Primary structure and activity of mouse methylmalonyl-CoA mutase

Michael F. WILKEMEYER, Ana Maria CRANE and Fred D. LEDLEY*

Howard Hughes Medical Institute, Department of Cell Biology and Department of Pediatrics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, U.S.A.

Methylmalonyl-CoA mutase (MCM) is an adenosylcobalamin-dependent enzyme that catalyses isomerization between methylmalonyl-CoA and succinyl-CoA (3-carboxypropionyl-CoA). Genetic deficiency of this enzyme in man causes an often fatal disorder of organic acid metabolism termed *mut* methylmalonicacidemia. We report cloning of a mouse MCM cDNA and the characterization of its primary structure and biological function. Mouse MCM in fibroblasts and crude liver extracts exhibits activity and reaction kinetics similar to those of the human enzyme. The predicted amino acid sequence of mouse MCM exhibits 94% identity with its human homologue and considerable identity with a prokaryotic MCM. Transfection of the mouse cDNA into cultured cells constitutes an active apoenzyme and can complement genetic deficiency of the apoenzyme in cells from patients with *mut* methylmalonicacidemia. These results establish that mouse MCM is homologous to human MCM in structure and function and provides a basis for using the mouse as a model for studying this enzyme and its deficiency state.

INTRODUCTION

Methylmalonyl-CoA mutase (MCM, EC 5.4.99.2) is an adenosylcobalamin-dependent mitochondrial-matrix enzyme that catalyses isomerization between methylmalonyl-CoA and succinyl-CoA (3-carboxypropionyl-CoA). This enzyme has been characterized extensively in man, where this reaction is involved in the degradation of several amino acids, odd-chain fatty acids and cholesterol via propionyl-CoA to the tricarboxylic acid cycle [1]. Genetic deficiency of this enzyme in man causes a disorder of organic acid metabolism termed *mut* methylmalonicacidemia (McKusick no. 251000) in which methylmalonate and abnormal metabolites of methylmalonate accumulate, leading to failure of many vital functions [1].

MCM has different functions in other species. In ruminants MCM is essential for gluconeogenesis from propionate, a pathway that may account for 60–80% of all glucose utilization [2,3]. In prokaryotes such as *Propionibacterium shermanii* [4], *Streptomyces erythreus* [5] and *Protaminobacter ruber* [6] and eukaryotes such as *Fasciola hepatica* [7] and *Spirometra mansonoides* [7], MCM catalyses the isomerization of succinyl-CoA to methylmalonyl-CoA during synthesis of propionate from tricarboxylic acid-cycle intermediates.

MCM has been purified from several species, including humans [8,9], *Prop. shermanii* [4,10], *Strep. erythreus* [5] and *Ascaris lumbricoides* [11]. We have recently reported cloning and sequencing of a human MCM cDNA [12,13]. MCM from *Prop. shermanii* has also been cloned and sequenced [14]. We now report the biochemical characterization, cloning and sequencing of mouse MCM and the expression of recombinant mouse MCM by gene transfer. These results provide a fundamental description of mouse MCM and establish a foundation for using the mouse as a model for studying the structure and function of MCM and the consequences of its deficiency state in humans.

MATERIALS AND METHODS

Assays of MCM activity *in vitro* and in live cultured cells

MCM was assayed *in vitro* by using the $\text{KMnO}_4/\text{HClO}_4$ procedure [15,16]. The assay reaction mixture contains 270 μg of extract, 100 mM-Tris/HCl buffer, pH 8.0, 300 μM -DL-methylmalonyl-CoA, 50000 d.p.m. of DL-[^{14}C]methylmalonyl-CoA (56 mCi/mmol) and 36 μM -adenosylcobalamin in a total volume of 250 μl and is incubated in the dark for 20 min at 30 °C. The reaction is stopped by addition of 50 μl of 2 M-HClO₄, incubation at 100 °C and 4 °C for 10 min each, and centrifugation. A sample (125 μl) of this supernatant is combined with 25 μl of 2 M-HClO₄ and 250 μl of 4% (w/v) KMnO_4 , and the mixture is incubated at 100 °C for 10 min and centrifuged. A 250 μl portion of this supernatant is placed in a glass scintillation vial containing 250 μl of propionate (99% free acid; Sigma Chemical Co.), evaporated to dryness, and its radioactivity counted in Aquasol (New England Nuclear). Succinyl-CoA formed by MCM is hydrolysed to succinate, the ^{14}C radioactivity of which is counted. Unchanged methylmalonyl-CoA is modified to propionate and evaporates during the final drying step. Experiments were performed in duplicate with a boiled blank.

MCM is measured in fibroblasts by incorporation of [^{14}C]propionate into trichloroacetic acid-precipitable material [16–18]. Confluent cells in a 25 mm well are incubated for 18 h in 0.5 ml of Puck's-G saline with 6×10^4 d.p.m. of [^3H]leucine/ml (150 Ci/mmol; Amersham-Searle) ([leucine] = 37 μM), 2×10^6 d.p.m. of [1- ^{14}C]propionate/ml (57 mCi/mmol; New England Nuclear), 100 μM -propionate, 15% (v/v) fetal-calf serum, 1.2 mg of NaHCO_3 /ml, 9.9 mg of glucose/ml and 0.01–1.0 μg of hydroxycobalamin/ml. Cells were washed three times in 0.13 M-NaCl/18 mM-sodium phosphate buffer, pH 8.0, and three times for 15 min with 5% (w/v) trichloroacetic acid at

Abbreviation used: MCM, methylmalonyl-CoA mutase.

* To whom correspondence should be addressed at: Department of Cell Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, U.S.A.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X57941.

Table 1. Oligonucleotides used for sequencing mouse MCM cDNA

Oligonucleotides complementary to the human (hMCM) or mouse (mMCM) cDNAs were used to sequence portions of the mouse cDNA clone.

