector. This is conditional on the targeting frequency of the locus being targeted. If the locus targets with poor frequencies, then substantial numbers of ES cell colonies surviving

G-418 or G-418 and GANC selection must be analyzed. The use of different selection genes may reduce the background of ES cell colonies randomly integrating the vector. If the gene being targeted is expressed in ES cells, then promoterless selection genes can be used. Promoterless NEO genes have been used to reduce the survival in culture of ES cells randomly integrating the vector. A promoterless NEO gene fused in-frame within the homologous sequences in the vector would result in the production of a fusion protein if the vector targets the transcriptionally active locus, resulting in the survival of the ES cell in G-418 selection. Random incorporation of this vector should result in no expression of the NEO gene; thus, this ES cell will not survive G-418 selection. Promoterless NEO genes have been used to target the polyoma middle T antigen in the genome of NIH 3T3 cells which resulted in a 100-fold improvement of selection enrichment by reducing the number of cells surviving selection that have random incorporation of vector DNA (Sedivy and Sharp, 1989). Another selection cassette has been designed as a "promoter trap" where the promoterless NEO gene is preceded by the adenovirus major late transcriptional splice acceptor site (Friedrich and Soriano, 1991). This expression cassette can be used in targeting vector constructs to enrich for selection of targeted cells in culture.

Recent advances have led to the development of techniques other than microinjection to introduce ES cells into blastocysts. Microinjection involves the injection of 10 to 15 ES cells into a blastocoel cavity, followed by the transfer of the blastocyst into pseudopregnant recipients. This microinjection technique requires a high level of expertise and expensive equipment. Recently, other techniques have been devised that do not require the use of expensive equipment. These co-culture techniques rely on the

ability of ES cells to aggregate with morulae and subsequently be efficiently internalized by the blastocyst stage *in vitro*. One method, the aggregation method, involves the sandwiching of a clump of ES cells between two eight cell morulas within a depression created by a darning needle (Hasty and Bradley, 1993; Nagy *et al.*, 1990). This technique also has been sucessfully used with tetraploid embryos (Nagy *et al.*, 1993). Another method, the co-culture method, involves the co-culture of eight cell embryos on a lawn of ES cells (Wood *et al.*, 1993). All of these methods have produced germline chimeras.

Replacement vs. Insertion Vector Design

Two different vector designs have been used for targeting DNA sequences in ES cells. The first and most commonly used is the replacement vector design, wherein the vector is linearized outside the region of homology and the homologous sequences within the vector replace the homologous sequences within the ES cell genome by a double reciprocal recombination or a gene conversion event. The second is the insertion vector design, wherein the vector is digested within an area of homology and gene targeting occurs with the insertion of the entire construct, including the vector backbone, integrating into the ES cell genome via a single reciprocal recombination. This results in the partial duplication of portions of the targeted gene. Insertion vectors have been used to target the hypoxanthine phosphoribosyltransferase (HPRT) gene, the *Hox* 2.6 gene and the cystic fibrosis gene (Hasty *et al.*, 1991a; Valancius and Smithies, 1991; Hasty *et al.*, 1991b; Dorin *et al.*, 1992; Kumar and Simons, 1993). Additionally, insertion vectors have been used to correct mutant HPRT gene in ES cell line (Doetschman *et al.*, 1987; Thompson *et al.*, 1989).

Comparisons of targeting frequencies between the two vector types initially indicated that insertion vectors target with up to 9-fold greater frequency than the replacement vectors at the HPRT locus (Hasty *et al.*, 1991b). These targeting events were done with a replacement vector containing only 1 kilobase (kb) of homology 3' to the NEO cassette. Another investigation, however, has shown that targeting frequency between replacement and insertion vectors is approximately the same when the shortest arms of the replacement vectors contain at least 3 kb of homology (Deng and Capecchi, 1992).

Homologous Recombination

The mechanism of homologous recombination in mammalian cells is poorly understood. The double-strand-break repair model is the most widely accepted model of how a vector can homologously recombine with endogenous sequence (Szostak *et al.*, 1983). In this model recombination is initiated with a double stranded break in the vector.

Homologous recombination in ES cells is very inefficient. Modifications of the vector can increase frequency of homologous recombination. Increasing the amount of homology to the targeted gene can increase targeting frequency. A 2-fold increase in homology increases the targeting frequency by 20-fold at the HPRT locus (Thomas and Capecchi, 1987). Mammalian recombination in ES cells seems to be most efficient when approximately 14 kb of homologous DNA is used in the targeting vector (Deng and Capecchi, 1992). The use of isogenic DNA (homologous sequence of DNA on the targeting vector cloned from 129/Sv mouse strain which is the same mouse strain from which ES cells are derived) also increases targeting frequency. Polymorphisms exist

between the different mouse strains which may decrease targeting frequency dramatically (Waldman and Liskay, 1988). Efficient homologous recombination requires perfect homology between the targeting vector and the endogenous sequence being targeted. Targeting frequencies at the retinoblastoma locus were 10 to 20 times higher when isogenic DNA was used as compared to nonisogenic DNA (Te Riele *et al.*, 1992). Vectors containing isogenic DNA targeted 5 to 10 times better than nonisogenic DNA containing vectors at the HPRT locus (Deng and Capecchi, 1992). Isogenic vectors targeted the creatine kinase M gene 25 times more efficiently (Deursen and Wieringa, 1992). Additionally, the introduction of a double stranded break in the vector DNA, either within or outside the area of targeting homology, can increase targeting frequency. Increased targeting due to a double stranded break has been demonstrated in yeast (Orr-Weaver et al., 1981) and in extrachromosomal recombination in mammalian cells (Kucherlapati *et al.*, 1984; Song *et al.*, 1985; Wong and Capecchi, 1986). Introduction of a double stranded break to linearize the vector increased targeting by 10- to 1000-fold (Kucherlapati *et al.*, 1984; Orr-Weaver *et al.*, 1981). It also has been speculated that actively transcribed genes in ES cells may be more efficient at homologous recombination than genes not actively expressed (Frohman and Martin, 1989; Nickoloff and Reynolds, 1990; Thomas and Capecchi, 1987).

Random Integration

Free DNA ends are considered recombinogenic for both random and targeted integration (Chang and Wilson, 1987). Random integration is proposed to occur by ligation of the vector ends to transient chromosomal breaks (Folger *et al.*, 1982; Thomas *et al.*, 1986; Wilson *et al.*, 1982). Random integration or end-to-end joining of DNA has

been proposed to occur by three different mechanisms. First, there is the homology independent mechanism, where protruding ends of the two DNA molecules are joined and the missing nucleotides are filled in. The next two mechanisms require short one to six bases of homology. During the initial alignment phase these short regions of homology pair with each other, followed by mismatch repair of the unpaired nucleotides (Roth *et al.*, 1985; Roth and Wilson, 1986). Also proposed are DNA alignment proteins which bind to DNA ends and align ends of a joining reaction so that the missing nucleotides on one DNA molecule can be filled in by DNA polymerases from the other DNA molecule.

Summary of Gene Targeting in ES cells

Gene targeting in ES cells is a powerful tool in developing mice with specific alterations in their genome. By using certain guidelines of vector design such as using isogenic DNA, using considerable amounts of homology in the targeting vector (i.e., 5-15 kb) and making sure that the short arm of homology is greater than 1 kb in length, one can usually optimize the frequency of targeted disruption for the gene of interest (Hasty and Bradley, 1993). Most gene targeting vectors use these guidelines. Even though a different locus may target with a dramatically different frequency when compared to another locus, targeting frequencies of different genes have generally been fairly efficient. According to recent publications, analysis of approximately 300 G-418 resistant colonies is generally sufficient to identify a targeted cell line (Fig. 2). Additionally, if positive-negative selection is used, the analysis of approximately 100 G-418 and GANC resistant colonies is usually sufficient to identify a targeted cell line (Fig. 2). Different genes target with different frequencies. Some genes target with low frequencies. It is unclear why these differences exist. Some loci may be more accessible to the targeting vector than

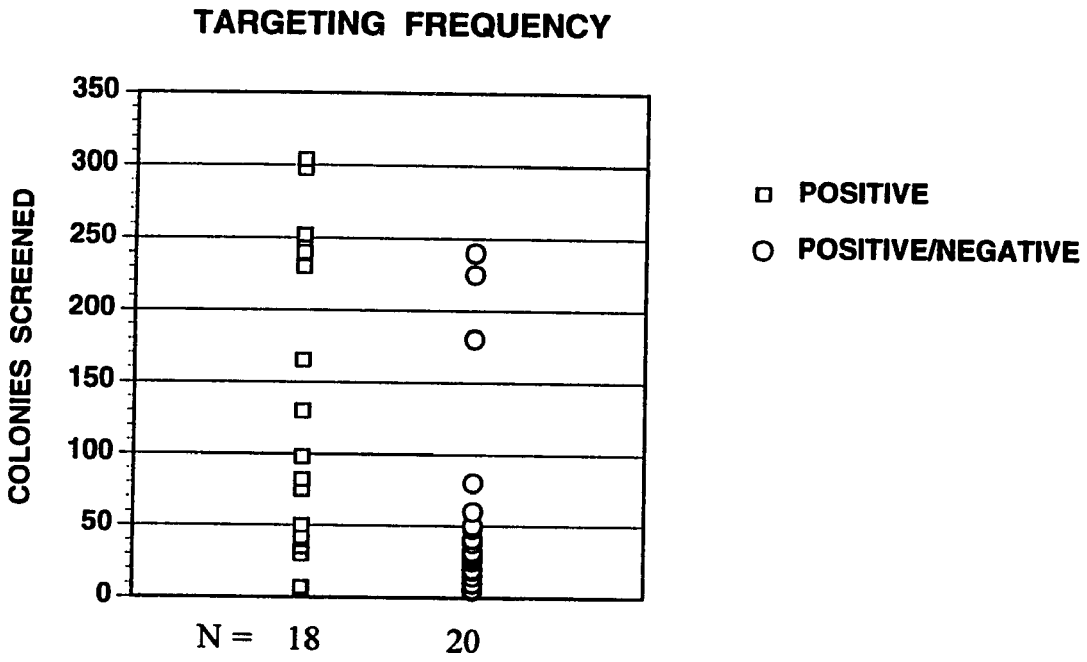
others. Transcriptionally active genes may be more susceptible to targeted disruption; however, this is controversial (Frohman and Martin, 1989; Nickoloff and Reynolds, 1990; Thomas and Rothstein, 1989). Repeat sequences within regions of vector homology may decrease targeting frequency. Nucleotide mismatches between vector homology and endogenous sequences may decrease targeting frequency. The reasons for differences in targeting frequencies between different loci are poorly understood. Fortunately, most loci seem to be targetable at fairly efficient frequencies, and this is a powerful tool for developing mice with specific genetic alterations.

Rationale of Dissertation

MCAD deficiency is a very common inborn error of metabolism associated with significant morbidity and mortality. Patients exhibiting clinical episodes have a wide range of metabolic problems affecting many biochemical pathways. The pathogenesis of MCAD deficiency is rather complex and poorly understood. Questions also exist regarding the proper management and treatment of MCAD deficient patients; for example, is carnitine supplementation necessary? Currently, no animal model exists for MCAD deficiency. An animal model would be useful in helping to better characterize the disease and in developing improved clinical treatment and management regimens. Furthermore, the influence of intracellular fatty acid levels on gene transcription of other acyl-CoA dehydrogenase genes and on gluconeogenic and urea cycle enzyme genes could be determined. A powerful tool existed to develop mouse models of human disease. We proposed to use ES cell technology to develop a MCAD deficient mouse model to better understand MCAD deficiency.

To develop a mouse model for MCAD deficiency using site-directed mutagenesis of ES cells, we had to clone mouse MCAD genomic fragments to be used for developing targeting vectors. Suitable probes had to be generated to screen mouse genomic libraries. The mouse MCAD cDNA would be cloned for use as a probe to screen mouse genomic libraries. The cDNA could also be used to better characterize MCAD genomic fragments. Mouse MCAD genomic fragments were to be cloned and characterized and used to develop MCAD targeting vectors. The cloning on mouse MCAD cDNA could provide important insights into the structure of the MCAD protein and its similarity when compared to other species and could highlight the possible important conserved regions of the protein. The cloning of the genomic gene and regulatory elements could provide a better understanding of transcriptional regulation of the mouse MCAD gene. Comparisons could be made to the human gene to determine similarities and differences in regulatory elements. Additionally, cDNA and genomic fragments could be used to definitively map the location of the mouse MCAD gene. Important information, therefore, could be obtained along the way to developing a mouse model for MCAD deficiency.

FIG. 2. Targeting frequency with positive and positive and negative selection. Targeting frequencies between different loci vary considerably. A review of published results on targeting frequencies, however, indicates screening approximately 100 ES cell colonies on positive and negative selection is usually sufficient to identify a targeted cell line. With ES cells on positive selection only, screening approximately 300 colonies is generally sufficient to identify a disrupted cell line.


TARGETING FREQUENCY
COLONIES SCREENED
350
300
250
200
150
100
50
0
N = 18 20
POSITIVE
POSITIVE/NEGATIVE

REFERENCES FOR INTRODUCTION

Bahary, N., Zorich, G., Pachter, J.E., Rudolph, L.L., and Friedman, J.M. (1991). Molecular genetic linkage maps of mouse chromosome 4 and 6. *Genomics* **11**: 33-47.

Bennett, M.J., Worthy, E., and Pollitt, R.J. (1987). The incidence and presentation of dicarboxylic aciduria. *J. Inherited Metab. Dis.* **13**: 241-242.

Bradley, A., Evans, M., Kaufman, M.H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309**: 255-256.

Bremer, J. (1983). Carnitine-metabolism and functions. *Physiol. Rev.* **63**: 1420.

Carter, M.E., Gulick, T., Moore, D.D., and Kelly, D.P. (1994). A pleiotropic element in the medium-chain acyl-coenzyme A dehydrogenase gene promoter mediates transcriptional regulation by multiple nuclear receptor transcription factors and defines novel receptor-DNA binding motifs. *Mol. Cell Biol.* **14**: 4360-4372.

Carter, M.E., Gulick, T., Raisher, B.D., Caira, T., Ladias, J.A.A., Moore, D.D., and Kelly, D.P. (1993). Hepatocyte nuclear factor-4 activates medium chain acyl-CoA gene transcription by interacting with a complex regulatory element. *J. Biol. Chem.* **268**: 13805-13810.

Catzeflis, C., Bachmann, C., Hale, D.E., Coates, P.M., Wiesmann, U., Colombo, J.P., Joris, F., and Deleze, G. (1990). Early diagnosis and treatment of neonatal medium-chain acyl-CoA dehydrogenase deficiency: report of two siblings. *Eur. J. Pediatr.* **149**: 577-581.

Chang, X.-B., and Wilson, J.H. (1987). Modification of DNA ends can decrease end joining relative to homologous recombination in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **84**: 4959-4963.

Coates, P.M. (1992). Historical perspective of medium-chain acyl-CoA dehydrogenase deficiency: a decade of discovery. In "New Developments in Fatty Acid Oxidation", eds. Coates PM and K Tanaka K. (Wiley-Liss, New York). pp.409-423.

Colle, E., Mamer, O.A., Montgomery, J.A., and Miller, J.D. (1983). Episodic hypoglycemia with omega-hydroxy fatty acid excretion. *Pediatr. Res.* **17**: 171-176.

DeChiara, T.M., Efstratiadis, A., and Robertson, E.J. (1990). A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. *Nature* **345**: 78-80.

Deng, C., and Capecchi, M.R. (1992). Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and targeting locus. *Mol. Cell. Biol.* **12**: 3365-3371.

Deursen, J.-V., and Wieringa, B. (1992). Targeting of the creatine kinase M gene in embryonic stem cells using isogenic and nonisogenic vectors. *Nucleic Acids Res.* **20**: 3815-3820.

Ding, J.H., Bross, P., Yang, B.Z., Iafolla, A.K., Millington, D.S., Roe, C.R., Gregersen, N., and Chen, Y.T. (1992). Genetic heterogeneity in MCAD deficiency: frequency of K329E allele and identification of three additional mutant alleles. In "New Developments in Fatty Acid Oxidation", eds. Coates PM and K Tanaka K. (Wiley-Liss, New York). pp. 479-488.

Doetschman, T.C., Gregg, R.G., Maeda, N., Hooper, M.L., Melton, D.W., Thompson, S., and Smithies, O. (1987). Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature* **330**: 576-578.

Doetschman, T.C., Maeda, N., and Smithies, O. (1988). Targeted mutation of the *HPRT* gene in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **85**: 8583-8587.

Dorin, J.R., Dickinson, P., Alton, E.W.F.W., Smith, S.N., Geddes, D.M., Stevenson, B.J., Kimber, W.L., Fleming, S., Clarke, A.R., Hooper, M.L., Anderson, L., Beddington, R.S.P., and Porteous, D.J. (1992). Cystic fibrosis in the mouse targeted insertional mutagenesis. *Nature* **359**: 211-215.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**: 154-156.

Folger, K.R., Wong, E.A., Geoffrey, W., and Capecchi, M. (1982). Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. *Mol. Cell. Biol.* **2**: 1372-1387.

Friedrich, G., and Soriano, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev.* **5**: 1513-1523.

Frohman, M.A., and Martin, G.R. (1989). Cut, paste, and save: new approaches to altering specific genes in mice. *Cell* **56**: 145-147.

Gregersen, N., Mortensen, P.B., and Kolvraa, S. (1983). On the biological origin of C6-C10-dicarboxylic and C6-C10-Omega-1-hydroxy monocarboxylic acids in human and rat

with acyl-CoA dehydrogenation deficiencies: in vitro studies on the omega- and omega-1-oxidation of medium-chain (C6-C12) fatty acids in human and rat liver. *Pediatr. Res.* **17**: 828.

Hasty, P., and Bradley, A. (1993). Gene targeting vectors for mammalian cells. In "Gene Targeting: A Practical Approach, ed. Joyner AL. (Oxford University Press Inc., New York). pp. 1-30.

Hasty, P., Ramirez-Solis, R., Krumlauf, R., and Bradley, A. (1991a). Introduction of a subtle mutation into the Hox-2.6 locus in embryonic stem cells. *Nature* **350**: 243-246.

Hasty, P., Rivera-Perez, J., Chang, C., and Bradley, A. (1991b). Target frequency and integration pattern for insertion and replacement vectors in embryonic stem cells. *Mol. Cell. Biol.* **11**: 4509-4517.

Ikeda, Y., Hale, D.E., Keese, S.M., Coates, P.M., and Tanaka, K. (1986). Biosynthesis of variant medium chain acyl-CoA dehydrogenase in cultured fibroblasts from patients with medium chain acyl-CoA dehydrogenase deficiency. *Pediatr. Res.* **20**: 843-847.

Ikeda, Y., Keese, S.M., Fenton, W.A., and Tanaka, K. (1987). Biosynthesis of four rat liver mitochondrial acyl-CoA dehydrogenases: In vitro synthesis, import into mitochondria, and processing of their precursors in a cell-free system and in cultured cell. *Arch. Biochem. Biophys.* **252**: 662-674.

Johnson, R.S., Sheng, M., Greenberg, M.E., Kolodner, R.D., Papaioannou, V.E., and Spiegelman, B.M. (1989). Targeting of nonexpressed genes in embryonic stem cells via homologous recombination. *Science* **245**: 1234-1236.

Joyner, A.L., Skarnes, W.C., and Rossant, J. (1989). Production of a mutation in mouse *En-2* gene by homologous recombination in embryonic stem cells. *Nature* **338**: 153-156.

Jung, S., Rajewsky, K., and Radbruch, A. (1993). Shutdown of class switch recombination by deletion of a switch region control element. *Science* **259**: 984-987.

Kelly, D.P., Gordon, J.I., Alpers, R., and Strauss, A.W. (1989). The tissue-specific expression and developmental regulation of two nuclear genes encoding rat mitochondrial proteins - medium chain acyl-CoA dehydrogenase and mitochondrial malate dehydrogenase. *J. Biol. Chem.* **264**: 18921-18925.

Kelly, D.P., Kim, J.-J., Billadello, J.J., Hainline, B.E., Chu, T.W., and Strauss, A.W. (1987). Nucleotide sequence of medium-chain acyl-CoA dehydrogenase mRNA and its expression in enzyme-deficient human tissue. *Proc. Natl. Acad. Sci. U.S.A.* **84**: 4068-4072.

Kelly, D.P., Whelan, A.J., Ogden, M.L., Alpers, R., Zhang, Z., Bellus, G., Gregersen, N., Dorland, L., and Strauss, A.W. (1990). Molecular characterization of inherited medium-chain acyl-CoA dehydrogenase deficiency. *Proc. Natl. Acad. Sci. U.S.A.* **87**: 9236-9240.

Kidd, J.R., Matsubara, Y., Castiglione, C.M., Tanaka, K., and Kidd, K.K. (1990). The locus for the medium-chain acyl-CoA dehydrogenase gene on chromosome 1 is highly polymorphic. *Genomics* **6**: 89-93.

Kim, J,-J.P., Wang, M., and Paschke, R. (1993). Crystal structures of medium-chain acyl-CoA dehydrogenase from pig liver mitochondria with and without substrate. *Proc. Natl. Acad. Sci. U.S.A.* **90**: 7523-7527.

