

Identification of the gene responsible for the *cblA* complementation group of vitamin B₁₂-responsive methylmalonic acidemia based on analysis of prokaryotic gene arrangements

C. Melissa Dobson*, Timothy Wai†, Daniel Leclerc‡, Aaron Wilson*, Xuchu Wu*, Carole Doré§¶, Thomas Hudson*§¶||, David S. Rosenblatt†*§||, and Roy A. Gravel*.*.*

*Department of Biochemistry and Molecular Biology, University of Calgary, Calgary, AB, Canada T2N 1N4; and Departments of §Human Genetics, †Biology, and ||Medicine, ‡McGill University Health Centre, and ¶Montreal Genome Centre, McGill University, Montreal, QC, Canada H3A 2T5

Communicated by Leon E. Rosenberg, Princeton University, Princeton, NJ, October 10, 2002 (received for review July 25, 2002)

Vitamin B₁₂ (cobalamin) is an essential cofactor of two enzymes, methionine synthase and methylmalonyl-CoA mutase. The conversion of the vitamin to its coenzymes requires a series of biochemical modifications for which several genetic diseases are known, comprising eight complementation groups (*cblA* through *cblH*). The objective of this study was to clone the gene responsible for the *cblA* complementation group thought to represent a mitochondrial cobalamin reductase. Examination of bacterial operons containing genes in close proximity to the gene for methylmalonyl-CoA mutase and searching for orthologous sequences in the human genome yielded potential candidates. A candidate gene was evaluated for deleterious mutations in *cblA* patient cell lines, which revealed a 4-bp deletion in three cell lines, as well as an 8-bp insertion and point mutations causing a stop codon and an amino acid substitution. These data confirm that the identified gene, *MMAA*, corresponds to the *cblA* complementation group. It is located on chromosome 4q31.1–2 and encodes a predicted protein of 418 aa. A Northern blot revealed RNA species of 1.4, 2.6, and 5.5 kb predominating in liver and skeletal muscle. The deduced amino acid sequence reveals a domain structure, which belongs to the AAA ATPase superfamily that encompasses a wide variety of proteins including ATP-binding cassette transporter accessory proteins that bind ATP and GTP. We speculate that we have identified a component of a transporter or an accessory protein that is involved in the translocation of vitamin B₁₂ into mitochondria.

Vitamin B₁₂ (cobalamin, Cbl) is one of the most complex vitamins found in nature (1). It acts as a coenzyme in variety of molecular rearrangements to generate oxidizable substrates. In mammals, Cbl is involved in two enzyme systems: methionine synthase (MS), a key enzyme in one carbon metabolism, and methylmalonyl-CoA mutase (MCM), an enzyme in the pathway of branched chain amino acid catabolism (Fig. 1). The conversion of the vitamin to its coenzyme forms requires a series of processing steps in cells. Following entry of Cbl into the lysosome as a complex with transcobalamin II, it is released by proteolytic digestion and transported into the cytosol across the lysosomal membrane. The free vitamin is reduced from cob(III)alamin to cob(II)alamin and cob(I)alamin, upon which it can be methylated to form MeCbl or adenosylated to AdoCbl. The reductive methylation of Cbl occurs on the MS protein and is catalyzed by MS reductase. The reactions leading to the synthesis of AdoCbl are less well understood. MCM is a mitochondrial enzyme, and the final steps of AdoCbl synthesis occur in the mitochondrion. The current view is that the first reductive step, cob(III)alamin to cob(II)alamin, occurs in the cytosol. Subsequently, Cbl, probably as cob(II)alamin, is transported into the mitochondrion, where the next reduction step, to cob(I)alamin, occurs followed by addition of the adenosyl group to form AdoCbl (2).

Blocks in any of the Cbl transport or modification steps usually cause severe disease in infancy or early childhood, although they may also present later in life (2). Depending on the nature of the metabolic block, affected patients may present with homocystinuria or methylmalonic aciduria or both. The combined defects are due to deficits in early steps of Cbl processing, whereas deficient activity of only MS or MCM indicates a block at a later step of Cbl metabolism or a primary defect of the apoenzyme. Eight complementation groups, *cblA* to *cblH*, have been identified that account for these disorders (Fig. 1). Two of the corresponding genes have been cloned: *MTRR* encoding MS reductase (*cblE* group; ref. 3) and *MTR* encoding MS (*cblG* group; refs. 4–6). Of the remaining complementation groups, three, *cblA*, *cblB*, and *cblH*, are uniquely involved in AdoCbl synthesis, whereas the remainder, *cblC*, *cblD*, and *cblF*, affect both MeCbl and AdoCbl synthesis (2).

Our aim was to identify the gene responsible for the *cblA* complementation group (7), a disorder that classically is associated with vitamin B₁₂-responsive methylmalonic aciduria (8). The biochemical features of this disorder have been described (9–13). Fibroblast cultures from *cblA* patients typically have decreased propionate incorporation into cellular protein and reduced synthesis of AdoCbl, whereas MeCbl synthesis remains normal (13). In contrast, extracts from the same fibroblast cultures are able to synthesize AdoCbl in the presence of chemical reducing agents, such as DTT (and ATP) (10–13). It has been proposed that *cblA* corresponds to a defect of a mitochondrial, NADPH-dependent aquacobalamin reductase (13, 14), or that it may have a block in mitochondrial binding or transport of Cbl (9, 10, 12).