Oligonucleotide no.	Sequence	Orientation	Sequence (5'...3')
6	hMCM 2258-2277	Sense	5'-CTTGATGATATTGAGAAAGTG-3'
15	hMCM 1740-1758	Anti-sense	5'-AGGGCATCTGTGATTCT-3'
17	hMCM 2145-2160	Anti-sense	5'-CCTCCACACATGACAA-3'
24	mMCM 1333-1353	Anti-sense	5'-CGTCATTTGTGTGCGACTCC-3'
26	mMCM 2393-2410	Anti-sense	5'-ACAGGGCCCGCAAGCATTCTAAGAGCTGCTG-3'
51	hMCM 1468-1488	Anti-sense	5'-ATTCTAGCTTGTCTTCGGGC-3'

4 °C. Precipitated material is resuspended in 1 ml 0.2 M-NaOH and its radioactivity counted with correction for ^3H and ^{14}C dual standards. ^{14}C incorporation into protein is commonly used as an assay for the presence or absence of the MCM holoenzyme in primary fibroblasts. ^3H incorporation normalizes for cell number and the rate of constitutive protein synthesis. Results are expressed as nmol of propionate/ μmol of leucine incorporated. Experiments were performed in triplicate.

Cloning and sequencing of mouse MCM

Segments of the human MCM cDNA clone hMCM26 [12,13] (Fig. 2a) were used as probes to screen for the mouse MCM cDNA. Approx. 4×10^5 plaques of a mouse liver cDNA library in $\lambda\text{gt}11$ (Stratagene) were screened and cDNA inserts were subcloned into the vector pGEM7zf(+) (Promega) [19,20]. Both strands of the mouse cDNA were sequenced by using subclones containing internal restriction fragments or terminal deletions introduced by *ExoIII* digestion (Fig. 2c) [21]. Double-stranded dideoxy sequencing [22] was performed with the SP6 and T7 promoter primers or synthetic oligonucleotide primers listed in Table 1. Sequence comparisons were made by using the on-line facilities of the Protein Identification Resource (National Biomedical Research Foundation, Washington, DC, U.S.A.).

Gene transfer and expression of mouse MCM

For expression in COS cells the mouse MCM cDNA was subcloned into the *EcoRI* site of the vector p91023B containing the adenovirus major late promoter [23]. COS cells [24] were transformed by $\text{Ca}_3(\text{PO}_4)_2$ co-precipitation [25], and protein extracts were harvested 48 h after transfection. Control cells were transfected with an anti-sense clone or a vector expressing a dihydropteridine reductase cDNA [26]. For expression in primary human fibroblasts, the MCM cDNA was subcloned into the *NotI* site of the vector pNassCMV [27], which contains the cytomegalovirus immediate early promoter. This clone was introduced via electroporation [28] into a primary fibroblast cell line GM50 (NIGMS Mutant Cell Repository) derived from a patient with *mut*⁰ methylmalonicacidemia. The GM50 cell line exhibits no enzyme activity and expresses decreased amounts of (human) MCM mRNA [16]. Electroporation was performed with a Gene Pulser (Bio-Rad Laboratories) after preincubation of 0.3 ml of cells at a density of $3 \times 10^6/\text{ml}$ for 5 min with 10–20 μg of DNA with a pulse of either 260V or 280V at a capacitance of 960 μF . Propionate incorporation was assayed after 48 h. Transformation efficiency was monitored by transforming cells with an identical vector containing *Escherichia coli* β -galactosidase (pCMV- βgal) [27] and scoring control plates for the number of cells staining with X-gal [29].

RESULTS AND DISCUSSION

Characteristics of mouse MCM

Mouse fibroblasts incorporate [^{14}C]propionate into trichloro-

Fig. 1. Propionate incorporation into primary mouse and human fibroblasts

Incorporation of [^{14}C]propionate into trichloroacetic acid-precipitable material by fibroblasts in culture is dependent upon their ability to metabolize propionate via methylmalonyl-CoA to the tricarboxylic acid cycle as evidenced by the absence of incorporation in cells genetically deficient in the MCM apoenzyme (*mut*⁰) or synthesis of the adenosylcobalamin cofactor (*cbl*). Propionate incorporation can be stimulated in *cbl* cells by the addition of excess hydroxycobalamin to the culture media. Mouse and human fibroblasts exhibit equivalent activity, which is not stimulated by cobalamin. Cell lines: Mouse, primary mouse fibroblast cell line; Human, mean of seven normal human cell lines; *mut*⁰, mutant human cell line GM50 with *mut*⁰ phenotype [16]; *cbl*, mutant human cell line HAR with *cbl* phenotype [16]. —, No added cofactor; +, hydroxycobalamin (1 $\mu\text{g}/\text{ml}$) added to medium. The means and standard deviations of triplicate samples are indicated.

acetic acid-precipitable material to a degree similar to normal human fibroblasts (Fig. 1). This activity is not stimulated by the addition of hydroxycobalamin, indicating that the availability of the cofactor is not a limiting factor. The absence of [^{14}C]propionate uptake in the *mut*⁰ cells and *cbl* cells demonstrates the requirement for the MCM apoenzyme and cobalamin cofactor in the metabolism of propionate. It should be noted that the propionate incorporation assay is not a linear measure of the MCM holoenzyme, since the enzyme is not limiting on the rate of propionate incorporation under normal conditions.

MCM enzyme activity assayed in crude extracts of mouse and human liver *in vitro* was linear as a function of protein. Mouse liver exhibited somewhat higher activity/mg of protein than a single control human liver that was available for this study, though this human liver exhibited more basal activity without exogenous adenosylcobalamin (Table 2). These results indicate that a larger fraction of MCM in the mouse liver is present as apoenzyme rather than holoenzyme. The apparent K_m of MCM in mouse and human liver extracts was determined for the substrate (methylmalonyl-CoA) and cofactor (adenosylcobalamin) and were indistinguishable (Table 2). Both the assay

Table 2. Properties of mouse and human MCM in crude liver extracts

Activities in crude extracts of mouse or human livers were measured as a function of extract, as a function of the concentration of cofactor ([adenosylcobalamin]) and as a function of the concentration of substrate ([methylmalonyl-CoA]). Activity is expressed as nmol of succinate formed/h per mg (–) without addition of exogenous adenosylcobalamin to crude extract and (+) with addition of 36 μ M-adenosylcobalamin to crude extract.