Kim, J.-J., and Wu, J. (1988). Structure of the medium-chain acyl-CoA dehydrogenase from pig liver mitochondria at 3-A resolution. *Proc. Natl. Acad. Sci. U.S.A.* **85**: 6677-6681.

Kolvraa, S., Gregersen, N., Christensen, E., and Hobolth, N. (1982). In vitro fibroblast studies in a patient with C6-C10-dicarboxylic aciduria: Evidence for a defect in general acyl-CoA dehydrogenase. *Clin. Chim. Acta* **126**: 53.

Kucherlapati, R.S., Eves, E.M., Song, K.-S., Morse, B.S., and Smithies, O. (1984). Homologous recombination between plasmids in mammalian cells can be enhanced by treatment of input DNA. *Proc. Natl. Acad. Sci. U.S.A.* **81**: 3153-3157.

Kumar, S., and Simons, J.P. (1993). The effects of terminal heterologies on gene targeting by insertion vectors in embryonic stem cells. *Nucleic Acids Res.* **21**:1541-1548.

Matsubara, Y., Kraus, J.P., Ozasa, H., Glassberg, R., Finnocchiaro, G., Ikeda, Y., Mole, J., Rosenberg, L.E., and Tanaka, K. (1987). Molecular cloning and nucleotide sequence of cDNA encoding the entire precursor of rat liver medium chain acyl-CoA dehydrogenase. *J. Biol. Chem.* **262**: 10104-10108.

Matsubara, Y., Narisawa, K., Miyabayashi, S., Tada, K., Coates, P.M., Bachmann, C., Elsas, I.I., Pollitt, R.J., Rhead, W.J., and Roe, C.R. (1990). Identification of a common mutation in patients with medium-chain acyl-CoA dehydrogenase deficiency. *Biochem. Biophys. Res. Comm.* **171**: 498-505.

Matsubara, Y., Narisawa, K., and Tada, K. (1992). Medium-chain acyl-CoA dehydrogenase deficiency: molecular aspects. *Eur. J. Pediatr.* **151**: 154-159.

McMahon, A.P., and Bradley, A. (1990). The *Wnt-1* (*int-1*) proto-oncogene is required for development of a large region of the mouse brain. *Cell* **62**: 1073-1085.

Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W., and Roder, J.C. (1993). Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **90**: 8424-8428.

Nichols, J., Evans, E.P., and Smith, A.G. (1990). Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity. *Development* **110**: 1341-1348.

Nickoloff, J.A., and Reynolds, R.J. (1990). Transcription stimulates homologous recombination in mammalian cells. *Mol. Cell. Biol.* **10**: 4837-4845.

Orr-Weaver, T.L., Szostak, J.W., and Rothstein, R.J. (1981). Yeast transformation: a model system for the study of recombination. *Proc. Natl. Acad. Sci. U.S.A.* **78**: 6354-6358.

Raisher, B.D., Gulick, T., Zhang, Z., Strauss, A.W., Moore, D.D., and Kelly, D.P. (1992). Identification of a novel retinoid-responsive element in the promoter region of the medium chain acyl-coenzyme A dehydrogenase gene. *J. Biol. Chem.* **267**: 20264-20269.

Rebouche, C.J., and Paulson, D.J. (1986). Carnitine metabolism and function in humans. *Ann. Rev. Nutr.* **6**: 41.

Robertson, E.J. (1987). Embryo-derived stem cell lines. In "Teratocarcinomas and Embryonic Stem Cell Lines, a Practical Approach," ed. Robertson EJ. (IRL Press, Oxford). pp. 71-112.

Roe, C.R., and Coates, P.M. (1989). Acyl-CoA dehydrogenase deficiencies. In "The Metabolic Basis of Inherited Disease," eds. Scriver C.R., Beaudet A.L., Sly W.S., and Valle D. (McGraw-Hill, Inc., New York). pp. 889-914.

Roth, D.B., Porter, T.N., and Wilson, J.H. (1985). Mechanisms of nonhomologous recombination in mammalian cells. *Mol. Cell. Biol.* **5**: 2599-2607.

Roth, D.B., and Wilson, J.H. (1986). Nonhomologous recombination in mammalian cells: role of short sequence homology in the joining reaction. *Mol. Cell. Biol.* **6**: 4295-4304.

Scriver, C.R., Kaufman, S., and Woo, S.L.C. (1989). The hyperphenylalaninemias. In: "The Metabolic Basis of Inherited Disease," eds. Scriver CR et al. (McGraw Hill, 6th ed., New York). pp. 495-546.

Sedivy, J.M. and Sharp, P.A. (1989). Positive genetic selection for gene disruption in mammalian cells by homologous recombination. *Proc. Natl. Acad. Sci. U.S.A.* **86**:227-231.

Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A., and Kucherlapati, R.S. (1985). Insertion of DNA sequences into the human chromosomal β-globin locus by homologous recombination. *Nature* **317**: 230-234.

Song, K.-Y., Chekuri, L., Rauth, S., and Kucherlapati, R. (1985). Effect of double stranded breaks on homologous recombination in mammalian cells and extracts. *Mol. Cell. Biol.* **5**: 3331-3336.

Soriano, P., Montgomery, C., Geske, R., and Bradley, A. (1991). Targeted disruption of the *c-src* proto-oncogene leads to osteopetrosis in mice. *Cell* **84**: 693-702.

Stanley, C.A. (1987). New genetic defects in mitochondrial fatty acid oxidation and carnitine deficiency. *Adv. Pediatr.* **34**: 59-86.

Stanley, C.A., Hale, D.E., and Coates, P.M. (1983). Medium-chain acyl-CoA dehydrogenase deficiency in children with nonketotic hypoglycemia and low carnitine levels. *Pediatr. Res.* **17**: 877-884.

Strauss, A.W., Duran, M., Zhang, Z., Alpers, R., and Kelly, D.P. (1990). Molecular analysis of medium-chain acyl-CoA dehydrogenase deficiency. In: "Fatty Acid Oxidation: Clinical, Biochemical, and Molecular Aspects," eds. Tanaka K, and Coates PM. (Alan Liss, Inc., New York). pp. 609-623.

Stumpf, D.A., Parker, W.D., and Angelini, C. (1985). Carnitine deficiency, organic acidemias, and Reye's syndrome. *Neurology* **35**: 1041.

Szostak, J.W., Orr-Weaver, T.L., and Rothstein, R.J. (1983). The double-strand-break repair model for recombination. *Cell* **33**: 25-35.

Te Riele, H., Maandag, E.R., and Berns, A. (1992). Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. *Proc. Natl. Acad. Sci. U.S.A.* **89**: 5128-5132.

Thomas, B.J., and Rothstein, R. (1989). Elevated recombination rates in transcriptionally active DNA. *Cell* **56**: 619-630.

Thomas, K.R., and Capecchi, M.R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**: 503-512.

Thomas, K.R., Folger, K.R., and Capecchi, M.R. (1986). High frequency targeting of genes to specific sites in the mammalian genome. *Cell* **44**: 419-428.

Thompson, S., Clarke, A.R., Pow, A.M., Hooper, M.L., and Melton, D.W. (1989). Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. *Cell* **56**: 313-321.

Thorpe, C. (1979). Acyl-CoA dehydrogenase from pig kidney. Purification and properties. *Biochemistry* **18**: 331-337.

Treem, W.R., Witzleben, C.A., Piccoli, D.A., Stanley, C.A., Hale, D.E., Coates, P.M., and Watkins, J.B. (1986). Medium-chain and long-chain acyl-CoA dehydrogenase deficiency: clinical, pathological and ultrastructural differentiation from Reye's syndrome. *Hepatology* **6**: 1270-1278.

Valancius, V., and Smithies, O. (1991). Testing an "In-Out" targeting procedure for making subtle genomic modifications in mouse embryonic stem cells. *Mol. Cell. Biol.* **11**: 1402-1408.

Waldman, A.S., and Liskay, R.M. (1988). Dependence of intrachromosomal recombination in mammalian cells on uninterrupted homology. *Mol. Cell. Biol.* **8**: 5350-5357.

Wilson, J.H., Berget, P.B., Pipas, J.M. (1982). Somatic cells efficiently join unrelated DNA segments end-to-end. *Mol. Cell. Biol.* **2**: 1258-1269.

Wong, E.A., and Capecchi, M.R. (1986). Analysis of homologous recombination in cultured cells in transient expression and stable transformation assays. *Somat. Cell. Mol. Genet.* **12**: 63-72.

Wood, S.A., Pascoe, W.S., Schmidt, C., Kemler, R., Evans, M.J., and Allen, N.D. (1993). Simple and efficient production of embryonic stem cell-embryo chimeras by coculture. *Proc. Natl. Acad. Sci. U.S.A.* **90**: 4582-4585.

Workshop on Molecular Aspects of MCAD Deficiency. (1992). Mutations causing medium-chain acyl-CoA dehydrogenase deficiency: a collaborative compilation of the data from 172 patients. In: "New Developments in Fatty Acid Oxidation," eds. Coates PM, and Tanaka K. (Wiley-Liss, New York). pp. 499-506.

Yokota, I., Tanaka, K., Coates, P.M., and Ugarte, M. (1990). Mutations in medium chain acyl-CoA dehydrogenase deficiency. *Lancet* **11**: 748.

Zhang, Z., Kelly, D.P., Kim, J.-J., Zhou, Y., Ogden, M.L., Whelan, A.J., and Strauss, A.W. (1992). Structural organization and regulatory regions of the human medium-chain acyl-CoA dehydrogenase gene. *Biochemistry* **31**: 81-89.

Zimmer, A., and Gruss, P. (1989). Production of chimaeric mice containing embryonic stem (ES) cells carrying a homeobox *Hox 1.1* allele mutated by homologous recombination. *Nature* **338**: 150-153.

MOLECULAR CLONING AND CHARACTERIZATION OF THE MOUSE MEDIUM-CHAIN ACYL-COA DEHYDROGENASE CDNA

RAVI J. TOLWANI, SUSAN C. FARMER, AND PHILIP A. WOOD

Department of Comparative Medicine, Schools of Medicine and Dentistry, University of Alabama at Birmingham

Genomics 23: 247-249, 1994

Abstract

Medium-chain acyl-CoA dehydrogenase (MCAD) is one of the four straight-chain length-specific dehydrogenases involved in the first step of fatty acid oxidation. Inherited defects of acyl-CoA dehydrogenases occur in humans, and MCAD deficiency is the most common. We have cloned the coding and 3' untranslated sequence of mouse MCAD cDNA. The mouse MCAD cDNA coding region is 1263 basepairs long with a 3' untranslated region of 576 basepairs and it encodes a 421 amino acid precursor protein. Comparing the nucleotide and deduced amino acid sequences of the mouse MCAD cDNA to rat and human MCAD cDNAs reveals considerable similarity between species. Amino acid residues involved in specific enzymatic functions such as α,β dehydrogenation are conserved, as well as specific amino acid substitutions previously identified in human MCAD deficiency.

Introduction

Fatty acid oxidation occurs primarily within the liver and cardiac and skeletal muscle to provide energy especially during fasting. ß-oxidation consists of a repeating circuit of four sequential steps in which acyl-CoA dehydrogenases catalyze the first step. The three straight-chain acyl-CoA dehydrogenases, short-chain (SCAD), medium-chain (MCAD), long-chain (LCAD), and very long-chain (VLCAD) acyl-CoA dehydrogenase, act on different acyl-CoA substrate chain lengths. MCAD acts on acyl-CoA chain lengths of C4-C14 (Finocchiaro *et al.*, 1987). MCAD precursor proteins, as other dehydrogenases, are encoded by a nuclear gene, assembled in the cytosol, and translocated into the mitochondrial matrix simultaneously with removal of the 25 amino acid leader peptide sequence. Mature protein monomers, each containing one molecule of flavin adenine

dinucleotide (FAD), form homotetramers inside the mitochondrial matrix to form the active enzyme (Ikeda *et al.*, 1987).

Inherited metabolic defects of each straight-chain acyl-CoA dehydrogenase have been identified. MCAD deficiency is by far the most common with a carrier frequency in Caucasians of 1:62 and a homozygote frequency estimated at 1 in 15,000 (Matsubara *et al.*, 1992). Patients with MCAD deficiency develop clinical episodes, often associated with fasting, characterized by hypoglycemia, disrupted ketogenesis, medium chain dicarboxylic aciduria and hyperammonemia (Roe and Coates, 1989). MCAD deficiency is associated with considerable mortality during early childhood.

We have cloned and characterized the mouse MCAD cDNA. An oligo(dT) and random-primed mouse liver λgt10 cDNA library (Clontech) was screened to clone the 3' fragment of the mouse MCAD cDNA. An 800 bp 5' fragment of rat MCAD cDNA was labeled with ^{32}P via random hexamer priming and used as a probe to screen approximately 300,000 plaques at a density of 10,000 plaques per 150mm plate. Three positive clones were carried successfully through four rounds of screening. DNA isolated from clone #1, containing the largest insert, was subcloned into pGEM11Zf(+) at the *Eco*R1 and *Bam*H1 sites. Various primers were designed within the cDNA insert to sequence the entire cDNA fragment by the dideoxy-chain termination method (Sequenase Ver. 2, U.S.B.). This cDNA fragment was 1478 basepairs (bp) long and consisted of coding sequence from nucleotide number 359 to number 1263 and 576 bp of 3' untranslated sequence ending 21 bp from a putative polyadenylation signal.

We generated the 5' end of MCAD cDNA with reverse transcription-PCR (Fig. 1). Using oligo (dT) cellulose, poly(A) RNA was purified from total RNA isolated from

the liver of BALB/cBy mice (Davis *et al.*, 1986). Poly(A) RNA was reverse transcribed (RT) with either MCAD primer 8 or random hexamer primers (cDNA Cycle Kit, Invitrogen). An 19 mer oligo (MCAD primer 1) was designed spanning nucleotide -7 to +12 based on the sequence of a mouse MCAD pseudogene (unpublished results). Either 15 μl of the product from the RT reaction using MCAD primer 8 or 15 μl of the product from the RT reaction using random hexamer primers were used in PCRs with 15pmol each of MCAD primer 1 and MCAD primer 7 (Fig. 2). PCR amplification was done in 1XPCR buffer (50 mM KCl, 10 mM Tris-HCl, 0.1% Triton X-100 and 1.5 mM MgCl) with 1.2 mM dNTPs and 4.5 U of *Taq* polymerase in a 100μl volume. The PCR cycle used was 95° for 4 minutes followed by 32 cycles of 94° for 1 minute, 57° for 1 minute and 72° for 1 minute concluding with 72° for 7 minutes. A 590 bp PCR product was generated and subcloned into pCRII (Invitrogen) and confirmed by sequencing to be the 5' end of MCAD cDNA (Fig 1).

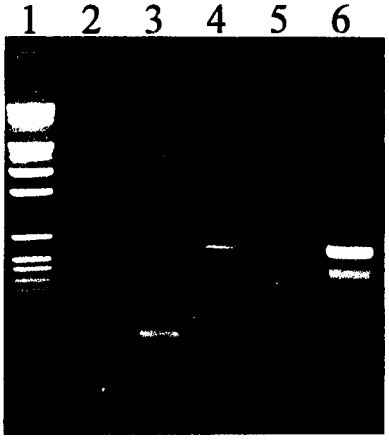
The mouse MCAD cDNA coding region is 1263 bp and encodes a 421 amino acid precursor protein similar to human and rat MCAD cDNAs (Kelly *et al.*, 1987; Matsubara *et al.*, 1987). The 3' untranslated region of mouse MCAD is 576 bp as compared to 578 bp of rat and 738 bp of human MCAD. Overall, mouse MCAD is 91% similar to the rat and 82% similar to the human MCAD gene at the nucleotide level over the coding region (Fig. 3). This demonstrates extensive conservation between the three species. Mouse sequence is 96% similar with rat and 87% similar with human at the deduced amino acid level (Fig. 4). We have previously reported comparable similarities between mouse, rat and human genes for short-chain acyl-CoA dehydrogenase (Kelly *et al.*, 1993).

A number of sites of conservation in the MCAD gene have been related to important functions. The most prevalent mutation causing MCAD deficiency is an A to G nucleotide replacement at position 985 of the MCAD cDNA. This results in the replacement of lysine by glutamic acid at amino acid 329 (K329E). Approximately 90% of affected patients are homozygous and another 7% are compound heterozygotes for this mutation (Workshop, 1992). The substitution of lysine to glutamic acid at 329 may inhibit homotetramer formation essential for MCAD activity (Kim *et al.*, 1993). This lysine is conserved in mouse MCAD. Other human mutations resulting in amino acid substitutions of arginine, methionine, cysteine, glycine and isoleucine at positions 53, 149, 244, 267 and 375, respectively, are also conserved across species (Workshop, 1992). Additionally, the glutamic acid at position 401, likely responsible for catalyzing the α-β dehydrogenation reaction, also is conserved in the mouse (3,7,10). The two amino acid residues methionine at 190 and tryptophan at 191, possibly involved with electron transfer flavoprotein, are also conserved (Kim *et al.*, 1993).

The cloned mouse MCAD cDNA, therefore, encodes a 421 amino acid precursor protein. The coding region is similar at the nucleotide level when compared with rat and human MCAD cDNAs. Conservation between these three species exists at specific amino acid residues where amino acid substitutions have resulted in human MCAD deficiency.

FIG. 1. Reverse transcription-PCR product from mouse poly(A) RNA isolated from the liver of BALB/cBy mice. Lane 1 = molecular weight markers; lane 2 = RT-PCR with MCAD primer 8 followed by PCR with MCAD primer 1 and 7; lane 3 = RT-PCR with MCAD primer 8 followed by PCR with MCAD primer 6 and 7; lane 4 = RT-PCR with random hexamer primers and subsequent PCR with MCAD primers 1 and 7; lane 5 = RT-PCR with random hexamer primers and subsequent PCR with MCAD primers 6 and 7; and lane 6 = positive control rat MCAD cDNA plasmid PCR with MCAD primers 1 and 7.

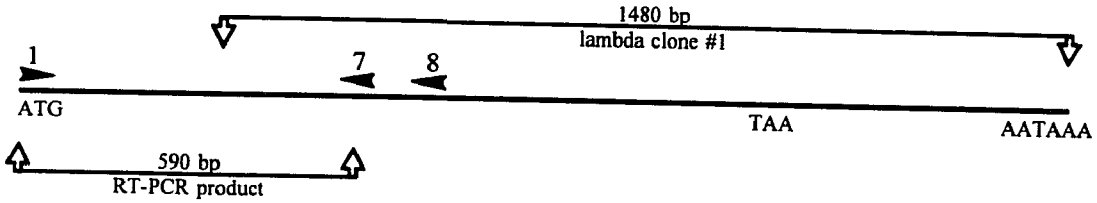
5' MCAD RT-PCR Product



1 = M.W.M.
2 = Primer 8 ▷ primer 1 and 7
3 = Primer 8 ▷ primer 6 and 7
4 = Random hexamers ▷ primer 1 and 7
5 = Random hexamers ▷ primer 6 and 7
6 = rat MCAD cDNA with primers 1 and 7

FIG. 2. Overview of generation of mouse medium-chain acyl-CoA dehydrogenase cDNA by library screening and reverse transcription - PCR. MCAD primers used for reverse transcription - PCR are illustrated.

Mouse Medium Chain Acyl-CoA Dehydrogenase cDNA - Cloning Strategy



RT-PCR Primers 5'→3'		MCAD sequence number
#1	AGGCACCATGGCAGCAGCG	-7→12
#7	TTGGTTATCCACATCTTCTG	562→581
#8	TGTATTCCCGGGTGTCGGC	676→695

FIG. 3. Comparison of nucleotide sequence of mouse (M), rat (R), and human (H) MCAD cDNA. Nucleotides which are conserved upon comparison with the mouse are illustrated with a line (-). The initiation codon, stop codon and putative polyadenylation signal are shaded. The mouse MCAD cDNA can be accessed at GenBank accession number U07159.