We have identified the gene responsible for the *cblA* defect through examination of bacterial gene clusters by using the National Center for Biotechnology Information (NCBI) utility, Clusters of Orthologous Groups of proteins (www.ncbi.nlm.nih.gov/COG). It was detected as one of the potential genes in an operon containing MCM. Several mutations, including two frameshift mutations, were identified in *cblA* patients, confirming the authenticity of the gene identification. Based on its predicted domain structure, we hypothesize that we have identified a component of a transporter required for the translocation of Cbl into mitochondria. We propose the gene name,

Abbreviations: MCM, methylmalonyl-CoA mutase; Cbl, cobalamin; AdoCbl, adenosylcobalamin; OHcbl, hydroxycobalamin; MS, methionine synthase; LAO, lysine-arginine-ornithine; MPTP, mitochondrial permeability transition pore; *MMAA*, methylmalonic acidemia linked to the *cblA* complementation group.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF524841–AF524846).

**To whom correspondence should be addressed. E-mail: rgravel@ucalgary.ca.

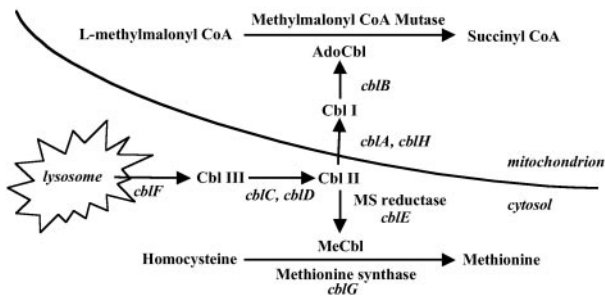


Fig. 1. Intracellular processing of Cbl (modified from ref. 2). Steps are identified according to complementation group names. Their gene (uppercase italics) and predicted functions are: *cblA* (*MMAA*), this study; *cblB*, cob(II)alamin adenosyltransferase; *cblC/D*, possibly cob(III)alamin reductase; *cblE* (*MTRR*), methionine synthase reductase; *cblF*, lysosomal Cbl efflux; *cblG* (*MTR*), methionine synthase; and *cblH*, possibly a component of Cbl transport or cob(II)alamin reductase.

MMAA, to denote methylmalonic acidemia linked to the *cblA* complementation group.

Materials and Methods

Patient Cell Lines. Fibroblast lines from six patients classified as members of the *cblA* complementation group were used in this study. Patient cell line WG1192 was reported to synthesize AdoCbl *in vitro* (11). Propionate uptake was 1.26 ± 0.10 nmol/mg protein/18 h without hydroxycobalamin (OHCbl) added to medium (13) compared to control values of 10.8 ± 3.2 nmol/mg protein/18 h. After incubation of cells in 25 pg/ml [⁵⁷Co]cyanocobalamin for 4 days, AdoCbl was at 5.0% of total Cbl as compared to normal levels of $15.29 \pm 4.2\%$. WG1776 is from a white female who presented at 7 mo with acidosis, ketosis, and lethargy. [¹⁴C]propionate uptake was 2.3 nmol/mg protein/18 h without and 4.6 nmol/mg protein/18 h with OHCbl in the culture medium (normal with OHCbl added, 10.9 ± 3.5 nmol/mg protein/18 h). AdoCbl was at 6.6% after [⁵⁷Co]cyanocobalamin incubation. WG1943 is from a 2-mo-old East Indian male with vomiting, lethargy, and Cbl-responsive methylmalonic acidemia. [¹⁴C]propionate uptake was 3.4 nmol/mg protein/18 h without and 9.0 nmol/mg protein/18 h with OHCbl in the culture medium. AdoCbl was at 6.8%. WG2014 is from a 3-mo-old white male who presented with fever, projectile vomiting, dehydration, generalized fine tremors, and Cbl-responsive methylmalonic acidemia. [¹⁴C]propionate uptake was 1.2 nmol/mg protein/18 h without and 7.1 nmol/mg protein/18 h with OHCbl in the culture medium. AdoCbl was at 7.5% (15). WG3080 is from a white female who presented at 7 days with lethargy, poor feeding, and seizures. The [¹⁴C]propi-

onate uptake was 0.6 nmol/mg protein/18 h without and 2.4 nmol/mg protein/18 h with OHCbl in the culture medium. AdoCbl was at 3.5%. There is no known consanguinity associated with any of the patient cell lines.

