ELSEVIER

Gene 240 (1999) 75–88

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Molecular cloning, expression and physical mapping of the human methionine synthase reductase gene ☆

Daniel Leclerc [a], Marie-Hélène Odièvre [a], Qing Wu [a], Aaron Wilson [a,b], Jack J. Huizenga [c], Rima Rozen [a,b,d], Stephen W. Scherer [c], Roy A. Gravel[a,b,d] *

[a] *Medical Research Council Group in Medical Genetics, The Montreal Children's Hospital Research Institute, Montreal, PQ, Canada H3H 1P3*
[b] *Department of Biology of the McGill University Health Centre, Montreal, PQ, Canada H3H 1P3*
[c] *Department of Genetics, The Hospital for Sick Children and the University of Toronto, Toronto ON, Canada M5G 1X8*
[d] *Departments of Human Genetics and Pediatrics of the McGill University Health Centre, Montreal, PQ, Canada H3A 1B1*

Received 10 June 1999; accepted 16 September 1999
Received by A. Dugaiczyk

Abstract

Methionine synthase reductase (EC 2.1.1.135) is a flavoprotein essential for maintenance of methionine synthase in an active state. We characterized the human gene for methionine synthase reductase (*MTRR*). The gene is approximately 34 kb and comprises 15 exons, varying in size from 43 to 1213 bp, and 14 introns whose sizes vary from 108 bp to 5 kb. The positions of several junctions are conserved between the *MTRR* gene and the *C. elegans* ortholog, as well as with the rat cytochrome P450 reductase gene. A 1.3 kb CpG island encompasses the 5′-flanking region and exon 1 and extends into intron 1. A short region including the transcription start site is sufficient to confer promoter activity, with a better outcome when accompanied by intron 1. The promoter region contains putative binding sites for Sp1, AP-1, AP-2 as well as CAAT motifs, but no consensus TATA box. Primer extension analysis revealed a single major transcription start site, located 137 bp upstream of the previously reported initiator ATG. An alternative splicing event involving a portion of exon 1 predicts that translation can potentially be initiated at two different ATG codons. The gene was physically assigned to a narrow area between markers WI1755 and D5S1957.

Keywords: Chromosome 5; Cobalamin; Gene structure; Gene transcription; Reductive activation; Vitamin B12

1. Introduction

Cobalamin-dependent methionine synthase catalyses the transfer of a methyl group from methyltetrahydrofolate to homocysteine, generating tetrahydrofolate and methionine. Methylcobalamin serves as an intermediate methyl carrier. Over time, the cob(I)alamin cofactor of methionine synthase becomes oxidized to the inactive form cob(II)alamin. In bacteria, reactivation of the enzyme proceeds by a reductive methylation that involves electron transfer from the FMN prosthetic group of flavodoxin and methyl transfer from *S*-adenosylmethionine. The reactivation process involves formation of a transient cob(I)alamin intermediate that is trapped by the methyl transfer from AdoMet (Jarrett et al., 1998). Another protein, flavodoxin reductase, containing a FAD prosthetic group, is necessary for regenerating the reduced flavodoxin in a NADPH-dependent reaction (Bianchi et al., 1993).

In humans, the reductive activation of methionine synthase was first recognized through the discovery of patients with a reduced level of methionine synthase activity, which could be fully restored in extract assays in the presence of a strong reducing agent. They proved to have a defect in a reducing system required for maintenance of the enzyme in a functional state (Rosenblatt et al., 1984). These patients are genetically distinct from those possessing primary methionine syn-

Abbreviations: cDNA, DNA complementary to RNA; PCR, polymerase chain reaction; UTR, untranslated region.

☆ The nucleotide sequences reported in this paper have been submitted to GenBank with Accession Nos. AF121202 to AF121214.

* Corresponding author. Present address: Departments of Cell Biology & Anatomy and Biochemistry & Molecular Biology, University of Calgary, Heritage Medical Research Building, Room 261, 3330 Hospital Drive NW, Calgary, Canada T2N 1N4.
Tel.: +1-403-220-3063; fax: +1-403-270-2772.
E-mail address: rgravel@ucalgary.ca (R.A. Gravel)

0378-1119/99/$ - see front matter
PII: S0378-1119(99)00431-X

thase deficiency and were assigned to the *cblE* complementation group of folate/cobalamin metabolism. Our desire to understand this inherited disease provided the impetus to clone the affected gene. Biochemical studies (Gulati et al., 1997) showed that *cblE* disease results from a defect in a NADPH-linked reducing system analogous to the flavodoxin/flavodoxin reductase required for maintaining methionine synthase activity in *E. coli*. However, because flavodoxins are not present in mammals, and given the occurrence of a single complementation group accounting for defects of the reductive–activation system, we hypothesized that the human counterpart would be a single protein containing the catalytic sites of both flavodoxin and flavodoxin reductase, as found, for example, in cytochrome P450 reductase (CPR). Using consensus sequences for binding sites for FMN, FAD and NADPH, we cloned a human cDNA corresponding to the 'methionine synthase reductase' reducing system required for maintenance of the methionine synthase in a functional state (Leclerc et al., 1998). To date, the sequences of only two eukaryotic methionine synthase reductases have been reported (human and *C. elegans*; Leclerc et al., 1998). The sequence of the mouse ortholog is very similar to the human counterpart (Wilson et al., 1999). The human gene, symbolized *MTRR*, was localized to 5p15.2–15.3 by fluorescence in-situ hybridization, consistent with results from somatic cell hybrid analysis and radiation hybrid analysis (Leclerc et al., 1998).