Comparative MCAD cDNA Sequence

```
                                                                                                                -7
MMCAD 5' untranslated                                                                                           AGGCACC
      1
MMCAD     ▒▒▒GCAGCAG CGTTCCGCAG AGGCTGCAGG GTCCTGAGAA GTGTTTCTCA TTTTGAGTGT CGAACACAAC
RMCAD     ---------- --C------- -----A--A- ---------- ----C----- ---------- ---G------
HMCAD     --------G- G----G-GC- -T-------- ---------- --A------G ----C-T--G A--T----G-
      71
MMCAD     ACTCGAAAGC GGCTCACAAG CAGGAGCCCG GATTAGGGTT TAGTTTTGAG TTGACGGAAC AGCAGAAAGA
RMCAD     --A-A---C- AT---T---- --------G- --C------- ---C--C--- --------G- ----------
HMCAD     -TA-A----- CAA--GAC-A -GT--A--A- -------A-- ---------- --C--C---- ----------
      141
MMCAD     GTTTCAAGCA ACTGCCCGCA AGTTTGCCAG AGAGGAGATT ATCCCCGTCG CCCCGGAATA TGACAAAAGC
RMCAD     -------A-- -T---T--G- ---------- ------A--A -----G---- ----A--C-- C--T------
HMCAD     A---------T -----T--T- -A-------- ------A--C ----AA--G- -TG-A----- ---T----CT
      211
MMCAD     GGGGAGTACC CGTTCCCTCT CATCAAAAGA GCCTGGGAAC TCGGCTTGAT CAACGCGCAC ATTCCGGAAA
RMCAD     -----A---- ---------- ------G--- ---------- -T--G----- ----A-A--- --------G-
HMCAD     --T--A--T- -AG----C-- A--T-G---- ---------- -T--T--A-- G---A-A--- -----A--G-
      281
MMCAD     GCTGCGGTGG CCTTGGCCTG GGAACGTTCG ATGCTTGTTT AATTACCGAA GAGTTGGCGT ATGGGTGTAC
RMCAD     -T--T----- T--------- -----T--T- ----G----- -----G---- --------A- ----------
HMCAD     A---T--A-- T-----A--T -----T--T- ---------- -----GT--- --A-----T- ----A-----
      351
MMCAD     AGGGGTGCAA ACTGCTATTG AAGCAAATTC TTTGGGGCAA ATGCCTGTGA TTCTTGCTGG AAATGATCAA
RMCAD     ---------G ---------- ---------- ---------- ---------- --A------- ----------
HMCAD     ------T--G ---------- ---G------ ---------- ------A-T- --A------- ----------
      421
MMCAD     CAAAAAAAGA AGTATTTGGG GAGGATGACG GAGCAGCCAA TGATGTGTGC TTACTGTGTG ACAGAGCCCT
RMCAD     --G--G---- ---------- ---------- --------G- ---------- C-----C--- -----A----
HMCAD     -----G---- ---------- ---A-----T ---G-----T ---------- ---T-----A -----A--TG
      491
MMCAD     CCGCAGGCTC TGATGTGGCG GCCATTAAGA CCAAAGCAGA GAAGAAGGGT GACGAGTATG TTATCAACGG
RMCAD     -A-------- ---------- -G-------- ---------- ---------- --T--A---- -C-----T--
HMCAD     GA-------- ------A--T -GT--A---- ---------- A-----A--A --T------A ----T--T--
      561
MMCAD     CCAGAAGATG TGGATAACCA ACGGGGGAAA GGCCAACTGG TATTTCTTGT TGGCACGTTC TAACCCAGAT
RMCAD     ---------- ---------- ---------- ---------- -----TG-A- --A-G--A-- ----------
HMCAD     T--------- ---------- ----A----- A--T--T--- -----T--A- ---------- -G-T------
      631
MMCAD     CCTAAAGTAC CCGCTAGTAA AGCCTTTACT GGATTCATTG TGGAAGCCGA CACCCCGGGA ATACATATTG
RMCAD     ---------- -T-------- ------C--C --------C- ----G----- ---------- -----C--C-
HMCAD     -------CT- -T----A--- ---------- ---------- -------A-- T-----A--- --T--G----
      701
MMCAD     GAAAAAAGGA ACTAAACATG GGCCAGCGAT GCTCTGACAC CAGAGGAATT GCTTTCGAAG ACGTCAGAGT
RMCAD     ---------- ---------- --T-----G- ---------- ---------C A-C------- -T-------
HMCAD     -G-G------ -T-------- ---------- -T--A--T-- T--------- -TC------- -T--G-A---
      771
MMCAD     GCCTAAGGAA AATGTGTTAA TCGGTGAAGG AGCAGGTTTC AAGATCGCAA TGGGTGCTTT TGATAGAACC
RMCAD     ---------- ---------- -T-------- ---------- -----T---- ----G----- ----------
HMCAD     ------A-TT -----T---- -T-----C-- ---T------ --AG-T---- ----A----- -----A----
      841
MMCAD     AGACCTACAG TCGCAGCTGG CGCTGTCGGG CTAGCCCAGA GAGCTCTAGA CGAAGCCACG AAGTATGCCC
RMCAD     --G--G--G- ----T----- T--------- ---------- ----C--G-- ------T--T ----------
HMCAD     ------GT-- -A--T----- T-----T--A T----A--A- -----T-G-- T-----T--C ----------
      911
MMCAD     TGGATAGGAA GACATTTGGA AAGCTGCTAG TGGAGCACCA AGGAGTTTCA TTTCTGCTCG CAGAAATGGC
RMCAD     ----C----- A--------- ---------- ---------- ---------- ---------- ----------
HMCAD     ----A----- A--T--C--- -----A--T- -A-------- --C-A-A--- ---A----G- -T--------
      981
MMCAD     GATGAAGGTT GAACTCGCTA GGCTCAGTTA CCAGAGAGCA GCCTGGGAGG TTGACTCCGG TCGCCGGAAC
RMCAD     ------A--- -----G--C- -A-------- ----C----- ---------- ---------- C---------
HMCAD     A-----A--- -----A---- -AA-G----- ---------- --T------- ----T--T-- ---T--A--T
     1051
MMCAD     ACTTACTATG CCTCGATTGC AAAGGCCTTT GCTGGAGACA TTGCCAATCA GCTAGCCACT GACGCCGTTC
RMCAD     --G----T-- ----T----- G--------- --------T- -------C-- ---C--T--C --T--T--G-
HMCAD     --C--T---- -T--T----- ------A--- --------T- ----A----- -T----T--- -TA--T--G-
     1121
MMCAD     AGATTTTCGG AGGCTATGGA TTCAACACAG AGTACCCTGT GGAGAAGCTG ATGAGGGACG CCAAGATCTA
RMCAD     ---------- ---------- --------T- -------A-- A--A------ ---C------ ----------
HMCAD     ----AC-T-- T---A----- --T--T---- -A--T----- A--A--A--A --------T- ----A-----
     1191
MMCAD     TCAGATTTAT GAAGGTACTG CACAAATTCA GAGGCTGATC ATAGCTCGTG AGCACATTGA AAAGTATAAA
RMCAD     ----------C ---------- ---------- ---------- ---------- ---------- ----------
HMCAD     ---------- ----------T ---------- A--A--T--T G----C----- -A-------- C-----C---
     1261
MMCAD     AATTT▒▒CAGG AATTACTATT GAACGATGCA TCACCCTCGT GTAACTAAGC TCCAAGCACT GTTGCTGCTT
RMCAD     -------▒▒▒
HMCAD     ------
     1331
MMCAD     CAGGGGAAAA GGGCTTTACT GTCTTCCCAA GGAAATGAGA TCAAAGACGA GTTTGGATCT GTGCAGCGGA
     1401
MMCAD     TTCCCATGGC GGAGGAACCT GTCTTCAGCT CTATGGTGAC CCTTTCTAGA TAGGTTTGGC TTTTGGACAA
     1471
MMCAD     TGATTGGTCC TTAGCCCCGA ATTGTGTTAG TTTGCTCTTT GATCACTTAA AATGGAAAAA CACCCTGGAC
     1541
MMCAD     TTTTAATGTT CATTCAAGTG ACAGGAAAGG CGGCTTGTCA AGGAAGAACT CATGATTCTA ACATAAACAC
     1611
MMCAD     TGAAAATTTG TGGTAGATTG GACACGTCAG ACTGTGACAT AGCAGCATTT CTGTGCTGAA CTGTTAATTT
     1681
MMCAD     TATAATTTTG TTATATTTGC TTTGTTTTGC ACAAAAGAGT AAAAAGTTTA TATTCATATT CTCCCATTAT
     1751
MMCAD     AAAACTAAAA CTTTTCTGGG AATCTTAGTA CTGAACGAGA ATTTTATTTG TCTTGTTTAA AATAATTC▒▒
     1821
MMCAD     ▒▒▒AGCTAAC CTTAACTCC  1839
```

FIG. 4. Comparison of the deduced amino acid sequence of mouse (M), rat (R) and human (H) MCAD. Amino acid residues found to be substituted in MCAD deficient patients are indicated (▼). These residues are conserved between species. A putative flavin adenine dinucleotide (FAD) binding site is shaded. Amino acid residues important in enzymatic function are shaded.

Amino Acid Sequence Comparison

```
      1                                                                  ▼
MMCAD MAAAFRRGCR VLRSVSHFEC RTQHSKAAHK QEPGLGFSFE LTEQQKEFQA TARKFAREEI
RMCAD ----L---YK ---------- -A--T-PSL- ---------- ---------T I---------
HMCAD ---G-G-C-- ----I-R-HW -S--T--NRQ R--------- F--------- ----------
      61
MMCAD IPVAPEYDKS GEYPFPLIKR AWELGLINAH IPESCGGLGL GTFDACLITE ELAYGCTGVQ
RMCAD -----D---- ---------- --------T- ---------- ---------- ----------
HMCAD -Q--A----T ----V---R- ------M-T- ---N------ --------S- ----------
      121                              ▼
MMCAD TAIEANSLGQ MPVILAGNDQ QKKKYLGRMT EQPMMCAYCV TEPSAGSDVA AIKTKAEKKG
RMCAD ---------- ----I----- ---------- ---------- ---------- Q---------
HMCAD ----G----- --I-I----- ---------- -E-L------ ---G------ G---------
      181
MMCAD DEYVINGQKM WITNGGKANW YFLLARSNPD PKVPASKAFT GFIVEADTPG IHIGKKELNM
RMCAD ----------▒ ▒--------- --V-T----- ---------- ---------- ----------
HMCAD ---I------▒ ▒--------- -------D-- --A--N---- ---------- -Q--R-----
      241▼                             ▼
MMCAD GQRCSDTRGI AFEDVRVPKE NVLIGEGAGF KIAMGAFDRT RPTVAAGAVG LAQRALDEAT
RMCAD ---------- T--------- ---------- ---------- ---------- ----------
HMCAD ---------- V----K---V -----D---- -V------K- --V------- ----------
      301                              ▼
MMCAD KYALDRKTFG KLLVEHQGVS FLLAEMAMKV ELARLSYQRA AWEVDSGRRN TYYASIAKAF
RMCAD ---------- ---------- ---------- ---------- ---------- --F-------
HMCAD ----E----- -------AI- -M-------- ----M----- ---------- ----------
      361                ▼
MMCAD AGDIANQLAT DAVQIFGGYG FNTEYPVEKL MRDAKIYQIY ▒GTAQIQRLI IAREHIEKYK
RMCAD ---------- ---------- ---------- ---------- ▒--------- ----------
HMCAD ---------- V----L--N- ---------- ---------- ▒--S------ V-----D---
      421
MMCAD N*
RMCAD -*
HMCAD -*
```

ACKNOWLEDGEMENTS

We thank Dr. Myron E. Hinsdale for his invaluable technical advice on RNA isolation and reverse transcription. We also thank Dr. Dan P. Kelly, Washington University, St. Louis for providing the rat MCAD cDNA probe. This work was supported by NIH grant R01-RR02599.

References

Davis, L.G., Dibner, M.D., Battey, J.F. (1986). "Basic Methods in Molecular Biology," Elsevier Science Publishing Co, Inc., New York.

Finocchiaro, G., Ito, M., and Tanaka, K. (1987). Purification and properties of short chain acyl-CoA, medium chain acyl-CoA, and isovaleryl-CoA dehydrogenases from human liver. *J. Biol. Chem.* **262**: 7982-7989.

Ghisla, S., Thorpe, C., Massey, V. (1984). Mechanistic studies with general acyl-CoA dehydrogenases and butryl-CoA dehydrogenase: evidence for the transfer of the β-hydrogen to the flavin N(5)-position as a hydride. *Biochemistry* **23**: 3154-3161.

Ikeda, Y., Keese, S.M., Fenton, W.A., and Tanaka, K. (1987). Biosynthesis of four rat liver mitochondrial acyl-CoA dehydrogenases: *in vitro* synthesis, import into mitochondria, and processing of their precursors in a cell-free system and in cultured cells. *Arch. Biochem. Biophys.* **252**: 662-674.

Kelly, C.L., Hinsdale, M.E., and Wood, P.A. (1993). Cloning and characterization of the mouse short-chain acyl-CoA dehydrogenase cDNA. *Genomics* **18**: 137-140.

Kelly, C.L., Kim, J.-J., Billadello, J.J., Hainline, B.E., Chu, T.W., and Strauss, A.W. (1987). Nucleotide sequence of medium-chain acyl-CoA dehydrogenase mRNA and its expression in enzyme-deficient human tissue. *Proc. Natl. Acad. Sci. U.S.A.* **84**: 4068-4072.

Kim, J.-J.P., Wang, M., and Paschke, R. (1993). Crystal structures of medium-chain acyl-CoA dehydrogenase from pig liver mitochondria with and without substrate. *Proc. Natl. Acad. Sci. U.S.A.* **90**: 7523-7527.

Matsubara, Y., Kraus, J.P., Ozasa, H., Glassberg, R., Finocchiaro, G., Ikeda, Y., Mole, J., Rosenberg, L.E., and Tanaka, K. (1987). Molecular cloning and nucleotide sequence of cDNA encoding the entire precursor of rat liver medium-chain acyl coenzyme A dehydrogenase. *J. Biol. Chem.* **262**: 10104-10108.

Matsubara, Y., Narisawa, K., Tada, K., Ikeda, H., Yeqi, Y., Danks, D.M., Green, A., and McCabe, E.R.B. (1992). Prevalence of K329E mutation in the medium-chain acyl-CoA dehydrogenase gene determined from Guthrie cards. In: "New Developments in Fatty Acid Oxidation," eds. Coates PM, and Tanaka K. (Wiley-Liss, New York). pp.453-462.

Roe, C.R., and Coates, P.M. (1989). Acyl-CoA dehydrogenase deficiencies. In: "The Metabolic Basis of Inherited Disease," eds. Scriver CR, Beaudet AL, Sly WS, and Valle D. (McGraw-Hill, Inc., New York). pp.889-914.

Workshop on Molecular Aspects of MCAD Deficiency. (1992). Mutations causing medium-chain acyl-CoA dehydrogenase deficiency: a collaborative compilation of the data from 172 patients. In: "New Developments in Fatty Acid Oxidation," eds. Coates PM, and Tanaka K. (Wiley-Liss, New York). pp. 499-506.

GENOMIC STRUCTURE, REGULATORY ELEMENTS AND CHROMOSOMAL LOCATION OF THE MOUSE MEDIUM-CHAIN ACYL-COA DEHYDROGENASE GENE

RAVI J. TOLWANI,[1] SUSAN C. FARMER,[1] DAVID M. KURTZ,[1]
KENNETH R. JOHNSON,[2] MURIEL T. DAVISSON,[2]
AND PHILIP A. WOOD[1]

[1]Department of Comparative Medicine, Schools of Medicine and Dentistry,
The University of Alabama at Birmingham

[2]The Jackson Laboratory, Bar Harbor, Maine

To be submitted to *Genomics*

Abstract

Medium-chain acyl-coenzyme A dehydrogenase (enzyme symbol-MCAD,; gene symbol-*Acadm*) is involved in the initial step of fatty acid β-oxidation within mitochondria. Inherited MCAD deficiency, an autosomal recessive disorder, occurs at high frequency in humans and is associated with considerable morbidity and mortality. In our efforts to better understand transcriptional regulation of acyl-CoA dehydrogenase genes and develop an animal model with MCAD deficiency via homologous recombination in embryonic stem cells, we have cloned and characterized the mouse MCAD gene (*Acadm)*. The *Acadm* gene spans approximately 25 kilobases and contains 12 exons. The promoter region is G:C-rich (60%) within 200 bp of CAP site) and does not contain a TATA box or CAAT box upstream of the transcriptional start site. A CpG island extends from sequences 5' of the transcriptional start site and extends into intron 1. Sequence analysis of the 5' regulatory region and a portion of intron one revealed several Sp1 consensus sites (GGGCGG or CCGCCC). Also present are hexamer DNA consensus sequences (RG(G/T)TCA), which may bind steroid/thyroid hormone nuclear receptors, and consensus elements of the mouse peroxisomal proliferator activated receptor (PPAR). Intracellular fatty acid intermediates may regulate transcription of the *Acadm* gene via PPAR. We also show that at least two *Acadm* processed pseudogenes are present in the mouse genome. We mapped the *Acadm* gene to the distal end of Chromosome 3, contrary to its mapping to Chromosome 8, as previously reported by others. The *Acadm* pseudogenes were mapped to mouse Chromosomes 8 and 11.

Introduction

Medium-chain acyl-CoA dehydrogenase (enzyme symbol MCAD; gene symbol *Acadm*) is one of four acyl-CoA dehydrogenase enzymes involved in the initial step of fatty acid β-oxidation. These enzymes catalyze the α-β dehydrogenation of acyl-CoA thioesters and transfer electrons to electron transfer flavoproteins (Roe and Coates, 1989). Specifically, MCAD is active in catalyzing the dehydrogenation of fatty acid thioesters that are C_4-C_{16} in length (Finocchiaro *et al.*, 1987). MCAD is transcribed in the nucleus, translated in the cytosol and translocated into the mitochondrial matrix. In the mitochondrial matrix the MCAD monomers, each containing one molecule of flavin adenine dinucleotide (FAD^+), assemble into homotetramers to gain enzymatic activity (Ikeda *et al.*, 1987).

Inherited MCAD deficiency occurs in humans. MCAD deficient patients develop clinical episodes during early childhood that are often associated with fasting. These patients can experience a wide range of metabolic derangements, including hypoglycemia, impaired ketogenesis, medium-chain dicarboxylic aciduria, hyperammonemia and secondary carnitine deficiency (Roe and Coates, 1989). Affected patients also may acquire fatty change in the liver and encephalopathy. MCAD deficiency is associated with considerable mortality. In one report, approximately 59% of infants between 15 and 26 months of age died during their first episode (Roe and Coates, 1989).

The prevalence of MCAD deficiency is estimated at 1 in 15,000 with a carrier frequency of 1:62 in the Caucasian population (Matsubara *et al.*, 1992). The most common mutation in MCAD-deficient patients is a point mutation resulting in the substitution of the nucleotide A with the nucleotide G at position 985 of the human

MCAD cDNA. This results in a glutamic acid to a lysine substitution at amino acid position 329. Approximately 89% of MCAD-deficient patients are homozygous and another 9-10% are compound heterozygotes for this point mutation (Workshop on Molecular Aspects of MCAD Deficiency, 1992).

Here we describe the isolation, characterization and mapping of the mouse *Acadm* gene and two *Acadm* processed pseudogenes. We analyzed the 5' regulatory regions for putative transcriptional control elements and made direct sequence comparisons to the human MCAD gene (*ACADM*) 5' regulatory region. The *Acadm* gene may be transcriptionally regulated by multiple nuclear receptors similar to the human MCAD gene (*ACADM*) and other genes associated with lipid metabolism (Raisher *et al.*, 1992; Carter *et al.*, 1993; Carter *et al.*, 1994; Gulick *et al.*, in press 1994; Ladias and Karathanasis, 1991; Ladias *et al.*, 1992; Rottman and Gordon, 1993; Rottman *et al.*, 1991). This study provides an insight into the putative mechanisms of transcriptional regulation of the *Acadm* gene. The *Acadm* gene fragments we have isolated are being used to develop an animal model with MCAD deficiency via gene targeting in embryonic stem cells.

Materials and Methods

Generation of Probes for Screening Genomic Libraries

Three separate DNA fragments were used as probes to screen mouse genomic libraries. The first probe was generated by polymerase chain reaction (PCR) as previously described (Wood *et al.*, 1992). This *Acadm* intron probe consists of a 1.5 kb mouse genomic fragment spanning from exon 6 to exon 7. The second and third probes were *Acadm* cDNA fragments composed of nucleotide numbers (nt) 359 to 1839 and nucleotide numbers -7 to 1839, respectively (Tolwani *et al.*, 1994; GenBank accession #U07159).

DNA probes were labeled with ^{32}P-dCTP via random hexamer priming (Feinberg and Vogelstein, 1983).

Screening of Genomic DNA Libraries

The 1.5 kb *Acadm* intron fragment was used as a probe to screen a λGEM11 mouse genomic library (Promega) produced from an ICR Swiss mouse. Approximately 300,000 plaques were screened at 10,000 plaques per 150 mm plate using described methods (Davis *et al.*, 1986). Additionally, mouse MCAD cDNA fragments were used as probes to screen a Lambda FIXII 129/Sv mouse genomic library (Stratagene) on two separate occasions. Approximately 1.2 X 10^6 plaques were screened at 10,000 plaques per 150 mm plate using the MCAD cDNA 359 to 1839 as a probe, and 2 X 10^6 plaques were screened at 50,000 plaques per plate using the MCAD cDNA fragment containing nucleotides -7 to 1839 as a probe. Positive phage clones were purified by four rounds of screening. Purified phage clones obtained were analyzed with two specific MCAD probes. A 5' fragment consisting of mouse MCAD cDNA nt -7 to 142 and a 3' fragment consisting of mouse nt 1013 to 1663 were generated by PCR using the full length mouse MCAD cDNA as a template (Tolwani *et al.*, 1994). These fragments were labeled and hybridized to the phage clones.

DNA from selected phage clones was isolated from purified phage plate lysates using the Lambda Prep System (Promega). An 18-kb insert from one phage clone obtained from the Lambda GEM11 library was freed from the phage arms by *Sfi*I digestion. The 3' *Sfi*I overhangs of the insert were blunt ended with Klenow fragment at 37°C for 15 minutes. *Not*I 12 mer linkers (Promega) were ligated to the blunt ends. After digestion with *Not*I, this entire 18-kb fragment was subcloned into the *Not*I site of

pGEM11Zf(+). Additionally, inserts from phage clones from the Lambda FIXII library were freed from the phage arms by *Not*I and subcloned into the *Not*I site of pGEM11Zf(+).