DNA and RNA Extraction and PCR Amplification. Genomic DNA was isolated as described (16) or by using Genomic-tip (Qiagen) as described in the manufacturer's protocol. PCR reactions were performed on genomic DNA with primers designed to amplify exons of the coding sequence and portions of the flanking introns (Table 1). Total cellular RNA was purified from a human liver hybridoma cell line (Huh-1) by using Trizol (Invitrogen). The total cDNA was amplified by using oligo-dT as described by the Omniscript protocol (Qiagen), followed by PCR amplification with RNA-specific primers (Table 1). The PCR products were purified with the QIAquick PCR purification kit (Qiagen) and cloned into pGEM (Promega) according to the supplier's instructions.

Heteroduplex Analysis and DNA Sequencing. Initial examination of PCR products was performed by heteroduplex analysis, as described (17). Heteroduplexes were obtained by mixing, melting, and reannealing the PCR products of mutant and control DNA before electrophoresis. Positive cell lines also underwent the same procedure without mixing with control to determine whether the mutation was on one or both alleles. The observation of a heteroduplex, with implication of a multibase mutation, was taken as indication that the correct gene was identified. DNA sequencing was accomplished in 96-well plates in 10- μ l reactions made up of 2 μ l of purified PCR product and 1 μ l of BigDye Terminator Cycle Sequencing Version 2.0 (Applied Biosystems) and analyzed with an ABI 3700 sequencer (Applied Biosystems).

Northern Blot. A multiple human tissue Northern blot (CLONTECH) was used for determining the level and size of mRNA species. A probe was generated from a 508-bp cDNA segment amplifying regions 380–887 bp of the full-length cDNA, using the Huh-1 cell line as the source of RNA. Detection of β -actin mRNA was used as control. All procedures followed the protocol supplied by CLONTECH.

Results

Identification of the *MMAA* cDNA. Initial efforts to identify sequences in model genomes that might represent an NAD(P)H and flavin-dependent aquacobalamin reductase were unsuccessful, owing to the excessive number of potential candidates. However, with the availability of the COG database, it became possible to narrow the search to genes in microbial organisms that were arranged in clusters with genes encoding Cbl-

Table 1. Splice junctions in the *MMAA* gene and primers used for exon/flanking sequence amplifications

Exon	Size, bp	PCR, bp	Sense* primer sequence	Antisense* primer sequence	Splice sites		Exon boundaries
					5'	3'	
1	≥ 19	ND	ND	ND		...GCGGgtg	≥ -85 to -68
2	506	515	tagggagggtcacaatcacattgagc	agactgacTACTCTGAAATGCTAGTG	tagGGAG...	...GTAGgtc	-67 to 439
3	123	397	tctctttccaccgtagGATTGTCTG	gatcatcttcaacttctggcaactg	tagGATT...	...GGTGgta	$440-562$
4	171	507	ggaactggctgataattgacccgta	cctaagtgttacagttctcctctggc	tagGATC...	...GTTGgtg	$563-733$
5	87	470	ggttgactgtgtgaccatgagtatg	gctggagcgatccatgtatattcagt	cagGTGT...	...GCAGgta	$734-820$
6	149	586	ctgccactcaatcatgttgacttg	tcatgactacacactccgtctttgtc	tagGGTA...	...AAAGgta	$821-969$
7	352	527	taagaattaactggcaggtatcagc	ctaacttcttcaagttttcaagcatg	tagGTAA...		$970-1317$
RNA†		1,359	cttcgggcggggaggtcacaatcaca	gtaaaattattatcacagatgaa			

*Splice sites were determined based on a comparison of the cDNA and genomic sequences. The coding sequences are uppercase, whereas intron and noncoding sequences are lowercase. ND, not done.

†The full-length coding sequence was amplified by RT-PCR from position -74 to 1285 , using the "RNA" primers.

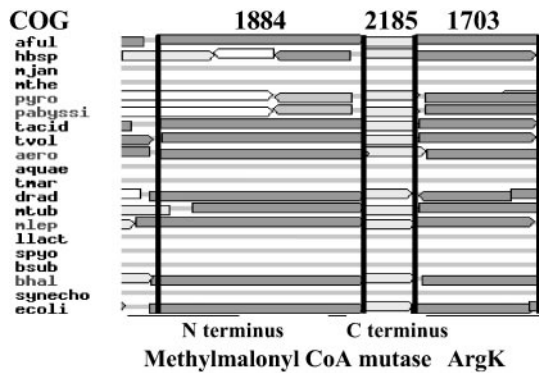


Fig. 2. Cluster of orthologous genes examining the MCM N terminus (COG1884) and C terminus (COG2185) and the genes identified as comprising COG1703. Bacteria discussed in this study include *Archaeoglobus fulgidus* (aful) and *Escherichia coli* K12 (ecol); others are listed at www.ncbi.nlm.nih.gov/COG.

dependent enzymes. A search for MCM yielded COG2185, identified as corresponding to the Cbl-binding, carboxyl-terminal domain of MCM. It is present as a member of a cluster in 16 of the 43 complete genomes represented in the database (Fig. 2). In 10 of these, COG1703 was found immediately adjacent to COG2185. COG1703 is characterized as “putative periplasmic protein kinase ArgK and related GTPases of G3E family.” COG1703 was found in the genomes of Archaea (five genera), Bacteria (five genera), and Eukaryota (one genus), as listed in the COG database. One member of this COG, namely ArgK, has been described as a coupled ATPase of the lysine-arginine-ornithine (LAO) transport system in *E. coli* (18). Given the proximity of COG1703 to MCM and its frequency in microbial genomes, we pursued it as a candidate for involvement in the AdoCbl pathway.