MTRR can be classified as a housekeeping gene, as it is expressed at relatively low levels in all cell types analysed (Leclerc et al., 1998). A predominant mRNA of about 3.6 kb was detected by Northern blot analysis. The cDNA-deduced protein, a novel member of the FNR family of electron transferases, contains 698 amino acids with a predicted molecular mass of 77 700. The authenticity of the cDNA sequence was confirmed by identification of mutations in *cblE* patients. RT-PCR-dependent heteroduplex analysis detected a 4 bp frameshift in two affected sibs and a 3 bp deletion in a third patient, deleting a highly conserved leucine residue (Leclerc et al., 1998). Additional mutation analysis in *cblE* disease revealed splice defects or other drastic changes yielding a truncated protein (unpublished). Enzyme polymorphisms have also been identified in the general population (Wilson et al., 1999). As discovered for methylenetetrahydrofolate reductase (Rozen, 1996), a mild deficiency of methionine synthase reductase could be associated with hyperhomocysteinaemia, a risk factor in cardiovascular disease and neural tube defects.

To facilitate the discovery and characterization of mutations and polymorphisms as well as to help understand the potential roles played by methionine synthase reductase, we report the molecular cloning and characterization of the gene for human methionine synthase reductase.

2. Materials and methods

2.1. Reagents

Restriction enzymes were from New-England Biolabs, Boehringer Mannheim and Gibco BRL. Taq DNA polymerase, calf intestinal phosphatase, T4 DNA ligase, Minimal Essential Medium (MEM), Optimem I, antibiotics-antimycotics and trypsin-EDTA were from Gibco-BRL. Magnum columns were from Genome Systems. The Wizard minipreps kit, butyryl-CoA and chloramphenicol acetyl transferase were from Promega. Qiafilter plasmid Midipreps, Maxipreps and Oligotex Direct mRNA Mini Kit were from Qiagen. Geneclean II and III kits were from Bio 101. The Hybond N+ nylon membrane was from Amersham. Oligonucleotide primers were synthesized by ACGT Corporation (Toronto) and are described in Table 1. Sequenase kits were from United States Biochemicals. Foetal bovine serum and iron-enriched serum were from Intergen. The electroporation apparatus was the Gene Pulser II from Biorad. 2,6,10,14 tetramethylpentadecane (TMPD) was from Sigma. Xylenes were from Fisher Scientific. Random primed DNA labeling kit and *O*-nitrophenyl-β-D-galactopyranoside were from Boehringer Mannheim. Scintillation liquid Betamax and α[^{32}P]dATP were from ICN Biomedicals. α[^{35}S]dATP and [^{3}H]chloramphenicol were from NEN. TA cloning kit (including pCR2.1) was from Invitrogen. Bluescript II SK− phagemid vector and Taq Extender were from Stratagene. pBLCAT3 and pBLCAT2 were kindly provided by G. Schütz (Luckow and Schütz, 1987), and pCMVβ was from C.T. Caskey (MacGregor and Caskey, 1989).

2.2. Screening of human genomic libraries

A P1-derived artificial chromosome (PAC) clone (104K2) was identified from a total human genomic library by hybridization screening with a human methionine synthase reductase cDNA probe containing the last 2 kb of the cDNA (clone 704947; Leclerc et al., 1998). PAC DNA was extracted using Magnum kb-100 columns, and digested with various restriction enzymes. Exon-containing fragments were identified by Southern blotting and hybridization with a 5′ end cDNA probe (RT-PCR product generated with primers AA468/Z597; Leclerc et al., 1998). YAC and BAC human genomic libraries were also screened using the same 5′ cDNA probe, and positive clones were used to perform physical mapping of the gene.