	Activity (nmol/h per mg)		Affinity constant (K_m)	
	–	+	Adenosyl- cobalamin	Methyl- malonyl-CoA
Mouse	15	275	0.20 μ M	0.27 mM
Human	42	170	0.24 μ M	0.35 mM

in vitro and propionate incorporation assay reflect the isomerization of methylmalonyl-CoA to succinyl-CoA. The reverse reaction by which succinate can be metabolized via succinyl-CoA and methylmalonyl-CoA to propionate is thought to play little (if any) role in higher eukaryotes [30,31] and was not assayed.

Nucleic acid sequence of mouse MCM cDNA

The longest mouse cDNA clone, moMCM209 (Fig. 2b), was sequenced completely in both directions as shown in Fig. 2(c). The sequence contains 3245 base residues (Fig. 3). A single long open reading frame begins with a Kozak consensus sequence (ACCAUG) [32] at position 47 and extends to a TAA termination codon at position 2291. The 5' untranslated region contains a prior ATG at position 37, which is interrupted by a TAA termination codon at position 58. A homologous reading frame exists in the human sequence [13]. The 3' untranslated region comprises 954 base residues and does not contain a poly(A) tail. A possible polyadenylation signal (AAATAAA) is present at base 3235, ten base residues before the end of the clone.

The nucleic acid sequence of mouse MCM exhibits profound similarity with human MCM. Comparison of the nucleic acid sequences of the mouse and human long open reading frames reveals 89 % identity. There is also considerable identity between the mouse and human sequences in the 5' untranslated region.

Predicted amino acid sequence of mouse MCM

The long open reading frame codes for a protein of 748 amino acid residues and a molecular mass of 82 891 Da. The fidelity of the predicted reading frame is supported by identity with 105/113 residues (93 % of peptide sequence from human MCM reported previously (Fig. 3) as well as extensive identity with the predicted amino acid sequence of human MCM (Fig. 4) [13]. The predicted mouse and human MCM sequences share 700 identities out of 748 possible matches (94 % identity) with the insertion of one break (Fig. 4 and Table 3). Mouse MCM preserves a strikingly asymmetrical amino acid composition, with proline, tryptophan and tyrosine residues disproportionately located in the *N*-terminal portion and cysteine residues located in the *C*-terminal portion.

There are few significant differences between the human and mouse apoenzymes. These include substitution of Cys-288 in mouse for Ser-288 in human and Ser-433 in mouse for Cys-433 in human. Since we did not observe any significant differences in the activity of the mouse and human enzymes, it is likely that these changes do not affect regions critical for structure or function. Most of the divergence between human MCM and mouse MCM is localized to the region representing the mitochondrial targeting sequence, where only 22/32 residues

(a) Human MCM cDNA **(b) Mouse MCM cDNA** **(c) Sequencing strategy**

Fig. 2. Schematic structure of human and mouse cDNA MCM and sequencing strategy for mouse MCM

(a) Human MCM cDNA [13] showing position of open reading frame, restriction sites, and the location of hMCM26b and hMCM26a' probes used to screen for mouse MCM. (b) Schematic of mouse cDNA clones showing position of open reading frame and restriction sites. (c) Scheme for sequencing clone moMCM209. Sequences were obtained from the full-length cDNA or subclones created by *ExoIII* terminal digestion, an *EcoRI*–*KpnI* fragment, a *ClaI*–*BamHI* fragment or an *EcoRI*–*XbaI* fragment. Primers included the T7 and SP6 promoter primers and synthetic oligonucleotides listed in Table 1. Key: E, *EcoRI*; B, *BamHI*; K, *KpnI*; H, *HindIII*; X, *XbaI*; C, *ClaI*.

(69 %) are identical. Despite this apparent divergence, the mouse protein preserves the characteristic features of a mitochondrial targeting sequence [33], including arginine residues (three residues), hydrophobic and hydrophilic residues (e.g. serine, three residues; leucine, seven residues), and an arginine residue two residues from the proteinase-cleavage site [34]. This pattern of identity with human MCM predicts that endopeptidase cleavage will occur after the leucine residue at position 30. Cleavage at this position would produce a mature enzyme with 718 residues and a molecular mass of 79 284 Da.

This degree of divergence between the mitochondrial targeting sequences of homologous proteins is not uncommon. For example, mouse and chicken aspartate aminotransferases exhibit 85 % overall identity in their amino acid sequences, but share only 14/29 % (48 %) identity in the mitochondrial targeting sequences, and mouse and human ornithine transcarbamoylases, which exhibit 92 % overall identity in their amino acid sequence, share only 22/32 (69 %) identity in their mitochondrial targeting sequence (results not shown).