The *A. fulgidus* sequence (GenBank accession no. AAB89957) from COG1703 was used as the template for a BLAST (TBLASTN) search of the human Expressed Sequence Tags (DBEST) database and Unigene clusters at NCBI. This search yielded 10 ESTs listed in Unigene cluster Hs.126216 and 18 others in Hs.21017. Align-

```

-84 CTTCGGTGGCTTCGGGCGGGGAGGTGACATFCACATTTAGGCCAAACGCATCCAGTGTTC
      M P M L L P H P H Q H F I 2
-24 TCTCCAGTACAAATAAAACGAATATGCCCAGTGCCTCCACATCCTCACCAGCATTTTC
      L R G L L R A P P R C Y H F I P H S S T 32
37 CFAAAAGCCCTTTAAGAGCACCTTCCGATGTTACCACCTTCATCTTTCACTCAAGTACT
      H L G S G I P C A Q P F N S L G L H C T 52
97 CATCCGGATCAGGAATCCCATGTCTCAGCCGTTTAACTTCTCTGGACTCCATTTGTACA
      K W M L L S D G L K R K L C V Q T T L K 72
157 AAGTGGATGCTGCTGTCAGATGGCTTAAAGAGAAAATATGTGTACAAACCACTTAAAG
      D H T E G L S D K E Q R F V D K L Y T G 92
217 GACCACACAGAAGGACTTTCGATAAAGAGCAAAAGATTTGGATAAACTTTTACTGTGT
      L I Q G Q R A C L A E A I T L V E S T H 112
277 TTAAATCCAGGCAAGGGCCCTTTAGCAGAGGCCATAACTCTTGTAGAAATCAACTCAC
      S R K K E L A Q V L L Q K V L L Y H R E 132
337 AGCAGGAAAAGGAGTTAGCCAGGCTTCTTCAGAAAATATTTACTTACCACAGAGAA
      Q E Q S N K G K P L A F R V G L S G P P 152
397 CAAGAAATCAATAAAGGAAAACCACTAGCATTTTCAGTAGGATGCTCTGGCCCCCT
      G A G K S T F I E Y F G K M L T E R G H 172
457 GGTCTGGAAAATCAACATTTATAGAATATTTGAAAATGCTTACTGAGAGAGGGCCAC
      K L S V L A V D P S S C T S G G S L G E 192
517 AAATATCTGTGCTAGCTGTGACCCCTTCTTCTTACTACTGTGGTGGATCCTTTAGT
      D K T R M T E L S R D M N A Y I R P S P 212
577 GATAAAACCGAATGACTGAGTTATCAAGAGATATGAATGCATACATGAGCCATCTCCT
      T R G T L G G V F R T T N E A I L L C E 232
637 ACTAGAGGAATTTAGAGCCGTGACAAGGACCAAAATGAGCTATTCTGTGTGAGAA
      G A G Y D I I L I E M A D L V A V T K S D 292
697 GGAGCGGATATGACATAATCTTATTGAAACCGTTPGCTGGGTGAGTCCGGATTTGCT
      V A D M V D M F V L L L P P A G G D E L 272
757 GTTGTGACATGGTTGACATGTTGTTTACTAGCCACCAGCAGGAGGAGGAGAGCTG
      Q G I K R G I I E M A D L V A V T K S D 292
817 CAGGATCAAAAGGGTATAATCGAGATGCGAGATGCTGAGCTGTAACTAAATCTGAT
      G D L I V P A R R I Q A E Y V S A L K L 312
877 GGAGACTTGTATTGCGCAGCTCGAAGGATACAAGCGGAATATGAGTGCCTGAAATTA
      L R K R S Q V W K P K V I R I S A R S G 332
937 CTCCGCAACCGTTCACAGCTCGGAAACCAAGTAACTGATTTCTGCCGAAAGTGA
      E C I S E M W D K M K D P Q D D L M L E 352
997 GAGGGATCTCTGAAATGCGGGAATAAAGATATTCACAGGCTAAATGCTTGGCAGT
      G E L T A K R R K Q Q K V W M W N L I Q 372
1057 GGGGAGCTGACTGCCAAACGACGGAAGCAACGAAAAGTTGGATGTGAAATCTCATTG
      E S V L E H F R T H P T V R E Q I P L K 392
1117 GAAAGTGTGTAGAGCATTTCAGGACCAACCCACAGTCCGGAAACAGATTCACCTCTG
      E Q K V L I G A L S P G L A A D F L L K 412
1177 GAACAAAAGGTTTCATTTGGGCCCTGTCGCCAGACTAGCAGCAGACTTCTTTGTAAAA
      A F K S R D *** 418
1237 GCTTTTAAAGCAGAGACTAATAAAATTCATCTCTGTATAAATTTTACATATCATTTCA
1297 TAAAGTATTTTAAAGAAAA

```

Fig. 4. Nucleotide and deduced amino acid sequence of human *MMAA*. The nucleotide and amino acid residues are numbered on the left and right margins, respectively. Two in-frame stop codons are indicated by *** immediately upstream and downstream of the coding sequence. A candidate polyA signal starting at 1,296 is underlined.

ment of the ESTs permitted construction of a composite 1,404-bp-long cDNA that contains a putative start codon with in-frame upstream stop codon at the 5' end and polyA tail at the 3' end (Figs. 3 and 4).