2.3. PCR, cloning, DNA sequencing and sequence analysis

Several putative intron–exon junctions were predictable by the occurrence of alternative splicing, splicing

Table 1
Oligonucleotides used in this study

Primers	Sequence	Orientation[a], location
SP6	5′-CGTCGACATTTAGGTGACACTATAGAG	S, PAC vector
AK47	5′-GTGACAAATTATGTTATCGCACATTG	A, intron 12
AK48	5′-CTGATGCCACTATTTAGTGATGTT	S, intron 11
AK50	5′-GATGCGCTGTAAAGTACTACAGTAAT	S, intron 7
AK53	5′-GTTACGATGATAGCACACACCCTT	A, intron 8
AL93	5′-TGTGTTCAGATGGAGAACTTAGGC	S, intron 4
AL94	5′-GGGTCAGGAATAAAATGGCAAGATG	A, intron 6
AL95	5′-GAAGACGGAGGGAAGAACTACGG	S, intron 14
AG124	5′-CACGCCGCGCCAGCCGGGCACGAACC	A, exon 1
AG126	5′-CAACCAATCCAATGTTGTAATTGAAGA	S, exon 5
AG127	5′-TCTTCAATTACAACATTGGATTGGTTG	A, exon 5
AD151	5′-CGAACAACAAATTCTTTCCACTTACC	S, exon 12
AD153	5′-GGTAAGTGGAAAGAATTTGTTGTTCG	A, exon 12
AH155	5′-AGGCTTTTGATCCTTTTGAGAAATAAA	S, exon 15
AH156	5′-ACTGAGTGGTGGCTGGCAAGAAG	A, exon 9
AH157	5′-CTCCAAGAACAACACCCAGATGG	S, exon 13
AJ283	5′-TCTCCCGACACCGAGGGGAAAC	A, upstream of *MTRR*
AD291	5′-TACATGGATTTTCTGCAGATCTTCAC	S, exon 2
AF331	5′-AGCTGTTTCGGTTTTTAGGTCATAC	A, exon 3
AF332	5′-TTTGCTCACCTGCGGTATGGGTT	S, exon 3
AF333	5′-AACAAGGCCAGCCAGCCTGTACA	A, exon 11
AF335	5′-CTTCTTGCCAGCCACCACTCAGT	S, exon 9
AF336	5′-CCACAACAGAGGTTCTGCGGAAG	S, exon 11
AI339	5′-AAGATTATGCCATTTCTTTCAGATCTTT	A, exon 15
AI340	5′-CCCAGACCATATTCGTGTGCAA	S, exon 10
AG361	5′-ATGCCATAAAAACCACTCTGCTGGTA	S, exon 6
AG362	5′-TACCAGCAGAGTGGTTTTTATGGCAT	A, exon 6
AG363	5′-CAGACTTTTCCTATCAGCCTGGAG	S, exon 7
AG364	5′-CTCCAGGCTGATAGGAAAAGTCTG	A, exon 7
AG365	5′-TTCATTTTTACCTGGTGTCTTGAAATC	S, exon 8
AG366	5′-GATTTCAAGACACCAGGTAAAAATGAA	A, exon 8
AI419	5′-GGATCCACATGCAACAGAATCAGA	S, upstream of *MTRR*
AI420	5′-CTAGTTTAGTGACTGCTTTGCAAGC	A, upstream of *MTRR*
AK480	5′-GCCGGGCACGAACCTGCACAACGCGACTTC	A, exon 1
AK485	5′-ACTGAATGTTCAAATGCAGAGATGG	S, intron 12
AK486	5′-ATCTAGATTATTGGCTTCTCGGTCC	A, intron 14
AL526	5′-CCAGTAAGCCATAGTTTGTTGACC	A, intron 5
AL527	5′-CCCCTGTGGATCTTGCGTAGTC	S, intron 5
AB587	5′-AACATCCAGCTTCATGGCCAGCAG	S, exon 14
AB588	5′-TTGCACAAGGGCATCATGTACATC	A, exon 15
Z594	5′-CTTGCACACGAATATGGTCTGGG	A, exon 10
1103C	5′-GGTGTATTCTGAATCACCGAGACC	A, exon 4
1902C	5′-ATAAACGACTTCAAGAGCTTGGAGC	S, exon 4
2101C	5′-AGGTTTGGCACTAGTAAAGCTGACT	A, exon 15

[a] S, sense; A, antisense.

abnormalities in some *cblE* patients and comparison with the *C. elegans MTRR* gene structure (GenBank Accession No Z35595). Relevant oligonucleotide primers were designed from the nucleotide sequence of a human *MTRR* cDNA (Leclerc et al., 1998) and, as the work was progressing, from appropriate intron or promoter sequences. PCR reactions, subcloning, and manual sequencing were performed as described previously (Leclerc et al., 1998). Automated sequencing was performed by Bio S&T (Montreal). PCR products were subcloned and sequenced at both ends to confirm their specificity. Genomic *MTRR* exons were all sequenced, and exon–intron boundaries were identified. The subclones encompass the whole of the *MTRR* gene and were designed to overlap each other. The sequence downstream of the polyadenylate addition sites was derived from a semi-specific PCR product obtained using primer AH155. The nucleotide sequence of 2.5 kb of the 5′-flanking region, the 15 exons, at least 80 bp of every intron at intron–exon junctions and 330 bp of the region downstream of the 3′UTR were determined. Partial sequences upstream of the *MTRR* transcription unit came from a PCR product involving primer SP6 (in the PAC vector) and antisense primer AI420. For