<----- PREDICTED MITOCHONDRIAL TARGETING SEQUENCE ----->

MetLeuArgAlaLysAsnGlnLeuPheLeuLeuSerProHisTyrLeuLysGlnLeuAsnIleProSerAlaSer

5' AACACTGACCGTTCTCATTTCTTTGGGAGTATTCATGCTTCACCATGTGAGAGCTAAGAATCAACTTTTTCCTATCGCCCATTCAGTCAAGCAGTAACATTCATCAGCTTC
 10 20 30 40 50 60 70 80 90 100 110 120

-----PREDICTED TERMINUS OF MATURE PEPTIDE-----

ArgTrpLysArgLeuHisGlnGlnProLeuHisProGluTrpAlaValLeuLysGlnLeuLysGlyLysAsnProGluAspLeuIleTrpHisThrProGluGlyIle
 5' CAGATGGAAACGCCCTTACAGCAGCAGCAACCCCTTACCCGAGAAATGGGCTGTACTGGCCAAAAGCAGCTGAAAGGCAAAACCCAGAGGACCTTATATGGCAGACCCGAGAAGGAT
 130 140 150 160 170 180 190 200 210 220 230 240

SerIleLysProLeuTyrSerArgAlaAspThrLeuAspLeuProGluGluLeuProGlyValLysProPheThrArgGlyProTyrProThrMetTyrThrTyrArgProTrpThrIle
 5' CTCTATAAGCCCTTATATTCAGGGCAGATCTGGAATACCCGAGAACTTCAGGAGTGAAGCATTACACGGGGACCATATCCTACCATGTATACCTATAGGCCCTGGACCAT
 250 260 270 280 290 300 310 320 330 340 350 360

ArgGlnTyrAlaGlyPheSerThrValGluGluSerAsnLysPheTyrLysAspAsnIleLysAlaGlyGlnGlnLeuSerValAlaPheAspLeuAlaThrHisArgGlyTyrAsp
 5' CCGTCAGTATGACGGCTTACTACTCTGGAAGAAACCAATAATCTATAAGCAATATTAAGCGTGGTCAGCAGGGTGTCTGTGTGCTTTCAGCTTGGCAACACATCGTGGTTATGA
 370 380 390 400 410 420 430 440 450 460 470 480

SerAspAsnProArgValArgGlyAspValGlyMetAlaGlyValAlaIleAspThrValGluAspThrLysIleLeuPheAspGlyIleProLeuGluLysMetSerValSerMetThr
 5' TTCAGACAACCCAGAGTTCGTGGAGATGTGGAATGGCTGGAGTGTCTATTGACACTGTAGAAGACCAAAATCTATTGATGGCATTCTTTAGAAAAATGTCCGTTTCCATGAC
 490 500 510 520 530 540 550 560 570 580 590 600

MetAsnGlyAlaValIleProValLeuAlaThrPheIleValThrGlyGluGlnGlyValProLysGluLysLeuThrGlyThrIleGlnAsnAspIleLeuLysGluPheMetVal
 5' TATGAACGGAGCTGTATCCAGCTCTGGCAACCTTTATAGTAACGGGAAGAGCAAGGTGTCGGAAGGAGAGCTCACTGGCACAATTGCAATGACATCTTAAAGAGTTTATGCT
 610 620 630 640 650 660 670 680 690 700 710 720

ArgAsnThrTyrIlePheProGluProSerMetLysIleIleAlaAspIlePheGlnTyrThrAlaGlnHisMetProLysPheAsnProIleSerIleSerGlyTyrHisMetGln
 5' CAGAAATACCTATATTTTCCCGCAGAGCGATCGATGAAATATTGCTGACATTTTCAATACACAGCAGAGCATGCCAAAATTAATGCAATTTGATTAGCGGGTACCATATGCA
 730 740 750 760 770 780 790 800 810 820 830 840

GluAlaGlyThrAspAlaIleLeuGluLeuAlaTyrThrIleAlaAspGlyLeuGluTyrCysArgThrGlyLeuGlnAlaGlyLeuThrIleAspGluPheAlaProArgLeuSerPhe
 5' GGAAGCAGGAGCTGATGCCATTTAGAACTGGCCTATACCATCGCAGATGGGTGGAGTCTGAGAGCTGGACTCCAGCGTGGACTCACAATGATGAATTTGACCAAGGTTGCTTT
 850 860 870 880 890 900 910 920 930 940 950 960

PheTrpGlyIleGlyMetAsnPheTyrMetGluIleAlaLysMetArgAlaGlyArgArgLeuTrpAlaHisLeuIleGluLysMetPheGlnProLysAsnSerLysSerLeuLeuLeu
 5' CTCTGGGGAATTGGAATGAACCTTCTACATGGAATAGCCAGATGCGGAGTGGGAGAGCTGTTGGGCTCACTTAATAGAAAAATGTTCCAGCCTAAAGCTTAAATCTCTTCTT
 970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080

ArgAlaHisCysGlnThrSerGlyTrpSerLeuThrGluGlnAspProTyrAsnAsnIleValArgThrAlaIleGluAlaMetAlaAlaValPheGlyGlyThrGlnSerLeuHisSer
 5' AAGAGCAGACTGCGCAGACTCGGGGTGGTCACTTACTGAGCAGGATCCCTTACAATACTTGTCCGACCTGCAATCGAAGCCATGGCAGCTGTGTTGGAGGAAACCCAGCTTTGGCATAG
 1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200

AsnSerPheAspGluAlaLeuGlyLeuProThrValLysSerAlaArgIleAlaArgAsnThrGlnIleIleIleGlnGluGluSerGlyIleProLysValAlaAspProTrpGlyGly
 5' CAACCTTTTATGAAGCGGTGGGTTTGGCCACTGTGAAAGTGGCCGAGTGTGCGGAACACAGAGATCATCTTCCAGGAAGATCTGGGATCCCAAGTGGGGGATCCCTGGGGAGG
 1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320

SerTyrMetMetGluSerLeuThrAsnAspValTyrGluAlaAlaLeuLysLeuIleTyrGluValGluGluMetGlyGlyMetAlaLysAlaValAlaGluGlyIleProLysLeuArg
 5' GTGGTACATGATGGAGTGGCTCACAATGACGTTTATGAGGCTGCTCTGAAGTTGATATGAAGTTGAAGAAATGGGTGGAAATGGCCAAAGCTGTAGCTGAAGGAATCCCTAAACTTGG
 1330 1340 1350 1360 1370 1380 1390 1400 1410 1420 1430 1440