The 5' and 3' ends of the composite human cDNA were

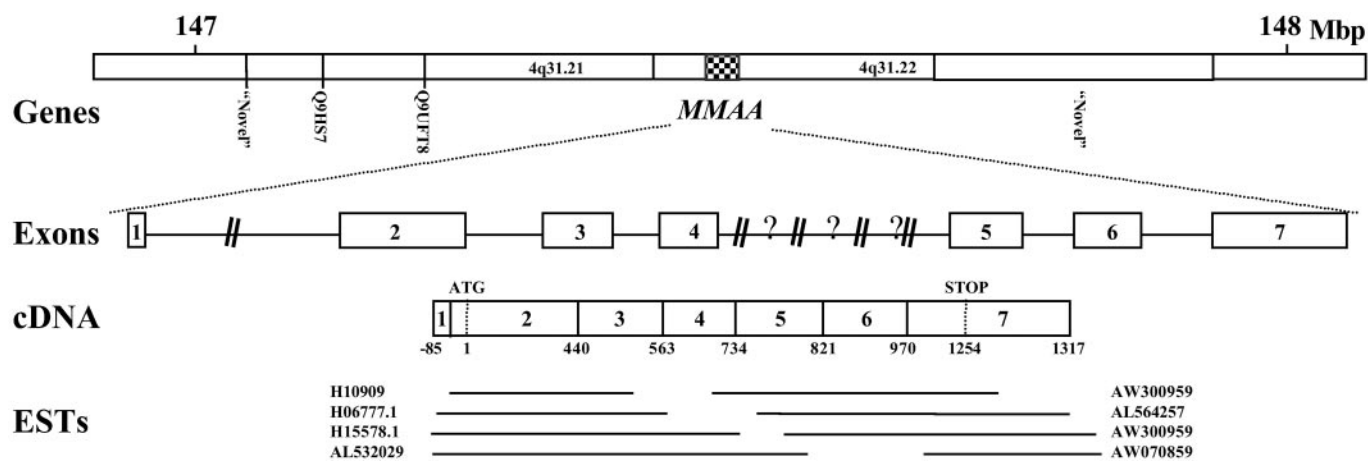


Fig. 3. Structural organization of the human *MMAA* gene. The top schematic (“Genes”) illustrates the flanking genes and map positions surrounding *MMAA* according to contigview in Ensembl (www.ensembl.org). Note that NCBI maps *MMAA* to the 147.36–147.38 megabase pairs (Mbp) location of chromosome 4. Both Ensembl and NCBI localize *MMAA* to 4q31.21–22. The next line (“Exons”) illustrates the location of exons (numbered boxes) and introns (lines) along the *MMAA* gene located in Contig NT_016606.11. Our independently confirmed sequences from genomic DNA were deposited in GenBank under the accession nos. AF524841–AF524846. Introns are drawn to scale, except where “?” is used to denote a region of unknown sequence and length and “//” is used to denote a longer region than illustrated. Next the cDNA (“cDNA”) is illustrated, with splice and start/stop codons indicated, showing that it was constructed from the numerous ESTs (“ESTs”) belonging to Unigene clusters Hs.126261 (10 ESTs) and Hs.21017 (18 ESTs) and verified by RT-PCR sequencing. Only some ESTs are illustrated here.

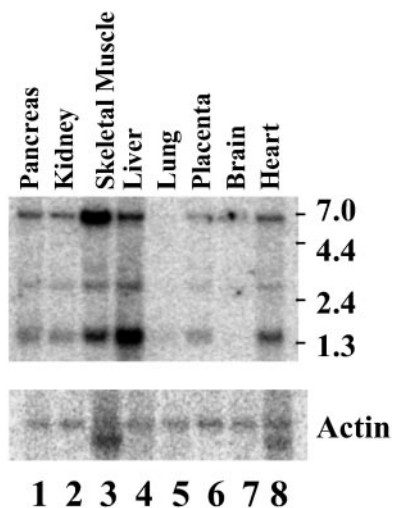


Fig. 6. Northern blot analysis of *MMAA* mRNA. A Northern blot of poly(A)⁺ RNA from the indicated human tissues. The membranes were hybridized with a ³²P-labeled 508-bp probe from *MMAA* mRNA (A) or a β -actin probe (B). Molecular size marker positions are indicated on the right.

human–mouse homology map (www.ncbi.nlm.nih.gov) that links human 4q31.3–4q35 with chromosome 8 of mouse.