DNA Sequence Analysis

Oligonucleotide primers were synthesized by the Oligonucleotide Shared Facility of the Comprehensive Cancer Center, The University of Alabama at Birmingham (U.A.B.). Some oligonucleotides were designed based on the nucleotide sequence of the mouse MCAD cDNA to determine the exon/intron boundaries of the MCAD gene and some were designed to sequence the 5' regulatory elements of the MCAD gene. These oligonucleotide primers were designed using the Gene Runner software (Hastings Software, Inc.) to be subsequently used for PCR amplification. All sequencing was done by the dideoxy-chain termination method (Sanger *et al.*, 1977) using Sequenase (Version 2.0, U.S.B.).

To determine if a CpG island exists within the 5' promoter region, a ration of observed/expected CpG content was determined. This ratio was calculated based on formula by Gardiner-Garden and Frommer (number of CpG X N)/(number of C X number of G), where N equals the total number of nucleotides in the analyzed sequence (Gardiner-Garden and Frommer, 1987). CpG islands are defined as regions of DNA with a (G+C) content greater than 50% and an observed/expected CpG greater than 0.6 spanning more than 200 basepairs (Gardiner-Garden and Frommer, 1987).

PCR Amplifications of Introns and Restriction Mapping

Intron sizes were partly determined by PCR amplifications using primers designed based on mouse MCAD cDNA. PCR amplifications were carried out in 100 μl volume

using 1X PCR buffer (50 mM KCl, 10 mM Tris-HCl, 0.1% Triton X-100 and 1.5 mM MgCl) with 1.2 mM dNTPs and 4.5 U of *Taq* polymerase. Templates were MCAD genomic fragments subcloned in pGEM11Zf(+) and mouse genomic DNA. Cycle conditions were 94°C/4 min followed by 32 cycles of 94°C/1 min, 57°C/1 min and 72°C/3 min followed by 72°C for 7 min. Approximately 40 μl of the reaction was separated by electrophoresis on 1% agarose gels for visual detection of the amplified products.

Amplified products were purified (DNA clean up kit - Promega) and digested with *Bam*HI, *Eco*RI and *Xho*I. Digested intron fragments were electrophoresed on 1% agarose gels to determine restriction sites present within each intron.

Additionally, the large genomic inserts within pGEM11Zf(+) were digested into smaller fragments with *Bam*HI, *Eco*RI and *Xho*I and subcloned. These smaller pieces were then digested with the above enzymes to complete the restriction map.

Primer Extension

Total RNA was extracted from heart of adult BALB/cBy mice by using RNAzol B method (Biotecx Laboratories, Inc.) using the manufacturer's instructions. Two oligonucleotides were designed: #1 (5'-GGCGTGGCCTAGGGAGTCC-3') and #2 (5'-GGGATCCGGTGCGGGGGCT-3'), were designed and used as primers. These primers were 5'-end-labeled with γ^{32}P-ATP (6,000 Ci/mmol) and T4 polynucleotide kinase in 50mM Tris-HCl pH 7.5, 10mM $MgCl_2$, 5mM DTT and 0.1mM spermidine by incubation at 37°C for 30 min. These labeled primers were annealed to 20 µg of heat denatured RNA in 1.0 M NaCl, 10mM Tris-HCl pH 8.0 and 1mM EDTA for 30 min. at 45°C. The hybridized RNA was washed in ethanol, precipitated and subsequently resuspended in

50mM Tris-HCl pH 7.6, 6mM $MgCl_2$, 100 mM NaCl, 10 mM DTT, 4 mM dNTPs, 20 units of RNasin and 20 units of AMV reverse transcriptase (Boehringer-Mannheim) and incubated for 30 min at 42°C. The extension product was phenol/chloroform extracted, precipitated and resuspended before analysis on 8% polyacrylamide/7 M urea sequencing gel.

Genomic Southern Blot

Genomic DNA was extracted from C57BL/6J, CAST/Ei and SPRET/Ei mice obtained from the Mouse Mutant Resource of The Jackson Laboratory (for description of strains see Johnson *et al.*, 1992). Extracted DNA (10 μg) was digested with *Bam*HI, *Eco*RI, *Pst*I, *Pvu*II, *Msp*I and *Taq*I, electrophoresed on 0.7% agarose gel and transferred to ZetaProbe (Bio-Rad Laboratories) membrane. The full-length mouse MCAD cDNA was used as a probe. Hybridization procedures were as previously described (Johnson *et al.*, 1992).

Genetic Mapping

To genetically map *Acadm* and related sequences in the mouse, we used previously described intersubspecific and interspecific backcross mapping panels (Johnson *et al.*, 1992). For the intersubspecific backcross, female F1 hybrids between an inbred strain of *Mus musculus castaneus* (CAST/Ei) and the standard laboratory strain C57BL/6J were backcrossed to C57BL/6J males (this cross is henceforth referred to as BCB). For the interspecific backcross, female F_1 hybrids between an inbred strain of *Mus spretus* (SPRET/Ei) and C57BL/6J were backcrossed to SPRET/Ei males (this cross is henceforth referred to as BSS). Each backcross panel has been typed for more than 300 polymorphic

loci distributed over all 19 autosomes and the X chromosome (Johnson *et al.*, 1994; unpublished K. Johnson).

Two specific probes were used to map the *Acadm* sequences, the *Acadm* intron probe generated by PCR as described earlier (Wood *et al.*, 1992) and the full length *Acadm* cDNA fragment including nucleotides -7 to 1839 (Tolwani *et al.*, 1994). To map DNA fragments that hybridized with the *Acadm* intron probe, we analyzed the segregation of a *Pvu*II RFLP among the BSS backcross progeny (Fig. 4). To map DNA fragments that hybridized with the *Acadm* cDNA probe, we analyzed the segregation of *Pst*I RFLPs among the BCB backcross progeny and the segregation of *Msp*I and *Taq*I RFLPs among BSS progeny (Fig. 4).

Genetic linkage was analyzed by comparing the segregation of *Acadm* hybridizing sequences with the segregation of previously typed reference loci in BSS and BCB backcrosses. The typing of these reference loci has been described previously (Johnson *et al.*, 1994). The computer program Map Manager (Manly, 1993) was used to perform linkage and haplotype analysis. Gene order on a chromosome was determined by minimizing the number of double crossover events required to explain observed haplotype distributions. The complete haplotype data for each backcross and chromosome have been deposited as Map Manager files in the Mouse Genome Database. The Accession Numbers for the BSS backcross are MGD-cREX-184 (Chr 3), MGD-cREX-185 (Chr 8) and MGD-cREX-186 (Chr 11). The Accession Number for the BCB backcross is MGD-cREX-187 (Chr 11).

Results

Isolation of the Mouse MCAD Gene

We initially used the mouse MCAD intron 1.5 kb PCR fragment consisting of DNA spanning from exon 6 to exon 7 and including intron 6 (Wood *et al.*, 1992) as a probe. Three phage clones were isolated from screening the Lambda GEM11 library, and the clone with the largest insert was subcloned into pGEM11Zf(+).

Mouse MCAD cDNA fragments were used to screen the Lambda FIXII mouse genomic library. Screening with cDNA nt 359 to 1839 as a probe produced eight phage clones. Seven of these clones were subcloned into the *Not*I site of pGEM11Zf(+). Screening with the entire cDNA nt. -7 to 1839 as a probe resulted in the isolation of 11 phage clones. These phage clones were analyzed by hybridization with 5' probe (MCAD cDNA nt. -7 to 142) and 3' probe (MCAD cDNA nt. 1013 to 1663). Clones number 9, 11, 13, 15, 18, 19 hybridized with the 5' and 3' probe, suggesting that these were pseudogenes. Clones 10, 12, 14, and 15 hybridized only to the 3' probe, suggesting they consisted of 3' genomic fragments. Two (clone 16 and 17) of the 11 clones obtained from this screening hybridized with only with the 5' probe, suggesting these clones were 5' MCAD genomic fragments. Inserts from clones 16 and 17 were subcloned into pGEM11Zf(+) at the *Not*I site.

Structure of Acadm Gene

The genomic fragments subcloned into pGEM11Zf(+) were analyzed by Southern blotting and sequencing. Primers were designed based on mouse MCAD cDNA sequence and used to determine the exon/intron boundaries of *Acadm* gene. DNA sequence analysis showed that the *Acadm* gene consisted of 12 exons. Exons 1 thru 7 were present in clone

#1, obtained from the ICR Swiss Lambda GEM11 library, while exons 8 thru 12 were present in clone #7, isolated from the 129/Sv Lambda FIXII library.

We further analyzed genomic fragments by PCR, sequencing and subcloning into smaller fragments to determine intron sizes and restriction maps. Amplification under stringent conditions of plasmid DNA containing cloned *Acadm* genomic fragments with primer pairs within consecutive exons resulted in strong PCR products of varying sizes. Subsequent PCR amplifications with the same primer sets using genomic DNA as a template resulted in products of equivalent sizes (data not shown).

To obtain a restriction map, the *Acadm* introns amplified by PCR were digested with *Bam*HI, *Eco*RI and *Xho*I. Additionally, the large genomic fragments within plasmid vectors were digested into smaller pieces and subsequently subcloned. These smaller fragments were analyzed by further sequencing and restriction digestions. Results obtained correlated with restriction digestions of the PCR derived introns. The restriction map of the *Acadm* locus is illustrated in Figure 1.

The entire *Acadm* gene spans 25 kb. There are 12 exons which range in size from 70 bp for exon 4 to 644 bp for exon 12. There is also a wide variation for intron sizes, ranging from 100 bp for intron 3 to 3.9 kb for intron 10.

Isolation and Characterization of Acadm Pseudogenes

Partial sequence analysis of inserts from clones 1, 2, 16 and 17 revealed similar sequences to the *Acadm* cDNA (Tolwani *et al.*, 1994). PCR and sequence analysis revealed the introns present within the *Acadm* functional gene were not present within these clones. Comparisons between these clones to the *Acadm* cDNA exhibit changes such as nucleotide substitutions, base insertions and base deletions. The 5' translation

start codon region of these clones contains nucleotide mismatches within the Kozak consensus sequence (Fig. 3). Homology to the functional *Acadm* gene ends 17 nucleotides and 46 nucleotides 5' of the translation start site in clone 1 and 16, respectively. Based on these results these phage clones contained *Acadm* processed pseudogenes. Additionally, hybridization of isolated phage clones 9, 11, 13, 15, 18 and 19 with both 5' and 3' specific probes suggested these clones were also *Acadm* processed pseudogenes. Based on hybridization and sequencing, 9 of the 20 phage clones isolated contained *Acadm* pseudogenes.

Determination of the Transcription Start Site

Primer extension experiments with mouse total heart RNA were done to determine the transcription start site of the *Acadm* gene. Several primers were designed within the 5' untranslated and translated region of *Acadm* gene since we suspected the presence of secondary structure within the message. Primer extension reactions from two of these primers (#1 and #2) revealed an identical transcription start site 196 nt. upstream of the translation start codon.

DNA Sequence Analysis of 5' Regulatory Region

We sequenced approximately 1600 bp 5' to the translation start elements and an additional 700 bp within intron 1 to locate regulatory motifs. The *Acadm* promoter region is GC rich with a G+C content of 60% within 200 bp of the transcription start site. No TATA or CCAAT DNA binding elements were located near the transcriptional start site. Most eukaryotic class II genes have a TATA element 20-30 bp upstream of the transcription start site. We located a region spanning from -300 to +670 (in reference to +1 as the transcription start site) where the G+C content is greater than 50% and the CpG

observed/expected content is greater than 6.25%, conforming to the definition of a CpG island (Gardiner-Garden and Frommer, 1987; Larsen *et al.*, 1992)). Several Sp1 binding sites (GGGCGG or CCGCCC) are also present at positions -109, +111, +147, +315, +318 in relation to the transcription start site located at position +1 (Fig. 6). The region at +315 located within intron 1 has an 11 bp GCCGCCCGCCC direct repeat.

The sequence GGCACACAGTCTTCT, closely homologous to the glucocorticoid receptor response element (GRE) GGTACANNNTGTTCT (Beato, 1989), is present at position -248. Additionally, there is a region with perfect homology to the Ap2 binding site GCCTGGGGG located at +270 within the 5' region of intron 1. A region with perfect consensus to sterol response element (SRE) is present at +495, also within intron 1. Additionally, a perfect consensus to insulin response element (IRE) is present at -109.

Also present are consensus elements for binding to nuclear receptors of the estrogen/thyroid hormone nuclear receptor subfamily of steroid/thyroid hormone receptor superfamily. Hexamer elements called the proximal (P) box with the consensus sequence RG(G/T)TCA are recognized to bind with retinoic acid receptors and thyroid receptors (Linney, 1992; Naar *et al.*, 1991; Umesono and Evans, 1989). These hexamer motifs are located at nt. -364, -358, -343, +42 and +56 in relation to the transcriptional start site. The region at -364 to -343 contains three P boxes, two in direct repeat conformation and two arranged in an inverted repeat separated by 4 basepairs. The region at +42 to +56 contains an inverted repeat separated by 8 basepairs.

Genetic Mapping

The *Acadm* intron probe detected only a single DNA fragment in *Bam*HI, *Eco*RI and *Pvu*II digests of DNA from all three parental strains: C57BL/6J, CAST/Ei and

SPRET/Ei (data not shown). The *Pvu*II RFLP between C57BL/6J and SPRET/Ei DNA (indicated in Fig. 4) was used to analyze *Acadm* segregation in BSS backcross progeny. The *Acadm* gene detected with the intron probe mapped to distal Chr 3 near *Mpmv9*. In contrast to the intron probe, the *Acadm* cDNA probe detected multiple hybridizing DNA fragments (Fig. 4). We analyzed the segregation among BSS backcross progeny of six *Msp*I and five *Taq*I C57BL/6J-specific DNA fragments (Fig 4.). Four of the six *Msp*I fragments cosegregated and mapped to the same locus as the *Acadm* gene that was mapped using the intron probe; however, the remaining two fragments mapped to different chromosomal locations. These two fragments thus represent pseudogene sequence (*Acadm-ps*) loci and not the authentic *Acadm* locus. We mapped one of them, designated *Acadm-ps1*, to Chr 11 and the other, designated *Acadm-ps2*, to Chr 8 (Fig. 4, Table 1). The five segregating *Taq*I fragments mapped to these same three chromosomal locations: three cosegregated and mapped to Chr 3, one to Chr 11 and one to Chr 8 (Fig. 4). Our results indicate, therefore, that the authentic murine *Acadm* gene is on distal Chr 3, and the *Acadm* pseudogenes are on Chrs 8 and 11 (Figs. 5a, 5b, 5c).

Discussion

The structure of the *Acadm* gene is very similar to that of the human gene (Zhang *et al.*, 1992). Both genes consist of 12 exons and the intron-exon boundaries are conserved. The mouse gene, however, spans 25 kb and is 55% the size of its human counterpart. The smaller size is primarily due to smaller introns 6 and 10, although there are other intron size variations between the mouse and human.

As in the human gene, the mouse promoter does not contain a TATA or CCAAT box. The mouse promoter region has a G+C content of 60% within 200 bp of the

transcription start site. Class II genes with GC rich promoter regions without a TATA-box or CAAT-box are usually constitutively expressed and are considered housekeeping genes (Smale and Baltimore, 1989; review Ye *et al.*, 1993). Similar promoters also have been found in some highly regulated genes like the c-*rel* gene (Hannink and Temin, 1990) and the *Ets1* transcription factor gene (Jorcyk *et al.*, 1991). We have determined that MCAD protein is present early in development in mouse embryonic stem cells (Wood *et al.*, 1992). MCAD is also expressed during fetal development as well as in adult tissues (Kelly *et al.*, 1989). Additionally, a 1000 bp CpG island exists encompassing the transcription start site, exon 1 and the 5' region of intron 1. CpG islands are usually associated with genes that are widely expressed such as the MCAD gene (Larsen *et al.*, 1992).

The *Acadm* promoter region contains several GC motifs, both upstream and downstream of the transcription start site, as expected in promoters which are TATA-less and have a high GC content (for review see Hernandez, 1993). Sp1 sites are usually located upstream of the transcriptional start site, but downstream Sp1 interactions can also stimulate promoter activity (Courey *et al.*, 1989). The binding of Sp1 to the GC motifs sites greatly stimulates transcription of other TATA-less promoters such as the lymphocyte-specific terminal deoxynucleotidyltransferase promoter (Smale and Baltimore, 1989). The Sp1-binding sites may serve a similar function in the *Acadm* gene.

Several factors have been shown to influence expression of the *Acadm* gene. *Acadm* expression in the rat was stimulated by fasting and glucagon administration (Nagao *et al.*, 1993). The amplitude of glucagon-enhanced *Acadm* expression in the rat liver and heart was greater when rats were previously fasted (Nagao *et al.*, 1993). Glucagon exerts

its effects via the cAMP signaling pathway. Since regulatory elements present in the rat may also be present in the mouse, we attempted to locate a cAMP response element (CRE) within the *Acadm* gene. A search of over 2390 bp encompassing the 5' regulatory region and portion of intron 1 revealed no CRE motif. However, a region with exact consensus to Ap2 element was located within the 5' region of the first intron. Ap2 sites, in addition to CRE, can also mediate a cAMP response. Additionally, administration of dexamethasome stimulated *Acadm* expression in rat heart and suppressed *Acadm* expression in rat liver (Nagao *et al.*, 1993). We located a sequence similar to the consensus for glucocorticoid response element (GRE) where glucocorticoids are thought to mediate transcriptional control of the *Acadm* gene (Table 2).

The human MCAD gene (*ACADM*) contains a complex regulatory region, called a nuclear receptor response element (NRRE-1), that is capable of interacting with numerous nuclear receptors. The NRRE-1 consists of four hexamer consensus sequences that bind various nuclear receptors. *ACADM* gene expression is regulated through binding to this region of retinoic acid receptors (RAR and RXR), hepatocyte nuclear factor 4 (HNF-4) and chicken ovalbumin upstream promoter transcription factor (COUP-TF) (Raisher *et al.*, 1992; Carter *et al.*, 1993; Carter *et al.*, 1994). We found a region at nucleotides -364 to -343 analagous to the human NRRE-1. As in the *ACADM* gene and other genes involved in lipid metabolism (Raisher *et al.*, 1992; Carter *et al.*, 1993; Carter *et al.*, 1994; Gulick *et al.*, in press, 1994; Ladias *et al.*, 1991; Ladias *et al.*, 1992; Rottman and Gordon, 1993; Rottman *et al.*, 1991), transcription of the *Acadm* gene may also be regulated by interaction with numerous nuclear receptors.

Intracellular fatty acid metabolites also can regulate the expression of the *ACADM* gene. The peroxisomal proliferator activated receptor (PPAR) mediates this regulation upon binding to the NRRE-1 of the *ACADM* gene (Gulick *et al.*, in press, 1994). Peroxisomal β-oxidation is also controlled by PPAR (Dreyer *et al.*, 1992). P box consensus elements present within the *Acadm* promoter region may be capable of interactions with PPAR.

In addition to the authentic *Acadm* gene, two *Acadm* processed pseudogenes were cloned and sequenced. Processed pseudogenes probably arise from retrotranscription of mRNA and insert at random sites in the genome. Interestingly, there have been no human *ACADM* pseudogenes cloned or detected by Southern analysis (personal communication, D. P. Kelly, Washington University, St. Louis). This suggests that the mouse pseudogenes were created after the evolutionary phylogenetic split of mouse and man. It would be interesting to determine if pseudogenes are present within other members of the Rodentidae such as the rat. The cloned *Acadm* processed pseudogenes were present within Chrs 8 and 11 as indicated by mapping. These pseudogenes contained sequence similarity within portions the 5' untranslated region closest to the translation start site of the mouse *Acadm* gene. There is no homology with the functional *Acadm* gene further upstream of this region. The absence of a promoter and the lack of introns within the pseudogenes suggest these processed pseudogenes are transcriptionally inactive. Even if these processed pseudogenes were transcribed, they would not translate a functional MCAD protein since the translation start site is mutated and there are premature stop codons present within the coding region.

FIG. 1. The structural map of the *Acadm* gene. The positions of exons 1 to 12 are indicated by shaded boxes with the 5' and 3' untranslated region depicted with open boxes. A restriction map with the enzymes *Bam*HI, *Eco*RI and *Xho*I is illustrated. The two overlapping phage clones used to characterize the *Acadm* gene are labeled at the bottom.

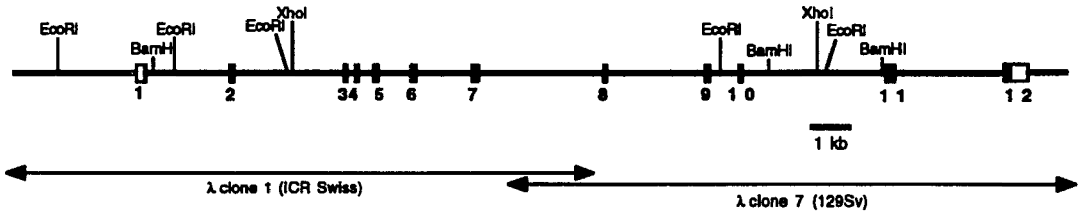

EcoRI
EcoRI
BamHI
EcoRI
XhoI
1
2
34
5
6
7
8
EcoRI
BamHI
XhoI
EcoRI
BamHI
9
10
11
12
1 kb
λ clone 1 (ICR Swiss)
λ clone 7 (129Sv)

FIG. 2. *Acadm* intron/exon organization. The exon-intron organization, including the nucleotide sequence of the splicing sites of the *Acadm* gene. The sizes of the 12 exons and 11 introns are indicated in basepairs. The consensus splice sequence is illustrated at the bottom. The * asterisks below the *Acadm* nucleotide sequence indicate the consensus splice sequences. The first exon, including the 5' untranslated region, is 224 bp long.