segments obtained by PCR, at least two independent subcloned PCR products were sequenced. Direct sequencing of PCR products was also performed for introns 2 and 3. The sizes of all but three introns (1, 7 and 13) were determined by long-distance PCR using oligonucleotide primers derived from the flanking regions. Introns 1, 7 and 13 were sized by DNA sequence data. Nucleic and deduced protein sequences from the 5′ region of the gene were analysed with the PBASE and TRANSPEP programs (PCGENE package, Intelligenetics), respectively. A sequence analysis for prediction of promoter elements was performed using the TESS interactive package (Schug and Overton, 1997) and the Transfac database (Heinemeyer et al., 1998). CpG islands were located as defined by Gardiner-Garden and Frommer (1987). Sequence comparisons were made between *MTRR* genomic sequence and sequences in the National Center for Biotechnology Information (NCBI) databases.

2.4. Preparation of promoter-cat reporter plasmids

A series of constructs were generated from sequences representing the expected *MTRR* promoter region and directionally cloned 5′ to the chloramphenicol acetyltransferase (*cat*) gene in the vector pBLCAT3 (Luckow and Schütz, 1987). For all *cat* constructs, the presence of upstream nonsense codons ensure that the methionine codon of the *cat* gene is the one suitable for initiation of translation.

The promoter-*cat* constructs, all derived from a 3.3 kb *Hin*dIII–*Pst*I fragment at the 5′ end of the *MTRR* gene were constructed as follows. A parent clone, pBamHI8b, was constructed by inserting a 8 kb *Bam*HI fragment of 104K2 PAC clone into the *Bam*HI site of pBluescript SK−. A derivative construct, pPstI6d, was obtained after *Pst*I digestion of pBamHI8b and subsequent religation of the 6 kb-remaining fragment. This clone contains the putative *MTRR* promoter, exon 1, the first intron and a part of exon 2. pPstI6d was used to create several reporter constructs, as follows. The 3.3 kb *Hin*dIII–*Pst*I fragment of pPstI6d was inserted into the *Hin*dIII–*Pst*I sites of pBLCAT3. The resulting construct was named pH-P. The 1.2 kb *Hin*dIII–*Avr*II fragment of pPstI6d was inserted into the *Hin*dIII–*Xba*I sites of pBLCAT3, and this construct was named pH-A. An intermediate, pCR2.1-HB, was constructed by inserting the 1.6 kb *Hin*dIII–*Bgl*I fragment of pPstI6d into pCR2.1 after modifying the ends with DNA polymerase. pCR2.1-HB was used to create two reporter constructs. The 1.7 kb *Bam*HI–*Xho*I fragment of pCR2.1-HB was inserted into the *Bam*HI–*Xho*I sites of pBLCAT3, resulting in a pH-B construct. The 0.4 kb *Avr*II–*Xho*I fragment of pCR2.1-HB was inserted into the *Xba*I–*Xho*I sites of pBLCAT3; this construct was named pA-B. The pH-P construct was cut with *Hin*dIII and *Afl*II. The ends were filled with DNA polymerase and religated, creating the pAfl-P construct. The pAfl-P construct was cut with *Bss*HII and *Xho*I. The ends were filled with DNA polymerase and religated. This construct was named pAfl-Bss. The pH-P construct was cut with *Hin*dIII and *Bsa*AI; the ends were filled with DNA polymerase and religated, creating the pBsa-P construct. In order to generate a construct lacking intron 1, we first performed a RT-PCR using two primers located in exons 1 and 4 (AL528/1103C). The PCR product was cut with restriction enzymes *Bss*HII (cuts in exon 1) and *Pst*I (cuts in exon 2) and was introduced into the *Bss*HII–*Pst*I site of pH-P previously digested with the same enzymes. This construct, lacking intron 1, was named pdintron. For all constructs, junctions were confirmed by sequencing. The plasmids were extracted by Qiafilter Maxiprep prior to transfection. Control plasmids included pBLCAT2 (Luckow and Schütz, 1987), containing the *herpes simplex* virus thymidine kinase promoter to drive the *cat* gene, as a positive control, and pBLCAT3 (Luckow and Schütz, 1987), containing no promoter, as a negative control. pCMVβ (MacGregor and Caskey, 1989), containing the *E. coli* β-galactosidase gene driven by a human cytomegalovirus promoter/enhancer, was used in cotransfections to assess the transfection efficiency.