Expression in Tissues. The 508-bp probe hybridized to three RNA species of \approx 1.4, 2.6, and 5.5 kb on the Northern blot, with highest levels of expression in liver and skeletal muscle. Skeletal muscle expressed the largest transcript predominantly, whereas in liver the smallest one predominated (Fig. 6). The 1.4-kb species is essentially identical to the predicted length of the composite cDNA.

Mutations in the *cbIA* Complementation Group. Fibroblast lines from patients from the *cbIA* complementation group were analyzed for mutations by PCR-heteroduplex analysis and DNA sequencing. Three cell lines, WG1192, WG2014, and WG3080, gave heteroduplexes on PAGE of PCR products obtained by amplification of exon 4 and flanking sequences (Fig. 7). The latter two cell lines showed heteroduplexes in the presence or absence of mixing with control DNA. These results suggest that WG1192 is homoallelic, and that WG2014 and WG3080 are heteroallelic for a multibase change within the amplified sequence. DNA sequencing revealed the mutation to be a 4-bp deletion, c.592.delACTG, which would disrupt reading frame and intro-

Patient	Allele 1	Allele 2
WG1192	c.592.del4	c.592.del4
WG1776	c.283.C>T (Q95X)	ND
WG1943	c.620.A>G (Y207C)	c.620.A>G (Y207C)
WG2014	c.592.del4	ND
WG3080	c.260.ins8	c.592.del4

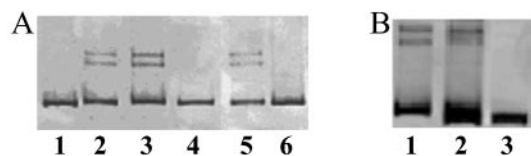


Fig. 7. Mutations reported here and heteroduplex tests. (A) Heteroduplex test for c.592.delACTG: lane 1, control cells; lane 2, WG2014 mixed with control cells in a 1:1 ratio; lane 3, WG2014 alone; lane 4, WG1192 alone; lane 5, WG1192 mixed with control; lane 6, WG1943 mixed with control cells in a 1:1 ratio. (B) Heteroduplex for c.259.delATAAACTT: lane 1, WG3080 alone; lane 2, WG3080 and control cells; lane 3, control cells alone.

duce a downstream stop codon in the transcribed sequence. WG3080 produced a second heteroduplex on PCR of exon 2 (Fig. 7). DNA sequencing revealed this change to be an 8-bp insertion, c.260.insATAAACTT. It would also result in a severely truncated protein. Given the presence of two different frameshift mutations on each allele of WG3080 and the 4-bp deletion on both alleles of WG1192, both cell lines would be null for a functional protein.

Two point mutations were also detected. Cell line WG1776 contains a premature stop codon, c.283.C>T (Q94X), on one allele, which would also disrupt protein expression. The second allele is unidentified. WG1943 contains a mutation, c.620.G>A, resulting in the amino acid substitution, Y207C, which was not identified in 70 control cell lines sequenced. This mutation is present on both alleles and suggests inactivation or instability of the resulting protein; although expression of the mutant protein would be required to clarify the effect of this substitution.

Discussion

We have combined the analysis of prokaryotic gene arrangements with the identification of mutations in affected patients to determine the identity of the disease-causing gene in patients from the *cbIA* complementation group. The key finding confirming the identity of the *MMAA* gene was the detection of heteroduplexes harboring mutations in the candidate gene for five *cbIA* cell lines. DNA sequencing revealed a 4-bp deletion and an 8-bp insertion in patient cell lines that would be expected to produce frameshifts and truncated proteins. Significantly, the 4-bp deletion was shown to be homoallelic in one cell line, whereas another was heteroallelic for both the deletion and insertion. These results demonstrate that in at least two patients, the mutations on both alleles are of such severity that the expected outcome is the complete absence of protein function. Based on the nature of mutations identified in these patients, we conclude that the *MMAA* gene is responsible for the *cbIA* complementation group. Given the similarity of the *cbIH* to *cbIA* complementation groups, we also tested a *cbIH* cell line for mutations, but found none, indicating that *cbIH* likely corresponds to a different gene.

It was anticipated that the *MMAA* gene would code for a Cbl reductase or perhaps a mitochondrial transporter of vitamin B₁₂. Early studies pointed to a mitochondrial defect because the reduction in AdoCbl synthesis was localized to mitochondria, and the activity was restored if the assay was conducted in cell extracts rather than whole cells (10). These findings are compatible with a block in mitochondrial transport, but no such transporter was identified (20, 21). Instead, Watanabe and coworkers (22, 23) identified an NADPH-dependent aquacobalamin reductase in rat liver and *Euglena* mitochondria, which was deficient in fibroblasts from one *cbIA* cell line (14). They took this observation as evidence that the *MMAA* gene might code for a reductase. Our results are consistent with mitochondrial targeting of the *MMAA* protein. It has a predicted leader sequence and signal cleavage site ($P = 0.998$; www.inra.fr/predotar; Fig. 5), but it is unlikely to be an aquacobalamin reductase because it does not contain domains suggestive of flavin or NAD(P)H-binding sites.