IleGluGluCysAlaAlaArgArgGlnAlaArgIleAspSerGlySerGluValIleValGlyValAsnLysTyrGlnLeuGluLysGluAspSerValGluValLeuAlaIleAspAsn
 5' CATTGAAGAATGTGCTGCCGGAAGACAGCTAGATAGATTCTGCTCTGAGGTAATTTGGAGTAATAAGTATCAGTTGGAAAAAGAGACTCTGTGGAGGCTCTGGCCATTGACAA
 1450 1460 1470 1480 1490 1500 1510 1520 1530 1540 1550 1560

ThrSerValArgLysLysGlnIleGluLysLeuLysLysIleLysSerSerArgAspGlnAlaLeuAlaGluGlnCysLeuSerAlaLeuThrGlnCysAlaAlaSerGlyAspGlyAsn
 5' CACTTCAGTCGCTAAGAAGCAGATTGAAAACTCAAGAAGATTAAATCCAGCAGGGATCAAGCTTTGGCTGAGCAGTGTCTCAGTGCACCTACCCAGTGTGCTGCCAGTGGAGATGGCAA
 1570 1580 1590 1600 1610 1620 1630 1640 1650 1660 1670 1680

IleLeuAlaLeuAlaValAspAlaAlaArgAlaArgCysThrValGlyGluIleThrAspAlaLeuLysLysValPheGlyGluHisLysAlaAsnAspArgMetValSerGlyAlaTyr
 5' TATTCTGGCTCTGGCAGTGCAGCTCGTGAAGATGTACAGTTGAGAAATACCGGATGCCTTGAAGAAAGTATTGGTGAGCATAAAGCTAATGATCGTATGCTGGAGTGCAGATA
 1690 1700 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800

ArgGlnGluPheGlyGluSerLysGlyIleThrSerAlaIleLysArgValAsnLysPheMetGluArgGluGlyArgArgProArgLeuLeuValAlaLysMetGlyGlnAspGlyHis
 5' TCGGCAGGATTTGGAGAAAGTAAAGAGATCAGATCTGCCATCAAGAGATTAATAATTCATGGAACGTGAAGGTGCGCAGACCTCGTCTTCTTGGCAAAAAATGGGCAAGATGGCCA
 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920

AspArgGlyAlaLysValIleAlaThrGlyPheAlaAspLeuGlyPheAspValAspIleGlyProLeuPheGlnThrProArgGluValAlaGlnHisAlaValAspAlaAspValHis
 5' TGACGGGGAGCCAGGTGATGCTGCTGCTGATCTTGGTTTGTGAGCTTATCAAGAACTCACCGCCCTGGGGCGGCAGATATCCTGTGATGTGTGGGGGGCTGATTCACCC
 1930 1940 1950 1960 1970 1980 1990 2000 2010 2020 2030 2040

AlaValGlyValSerThrLeuAlaAlaGlyHisLysThrLeuValProGluLeuIleLysGluLeuThrAlaLeuGlyArgProAspIleLeuValMetCysGlyGlyValIleProPro
 5' TGCTGTGGGTGTGAGCAGCACTTGTGCTGCTGCTGATCAAAACCCCTCGTTCTGAGCTTATCAAGAACTCACCGCCCTGGGGCGGCAGATATCCTGTGATGTGTGGGGGGCTGATTCACCC
 2050 2060 2070 2080 2090 2100 2110 2120 2130 2140 2150 2160

GlnAspTyrGluPheLeuTyrGluValGlyValSerAsnValPheGlyProGlyThrArgIleProArgAlaAlaValGlnValLeuAspAspIleGluLysCysLeuAlaGluLysGln
 5' ACAGGATTAATGATTTCTGATGAAGTGTGTTTCCAAAGTGTGTCCTGGAACCCGGATTCGAAGAGCTGTCTCAAGTGTCTGATGATTTGCAAGATGTTTGGCAGAGAGCA
 2170 2180 2190 2200 2210 2220 2230 2240 2250 2260 2270 2280

GlnSerVal***

5' GCAGTCTGTGTAACATCAGGATCTTTAGTTTACTGTAAGCAGCAGTCTGTGTAACATCTGGATCTTTAGTTATGACCCGAAGAGGTAGTGGTAAATCCGTTTCGGTCAGGTAATCCGAG
 2290 2300 2310 2320 2330 2340 2350 2360 2370 2380 2390 2400

5' CAGCTCTTGAATGCTTGGCTTTGAAACCTCAGATTAAAAGATTCTAAGTATAAATGTGAGTCAATGCAATACACATATTTAAAGAACTCAGCTTCAAAAAAACCCTCTGA
 2410 2420 2430 2440 2450 2460 2470 2480 2490 2500 2510 2520

5' AAATCATTATACCAATATGTTTAAAGTATATTCTAGAAAAATAAAAATGGTATACCTATTAGAAATAGCTTACAATTTCTTAAATAATTAAGCTTTTGTGTTTATCTGTTT
 2530 2540 2550 2560 2570 2580 2590 2600 2610 2620 2630 2640

5' ACATGGGCTCTTACTATGACAGCTGCTGCTGCAACTAACTCTGTAGAAGCAGTCTGGCCCTCAAACTCACAGAGATCCACCTGCCCTCTGCTCTAGTACAGGATTAAGGTATGT
 2650 2660 2670 2680 2690 2700 2710 2720 2730 2740 2750 2760

5' GCCTCGATGCTGCAAAATTATTAACCTTAAAAATTAATAAAATTTTAAATTTAACGGATTCTTCCTTCAATTTAGTGTGATTGTGTACTAGAACTCAAGTTCCTGTGTTAATTT
 2770 2780 2790 2800 2810 2820 2830 2840 2850 2860 2870 2880

5' AATAAAAAATGACATTATCTTACAGAAATAGTGGAGGTTATTTGGATGGATCTCAGCGAGACCTATTGCTTGAATAAAAGTAGTGTATACCCCGCTGTACAATGGAATAG
 2890 2900 2910 2920 2930 2940 2950 2960 2970 2980 2990 3000