2.5. Cell culture, DNA transfection, and enzyme assays

Human myotubes were prepared as described previously (Shoubridge et al., 1996). COS-7 cells were grown to sub-confluence in MEM supplemented with 10% serum (5% foetal bovine serum and 5% iron enriched calf serum) and antibiotics-antimycotics. DNA transfection was performed by electroporation using 10 μg of promoter-*cat* fusion plasmid and 10 μg of pCMVβ control plasmid for monitoring the transfection efficiency. In addition, pBLCAT3 and pBLCAT2 were transfected in each set of experiments as a negative and positive control, respectively. Each transfection was carried out at least three times. Eight-hundred microliters of COS-cell suspension (20×10^6 cells/ml in OPTIMEM I supplemented with 5% foetal calf serum) were transferred into an electroporation cuvette (4 mm gap) containing plasmids. After 15 min incubation on ice, the cuvette was pulsed at 950 μF, 280 V. At 72 h post-transfection, the cells were scraped and the pellets were assayed for CAT activity (Kingston and Sheen, 1998) and β-galactosidase activity (An et al., 1982). The cell extracts were prepared by the 'freeze–thaw' method (Kingston and Sheen, 1998) and the cytoplasmic extract was heated at 65°C for 10 min (for CAT assays only). CAT activity was determined by phase-extraction assay using [^{3}H]chloramphenicol as substrate (Kingston and Sheen, 1998). The butyrylated products were separated from the unbutyrylated products by organic extraction

in 2:1 (v/v) TMPD/xylenes and quantified by liquid scintillation counting. CAT activity was normalized to β-galactosidase activity to account for variations in transfection efficiency.

2.6. Primer extension

Transcription initiation sites were determined by primer extension analysis following standard procedures and confirmed by RT-PCR. Poly(A)$^+$ RNA was isolated from cultured human fibroblasts or from human myotubes using the Oligotex Direct mRNA Mini Kit method. In primer extension, antisense oligonucleotides AG124 or AK480 were used as gene-specific oligonucleotides. The extension products were loaded onto a polyacrylamide sequencing gel beside sequencing reactions of the 5′ region of the *MTRR* gene, accomplished with the same primers. Sequencing reactions were performed using dGTP or dITP, to facilitate sequencing of GC-rich regions.

3. Results

3.1. Isolation and structural organization of the human MTRR gene

An initial survey of the NCBI gss database showed that BAC3095C7 (GenBank Accession No AQ139528) contains a portion of *MTRR* intron 1, exon 2 and downstream sequence. Screening of human YAC, BAC and PAC genomic DNA libraries with probes in the *MTRR* coding sequence additionally yielded three YAC, four BAC and one PAC clone containing (PCR experiments, data not shown) the extreme 5′ and 3′ ends of the published cDNA (Leclerc et al., 1998). Since this suggested that they all contained the entire transcription unit, we made use of the smallest clone, PAC104K2, for primary identification of all junctions and for subcloning a segment encompassing the *MTRR* promoter. Purified DNA from 104K2 was digested with *Bam*HI, and an 8 kb segment, which was shown to be positive by Southern blotting with a probe corresponding to the 5′ region of the cDNA, was subcloned into the pBluescritpII SK− vector to generate pBamHI8b. A 3300 bp *Hin*dIII–*Pst*I portion of the insert was completely sequenced (pPstI6d) and was subsequently shown to contain the promoter, exon 1, intron 1, as well as a part of exon 2.

A 54 bp sequence representing the most 5′ untranslated region of the cDNA sequence previously reported (Leclerc et al., 1998) was contiguously identified as the 3′ end of exon 1. Therefore, the previously reported initiating methionine codon (Leclerc et al., 1998) is in exon 2. The nucleotide sequences of all exons and their boundaries were determined (GenBank Accession Nos AF121202 to AF121213; Table 2). All exon–intron boundaries conserve the AG–GT rule (Table 2) and maintain the open reading frame as predicted by the cDNA sequence (Leclerc et al., 1998). In addition, other consensus elements of splice junctions occur with a high frequency. For example, 13 of the 14 exons show 5′ splice sites that end with G, and eight of them end with AG. The exons range in size from 43 bp (exon 10) to 1213 bp (exon 15). The introns range in size from 108 bp (intron 13) to 5 kb (intron 5). The human *MTRR* gene was found to encompass approximately 34 kb (Fig. 1), which is of the average size for vertebrate protein coding genes. A member of the AluSb2 subfamily was identified in intron 10 (GenBank Accession No. AF121209). Exon 15 contains 145 bp encoding the C-terminal part of the enzyme as well as 1068 bp of the 3′UTR region predicted from the cDNA sequence, encompassing the putative polyadenylation signal. The sequences downstream of the polyadenylate addition sites, which were previously identified in multiple clones (Leclerc et al., 1998), are GT- and T-rich (GenBank Accession No. AF121213). This is a general, though not universal, property of 3′ ends of RNA polymerase II-dependent genes (Proudfoot and Whitelaw, 1988).

We showed previously that the protein sequence of human MTRR is similar to its *C. elegans* ortholog, as well as to CPR (Leclerc et al., 1998). When the positions and types of the introns in the human *MTRR* gene were compared to those of the *C. elegans* orthologous gene (The *C. elegans* sequencing consortium, 1998) and to the rat CPR gene (Porter et al., 1990), it was revealed that seven and four junctions were conserved, respectively (Fig. 2). The MTRR protein also has a good, albeit lower, global homology with the three isologs of nitric oxide synthase (NOS) (Leclerc et al., 1998), another member of the FNR family of electron transferases. The genomic structures of the human NOS II (Chartrain et al., 1994) and human NOS III (Marsden et al., 1993) are both very similar and show 12/13 conserved junctions in the FMN-FAD-NADPH domains, but only one of these junctions (intron 19 in NOS genes) is conserved in the human *MTRR* gene (intron 7, data not shown).