In considering a possible role for the *MMAA* gene, we first examined the published work on ArgK, the *E. coli* protein that is 61% similar to *MMAA*. The *argK* gene was initially identified as the site of mutations resulting in a deficiency in uptake of lysine (L), ornithine (O), and arginine (A) (18). The affected gene, *argK*, was expressed and shown to code for a protein that has intrinsic ATPase activity and could phosphorylate the LAO and AO periplasmic binding proteins *in vitro* (18). However, a revertant of the original mutant had restored ATPase activity and LAO uptake but remained defective in the protein kinase activity. It was concluded that ArgK is an essential, common

accessory to the LAO and AO transport systems through its ATPase activity, but that its putative kinase activity was not involved (18). The LAO/AO and arginine periplasmic binding proteins have been shown to be associated with two independent inner membrane transport systems, HisP and ArtP. Both are ATPases with ATP-binding cassette signatures (24, 25). This challenges the notion that ArgK, which lacks an ATP-binding cassette signature, is an LAO/AO transport-associated ATPase.

Therefore, with the possibility that ArgK might be involved with vitamin B₁₂ transport in microbes, we reviewed the components of the vitamin B₁₂ transport system in *E. coli* (26). Vitamin B₁₂ is transported across the inner membrane by two identical membrane-spanning components, BtuC, and an associated ATPase, BtuD. A comparison of the sequence of BtuD with that of ArgK shows them to be distinct and related only in the Walker A motif, commonly found in ATP-binding proteins. An extensive phylogenetic analysis of nucleotide-binding proteins included ArgK as a member, proposing that it belongs to one of four subfamilies comprising the G3E family of P-loop GTPases (27). The G3E branch of the phylogenetic tree is defined by the specific sequence signatures associated with nucleotide binding, including Walker A and Walker B motifs, a Mg²⁺-binding aspartate residue and a sequence specific to GTP defined by an [N/T]KxD motif (Fig. 5). In addition to ArgK, the G3E family comprises three additional subfamilies represented by: UreG, an accessory urease subunit involved in the assembly of the metallocentre of nickel metalloenzymes; HypB, a hydrogenase expression protein that is also involved in metal binding and metalloenzyme assembly; and, interestingly, CobW, a protein in the biosynthetic pathway of vitamin B₁₂ whose role remains undefined. It is significant that all of these proteins are involved in the assembly of metalloenzymes which, conceptually, reinforces the idea of a similar role for ArgK in Cbl metabolism. Nonetheless, although the proximity of *argK* to the MCM gene in bacterial gene clusters has been recognized previously, efforts to identify a role for the ArgK protein in this area of metabolism have been unsuccessful (28, 29).

In sum, the studies on ArgK do not readily point to a function for the protein and offer only limited insight into predicted function of the MMAA protein. Nevertheless, our knowledge of

the biochemical abnormalities of *cblA* patients strongly suggests that MMAA and, by extension, ArgK have specific involvement in vitamin B₁₂ utilization. An interesting aspect of previous biochemical studies included the finding that vitamin B₁₂ uptake in isolated rat liver mitochondria depended on mitochondrial swelling mediated by calcium or phosphate addition (11, 30). This appears to occur through opening of the mitochondrial permeability transition pore (MPTP) that is permeable to molecules of <1,500 Da, a value bordering on the molecular mass of vitamin B₁₂ (1,350 Da) (31). Acute opening of the MPTP causes mitochondrial swelling and apoptosis and can be initiated by the addition of calcium or phosphate salts (32). We speculate that vitamin B₁₂ uptake might occur via MPTP pores, opened within a physiological range, which might be stimulated by the ATP or GTPase activity of the associated MMAA protein. A search for transmembrane domains was equivocal for human MMAA, and negative for MMAA homologs in other species. If MMAA is an accessory protein, it would likely be on the matrix side of the MPTP. As a mechanism, this conjecture would obviate the need for a unique transport protein, for which there seems to be little evidence, but would offer the possibility of a requirement for calcium/phosphate stimulation and ATP dependence as originally detected (11). Further, if the system is inherently leaky so that there is a low level of transport in the absence of activation, then this might provide an explanation for the vitamin B₁₂ responsiveness in these patients (8). In conclusion, we propose a role for MMAA in the transport of vitamin B₁₂ into the mitochondrion, either as a component of a transport system or as a soluble accessory protein required for translocation across a nonspecific membrane pore, possibly the MPTP.

We acknowledge the contributions of Pierre Lepage (Genome Quebec Innovation Centre, Montreal) for aiding in sequencing, Nora Matiaszuk and Gail Dunbar (Royal Victoria Hospital, Montreal) for growing cell cultures, and David Watkins (Royal Victoria Hospital) for helpful discussions. These studies were supported by grants from the Canadian Institute for Health Research, the National Institutes of Health—National Heart and Lung Institute, and the March of Dimes. Scholarship support to C.M.D. was provided by the Canadian Institute for Health Research and the Alberta Heritage Foundation for Medical Research.