5' GGGACATGCTAGATAGGCTTCCAGGATGTGTGAGCTCCTCAGATGGGAATGATATTGTTAATGTGAGGTTCAAACTCTGGAAAAGATGTGTGTGACGCTCTCTTAAATGAATGTTGAT
 3010 3020 3030 3040 3050 3060 3070 3080 3090 3100 3110 3120

5' GGACAGATGCAATTTGCTATAATGACAGTCTAAATCTATTACATAAAGTAAATGAATTTTCAGATCTAGAAACAGCGTGTATGATACACTTTCTTAAATATAAGTTATAAAAAATAA
 3130 3140 3150 3160 3170 3180 3190 3200 3210 3220 3230 3240

5' AAAAT 3'

3245

Fig. 3. Nucleotide sequence of mouse MCM cDNA and translation of long open reading frame

The first AUG at position 37 precedes a short reading frame that is interrupted by a TAA termination codon at position 58. The AUG at position 47 is followed by a long open reading frame through to a TAA at position 2291. Segments of amino acid sequence determined from purified human MCM are underlined and identities with the predicted murine sequence are in italics. Cleavage of the mitochondrial targeting sequence is predicted to occur after Leu-30 based on identity between the mouse and human sequences. The most likely polyadenylation signal begins at position 3237.

	1	10	20	30	40	50	60	70	80	90	100	110	120	130	140
MOCH	1	MLRAKRLFLSLPHYLAKLHPSASRW--	KRLLRQQPLHPEMAVLA	AKQKGNPEDLTWETPEGISIK	PLYSRADTLDLPEELPGV	KPTTRGPTPTMTTTPFT	TRQYAGFSTVESK	KFYDMILAGQQLSVAFDL							
BMCH		MLRAKRLFLSLPHYLAKLHPSASRW--	KRLLRQQPLHPEMAVLA	AKQKGNPEDLTWETPEGISIK	PLYSRADTLDLPEELPGV	KPTTRGPTPTMTTTPFT	TRQYAGFSTVESK	KFYDMILAGQQLSVAFDL							
MOCH	150	160	170	180	190	200	210	220	230	240	250	260	270	280	
MOCH		ATBEGYDSNIFKVEDVAGAGVAIDIV	EDTKILFDGIPLEDSVSHNGAVIP	VLATFIVTGEQQVPEKELNG	TIQNDILKEFWHTYIFPP	PEPSMKIADIPQTIAQRMK	FPFISISGTHQZAGTDAIL	ELAT							
BMCH		ATBEGYDSNIFKVEDVAGAGVAIDIV	EDTKILFDGIPLEDSVSHNGAVIP	VLATFIVTGEQQVPEKELNG	TIQNDILKEFWHTYIFPP	PEPSMKIADIPQTIAQRMK	FPFISISGTHQZAGTDAIL	ELAT							
MOCH	290	300	310	320	330	340	350	360	370	380	390	400	410	420	
MOCH		IADLECYCRITGLQAGLTIDEFAPRL	SFTWIGMFTETIADPRAGERL	MAHLTERFQPKRSKILLRAHQ	TSKSLTEQDPYNNIVRTATEA	MAAVFGGTQSLKSNFDEAL	GLPTVESARIARNTQII	IQEESGIFK							
BMCH		IADLECYCRITGLQAGLTIDEFAPRL	SFTWIGMFTETIADPRAGERL	MAHLTERFQPKRSKILLRAHQ	TSKSLTEQDPYNNIVRTATEA	MAAVFGGTQSLKSNFDEAL	GLPTVESARIARNTQII	IQEESGIFK							
MOCH	430	440	450	460	470	480	490	500	510	520	530	540	550	560	
MOCH		VADPWGSSHPHESLTHTDVTAAIKL	ITYEVEEPGGMAKAVABGIP	KLRIEBCAARQARIDSSSEV	TVGVNKKQLKEDSVVEVL	AIDMTSVKQIETKIKKES	SDQALADQCLSALTQCA	ASGDMILALAVDAARBC							
BMCH		VADPWGSSHPHESLTHTDVTAAIKL	ITYEVEEPGGMAKAVABGIP	KLRIEBCAARQARIDSSSEV	TVGVNKKQLKEDSVVEVL	AIDMTSVKQIETKIKKES	SDQALADQCLSALTQCA	ASGDMILALAVDAARBC							
MOCH	570	580	590	600	610	620	630	640	650	660	670	680	690	700	
MOCH		TVGEITDALKVPEGEHKAIDRW	NSGAYTRQFGESEKITSAT	KRVNKPHEDEGRPELLVA	RMQQGHRGAKVIATG	FDLGEFDVDIGFLPQT	PREVAQAVADVHAGV	STLAAGHKLVPEL	IKELTALGERFDILLVH						
BMCH		TVGEITDALKVPEGEHKAIDRW	NSGAYTRQFGESEKITSAT	KRVNKPHEDEGRPELLVA	RMQQGHRGAKVIATG	FDLGEFDVDIGFLPQT	PREVAQAVADVHAGV	STLAAGHKLVPEL	IKELTALGERFDILLVH						
MOCH	710	720	730	740	750										
MOCH		CGGVIPRQDYEEFLYEVGSN	VFGRTRIFRAAVQVLD	DIKKCLAEKQSV											
BMCH		CGGVIPRQDYEEFLYEVGSN	VFGRTRIFRAAVQVLD	DIKKCLAEKQSV											

Fig. 4. Alignment of mouse and human MCM protein sequences

Residues which are not identical are indicated by *. The mitochondrial targeting sequence of human MCM and the predicted targeting sequence of mouse MCM are underlined. The quantitative degree of similarity calculated from this alignment is shown in Table 3.