3.2. Minisatellite sequence in intron 6

We identified tandem repeats of a 29 bp consensus core sequence (minisatellite) close to the 5′ splice site in intron 6 (Fig. 3). No sequence identical to this minisatellite was found by surveying GenBank. We analyzed a panel of DNA samples representing 98 alleles using primers AG361–AL94 to generate PCR products containing the minisatellite repeat in intron 6. Three allele types were observed, with PCR products of 292, 408 and 437 bp (data not shown). Representative products

Table 2
Exon–intron structure of the human *MTRR* gene

Exon	Size (bp)	Range of amino acids[a]	Intron	Approximate size (bp)	5′ junction/cDNA/3′ junction[b]
1	112	Untranslated			Promoter/GGAGCTGCGTGG/gtaagc
			1	1552	
2	154	MET[1]–LYS[43]			tttcag/TTTCACGATAAG/gttaga
			2	2500	
3	154	TYR[44]–LEU[94]			ttgtag/TATGACTACTGG/gtaatg
			3	1800	
4	118	GLY[95]–GLY[134]			ctctag/GTCTCGTGTAGG/gtaagg
			4	3000	
5	379	LEU[135]–GLN[260]			gtccag/TTTAGAGGCCAG/gtaagg
			5	5000	
6	123	GLU[261]–SER[301]			ttcaag/GAGGAAATTTCA/gtaagt
			6	2100	
7	154	ASN[302]–LYS[352]			ctctag/AATACAAGAAAG/gtaaca
			7	760	
8	89	GLY[353]–LYS[382]			gaaaag/GAGCTAAAAAAG/gtattt
			8	2500	
9	181	ALA[383]–LEU[442]			ttgtag/GCATTTTGCTCG/gtgagt
			9	2300	
10	43	GLU[443]–SER[457]			ttctag/AACATCTGCAAG/gtacta
			10	1300	
11	187	SER[458]–LYS[519]			ttgcag/CTCAAGCCTAAG/gtaaga
			11	3000	
12	119	ILE[520]–ARG[559]			tttcag/ATATCCACATAG/gtatgt
			12	1200	
13	93	GLU[560]–ARG[590]			acttag/AGAGAAATTCAG/gtattg
			13	108	
14	183	GLU[591]–GLY[651]			tttcag/AAAAGAGTGTGG/gtgagt
			14	3000	
15	1213	ASP[652]–OCH[699]			ttctag/AGATGCAATCTT[c]

[a] Considering the initiator ATG previously described in Leclerc et al. (1998).
[b] Exon sequences are in upper-case letters, and intron sequences are in lower-case letters.
[c] These bases are followed by polyadenylate in mRNA.

were subcloned and sequenced and shown to contain five, nine or 10 repeats. The product containing five repeats is made up of the first five repeat sequences, whereas that containing nine repeats lacks the eighth unit. The observed frequencies of the five, nine and 10 repeat units are 2, 79 and 19%, respectively. In *cblE* patients (11 individuals), analysis of the intron 6 sequence revealed the presence of the nine and/or 10 repeat units in different individuals (data not shown).

3.3. Sequence of the human MTRR 5′ region

The nucleotide sequence of 1.6 kb of the 5′-flanking region, exon 1 and a part of intron 1 are shown in Fig. 4. The proximal 546 bp of the 5′-flanking region show a high G+C content (50–75%) and constitute a CpG island that extends into the gene (Fig. 5). The 5′ flanking region of the gene lacks a TATA box, but four Sp1 boxes were found at positions −377, −210, −53 and –34, where the search for Sp1 sites was restricted to the conservatively defined consensus GC box, GGGCGG, although the TESS software (Schug and Overton, 1997) detects many other sites in the promoter region that could potentially bind Sp1. Two inverted CAAT boxes are present at positions −19 and 11; four AP-2 sites [consensus sequence CCC(A/C)N-(G/C)(G/C)(G/C)] are at positions −486, −302, −239 and −168; one AP-1 site (TGATTGG) is found at position −21; there are also several other potential binding sites for regulatory elements, including eight potential glucocorticoid response elements (GRE).

3.4. Identification of the MTRR transcription initiation sites

To better define the 5′ boundary of the transcription unit, primer extension studies were performed (Fig. 6) with ^{32}P-end-labelled primers (AG124 and AK480) from the 5′untranslated region of *MTRR* using control tRNA (lane B) or mRNA from normal fibroblasts (lane F) or from human myotubes (lane M). The same primers were used to generate dideoxy sequencing ladders (lanes GATC), which were co-electrophoresed with the primer extension reactions.