EXON-INTRON ORGANIZATION OF THE MOUSE MCAD GENE

3' INTRON SPLICE SITE		EXON			5' INTRON SPLICE SITE	
	#1		224 bp	AGGCTGCAGG	gtgagc	2450 bp
ttttttccag	#2	GTCCTGAGAA	88 bp	TTTAGTTTTG	gtatgt	3100 bp
tgttcccatg	#3	AGTTGACGGA	98 bp	AAGCGGGGAG	gtgggt	103 bp
ctcctttcag	#4	TACCCGTTCC	70 bp	GAAAGCTGCG	gtaagc	400 bp
gacgctgtag	#5	GTGGCCTTGG	101 bp	TTCTTTGGGG	gtaagt	900 bp
cccttggtag	#6	CAAATGCCTG	81 bp	AATGATGTGT	gtgagt	1500 bp
gatatccaag	#7	GCTTACTGTG	171 bp	AGGCCAACTG	gtaggc	3500 bp
tgtgcgttag	#8	GTATTTCTTG	109 bp	TGGAAAAAAG	gtaaac	2650 bp
ttcactctag	#9	GAACTAAACA	141 bp	CAGACCTACA	gtaagc	700 bp
catgttctag	#10	GTCGCAGCTG	96 bp	GCTAGTGGAG	gtaatc	3900 bp
ttaattctag	#11	CACCAAGGAG	249 bp	GATCTATCAG	gtaagc	2900 bp
tttatctttc	#12	ATTTATGAAG	645 bp	CCTTAACTCC	poly A addition site	

consensus splicing sequence

ccccccncag	G	AAG	gtaagt
ttttttt		C	g

FIG. 3. Nucleotide sequence of the 5' region of the two *Acadm* processed pseudogenes cloned by library screening. The *Acadm* gene sequence is compared with the nucleotide sequence of the two processed pseudogenes. The nucleotide sequence of the *Acadm* cDNA is illustrated. The pseudogenes do not contain any intron sequences. The Kozak consensus translational start site is present below. Asterisks * indicate nucleotide similarities, while carots ^ indicate nucleotide mismatches of the Kozak consensus sequence.

Nucleotide Sequence of Mouse MCAD Gene and MCAD Pseudogenes

```
                                    exon 1                                      intron 1
MCAD gene                aggagccatc ATGGCAGCAG GGTTCCGCAG AGGCTGCAGG gtgagcgggt
                               *    ****  *
                                    exon 1                                      exon 2
Pseudogene (clone 1)     aggagccagc acggcagcgg cgtt-cgcag aagctgcagg gtcctga-aa
                               *    ^^**  *
                                    exon 1                                      exon 2
Pseudogene (clone 16)    aggagccata ccggcagc-g cgttgc-cag aagctgcagg gtcctgagaa
                               *    ^^**  *
                                    exon 1                                      exon 2
MCAD cDNA                           ATGGCAGCAG GGTTCCGCAG AGGCTGCAGG GTCCTGAGAA
                                    ****  *
     Kozak Consensus Sequence:  A-- ATGG--G
```

FIG. 4. Southern blot for RFLPs between parental mouse strains C57BL/6J (lanes marked B), CAST/Ei (lanes marked C) and SPRET/Ei (lanes marked S) that were detected using the entire *Acadm* cDNA as a hybridization probe. DNA fragments indicated by arrows in lanes marked B were mapped using the BSS backcross; fragments indicated by arrows in lanes marked C were mapped using the BCB backcross. DNA fragments marked by the letter a mapped to the *Acadm* gene locus on Chr 3, fragments marked b mapped to the *Acadm-ps1* locus on Chr 11 and fragments marked c mapped to the *Acadm-ps2* locus on Chr 8. The *Pvu*II fragments marked by d in lanes B, C and S were the only fragments detected with the *Acadm* intron probe. Using the BSS backcross, the *Pvu*II fragment marked by d in lane B was mapped to the *Acadm* gene locus on Chr 3. *Hind*III-digested λDNA fragment sizes are marked by arrows and correspond to 9.4, 6.6, 4.4, 2.3, 2.0 and 0.6 kb.

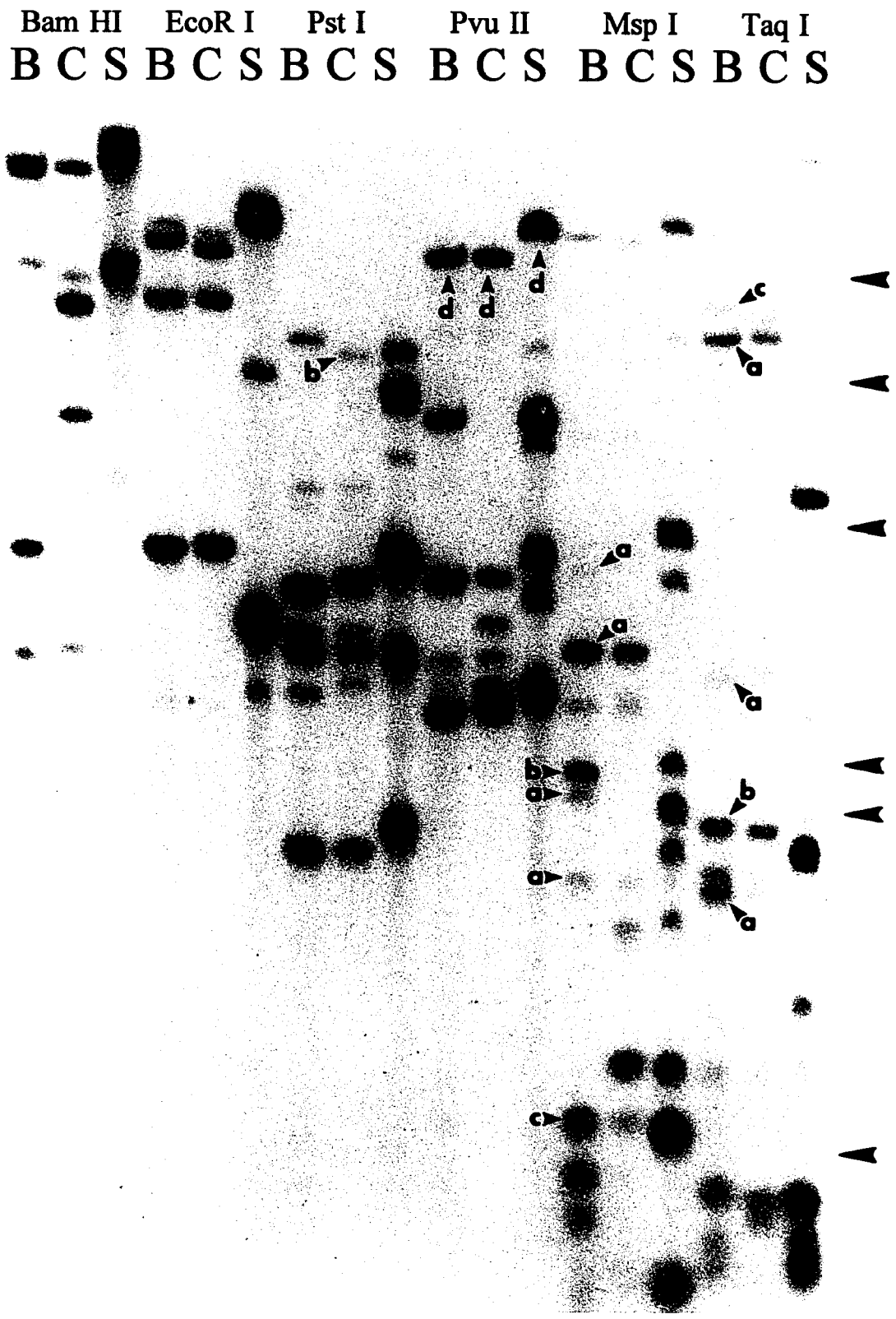
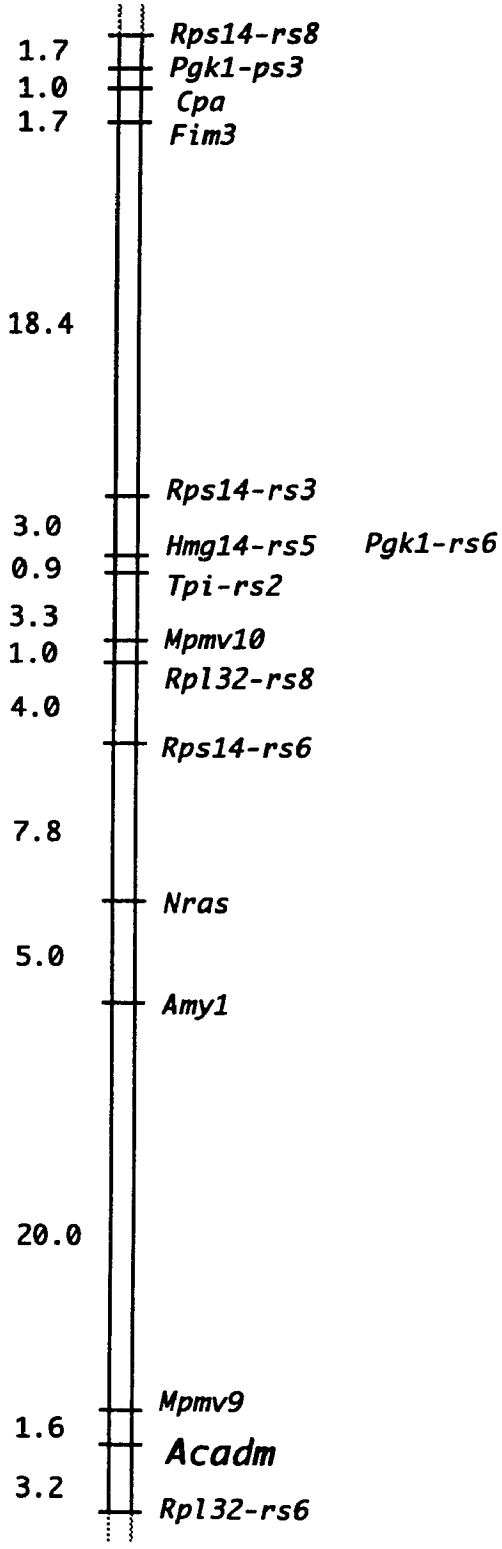
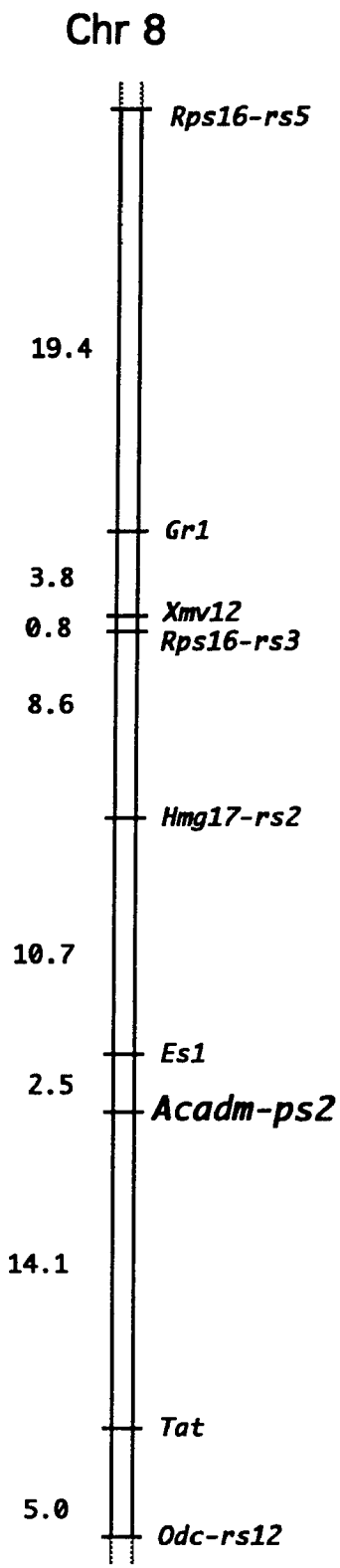

Bam HI
EcoR I
Pst I
Pvu II
Msp I
Taq I
B C S B C S B C S B C S B C S B C S
a
b
c
d

FIG. 5. Chromosomal locations of the *Acadm* gene and pseudogenes as determined from analysis of the BSS backcross. 5a) The *Acadm* is located at the distal end of Chr 3 near *Mpmv9*. 5b) and 5c) The two *Acadm* processed pseudogenes are located within Chr 8 and 11.


Chr 3
Rps14-rs8
1.7
Pgk1-ps3
1.0
Cpa
1.7
Fim3
18.4
Rps14-rs3
3.0
Hmg14-rs5
Pgk1-rs6
0.9
Tpi-rs2
3.3
Mpmv10
1.0
Rpl32-rs8
4.0
Rps14-rs6
7.8
Nras
5.0
Amy1
20.0
Mpmv9
1.6
Acadm
3.2
Rpl32-rs6


Chr 8
Rps16-rs5
19.4
Gr1
3.8
Xmv12
0.8
Rps16-rs3
8.6
Hmg17-rs2
10.7
Es1
2.5
Acadm-ps2
14.1
Tat
5.0
Odc-rs12

Chr 11

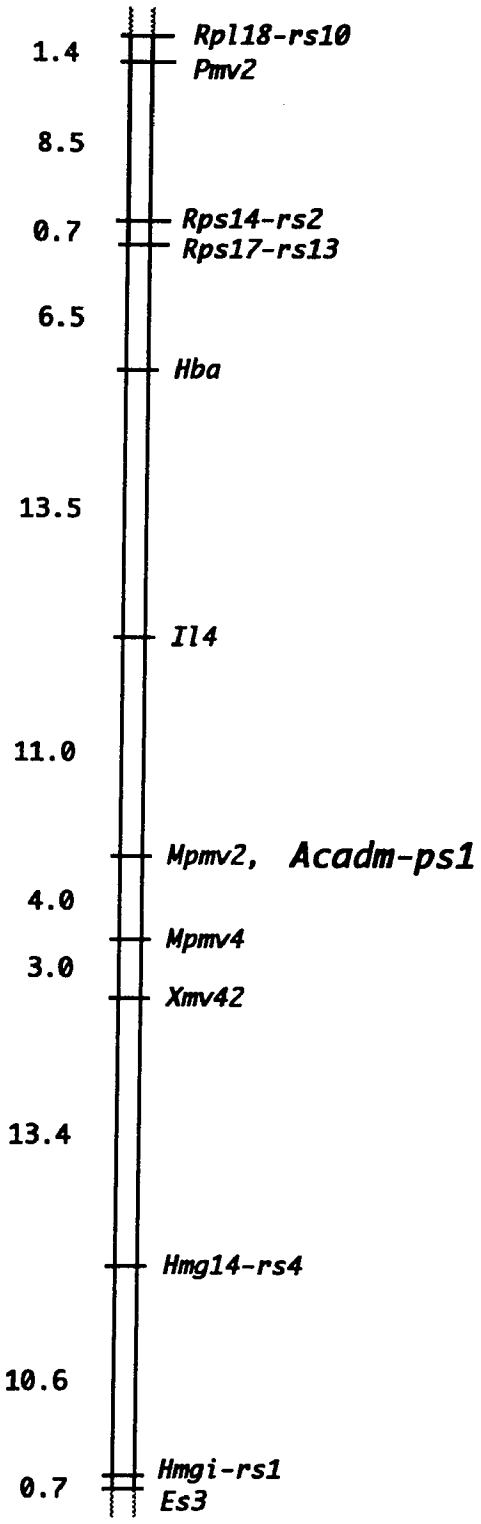

Rpl18-rs10
1.4
Pmv2
8.5
Rps14-rs2
0.7
Rps17-rs13
6.5
Hba
13.5
Il4
11.0
Mpmv2, Acadm-ps1
4.0
Mpmv4
3.0
Xmv42
13.4
Hmg14-rs4
10.6
Hmgi-rs1
0.7
Es3

FIG. 6. The 5' regulatory region of the *Acadm* gene. Nucleotide sequence of the 5' regulatory region including exon 1 and portion of intron 1 of the *Acadm* gene. The principal transcriptional start site (TSS) as determined by primer extension is illustrated by a vertical arrow and labeled +1. Translated Exon 1 sequence is capitalized. The sequence of putative regulatory elements is highlighted and labeled. Arrows indicate consensus sequences to P Boxes. The region labeled NRRE for nuclear receptor response element may be capable of interacting with multiple nuclear receptors. Several Sp1 binding sites are labeled along with consensus sequences to the glucocorticoid response element (GRE), insulin response element (IRE) and Ap2 element (Ap2).

5′ Regulatory Region of the Mouse MCAD Gene

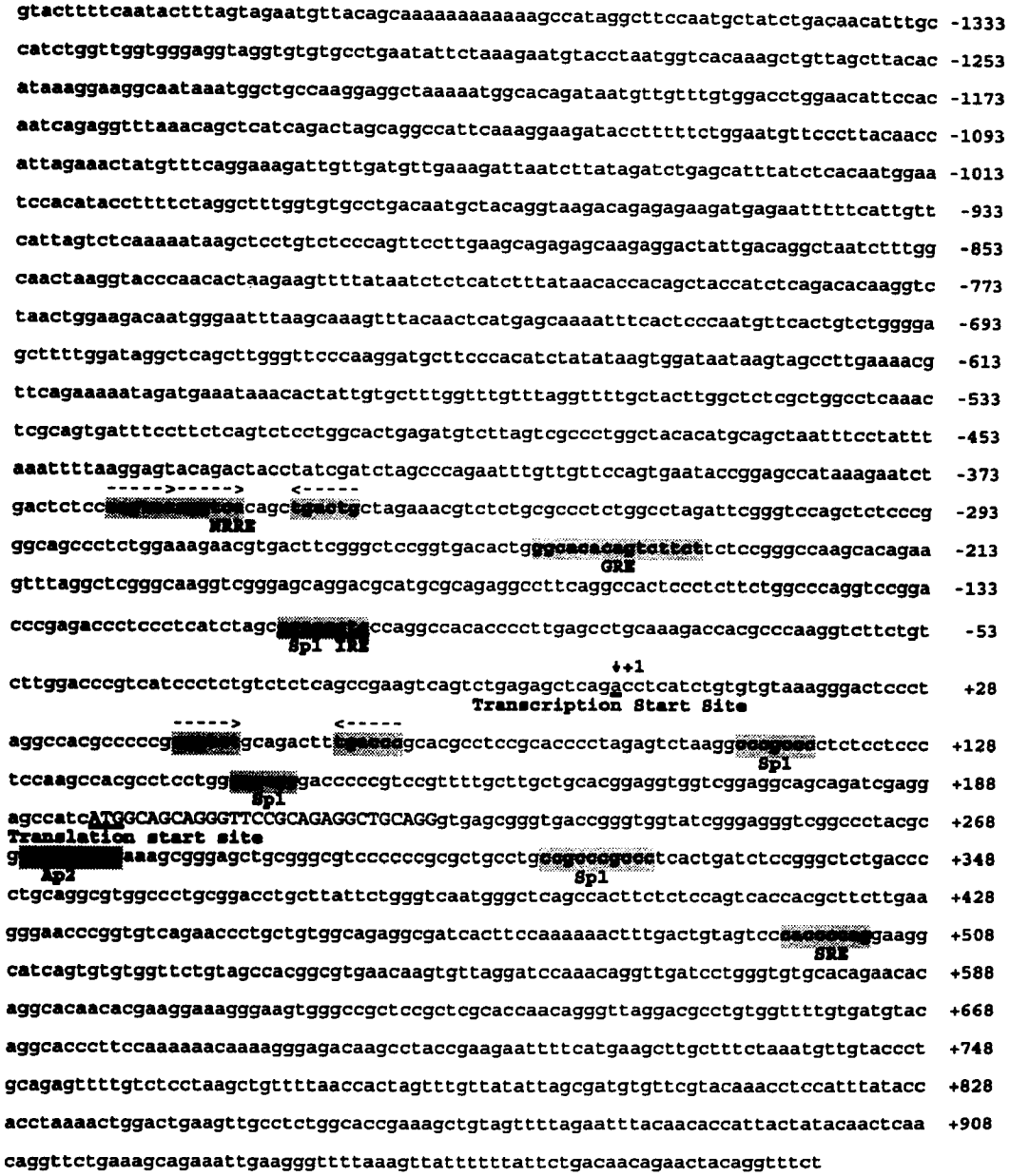


Table 1. RFLP linkage analysis

Locus Pair	r	N	%r	SE(%r)
BSS backcross:				
Chr 3				
Nras:Amy1	6	119	5.04	2.01
Amy1:Mpmv9	12	60	20.00	5.16
*Mpmv9:**Acadm***	1	61	1.64	1.63
***Acadm**:Rpl32-rs6*	4	127	3.15	1.55
Chr 8				
Hmg17-rs2:Es1	15	140	10.71	2.61
*Es1:**Acadm-ps2***	3	120	2.50	1.43
***Acadm-ps2**:Tat*	15	106	14.15	3.39
Chr 11				
Il4:Mpmv2	15	136	11.03	2.69
*Mpmv2:**Acadm-ps1***	0	125	0.00	0.00
***Acadm-ps1**:Mpmv4*	5	125	4.00	1.75
Mpmv4:Xmv42	4	132	3.03	1.49
Xmv42:Hmg14-rs4	18	134	13.43	2.95
BCB backcross:				
Chr 11				
*Il4:**Acadm-ps1***	21	142	14.79	2.98
***Acadm-ps1**:Rps6-rs6*	20	137	14.60	3.02
Rps6-rs6:Hmg14-rs4	6	135	4.44	1.77

Backcross analysis of gene order and recombination frequencies. Previously mapped loci were used to position *Acadm* and two *Acadm* related sequences, *Acadm-ps1* and *Acadm-ps2*. Gene order was determined by minimizing double crossover events; loci on each chromosome are listed from centromere to distal telomere.