- Raux, E., Schubert, H. L. & Warren, M. J. (2000) *Cell. Mol. Life Sci.* **57**, 1880–1893.
- Rosenblatt, D. S. & Fenton, W. A. (2001) in *The Metabolic and Molecular Bases of Inherited Disease*, eds Scriver, C. R., Beaudet, A. L., Valle, D. & Sly, W. S. (McGraw-Hill, New York), pp. 3897–3933.
- Leclerc, D., Odievre, M., Wu, Q., Wilson, A., Huizenga, J. J., Rozen, R., Scherer, S. W. & Gravel, R. A. (1999) *Gene* **240**, 75–88.
- Leclerc, D., Campeau, E., Goyette, P., Adjalla, C. E., Christensen, B., Ross, M., Eydoux, P., Rosenblatt, D. S., Rozen, R. & Gravel, R. A. (1996) *Hum. Mol. Genet.* **5**, 1867–1874.
- Li, Y. N., Gulati, S., Baker, P. J., Brody, L. C., Banerjee, R. & Kruger, W. D. (1996) *Hum. Mol. Genet.* **5**, 1851–1858.
- Chen, L. H., Liu, M. L., Hwang, H. Y., Chen, L. S., Korenberg, J. & Shane, B. (1997) *J. Biol. Chem.* **272**, 3628–3634.
- Gravel, R. A., Mahoney, M. J., Ruddle, F. H. & Rosenberg, L. E. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3181–3185.
- Matsui, S. M., Mahoney, M. J. & Rosenberg, L. E. (1983) *N. Engl. J. Med.* **308**, 857–861.
- Mahoney, M. J. & Rosenberg, L. E. (1971) *J. Lab. Clin. Med.* **78**, 302–308.
- Mahoney, M. J., Hart, A. C., Steen, V. D. & Rosenberg, L. E. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2799–2803.
- Fenton, W. A. & Rosenberg, L. E. (1978) *Annu. Rev. Genet.* **12**, 223–248.
- Cooper, B. A., Rosenblatt, D. S. & Watkins, D. (1990) *Am. J. Hematol.* **34**, 115–120.
- Watkins, D., Matiaszuk, N. & Rosenblatt, D. S. (2000) *J. Med. Genet.* **37**, 510–513.
- Watanabe, F., Saido, H., Yamaji, R., Miyatake, K., Isegawa, Y., Ito, A., Yubisui, T., Rosenblatt, D. S. & Nakano, Y. (1996) *J. Nutr.* **126**, 2947–2951.
- Gordon, B. A. & Carson, R. A. (1976) *Can. Med. Assoc. J.* **115**, 233–236.
- Wilson, A., Leclerc, D., Rosenblatt, D. S. & Gravel, R. A. (1999) *Hum. Mol. Genet.* **8**, 2009–2016.
- Leclerc, D., Wilson, A., Dumas, R., Gafuik, C., Song, D., Watkins, D., Heng, H. H., Rommens, J. M., Scherer, S. W., Rosenblatt, D. S., et al. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 3059–3064.
- Celis, R. T., Leadlay, P. F., Roy, I. & Hansen, A. (1998) *J. Bacteriol.* **180**, 4828–4833.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410.
- Mellman, I. S., Youngdahl-Turner, P., Willard, H. F. & Rosenberg, L. E. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 916–920.
- Kolhouse, J. F. & Allen, R. H. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 921–925.
- Watanabe, F. & Nakano, Y. (1997) *Methods Enzymol.* **281**, 295–305.
- Watanabe, F., Nakano, Y. & Kitaoka, S. (1987) *Biochem. J.* **247**, 679–685.
- Wissenbach, U., Six, S., Bongaerts, J., Ternes, D., Steinwachs, S. & Uden, G. (1995) *Mol. Microbiol.* **17**, 675–686.
- Higgins, C. F. & Ames, G. F. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 6038–6042.
- Locher, K. P., Lee, A. T. & Rees, D. C. (2002) *Science* **296**, 1091–1098.
- Leipe, D. D., Wolf, Y. I., Koonin, E. V. & Aravind, L. (2002) *J. Mol. Biol.* **317**, 41–72.
- Haller, T., Buckel, T., Retej, J. & Gerlt, J. A. (2000) *Biochemistry* **39**, 4622–4629.
- Bobik, T. A. & Rasche, M. E. (2001) *J. Biol. Chem.* **276**, 37194–37198.
- Fenton, W. A., Ambani, L. M. & Rosenberg, L. E. (1976) *J. Biol. Chem.* **251**, 6616–6623.
- Chernyak, B. V. & Bernardi, P. (1996) *Eur. J. Biochem.* **238**, 623–630.
- Letuve, S., Druilhe, A., Grandsaigne, M., Aubier, M. & Pretolani, M. (2002) *Am. J. Respir. Cell Mol. Biol.* **26**, 565–571.