A family of six extension products was identified, terminating with the local sequence GGAGCTT (Fig. 6,

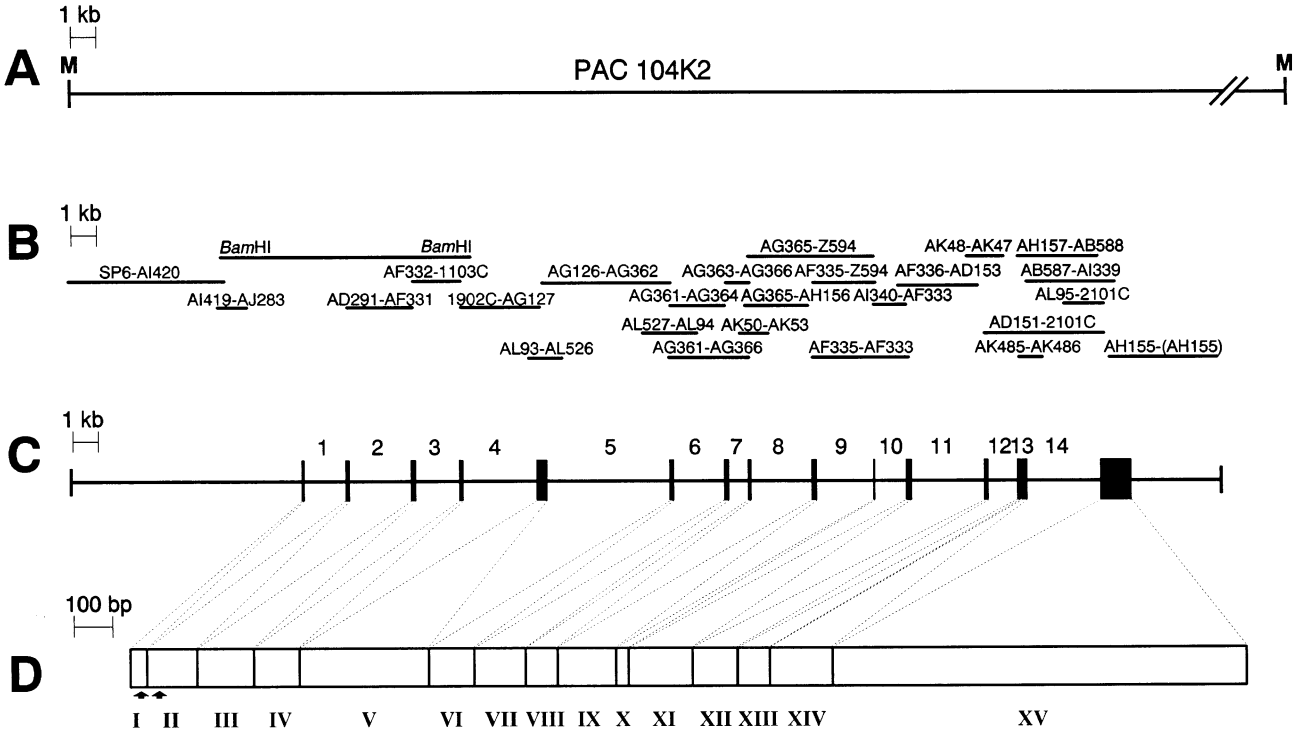

Fig. 1. Structural organization of the human methionine synthase reductase gene. (**A**) Genomic clone PAC104K2 containing the entire gene. The letter M shows the *Mbo*I site present at the insertion site in the vector. (**B**) Segments generated by PCR and 8 kb *Bam*HI fragment derived from clone PAC104K2. The PCR fragments and the *Bam*HI segment are represented as lines. Oligonucleotides are indicated above lines and are described in Table 1. The primer in parentheses designates a mispriming outcome that generated a valid internal sequence. (**C**) In the gene structure, exons are indicated by solid boxes. Introns are numbered in arabic numerals. (**D**) Methionine synthase reductase mRNA; exons are numbered in roman numerals. The arrowheads under exons 1 and 2 show the two potential initiation codons predicted from alternative splicing.