Table 2. 5' regulatory elements

Binding Site	**Element**	**Sequence in *Acadm* Gene and Location**	
GGGCGG or CCCGCC	GC Box	CCCGCC	-109
		CCCGCC	+111
		GGGCGGG	+147
		CCGCCC	+315
		CCCGCC	+318
GCCTGGGGG	Ap2	GCCTGGGGG	+270
RG(G/T)TCA	P Box	AAGTAA	-364
		AGGTCA	-358
		CAGTCA	-343
		AGGTCT	+42
		GGGTCA	+56
GGTACAnnnTGTTCT	GRE	GGCACACAGTCTTCT	+270
STGSSGYG	SRE	CTGGGGTG	+495
CCCGCCTC	IRE	CCCGCCTC	-109

Identification and localization of potential *cis*-acting regulatory elements and *trans*-acting factors within the promoter region of the *Acadm* gene. GRE = glucocorticoid response element; IRE = insulin response element; SRE = sterol response element.

The mouse *Acadm* gene was mapped previously by others to Chr 8 by using a 1000 bp cDNA fragment as a probe (Bahary *et al.*, 1991). This fragment was produced by PCR amplification from a rat cDNA library using oligonucleotides designed according to the sequences of human and rat cDNA. RFLP analysis with C57BL/6J, DBA/2J and *M. spretus* mice demonstrated linkage to markers on Chr 8 (Bahary *et al.*, 1991). The *ACADM*, on the other hand, has been mapped to Chr 1p31 (Kidd *et al.*, 1990). The purported map location of the mouse *Acadm* gene on Chr 8 is in a region that contains homology to human Chr 16q but not to human Chr 1. Most markers from proximal human 1p have been mapped to mouse Chr 3 (Nadeau *et al.*, 1992). These results, along with the cloning of *Acadm* pseudogenes, prompted us to reexamine the chromosomal location of *Acadm*.

We mapped the *Acadm* gene to the distal end of mouse Chr 3 using the *Acadm* intron fragment as a probe. This extreme distal region of mouse Chr 3 recently has been shown to harbor the mouse retinal pigment epithelium gene *Rpe65* whose human homolog, *RPE65*, maps to Chr 1p31 (Copeland *et al.*, 1993). The human homolog of the mouse *Acadm* gene, *ACADM*, also maps to Chr 1p31 (Kidd *et al.*, 1990) and, thus, confirms and extends this region of conservation between very distal mouse Chr 3 and human Chr 1p31. Our results indicate that the authentic *Acadm* gene is on distal Chr 3 and that Bahary and co-workers (1991) mapped an *Acadm* pseudogene to Chr 8, corresponding to what we have designated *Acadm-ps2*.

The presence of many putative regulatory elements in the promoter region suggests the *Acadm* gene may be highly regulated. Functional analysis of the regulatory elements will give important insight into mechanisms of transcriptional control. Understanding

factors such as the role of fatty acid intermediates on transcriptional control will provide important insight into the pathogenesis of MCAD deficiency. The availability and characterization of the *Acadm* gene are a step toward the development of an animal model of MCAD deficiency which may also be helpful in better understanding the pathogenesis of this disease.

Acknowledgements

We thank Dr. Dan P. Kelly, Washington University, St. Louis, for advice and preprints of his related work. We also thank Dr. Myron E. Hinsdale for invaluable advice on work with RNA. We thank Dr. Tim Townes for advice on the primer extension experiments. Thanks go to Dr. Jeff Engler, Director of the Oligonucleotide Shared Facility, Comprehensive Cancer Center, U.A.B., for preparation of the oligonucleotide primers. This work was supported by N.I.H. grants RR-02599 (P.A.W.), GM46697 (K.R.J.), RR01183 (M.T.D.) and T-32-RR-07003 (R.J.T.).

References

Bahary, N., Zorich, G., Pachter, J.E., Leibel, R.L., and Friedman, J.M. (1991). Molecular genetic linkage maps to mouse chromosome 4 and 6. *Genomics* **11**: 33-47.

Beato, M. (1989) Gene regulation by steroid hormones. *Cell* **56**: 335-344.

Carter, M.E., Gulick, T., Moore, D.D., and Kelly, D.P. (1994). A pleiotropic element in the medium-chain acyl coenzyme A dehydrogenase gene promoter mediates transcriptional regulation by multiple nuclear receptor transcription factors and defines novel receptor-DNA binding motifs. *Mol. Cell. Biol.* **14**: 4360-4372.

Carter, M.E., Gulick, T., Raisher, B.D., Caira, T., Ladias, J.A.A., Moore, D.D., and Kelly, D.P. (1993). Hepatocyte nuclear factor-4 activates medium chain acyl-CoA dehydrogenase gene transcription by interacting with a complex regulatory element. *J. Biol. Chem.* **268**: 13805-13810.

Copeland, N.G., Jenkins, N.A., Gilbert, D.J., Eppig, J.T., Maltais, L.J., Miller, J.C., Dietrich, W.F., Weaver, A., Lincoln, S.E., Steen, R.G., Stein, L.D., Nadeau, J.H., and

Lander, E.S. (1993) A genetic linkage map of the mouse: Current applications and future prospects. *Science* **262**: 57-66.

Courey, A.J., Holtzman, D.A., Jackson, S.P., and Tigan, R. (1989). Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. *Cell* **59**: 827-836.

Davis, L.G., Dibner, M.D., and Battey, J.F. (1986). "Basic Methods in Molecular Biology," Elsevier Science Publishing Co, Inc., New York.

Dreyer, C., Krey, G., Keller, H., Givel, F., Helftenbein, G., and Wahli, W. (1992). Control of the peroxisomal B-oxidation pathway by a novel family of nuclear hormone receptors. *Cell* **68**: 879-887.

Feinberg, A.P., and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**: 6-13.

Finocchiaro, G., Ito, M., and Tanaka, K. (1987). Purification and properties of short chain acyl-CoA, medium chain acyl-CoA, and isovaleryl-CoA dehydrogenases from human liver. *J. Biol. Chem.* **262**: 7982-7989.

Gardiner-Garden, M., and Frommer, M. (1987). CpG islands in vertebrate genomes. *J. Mol. Biol.* **196**: 261-282.

Gulick, T., Cresci, S., Caira, T., Moore, D.D., and Kelly, D.P. (1994). The peroxisome proliferator activated receptor regulates mitochondrial fatty acid oxidative enzyme gene expression. *Proc. Natl. Acad. Sci. U.S.A.*, In press.

Hannink, M., and Temin, H.M. (1990). Structure and autoregulation of the c-rel promoter. *Oncogene* **5**: 1843-1850.

Hernandez, N. (1993). TBP, a universal eukaryotic transcription factor. *Genes Dev.* **7**: 1291-1308.

Ikeda, Y., Keese, S.M., Fenton, W.A., and Tanaka, K. (1987). Biosynthesis of four rat liver mitochondrial acyl-CoA dehydrogenases: in vitro synthesis, import into mitochondria, and processing of their precursors in a cell-free system and in cultured cells. *Arch. Biochem. Biophys.* **252**: 662-674.

Johnson, K.R., Cook, S.A., Davisson, M.T. (1992). Chromosomal localization of the murine gene and two related sequences encoding high-mobility-group I and Y proteins. *Genomics* **12**: 503-509.

Johnson, K.R., Cook, S.A., Davisson, M.T. (1994). Identification and genetic mapping of 151 dispersed members of 16 ribosomal protein multigene families in the mouse. *Mamm. Genome*, In press.

Jorcyk, C.L., Watson, D.K., Mavrothalassitis, G.J., and Papas, T.S. (1991). The human Ets-1 gene: Genomic structure, promoter characterization and alternative splicing. *Oncogene* **6**: 523-532.

Kelly, D.P., Gordon, J.I., Alpers, R., and Strauss, A.W. (1989). The tissue-specific expression and developmental regulation of two nuclear genes encoding rat mitochondrial proteins - medium chain acyl-CoA dehydrogenase and mitochondrial malate dehydrogenase. *J. Biol. Chem.* **264**: 18921-18925.

Kidd, J.R., Matsubara, Y., Castiglione, C.M., Tanaka, K., and Kidd, K.K. (1990). The locus for the medium-chain acyl-CoA dehydrogenase gene on chromosome 1 is highly polymorphic. *Genomics* **6**: 89-93.

Ladias, J.A.A., Hadzopoulou-Chadaras, M., Dardassis, D., Cardot, P., Cheng, J., Zannis, V., and Cladaras, C. (1992). Transcriptional regulation of human apolipoprotein genes ApoB, ApoCIII, and ApoA1 by members of the steroid hormone receptor superfamily HNF-4, ARP-1, EAR-2, and EAR-3. *J. Biol. Chem.* **267**: 15849-15860.

Ladias, J.A.A., and Karathanasis, S.K. (1991). Regulation of the apolipoprotein A1 gene by ARP-1, a novel member of the steroid receptor superfamily. *Science* **251**: 561-565.

Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. (1992). CpG islands as gene markers in the human genome. *Genomics* **13**: 1095-1107.

Linney, E. (1992). Retinoic acid receptors: transcription factors modulating gene regulation, development, and differentiation. *Curr. Top. Dev. Biol.* **27**: 309-351.

Manly, K.F. (1993). A Macintosh program for storage and analysis of experimental genetic mapping data. *Mammal. Genome* **4**: 303-313.

Matsubara, Y., Narisawa, K., Tada, K., Ikeda, H., Yeqi, Y., Danks, D.M., Green, A., and and McCabe, E.R.B. (1992). Prevalence of K329E mutation in the medium-chain acyl-CoA dehydrogenase gene determined from Guthrie cards. In "New Developments in Fatty Acid Oxidation" (P.M. Coates, and K. Tanaka, Eds.), pp.453-462, Wiley-Liss, New York.

Naar, A.M., Boutin, J.M., Lipkin, S.M., Yu, V.C., Holloway, J.M., Glass, C.K., and Rosenfeld, M.G. (1991). The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. *Cell* **65**: 1267-1279.

Nadeau, J.H., Davisson, M.T., Doolittle, D.P., Grant, P., Hillyard, A.L., Kosowsky, M., and Roderick, T.H. (1992). Comparative map for mice and humans. *Mamm. Genome* **3**: 480-536.

Nagao, M., Parimoo, B., and Tanaka, K. (1993). Developmental, nutritional, and hormonal regulation of tissue-specific expression of the genes encoding various acyl-CoA

dehydrogenase and alpha-subunit of electron transfer flavoprotein in rat. *J. Biol. Chem.* **268**: 24114-24124.

Raisher, B.D., Gulick, T., Zhang, Z., Strauss, A.W., Moore, D.D., and Kelly, D.P. (1992). Identification of a novel retinoid-response element in the promoter region of the medium chain acyl-Coenzyme A dehydrogenase gene. *J. Biol. Chem.* **267**: 20264-20269.

Roe, C.R., and Coates, P.M. (1989). Acyl-CoA dehydrogenase deficiencies. In "The Metabolic Basis of Inherited Disease" (C.R. Scriver, A.L. Beaudet, W.S. Sly and D. Valle, Eds.), pp.889-914, McGraw-Hill, Inc., New York.

Rottman, J.N., and Gordon, J.I. (1993). Comparison of the patterns of expression of rat intestinal fatty acid binding protein/human growth hormone fusion genes in cultured intestinal epithelial cell lines and in the gut epithelium of transgenic mice. *J. Biol. Chem.* **268**: 11994-12002.

Rottman, J.N., Widom, R.L., Nadal-Ginard, B., Mahdavi, V., and Karathanasis, S.K. (1991). A retinoic acid-responsive element in the apolipoprotein A1 gene distinguishes between two different retinoic acid response pathways. *Mol. Cell. Biol.* **11**: 3814-3820.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain termination inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **75**: 5463-5467.

Smale, S.T., and Baltimore, D. (1989). The "Initiator" as a transcriptional control element. *Cell* **57**: 103-113.

Tolwani, R.J., Farmer, S.C., and Wood, P.A. (1994). Molecular cloning and characterization of the mouse medium-chain acyl-CoA dehydrogenase cDNA. *Genomics* **23**: 247-249, 1994

Umesono, K., and Evans, R.M. (1989). Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell* **57**: 1139-1146.

Wood, P.A., Farmer, S.C., Tolwani, R.J., Warren, J.R., Steinkampf, M.P., Johnson, L.R., Mountz, J.D., and Kelly, D.P. (1992). Molecular studies of mouse medium- and long-chain acyl-CoA dehydrogenase genes for site-directed mutagenesis of embryonic stem cell. In "New Developments in Fatty Acid Oxidation" (P.M. Coates and K. Tanaka, Eds.), pp. 151-160, Wiley-Liss, New York.

Workshop on Molecular Aspects of MCAD Deficiency. (1992). Mutations causing medium-chain acyl-CoA dehydrogenase deficiency: a collaborative compilation of the data from 172 patients. In "New Developments in Fatty Acid Oxidation" (P.M. Coates and K. Tanaka, Eds.), pp. 499-506, Wiley-Liss, New York.

Ye, K., Dinarello, C.A., and Clark, B.D. (1993). Identification of the promoter region of human interleukin 1 type 1 receptor gene: Multiple initiation sites, high G+C content, and constitutive expression. *Proc. Natl. Acad. Sci. U.S.A.* **90**: 2295-2299.

Zhang, Z., Kelly, D.P., Kim, J.-J., Zhou, Y., Ogden, M.L., Whelan, A.J., and Strauss, A.W. (1992). Structural organization and regulatory regions of the human medium-chain acyl-CoA dehydrogenase gene. *Biochemistry* **31**: 81-89.

TARGETED DISRUPTION OF THE MOUSE MEDIUM-CHAIN ACYL-COA DEHYDROGENASE GENE

RAVI J. TOLWANI, SUSAN C. FARMER, CARL A. PINKERT, DOUG A. HAMM, AND PHILIP A. WOOD

Department of Comparative Medicine, Schools of Medicine and Dentistry
The University of Alabama at Birmingham

To be submitted to *Genomics*

Abstract

Medium-chain acyl-CoA dehydrogenase (MCAD) is involved in the initial step of fatty acid β-oxidation and is essential for complete fatty acid oxidation. MCAD deficiency exists in humans as an autosomal recessive disorder with an incidence of disease estimated at 1:15,000 (Matsubara *et al.*, 1992). Patients with MCAD deficiency, under fasting conditions, develop a wide range of metabolic derangements which often lead to death (Roe and Coates, 1989). The pathogenesis of the biochemical problems is poorly understood and certain aspects of patient management are controversial. An animal model for MCAD deficiency could be very important for improving understanding of disease pathogenesis and for better therapeutic regimens. We report here the generation of embryonic stem (ES) cell lines disrupted at the MCAD locus by an insertion targeting vector. This vector was designed to duplicate exons 8, 9 and 10 at the MCAD locus. The codon reading frame is altered upon translation of any mRNA transcribed from this duplicated region, resulting in the formation of premature stop codons. The resulting monomer would be without the C terminal α-helixes including a major portion of the H and the entire I α-helixes responsible for the intersubunit contacts for forming the functional homotetramer (Kim and Wu, 1988; Kim *et al.*, 1993). This mutation resembles the common human mutation where the mutant MCAD monomer is unable to assemble into functional homotetramers (Kim *et al.*, 1993). After electroporation and selection, ES cells with the desired recombination were detected by polymerase chain reaction dependent on gap repair and subsequently confirmed by Southern blot hybridization. Subsequently, disrupted ES cells were used to generate chimeric mice.

INTRODUCTION

β-oxidation of fatty acids provides energy, especially during fasting conditions. Fatty acid oxidation occurs in the mitochondria and consists of a repeating circuit of four sequential steps. There are four straight chain acyl-CoA dehydrogenases involved in the initial step. Medium-chain acyl-CoA dehydrogenase (MCAD), specifically, is responsible for catalyzing the dehydrogenation of fatty acid thioesters between C_4 and C_{16} carbons in length (Roe and Coates, 1989). MCAD is transcribed in the nucleus, synthesized in the cytosol and translocated into the mitochondrial matrix (Matsubara *et al.*, 1987; Kelly *et al.*, 1987; Ikeda *et al.*, 1987). Once inside the mitochondrial matrix, the MCAD monomers are assembled into homotetramers to gain enzymatic activity (Ikeda *et al.*, 1987).

MCAD activity is essential for complete fatty acid oxidation. Inherited MCAD deficiency exists in humans as an autosomal recessive disorder. MCAD deficiency was first described by Kolvraa and associates in 1982 and has been described in hundreds of patients (Kolvraa *et al.*, 1982; Roe and Coates, 1989; Coates, 1992; Matsubara *et al.*, 1992; Ding *et al.*, 1992; Workshop on Molecular Aspects of MCAD Deficiency, 1992). The carrier frequency in the Caucasian population has been estimated to be between 1 in 50 to 80 with an incidence of clinical disease expected at around 1 in 15,000 (Matsubara *et al.*, 1992; Bennett *et al.*, 1987; Roe and Coates, 1989).

The most common mutation in MCAD deficient patients results from a nucleotide substitution of deoxyadenosine to deoxyguanosine at position 985 of the cDNA (Roe and Coates, 1989; Ding *et al.*, 1992; Matsubara *et al.*, 1992; Coates, 1992; Kelly *et al.*, 1987; Strauss *et al.*, 1990; Workshop on Molecular Aspects of MCAD Deficiency, 1992). This

results in the substitution of the basic amino acid lysine for the acidic amino acid glutamic acid at position 304 of the mature monomer. Amino acid 304 is located in an H α-helix which is in part responsible for the intersubunit contacts of the MCAD monomer to form homotetramers. The change in charge in this region is thought to interfere with MCAD tetramer formation which is essential for MCAD activity (Kim and Wu, 1988; Kim *et al.*, 1993). Approximately 90% of patients are homozygous and an additional 7% are compound heterozygotes for this point mutation (Workshop on Moledular Aspects of MCAD Deficiency, 1992).

MCAD deficient patients exhibit clinical episodes often associated with fasting. Patients present with signs usually during the first two years of life. Patients can develop hypoglycemia, disrupted ketogenesis, hyperammonemia, aciduria, secondary carnitine deficiency, fatty change of the liver and encephalopathy (Roe and Coates, 1989). It is estimated that approximately 59% of patients between 15 to 26 months of age die during their first clinical episode (Roe and Coates, 1989).

The pathogenesis of the wide range of metabolic disturbances in MCAD deficiency is poorly understood and certain aspects of patient management are controversial. No animal model for MCAD deficiency currently exists. An animal model for MCAD deficiency could be helpful in better understanding the pathogenesis of MCAD deficiency and in developing better management regimens for human patients. To gain further insight into the mechanisms of this disease, we sought to develop a mouse model of MCAD deficiency by gene targeting in embryonic stem (ES) cells. Site directed mutagenesis in ES cells provides a means of developing mouse models of human disease (for reviews see Capecchi, 1989; Koller and Smithies, 1992). Here we describe the

development of ES cell lines with targeted disruption of the MCAD gene. These ES cells have been used to generate chimeras and ultimately develop homozygous MCAD deficient mice.

MATERIALS AND METHODS

Library Screening

Mouse MCAD cDNA fragments (Tolwani *et al.*, 1994b) labeled with ^{32}P-dCTP were used as a probe to screen a Lambda FIX II 129 Sv mouse genomic library. Genomic inserts from purified phage clones were digested free from the phage arms by *Not*I digestion and subcloned into the *Not*I site of calf intestinal alkaline phosphatase treated pGEM11Z(f)+ (Promega) (Tolwani *et al.*, 1994a). These clones were further analyzed by restriction digestions with *Ava*I, *Bam*HI, *Bgl*I, *Bgl*II, *Bst*EII, *Bst*XI, *Cla*I, *Dra*I, *Eco*RI, *Hind*III, *Kpn*I, *Msc*I, *Mro*I, *Nsi*I, *Nsp*I, *Pst*I, *Pvu*I, *Pvu*II, *Sac*I, *Sma*I, *Sph*I, *Sty*I, *Xba*I and *Xho*I. Digested DNA was electrophoresed on 1% agarose gels, blotted and probed with various mouse MCAD cDNA fragments. Additionally, the exon/intron boundaries were sequenced with the dideoxy chain termination method with Sequenase (U.S.B.).