Table 3. Sequence similarity between MCMs from mouse human and *Prop. shermanii*

The Table shows the number of amino acid identities and possible matches (top diagonal) and the statistical significance of the similarity score in standard deviations (bottom diagonal) between optimal alignments mouse MCM (moMCM), human MCM (HMCM) and the α -chain (MUTB) and β -chain (MUTA) of *Prop. shermanii* MCM. Alignments and significance score were calculated by using the program ALIGN. The significance scores compare the degree of linear sequence similarity to that obtained by comparing randomized sequences with the same amino acid composition. Data comparing the *Prop. shermanii* sequences with human MCM are from Leadley & Ladley [35].

	moMCM	HMCM	MUTA	MUTB
	Identities/possible matches			
moMCM	–	700/748	164/632	411/720
HMCM	> 100	–	139/619	422/704
MUTA	17	13	–	140/629
MUTB	64	79	14	–
	Standard deviations from random			

Mouse MCM, like human MCM, also exhibits considerable sequence similarity to prokaryotic MCM from *Prop. shermanii* (Table 3). Profound sequence similarity between human and *Prop. shermanii* MCM has been described previously [35]. *Prop. shermanii* MCM is a heterodimer composed of two polypeptides, an α -subunit (MUTB) with a molecular mass of 79 kDa, and a β -subunit (MUTA) with a molecular mass of 67 kDa [10,14]. The human MCM monomer exhibits 60% identity with the MUTB gene and 22% identity with the MUTA gene (Table 3) [35]. The mouse protein exhibits quantitatively sequence similarity to the *Prop. shermanii* proteins, with 57% identity with the MUTB gene and 26% identity with the MUTA gene (Table 3). Thus the structure-function determinants of the apoenzyme are tightly constrained, despite the fact that the eukaryotic and prokaryotic enzymes have different quaternary structure and catalyse opposite reactions under physiological conditions. The uniformly high degree of sequence conservation between these proteins makes it impossible to pinpoint specific regions that are of particular functional importance.

Expression of MCM activity from the mouse MCM cDNA

In order to demonstrate the authenticity and function of the mouse cDNA clone, the cDNA was introduced into COS cells by using the p91023B expression vector. COS cells exhibit endogenous MCM activity that was not affected by transformation with anti-sense MCM or a dihydropteridine reductase expression vector (Table 4). MCM apoenzyme activity was quantitatively increased by transfection with mouse MCM in the sense orientation, indicating that the moMCM209 clone encodes an active enzyme. The endogenous MCM activity in COS cells precludes more detailed analysis of the recombinant gene product.

In order to have a better system in which to study the MCM cDNA and to demonstrate that the mouse MCM cDNA was congruous with the locus defective in *mut* forms of MMA, the mouse MCM cDNA was introduced into a human *mut*⁰ fibroblast line (GM50) by using the vector pCMV-moMCM. GM50 cells electroporated with the pCMV-moMCM sense construct showed an increase in MCM activity, as measured by the [¹⁴C]propionate uptake assay, to essentially normal levels (Fig. 5). These data confirm the authenticity of the mouse MCM cDNA clone and confirm somatic-cell complementation data that demonstrated that *mut* MMA arises from defects in the MCM apoenzyme.

Table 4. MCM activity in COS cells transformed with mouse MCM and controls

The full-length MCM cDNA in the vector p91023B was introduced into COS cells by Ca₃(PO₄)₂ co-precipitation and soluble protein extracts were prepared after 72 h. Activity is expressed in nmol of succinate formed/h per mg in the sample minus a boiled blank. Controls were non-transfected COS cells and COS cells transfected with MCM in anti-sense orientation or a clone expressing dihydropteridine reductase. Transformed COS cells were transformed with mouse MCM expression vector. The means and ranges for seven separate transfection experiments are shown.

Sample	Mean activity (range) (nmol/h per mg)
Controls	
COS (non-transfected)	24.5
COS (transfected)	23.6 (19.7–27.8)
Transformed	46.0 (38.7–53.5)

Fig. 5. Complementation of MCM enzyme activity in MCM-deficient fibroblasts

The full-length mouse MCM cDNA in the vector pNassCMV (pCMV-moMCM) was introduced into human primary fibroblasts genetically deficient in the MCM apoenzyme, and propionate incorporation was measured 48 h after transformation. Activity in normal fibroblasts was not affected by transformation with the control plasmid pCMV-βGal. Activity in *mut*⁰ cells was reconstituted to normal levels by transformation with the pCMV-moMCM (sense) vector but was not affected by transformation with pCMV-βGal or pCMV-moMCM (anti-sense). Key: Normal cells, normal human fibroblast; *mut*⁰ cells, GM50 cell line exhibiting *mut*⁰ phenotype; βG, transformation performed with 10 μg of pCMV-βGal; MCM(10), transformation performed with 10 μg of pCMV-moMCM (sense); MCM(20), transformation performed with 20 μg of pCMV-moMCM (sense); MCM(AS) transformations performed with 20 μg of pCMV-moMCM (anti-sense). The means and standard deviations (small black bars) for triplicate samples for electroporation pulses of 260 V (■) and 280 V (■) are shown. Basal activity in the absence of electroporation is also shown (■).

Conclusion

Several lines of investigation are converging on a description of structure-function determinants and biological significance of MCM. The cloning and expression of human, *Prop. shermanii*, and now mouse MCM have provided the primary structure of MCM and a system for studying mutations. Other work has identified several mutations that give rise to variant phenotypes

of *mut* methylmalonicacidaemia, including complete MCM deficiency [36], failure of post-translational processing of MCM [37] and decreased MCM activity due to an altered K_m (adenosylcobalamin) (R. Jansen & F. D. Ledley, unpublished work). This information, together with structural data from the recently crystallized *Prop. shermanii* MCM [38], should enable critical analysis of the structure and function of the MCM apoenzyme. In addition, the demonstrated structural and functional homology between the mouse and human enzymes suggests that the mouse may be a useful model system for studying the physiological role of MCM and constructing an animal model of MCM deficiency that might mimic the phenotype of the human deficiency state.

We thank Dr. Grant MacGregor and Ruud Jansen for their assistance, Dr. David George and Dr. Lois Hunt of the Protein Identification Resource for assistance with sequence comparisons, and Tammy Reid for preparation of the manuscript. F. D. L. is an Assistant Investigator of the Howard Hughes Medical Institute. This work was supported by National Institutes of Health Grants HD-24186 to (F. D. L.) and HD-24064 (to Baylor Mental Retardation Research Center).

REFERENCES

- Rosenberg, L. E. & Fenton, W. A. (1989) in the Metabolic Basis of Inherited Disease (Scriver, C. R., Beaudet, A. L., Sly, W. S. & Valle, D., eds.), 3rd edn., pp. 822–844, McGraw-Hill, New York
- Leng, R. A., Steel, J. W. & Luick, J. R. (1967) *Biochem. J.* **103**, 785–790
- Bergman, E. N., Walter, E., Roe, E. & Kon, K. A. (1966) *J. Physiol. (London)* **211**, 793–799
- Murthy, V. V., Jones, E., Cole, T. W., Jr. & Johnson, J., Jr. (1977) *Biochim. Biophys. Acta* **483**, 487–491
- Hunaiti, A. A. & Kolattukudy, P. E. (1984) *Antimicrob. Agents Chemother.* **25**, 173–178
- Ueda, S., Sato, K. & Shimizu, S. (1978) *J. Nutr. Sci. Vitaminol.* **24**, 477–489
- Pietrzak, S. M. & Saz, H. J. (1981) *Mol. Biochem. Parasitol.* **3**, 61–70
- Kolhouse, J. F., Utley, C. & Allen, R. H. (1980) *J. Biol. Chem.* **255**, 2708–2712
- Fenton, W. A., Hack, A. M., Willard, H. F., Gertler, A. & Rosenberg, L. E. (1982) *Arch. Biochem. Biophys.* **214**, 815–823
- Francalanci, F., Davis, N. K., Fuller, J. Q., Murfitt, D. & Leadlay, P. F. (1986) *Biochem. J.* **236**, 489–494
- Han, Y. S., Bratt, J. M. & Hogenkamp, H. P. (1984) *Comp. Biochem. Physiol. B* **78**, 414–445
- Ledley, F. D., Lumetta, M., Nguyen, P. N., Kolhouse, J. F. & Allen, R. H. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3518–3521
- Jansen, R., Kalousek, F., Fenton, W. A., Rosenberg, L. E. & Ledley, F. D. (1989) *Genomics* **4**, 198–205
- Marsh, N., McKie, N., Davis, N. K. & Leadlay, P. F. (1989) *Biochem. J.* **260**, 345–352
- Kolhouse, J. F., Stabler, S. P. & Allen, R. H. (1987) *Methods Enzymol.* **166**, 407–414
- Ledley, F. D., Crane, A. M. & Lumetta, M. (1990) *Am. J. Hum. Genet.* **46**, 539–547
- Willard, H. F., Ambani, L. M., Hart, A. C., Mahoney, M. J. & Rosenberg, L. E. (1976) *Hum. Genet.* **34**, 277–283
- Morrow, G., Revsin, B., Mathews, C. & Giles, H. (1976) *Clin. Genet.* **10**, 218–221
- Woo, S. L. C. (1982) *Methods Enzymol.* **68**, 390–394
- Berger, S. B. & Kimmel, A. R. (1987) *Guide to Molecular Cloning Techniques*, Academic Press, New York
- Henikoff, S. (1984) *Gene* **28**, 351–359
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Kaufman, R. J. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 689–693
- Mellon, P., Parker, V., Gluzman, Y. & Maniatis, T. (1981) *Cell* **27**, 279–288
- Chen, C. & Okayama, H. (1987) *Mol. Cell. Biol.* **7**, 2745–2752
- Lockyer, J., Cook, R. G., Milstien, S., Kaufman, S., Woo, S. L. C. & Ledley, F. D. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 3329–3333
- MacGregor, G. & Caskey, C. T. (1989) *Nucleic Acids Res.* **17**, 2365
- Shigekawa, K. & Dower, W. J. (1988) *BioTechniques* **6**, 742–751
- MacGregor, G. R., Nolan, G. P., Fiering, S. & Herzenberg, L. A. (1990) in *Methods in Molecular Biology* (Murray, E. & Walker, J. M., eds.), vol. 7, Humana Press, Clifton, in the press
- Giudici, T. A., Chen, R. G., Oizumi, J., Shaw, K. N., Ng, W. G. & Donnell, G. N. (1986) *Biochem. Med. Metab. Biol.* **35**, 384–398
- van der Westhuyzen, J., Cantrill, R. C., Fernandes-Costa, F. & Metz, J. (1983) *J. Nutr.* **113**, 531–537
- Kozak, M. (1986) *Cell* **44**, 283–292
- Roise, D. & Schatz, G. (1988) *J. Biol. Chem.* **263**, 4509–4511
- Kalousek, F., Hendrick, J. P. & Rosenberg, L. E. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7536–7540
- Leadlay, P. F. & Ledley, F. D. (1989) in *Cobalamin '88* (Proc. Int. Symp. Biomedicine and Physiology of Vitamin B12 1st) (Linnell, J. C. & Bhatt, H. R., eds.), pp. 343–354, Children's Medical Charity, London
- Jansen, R. & Ledley, F. D. (1990) *Am. J. Hum. Genet.*, in the press
- Ledley, F. D., Jansen, R., Nham, S.-U., Fenton, W. E. & Rosenberg, L. E. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3147–3150
- Marsh, N., Leadlay, P. F. & Evans, P. R. (1988) *J. Mol. Biol.* **200**, 421–422

Received 23 January/27 April 1990; accepted 3 May 1990