Construction of Targeting Vectors

A neomycin cassette containing the phosphoglycerate kinase promoter and bovine growth hormone polyadenylation signal (Soriano *et al.*, 1991) was digested free with *Xho*I and subcloned into the *Sal*I site of pGEM11Z(f)+. The plasmid was digested with *Eco*RI and the overhangs were filled in using 2.5 units of Klenow fragment and 2mM dNTPs for 15 min at 37°C in 2mM Tris pH 8.0 and 10mM $MgSO_4$ buffer. Subsequent ligation of the blunt ends recircularized the vector and destroyed the *Eco*RI site within the polylinker

of the pGEM11Z(f+) plasmid. Next, an eight kilobase (kb) MCAD genomic fragment was directionally cloned into the *Not*I and *Xho*I sites of pGEM11Zf(+) to complete the construction of MCAD insertion vector (IV) 1. MCAD IV1 DNA was prepared by CsCl banding technique and linearized with *Bam*HI and *Eco*RI restriction digestion.

To design MCAD IV2, MCAD IV1 was digested with *Bam*HI and *Eco*RI. The vector was purified free from the 1.3 kb *Bam*HI/*Eco*RI fragment by gel purification. Three oligonucleotides were designed: #1 5'-AATTGTCGACA-3'; #2 5'-GATCGTCGACA-3'; and #3 5'-TCGATGTCGAC-3'. The three oligonucleotides were used in a ligation reaction with the digested and purified vector. The intent was to ligate the long arm to the short arm of homology to remove the 1.3 kb fragment and destroy the *Bam*HI and *Eco*RI sites and create a new and unique *Sal*I site. The MCAD IV2, therefore, is linearized for electroporation with *Sal*I digestion.

Cell Culture and Transfection

SNL76/7 feeder cells (McMahon and Bradley, 1990) were grown to confluency on triple flasks (Nunc) using Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% newborn calf serum (Gibco). Mitomycin-C treatments were done by adding 0.5 mg mitomycin powder (Sigma) to 50 ml of media per each triple flask. These flasks were rinsed with phosphate buffered saline (PBS) three times after incubation for 2 hours at 37°C with mitomycin-C. These cells were then trypsinized and frozen at 5 X 10^6 cells/ml. Approximately one frozen vial of mitomycin-C treated SNL76/7 was seeded on 2-100mm gelatin coated plates.

AB1 ES cells (McMahon and Bradley, 1990) were grown on a confluent mitomycin-C treated SNL76/7 feeder layer at 37°C and 10% CO_2 in DMEM medium

(Cellgro) supplemented with 15% fetal calf serum (Gibco), 10^{-4}M β-mercaptoethanol, non-essential amino acids and nucleosides. ES cells were passaged every 2 days.

AB1 ES cells were trypsinized and resuspended at a density of 10^7 cells in 0.8 ml of PBS and placed in a 0.4 cm cuvette with 25 μg of MCAD targeting vector. MCAD IV1 was prepared for electroporation by digestion with *Bam*HI and *Eco*RI, extracted with phenol and chloroform, ethanol precipitated and resuspended in sterile water at a concentration of 1 μg/ml. MCAD IV2 was digested with *Sal*I and subsequently purified. The electroporations were done at 0.23 kV and 575 μF using a Bio-Rad gene pulser. After a 10 minute recovery, the electroporated cells were plated onto 6 100mm dishes containing mitomycin-C inactivated SNL76/7 feeder cell layer. A nonelectroporated ES cell control plate was also seeded at the same time. Selection with G-418 (Geneticin-GIBCO) at 350 μg/ml of media was applied 24 hours post electroporation and continued for 10 days. Resistant colonies were transferred to individual wells of a 24 well microtiter dish seeded with mitomycin-C inactivated SNL76/7 cells. These colonies were trypsinized individually after 2 days to disperse ES cells over the entire well and then grown for an additional 3 days. At this time the wells were trypsinized and approximately 80% of the cells were frozen. The remaining cells in the wells were expanded again for 4-5 days and used for DNA preparation. A total of 3 separate experiments were completed with MCAD IV1 electroporated in the first experiment and MCAD IV2 electroporated in subsequent experiments.

Analysis of ES Cell Colonies

Genomic DNA was isolated from expanded cells in 24 well microtiter plates. Cells from each well were incubated individually in 200 μl of lysis solution (0.5% SDS,

50mM NaCl, 25mM EDTA and 200 μg/ml protinase K) overnight at 55°C. This was followed by the addition of 80 μl of saturated NaCl solution, shaking for 20 seconds and centrifugation for 15 minutes. Approximately 500μl of 100% ethanol was added to the supernatant in a new tube. The precipitated DNA was removed with glass pipettes, washed in 70% ethanol and pelleted, followed by resuspension in 25μl of sterile water.

Genomic DNA from resistant clones was screened for homologous recombinants by polymerase chain reaction (PCR). Initially, the samples were amplified with 15 pmol each of primers derived from MCAD sequence from exon 10 (5'-GACATTTGGAAAGCTGCTAGTG-3') and from the PGKNeopA sequence (5'-GACTGCCTTGGGAAAAGCGCCTC-3'). PCR reactions were done in 100μl final volume in 1X PCR buffer (50mM KCl, 10mM Tris-HCl pH 9.0, 0.1% Triton X-100 and 1.5mM $MgCl_2$) with 1.2 mM dNTPs and 4.5 U of *Taq* polymerase. Amplifications were at 94°C for 1 minute, 57°C for 1 minute and 72°C for 1 minute for 40 cycles. Ten microliters of the product were amplified a second time in a 100μl final volume with the nested primers derived from intron 10 sequence (5'-CTGTGTGGAGACCAGAGGACC-3') and nested PGKNeopA sequence (5'-GAAAAGCGCCTCCCCTACCCG-3'). Approximately 40μl of the final nested PCR products were separated by electrophoresis on a 1% agarose gel to visualize the PCR product.

Southern blots were done to determine and confirm homologous recombination. Approximately 5μg of genomic DNA from experiment #1 was digested with *Msc*I, separated by electrophoresis on a 1% agarose gel and transferred by capillary transfer to Zeta probe membrane (Bio-Rad) and subsequently UV crosslinked. Blotted membranes were hybridized with the 1.6 kb PGKNeopA probe labeled with ^{32}P via random hexamer

priming (Fineberg and Vogelstein, 1983). Genomic DNA from experiments #2 and #3 was digested with *Eco*RI, electrophoresed and blotted. These filters were hybridized with several different DNA probes: a 850 bp PCR product derived from exon 10 to intron 10; a 800 bp PCR product derived from intron 10; and a 550 bp PCR product including portions of intron 9 and exon 10. Hybridization was done in 0.5 M Na_2PO_4 pH7.2, 7% SDS for 24 hours, followed by washing twice in 40mM Na_2PO_4 and 0.5% SDS at 65°C for 45 min.

Generation of Chimeric Mice

Approximately 10-15 ES cells were injected into the blastocoele cavity of 3.5 day blastocysts from C57BL/6J mice (Hasty and Bradley, 1993). These blastocysts were surgically implanted to pseudopregnant ICR mice 2.5 days postcoitum. Pups born were examined for chimerism by determining if agouti pigmentation was present on a black background coat color 5-10 days after birth. The chimeric mice were subsequently mated with wild type mice C57BL/6J mice to generate pups heterozygous for the MCAD mutation.

Results

Vector Construction

A 16 kb MCAD genomic fragment including exons 8 to 12 was cloned from a Lambda FIXII 129 Sv mouse genomic library (Stratagene) (Tolwani *et al.*, 1994a; unpublished results). A 129 Sv mouse genomic library was used since isogenic DNA in targeting vectors increases targeting frequencies due to potentially reduced polymorphisms between vector and target DNA (Waldman and Liskay, 1988; Te Riele *et al.*, 1992; Deng and Capecchi, 1992; Deursen and Wieriga, 1992). Approximately 8 kb of MCAD

homology which included exons 8 to 10 was used to construct the MCAD insertion vector. The vector was designed so it could be linearized by digestions with *Eco*RI and/or *Bam*HI present within the region of MCAD homology. Since these sites must be unique to the vector, other *Eco*RI and *Bam*HI sites within the vector polylinker were initially removed. The *Eco*RI site within the pGEM11Z(f)+ polylinker was destroyed by Klenow fragment and blunt end ligation. The *Bam*HI site within the plasmid polylinker was removed when the 8 kb MCAD genomic fragment was cloned into the *Not*I and *Xho*I sites of pGEM11Z(f)+. The MCAD IV1 was linearized by *Eco*RI and *Bam*HI digestions, resulting in removal of a 1.3 kb fragment encompasing exon 10.

MCAD IV2 was designed to remove the 1.3 kb *Bam*HI to *Eco*RI fragment from the region of homology from MCAD IV1 and create a new and unique *Sal*I site used to linearize the vector to prepare for electroporation.

Identification of Disrupted Cell Line

Each experiment resulted in approximately 2,000 resistant colonies surviving selection with G-418 at 350 μg/ml of medium for 10 days. No ES colonies survived selection from control ES cells plates which were not electroporated. Approximately 150 to 250 ES cell colonies were analyzed for targeted insertion with each experiment. In experiment #1, MCAD IV1 was electroporated after it was linearized with *Bam*HI and *Eco*RI digestions. G-418 resistant ES colonies were pooled in groups of 1 to 3 colonies before genomic DNA isolation, digestion with *Msc*I and transfer to Zeta probe membrane and were hybridized with the Neo probe. The Neo probe hybridized with a 4.3 Kb fragment, expected upon targeted insertion, from *Msc*I digested genomic DNA from 6 out of 155 colonies analyzed (FIG. 3B).

MCAD IV2 was electroporated in experiments #2 and #3. Approximately 150 G-418 resistant colonies were analyzed in experiment #2. A 850 bp probe generated by PCR from exon 10 to intron 10 was used. This DNA fragment is not present within MCAD IV2. The entire 1.3 kb fragment deleted from the vector was not used as a probe due to the presence of repeat sequences (data not shown). This PCR probe hybridized to a 3.1 kb fragment on blots containing *Eco*RI digested genomic DNA. This probe did not hybridize to a 13.2 kb band indicating targeted integration.

Approximately 2μg of genomic DNA obtained from G-418 resistant ES colonies from experiment #3 were amplified by PCR with the initial primer set and subsequently by the nested primer set. Conditions were worked out with a positive control vector to detect PCR product by gel electrophoresis at a sensitivity of 0.05 haploid genomes in 1 μg of genomic DNA (data not shown). Approximately 20 out of 225 colonies amplified by PCR gave the 2.1 kb expected PCR product (FIG. 3A). On Southern blot analysis with the exon 10 external probe, at least 2 of these 20 colonies revealed a 13.2 kb band in addition to the 3.1 kb endogenous band indicating targeted insertion of the vector at the MCAD locus (FIG. 3C).

Microinjection and Chimera Generation

We first determined whether our stock AB1 ES cell line was pleuripotent. AB1 ES cells maintained with no selection were injected into C57BL\6J blastocysts and produced chimeric mice. The best male chimera produced all agouti offspring indicating germ-line transmission of our stock AB1 ES cell line.

ES cells microinjected from clone #41 derived from experiment #1 resulted in 2 male chimeras and 2 female chimeras. These mice have been mated with wild type

C57BL\6J to determine if these mice are germline chimeras. Most recently, other cell lines from experiment #3 have been sucessfully injected into 185 C57BL\6J blastocysts which have been transfered into 18 recipient ICR mice. We have generated chimeric male mice from these last experiments. These mice are being mated with wild type C57BL/6J mice to generate mice heterozygous with MCAD deficiency. Heterozygous MCAD deficient mice will be eventually used to generate mice homozygous for MCAD deficiency.

FIG. 1. Strategy for disruption of the MCAD locus. MCAD IV1 and MCAD IV2, digested with *Eco*RI and *Bam*HI or *Sal*I, respectively, integrate at the endogenous MCAD locus after gap repair of the deleted 1.3 kb exon 10 region. This results in duplication of exon 8, 9 and 10 at the disrupted allele. Small arrows indicate positions of PCR primers used to screen for homologous recombination. Amplification of this region produces the 2.1 kb diagnostic PCR product as illustrated. PCR positive ES cell colonies were then analyzed by Southern blots. Upon targeted disruption probe A, the entire PGKNeopA gene, hybridizes to a 800 bp and a 4.3 kb fragment with *Msc*I digested genomic DNA. Probe B, a DNA fragment consisting of a portion of intron 10 which is not present in the vector, confirms targeted insertion by hybridizing to a 13.2 kb fragment, along with the endogenous 3.1 kb band with EcoRI digested genomic DNA.

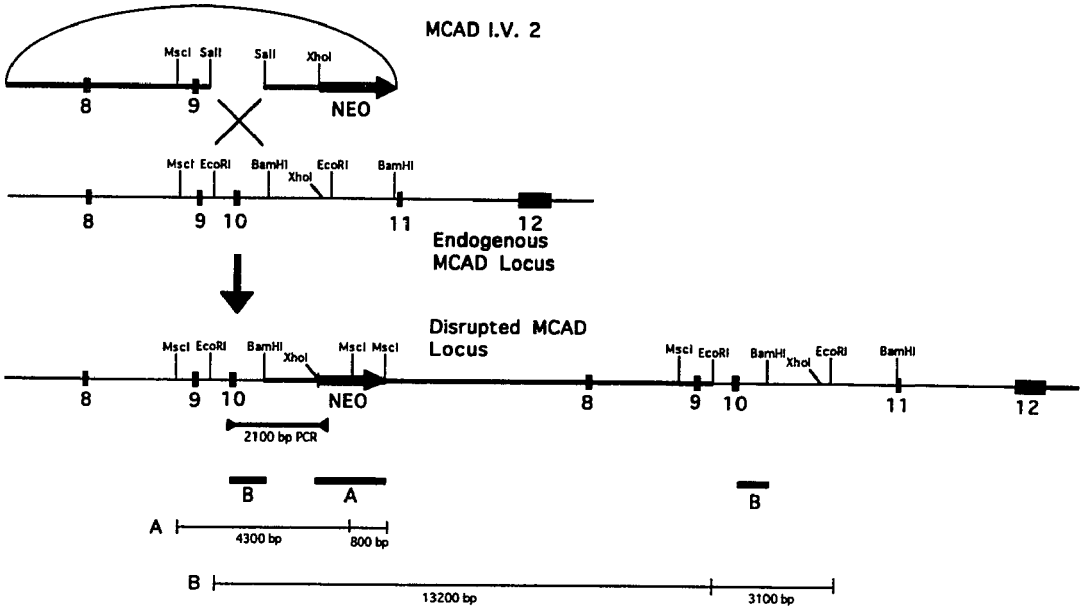

MCAD I.V. 2
MscI SalI
SalI XhoI
8
9
NEO
MscI EcoRI BamHI XhoI EcoRI BamHI
8
9 10
11
12
Endogenous
MCAD Locus
Disrupted MCAD
Locus
MscI EcoRI BamHI XhoI MscI MscI
8
9 10
NEO
2100 bp PCR
MscI EcoRI BamHI XhoI EcoRI BamHI
8
9 10
11
12
B
A
B
A
4300 bp
800 bp
B
13200 bp
3100 bp

FIG. 2. Structure of the MCAD monomer before and after targeted disruption. a) The normal MCAD message produces a MCAD monomer including α-helixes A through L within the N-terminal and C-terminal domains and β-sheets 1 to 7 within the middle domain. b) The mutated message after targeted insertion of exon 8, 9, and 10 (highlighted region). The C-terminal domain is shortened due to premature stop codons produced upon translation of the duplicated exon 8 due to a frameshift mutation. This eliminates the major portion of α-helix H and α-helix I to L entirely. This truncated monomer is unable to assemble into functional homotetramers due to loss of α-helixes responsible for intersubunit contacts.

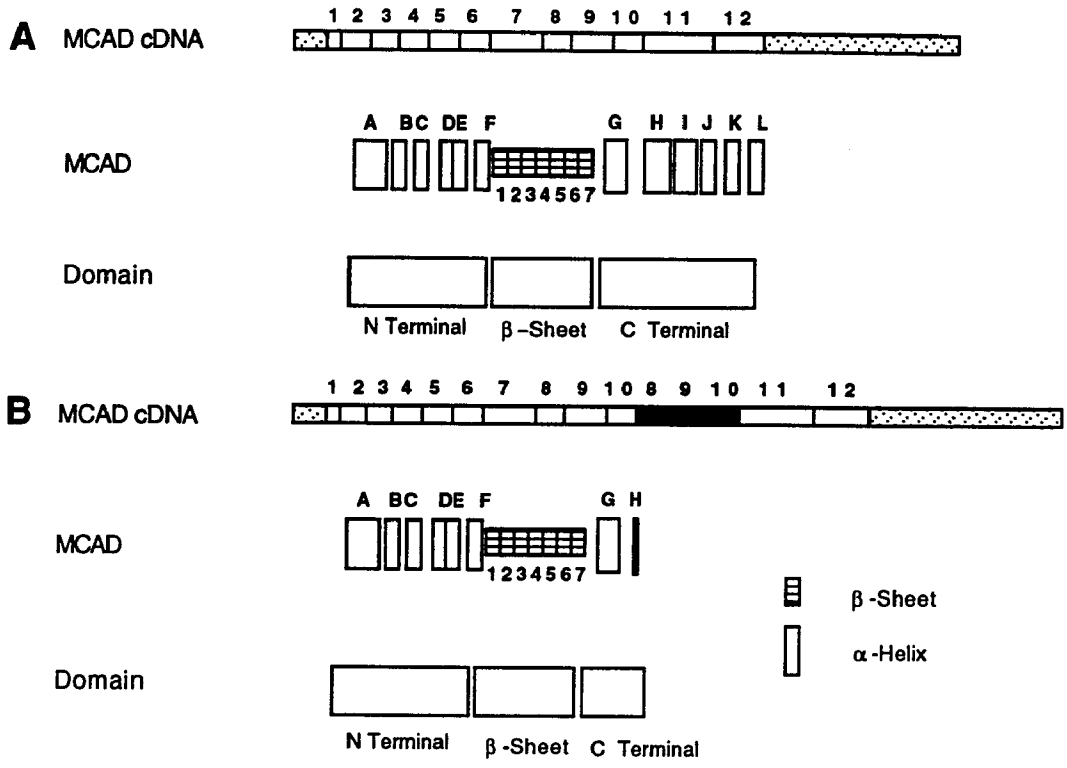
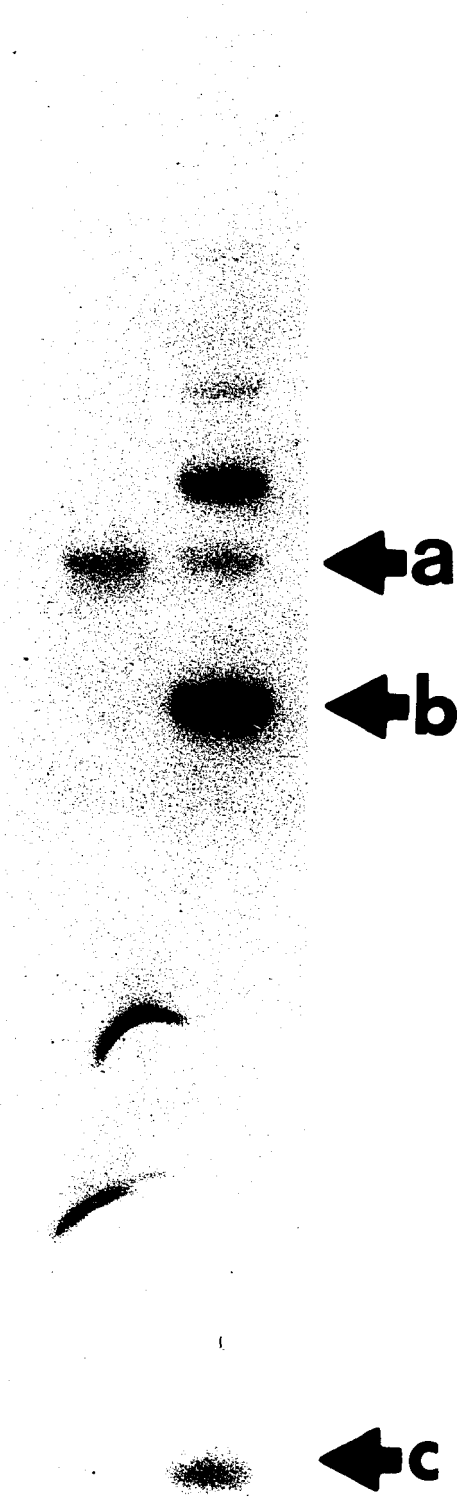

A
MCAD cDNA
1 2 3 4 5 6 7 8 9 10 11 12
MCAD
A BC DE F G H I J K L
1234567
Domain
N Terminal
β-Sheet
C Terminal
B
MCAD cDNA
1 2 3 4 5 6 7 8 9 10 8 9 10 11 12
MCAD
A BC DE F G H
1234567
β-Sheet
α-Helix
Domain
N Terminal
β-Sheet
C Terminal

FIG. 3. Analysis of ES cell colonies by PCR and Southern blot hybridization. a) Amplification of genomic DNA from G-418 resistant ES cell colonies by initial and nested PCR primer sets. Lane 1 = pGEM and Lambda *Hind*III (Promega) molecular weight markers; lane 2 = amplification from positive control template results in 2.1 kb PCR product; lane 3 = amplification from clone #41; lane 4 = amplification from clone #44-3. Arrow indicates diagnostic 2.1 kb PCR product. b) Genomic DNA from MCAD IV1 and MCAD IV2 electroporated G-418 resistant ES cell colonies digested with *Eco*RI, electrophoresed on 1% agarose gel (individual colonies), transferred and probed with external PCR derived probes. Targeted integration of the vectors results in the hybridization of a 13.2 kb fragment (arrow colony #41 and colony #44-3), along with the 3.1 kb endogenous fragment. c) Genomic DNA from MCAD IV1 electroporated G-418 resistant ES cell colonies was digested with *Msc*I, electrophoresed on 1% agarose gel (1-3 colonies per lane), transferred and probed with the NEO probe. Targeted insertion of the MCAD IV1 should produce a 4.3 kb band (arrow a), along with an 800 bp band as seen in lanes 1 and 2 (clones #41 and #42, respectively). Integration of head-to-tail concatamers without gap repair should produce a 3.2 kb band (arrow b) as seen in lane 2 (clone #42). Lane 1 contains genomic DNA from a single ES cell colony that appears to be targeted, while lane 2 contains genomic DNA from multiple ES colonies, of which one appears targeted.

MW
Positive Control
41
44-3
2.1 kb

43-4
44-3
41
13.2 kb
3.1 kb


a
b
c

DISCUSSION

Certain guidelines were used to increase the probability of disrupting the MCAD gene with insertion vectors. Increasing the amount of homology by 2-fold increased the targeting frequency by 20-fold at the HPRT locus (Thomas and Capecchi, 1987). The MCAD insertion vectors were designed to contain approximately 7 kb of homology. We used isogenic DNA in the vector to increase targeting frequency. Additionally, better targeting frequencies are expected when the short arm of the linearized vector is greater than 1 kb in length (Hasty and Bradley, 1993). The MCAD insertion vectors were designed so that the short arm was 1.2 kb in size. These characteristics should maximize targeting frequencies at the MCAD locus. We have previously shown that MCAD is expressed in ES cells (Wood *et al.*, 1992). Actively transcribed genes in ES cells may be more efficient at homologous recombination than a locus not transcriptionally active (Frohman and Martin, 1989; Nickoloff and Reynolds, 1990; Thomas and Rothstein, 1989). We obtained a good targeting frequency at the MCAD locus with 9 targeted cell lines from a total of approximately 525 G-418 resistant ES cell colonies analyzed.

MCAD IV2 was designed to improve screening and confirmation of a targeted insertion. Removal of the 1.3 kb piece of homology encompassing exon 10 from the vector was designed to eliminate false positive PCR signals when screening for homologous recombination. MCAD IV1 could potentially give false positive PCR signals resulting from random integration of vector incompletely digested with *Bam*HI endonuclease. The elimination of the 1.3 kb fragment from MCAD IV2 would eliminate false positive PCR signals. Positive PCR signals would only occur if gap repair occurred during targeted integration.

The duplication of the exon 10 region at the site of targeted insertion results from gap repair of this deleted region within the vector. Gap repair is thought to proceed via the double strand break repair model where recombination is initiated with a double stranded break in the vector (Szostak *et al.*, 1983). Repair of the deleted region within the vector is thought to occur after alignment of the vector with endogenous homologous sequences and strand invasion of the ends of the vector to form D loops. The deleted region in the vector is repaired preceding targeted insertion with endogenous exon 10 region sequences within the ES cell used as a template. PCR screening based on gap repair provides a good screen for targeted integration since gap repair only occurs preceding targeted integration.

We unexpectedly obtained several false positive PCR signals during screening of G-418 resistant ES colonies electroporated with MCAD IV2. It is reasonable to assume that these PCR false positives were due to contamination with PCR positive control vectors. Attempts to avoid contamination by using a separate room with PCR dedicated pipetters, filter pipet tips and separate solutions were not entirely successful. Double amplification of target sequences with initial and then nested primers may have contributed to DNA contamination.

We designed the MCAD targeting vectors to produce a mutation with a similar mechanism as the most common mutation of human MCAD deficient patients. The insertion vectors were designed to duplicate exon 8, 9 and 10 at the MCAD locus. This mutation is designed to alter the codon reading frame upon translation of this duplicated region. Translation of the duplicated exon 8 region results in the formation of premature stop codons resulting in a truncated MCAD monomer. Specifically, the first premature

stop codon is formed after translation of only 7 amino acids from the duplicated exon 8. This results in an MCAD monomer minus α-helixes of the C-terminal domain (See FIG. 2). Most important, this would result in elimination of the major portion of α-helix H and the entire α-helix I, both responsible for making the subunit contacts to form the functional MCAD homotetramer (Kim *et al.*, 1993). This would interfere with normal MCAD homotetramer formation abolishing MCAD activity similar to the K329E mutation in humans (Kim *et al.*, 1993; Roe and Coates, 1989).

The only mouse model with a defect in a straight chain acyl-CoA dehydrogenase enzyme is the BALB/cByJ mouse. This mouse has short-chain acyl-CoA dehydrogenase (SCAD) deficiency and mimics the short-chain organic acidemia and aciduria of human SCAD deficient patients (Armstrong *et al.*, 1993; Roe and Coates, 1989). By extension, the MCAD deficient mouse should mimic MCAD deficiency in humans. Our efforts at cloning and characterization of the mouse MCAD gene, design and construction of targeting vectors and transfection experiments have resulted in the identification of ES cell lines disrupted at the MCAD locus, a significant step toward developing an MCAD deficient mouse. Blastocyst injections have produced chimeric mice with a specifically altered genotype that would be expected to produce an MCAD mRNA unable to be translated into a complete MCAD monomer. A mouse model for MCAD deficiency would be valuable for better understanding the pathogenesis of the wide range of metabolic problems human patients exhibit. The effect of MCAD deficiency, especially during fasting, on blood glucose and ammonia levels could be determined. MCAD deficient mice could be evaluated to determine if similar biochemical abnormalities occur as in the human. The effects of intracellular fatty acid metabolites on the transcriptional

regulation of the MCAD gene via peroxisomal proliferator activated receptors, as reported in the human MCAD gene (Gulick *et al.*, 1994, in press), can be studied in the mouse. Studies determining the effect MCAD deficiency on transcriptional regulation of other genes such as other acyl-CoA dehydrogenases, gluconeogenic enzymes and urea cycle enzymes could be carried out readily. Questions concerning importance of carnitine supplementation and dietary regulation of certain fatty acids could be answered with the MCAD deficient mouse. The mouse model may also be valuable for determining efficacy of novel therapeutic approaches such as somatic gene therapy to specific organs.

ACKNOWLEDGEMENTS

We thank Dr. Allan Bradley, Baylor College of Medicine, for providing the AB1 ES cells and SNL76/7 feeder cells. We thank Drs. Allan Bradley, Paul Hasty and Dan Kelly (Washington University, St. Louis) for helpful discussion and advice. Oligonucleotides were synthesized by the Oligonucleotide Shared Facility, Comprehensive Cancer Center, University of Alabama at Birmingham, Dr. Jeff Engler, director. This work is supported by N.I.H. grants R01-RR02599 (P.A.W.) and T32-RR-07003 (R.J.T.).

REFERENCES

Armstrong, D.L., Masiowski, M.L., and Wood, P.A. (1993). Pathologic characterization of short-chain acyl-CoA dehydrogenase deficiency in BALB/cByJ mice. *Amer. J. Med. Genet.* **47**: 884-892.

Bennett, M.J., Worthy, E., and Pollitt, R.J. (1987). The incidence and presentation of dicarboxylic aciduria. *J. Inherit. Metab. Dis.* **13**: 241-242.

Capecchi, M.R. (1989). Altering the genome by homologous recombination. *Science* **244**: 1288-1292.

Coates, P.M. (1992). Historical perspective of medium-chain acyl-CoA dehydrogenase deficiency: a decade of discovery. In "New Developments in Fatty Acid Oxidation" (P.M. Coates and K. Tanaka, Eds.), pp. 409-423, Wiley-Liss, New York.

Deng, C., and Capecchi, M.R. (1992). Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and targeting locus. *Mol. Cell. Biol.* **12**: 3365-3371.

Deursen, J.-V., and Wieriga, B. (1992). Targeting of the creatine kinase M gene in embryonic stem cells using isogenic and nonisogenic vectors. *Nucleic Acids Res.* **20**: 3815-3820.

Ding, J.H., Bross, P., Yang, B.Z., Iafolla, A.K., Millington, D.S., Roe, C.R., Gregersen, N., and Chen, Y.T. (1992). Genetic heterogeneity in MCAD deficiency: frequency of K329E allele and identification of three additional mutant alleles. In "New Developments in Fatty Acid Oxidation" (P.M. Coates and K. Tanaka, Eds.), pp. 479-488, Wiley-Liss, New York.

Fineberg, A.P., and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**: 6-13.

Frohman, M.A., and Martin, G.R. (1989). Cut, paste, and save: new approaches to altering specific genes in mice. *Cell* **56**: 145-147.

Gulick, T., Cresci, S., Caira, T., Moore, D., and Kelly, D.P. (1994). The peroxisome proliferator activated receptor regulates mitochondrial fatty acid oxidative enzyme gene expression. *Proc. Natl. Acad. Sci. U.S.A.*, In press.

Hasty, P., and Bradley, A. (1993). Gene targeting vectors for mammalian cells. In "Gene Targeting: A Practical Approach" (A.L. Joyner, Ed.), pp. 1-30, Oxford University Press Inc., New York.

Ikeda, Y., Keese, S.M., Fenton, W.A., and Tanaka, K. (1987). Biosynthesis of four rat liver mitochondrial acyl-CoA dehydrogenases: In vitro synthesis, import into mitochondria, and processing of their precursors in a cell-free system and in cultured cells. *Arch. Biochem. Biophys.* **252**: 662-674.

Kelly, D.P., Kim, J.-J., Billardello, J.J., Hainline, B.E., Chu, T.W., and Strauss, A.W. (1987). Nucleotide sequence of medium-chain acyl-CoA dehydrogenase mRNA and its expression in enzyme-deficient human tissue. *Proc. Natl. Acad Sci. U.S.A.* **84**: 4068-4072.

Kim, J.-J.P., Wang, M., and Paschke, R. (1993). Crystal structures of medium-chain acyl-CoA dehydrogenase from pig liver mitochondria with and without substrate. *Proc. Natl. Acad. Sci. U.S.A.* **90**: 7523-7527.

Kim, J.-J., and Wu, J. (1988). Structure of the medium-chain acyl-CoA dehydrogenase from pig liver mitochondria at 3-A resolution. *Proc. Natl. Acad. Sci. U.S.A.* **85**: 6677-6681.

Koller, B.H., and Smithies, O. (1992). Altering genes in animals by gene targeting. *Annu. Rev. Immunol.* **10**: 705-730.

Kolvraa, S., Gregersen, N., Christensen, E., and Hobolth, N. (1982). In vitro fibroblast studies in a patient with C6-C10-dicarboxylic aciduria: Evidence for a defect in general acyl-CoA dehydrogenase. *Clin. Chim. Acta* **126**: 53.

Matsubara, Y., Kraus, J.P., Ozasa, H., Glassberg, R., Finnocchiaro, G., Ikeda, Y., Mole, J., Rosenberg, L.E., and Tanaka, K. (1987). Molecular cloning and nucleotide sequence of cDNA encoding the entire precursor of rat liver medium chain acyl-CoA dehydrogenase. *J. Biol. Chem.* **262**: 10104-10108.

Matsubara, Y., Narisawa, K., and Tada, K. (1992). Medium-chain acyl-CoA dehydrogenase deficiency: molecular aspects. *Eur. J. Pediatr.* **151**: 154-159.

McMahon, A.P., and Bradley, A. (1990). The Wnt-1 (int-1) protooncogene is required for development of a large region of the mouse brain. *Cell* **62**: 1073-1085.

Nickoloff, J.A., and Reynolds, R.J. (1990). Transcription stimulates homologous recombination in mammalian cells. *Mol. Cell. Biol.* **10**: 4837-4845.

Roe, C.R., and Coates, P.M. (1989). Acyl-CoA dehydrogenase deficiencies. In "The Metabolic Basis of Inherited Disease" (C.R. Scriver, A.L. Beaudet, W.S. Sly and D. Valle, Eds.), pp. 889-914, McGraw-Hill, Inc., New York.

Soriano, P., Montgomery, C., Geske, R., and Bradley, A. (1991). Targeted disruption of the c-src proto-oncogene leads to osteopetrosis in mice. *Cell* **64**: 693-702.

Strauss, A.W., Duran, M., Zhang, Z., Alpers, R., and Kelly, D.P. (1990). Molecular analysis of medium-chain acyl-CoA dehydrogenase deficiency. In "Fatty Acid Oxidation: Clinical, Biochemical, and Molecular Aspects" (K. Tanaka and P.M Coates, Eds.), Alan Liss, Inc., New York.

Szostak, J.W., Orr-Weaver, T.L., and Rothstein, R.J. (1983). The double-strand-break repair model for recombination. *Cell* **33**: 25-35.

Te Riele, H., Maandag, R., and Berns, A. (1992). Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. *Proc. Natl. Acad. Sci. U.S.A.* **89**: 5128-5132.

Thomas, B.J., and Rothstein, R. (1989). Elevated recombination rates in transcriptionally active DNA. *Cell* **60**: 921-928.

Thomas, K.R., and Capecchi, M.R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**: 503-512.

Tolwani, R.J., Farmer, S.C., Kurtz, D.M., Johnson, K.R., Davisson, M.T., and Wood, P.A. (1994a) Genomic structure, regulatory elements and chromosomal location of the mouse medium chain acyl-CoA dehydrogenase gene. This volume, p. 49.

Tolwani, R.J., Farmer, S.C., and Wood, P.A. (1994b). Molecular cloning and characterization of the mouse medium-chain acyl-CoA dehydrogenase cDNA. *Genomics* **23**: 247-249.

Waldman, A.S., and Liskay, R.M. (1988). Dependence of intrachromosomal recombination in mammalian cells on uninterrupted homology. *Mol. Cell. Biol.* **8**: 5350-5357.

Workshop on Molecular Aspects of MCAD Deficiency. (1992). Mutations causing medium-chain acyl-CoA dehydrogenase deficiency: a collaborative compilation of the data from 172 patients. In "New Developments in Fatty Acid Oxidation" (P.M. Coates, and K. Tanaka, Eds.), pp. 499-506, Wiley-Liss, New York.

SUMMARY AND CONCLUSIONS

Many important steps have been taken along the way toward the development of an MCAD deficient mouse model. The entire mouse MCAD cDNA has been cloned and characterized. This cDNA has been used to clone the mouse MCAD genomic gene. The authentic MCAD gene has been definitively mapped to chromosome 3. Insertion vectors have been designed using MCAD genomic fragments and used to develop an ES cell line disrupted at the MCAD locus. Chimeric mice derived from these ES cells have been produced.

The mouse MCAD cDNA is remarkably similar at the nucleotide and deduced amino acid level to the human and rat MCAD cDNAs. The structure of the MCAD monomer, therefore, is probably similar among all three species. This conservation among species implies important functional regions have been maintained to preserve enzymatic activity. All amino acid residues where substitutions have resulted in MCAD deficiency in humans are conserved between species, suggesting important functions of these amino acids. Amino acid residues putatively involved in α,β-dehydrogenation and in direct interactions with flavine adenine dinucleotide also are conserved.

The structure of the mouse MCAD gene has been determined. The promoter region does not contain a TATA or CCAAT box and is GC rich with several putative Sp1 binding regions suggesting constitutive expression of this gene. We have also looked at putative regulatory elements within this region. An imperfect consensus sequence for a

putative glucocorticoid response element has been identified within this region. Previous work has shown that dexamethasone administration stimulates MCAD expression in the rat (Nagao *et al.*, 1993). Putative motifs arranged in inverted repeats and direct repeats for interactions with steroid nuclear receptors also have been identified. This region is similar to the nuclear receptor response element (NRRE-1) within the human MCAD promoter region which interacts with various receptors such as retinoic acid receptors, hepatocyte nuclear factor 4 and chicken ovalbumin upstream promoter transcription factor (Raisher *et al.*, 1992; Carter *et al.*, 1993; Carter *et al.*, 1994). Fatty acid metabolites also regulate expression of the human MCAD gene via the peroxisomal proliferator activated receptor (PPAR) (Gulick *et al.*, 1994, in press). Fasting induces MCAD expression in the rat (Nagao *et al.*, 1993), possibly via the same mechanism. The mouse promoter region also contains consensus sequences for binding PPAR, suggesting fatty acid metabolites may indirectly play a role in mouse MCAD expression. The mouse MCAD gene, therefore, may be regulated by interactions with numerous nuclear receptors similar to other genes involved in lipid metabolism. Further functional studies of the promoter region could provide important insights into mechanisms of transcriptional control of the mouse MCAD gene.

Two mouse MCAD processed pseudogenes were cloned and characterized. Bahary and co-workers (1991) had previously mapped the MCAD gene to mouse chromosome 8, a region that contained no homology to human chromosome 1 where the human MCAD gene was mapped (Kidd *et al.*, 1990). Based on these data, we were prompted to reexamine the chromosomal location of the MCAD gene. RFLP linkage analysis mapped the authentic MCAD gene to the distal end of mouse chromosome 3, a region

with homology to human chromosome 1. The 2 processed MCAD pseudogenes were mapped to chromosome 8 and 11. Bahary and co-workers mapped the MCAD processed pseudogene on mouse chromosome 8 and mistakenly labeled it as the MCAD gene.

The cloning and characterization of the mouse MCAD gene allowed the design and construction of MCAD insertion type targeting vectors used for ES cell gene transfer experiments that led to the generation of MCAD disrupted ES cell lines. Chimeric mice with the specifically altered genotypes have been produced. We designed the MCAD insertion vectors to produce a mutation that would inhibit MCAD monomers from forming functional homotetramers. The net result is similar to the most common human mutation which produces MCAD monomers that are unable to acquire enzymatic activity by their inability to assemble into homotetramers (Kim *et al.*, 1993). An MCAD deficient mouse would be valuable in further investigation of pathogenetic mechanisms of MCAD deficiency. The effect of MCAD deficiency and effect of intracellular fatty acid metabolite levels on transcriptional control of the MCAD gene, possibly via the peroxisomal proliferator activated receptor, can be studied. Influence on other acyl-CoA dehydrogenase genes, along with genes involved in related metabolic pathways such as the gluconeogenic enzymes and urea cycle enzymes, also could be evaluated. The deficient mouse could be used to answer questions concerning importance of carnitine supplementation and dietary regulation and the development of novel therapeutic approaches.

LIST OF REFERENCES

Bahary, N., Zorich, G., Pachter, J.E., Leibel, R.L., and Friedman, J.M. (1991). Molecular genetic linkage maps to mouse chromosome 4 and 6. *Genomics* **11**: 33-47.

Carter, M.E., Gulick, T., Moore, D.D., and Kelly, D.P. (1994). A pleiotropic element in the medium-chain acyl coenzyme A dehydrogenase gene promoter mediates transcriptional regulation by multiple nuclear receptor transcription factors and defines novel receptor-DNA binding motifs. *Mol. Cell. Biol.* **14**: 4360-4372.

Carter, M.E., Gulick, T., Raisher, B.D., Caira, T., Ladias, J.A.A., Moore, D.D., and Kelly, D.P. (1993). Hepatocyte nuclear factor-4 activates medium chain acyl-CoA dehydrogenase gene transcription by interacting with a complex regulatory element. *J. Biol. Chem.* **268**: 13805-13810.

Gulick, T., Cresci, S., Caira, T., Moore, D.D., and Kelly, D.P. (1994). The perixosomal proliferator activated receptor regulates mitochondrial fatty acid oxidative enzyme gene expression. *Proc. Natl. Acad. Sci. U.S.A.* In press.

Kidd, J.R., Matsubara, Y., Castiglione, C.M., Tanaka, K., and Kidd, K.K. (1990). The locus for the medium-chain acyl-CoA dehydrogenase gene on chromosome 1 is highly polymorphic. *Genomics* **6**: 89-93.

Kim, J.-J.P., Wang, M., and Paschke, R. (1993). Crystal structure of medium-chain acyl-coA dehydrogenase from pig liver mitochondria with and without substrate. *Proc. Natl. Acad. Sci. U.S.A.* **90**: 7523.

Nagao, M., Parimoo, B., and Tanaka, K. (1993). Developmental, nutritional, and hormonal regulation of tissue-specific expression of the genes encoding various acyl-CoA dehydrogenase and alpha-subunit of electron transfer flavoprotein in rat. *J. Biol. Chem.* **268**: 24114-24124.

Raisher, B.D., Gulick, T., Zhang, Z., Strauss, A.W., Moore, D.D., and Kelly, D.P. (1992). Identification of a novel retinoid-response element in the promoter region of the medium chain acyl-coenzyme A dehydrogenase gene. *J. Biol. Chem.* **267**: 20264-20269.

GRADUATE SCHOOL
UNIVERSITY OF ALABAMA AT BIRMINGHAM
DISSERTATION APPROVAL FORM

Name of Candidate Ravi J. Tolwani

Major Subject Molecular and Cellular Pathology

Title of Dissertation Cloning, Characterization and Targeted Disruption of the Mouse Medium-Chain Acyl-CoA Dehydrogenase Gene

Dissertation Committee:

Philip A. Wood, Chairman

Steve Qualls

Ian Jones

Jerry Thompson

Kevin Dysart

Director of Graduate Program Thomas M. Town

Dean, UAB Graduate School Joan F. Lorden

Date 12/15/94