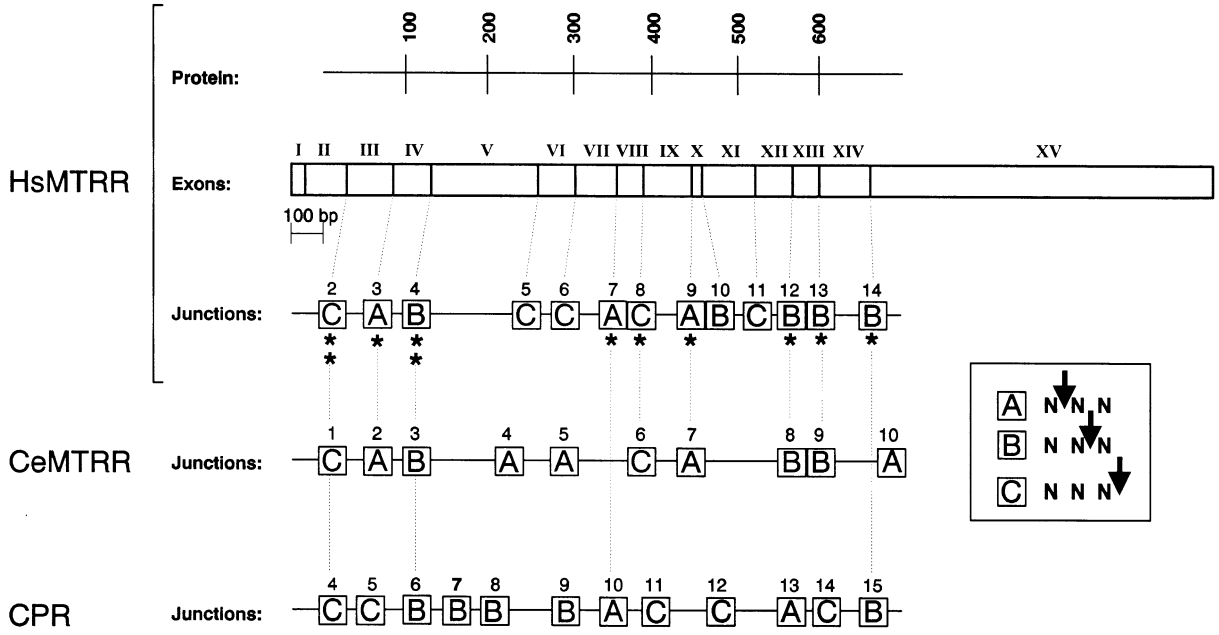

Fig. 2. Homology between *MTRR* gene and other related genes. The human *MTRR* gene (HsMTRR) structure is compared to the *C. elegans* ortholog gene (CeMTRR) and rat cytochrome P450 reductase gene (CPR). Junctions are represented by square boxes identified as A, B or C. Type A introns are interrupted between two codons. Type B introns interrupt a codon between the first and second nucleotides, and type C introns interrupt a codon between the second and third nucleotides. Numbers above junctions identify the relevant introns. The presence of one or two stars under the junctions in the human *MTRR* gene indicates that a similar junction occurs, at an equivalent amino acid position, in one or both of the other genes, respectively.

```
                    ----------AG361---------->
  Exon 6    ...  ATGCCATAAAAACCACTCTGCTGGTAGAATTGGACATTTCA//

Intron 6     1   GTAAGTTGCAAAATTTATTTCTCAGCA

Intron 6    28   CAGTAATATTGTCAGGATGTGCAGAAA

Intron 6    55   GAGTTGTGTGGTGCTTGGTTATATATGCT  (repeat # 1)
Intron 6    84   GAGTGGTGTGCTGCTTGGTTACATATGCT  (repeat # 2)
Intron 6   113   GAGTTGTGTGGTGCTTGGTTACATATGCT  (repeat # 3)
Intron 6   142   GAGTTGTGTGGTGCTTGGTTATGTATGCT  (repeat # 4)
Intron 6   171   GAGTTGTGTGGTGCTTGGTTATACATGCT  (repeat # 5)
Intron 6   200   GAGTTGTGTGGTGCTTGGTTACATATGCT  (repeat # 6)
Intron 6   229   GAGTTGTGTGGTGCTTGGTTACATATGCT  (repeat # 7)
Intron 6   258   GAGTTGTGTGGTGCTTGGTTACATATGCT  (repeat # 8)
Intron 6   287   GAGTTGTGTGGTGCTTGGTTATGTATGCT  (repeat # 9)
Intron 6   316   GAGTTGTGTGGTGCTTGGTTATACATGCT  (repeat #10)
                  **** ***** **********   *****

Intron 6   345   CAGAAGAATGAGAGGTCTCTGACAGGACATCTTGCCATTTTATTCCTGACCC
                                          <----------AL94----------
```

Fig. 3. Alignment of tandem repeat sequences (minisatellite) in intron 6. The numbers indicate the positions of nucleotides at the 5′ ends of sequences on individual lines, numbered from the beginning of intron 6. The consensus sequence is described by stars below the repeat sequences.

Fig. 4. Nucleotide sequence of 1.6 kb of the 5′-flanking region, exon 1 and a part of intron 1 of the human *MTRR* gene. Potential control elements are identified below the sequence and are underlined. Restriction enzyme sites used for the promoter-*cat* reporter plasmid constructs are double-underlined. Transcription initiation sites detected by primer extension are indicated by asterisks. The uppermost nucleotide residue detected by primer extension is numbered +1, and nucleotides preceding it are indicated by negative numbers. Exon 1 is indicated by upper-case letters, and the first potential initiating methionine codon is identified (MET). The 26 bases of exon 1 involved in alternative splicing are represented by upper-case bold letters.

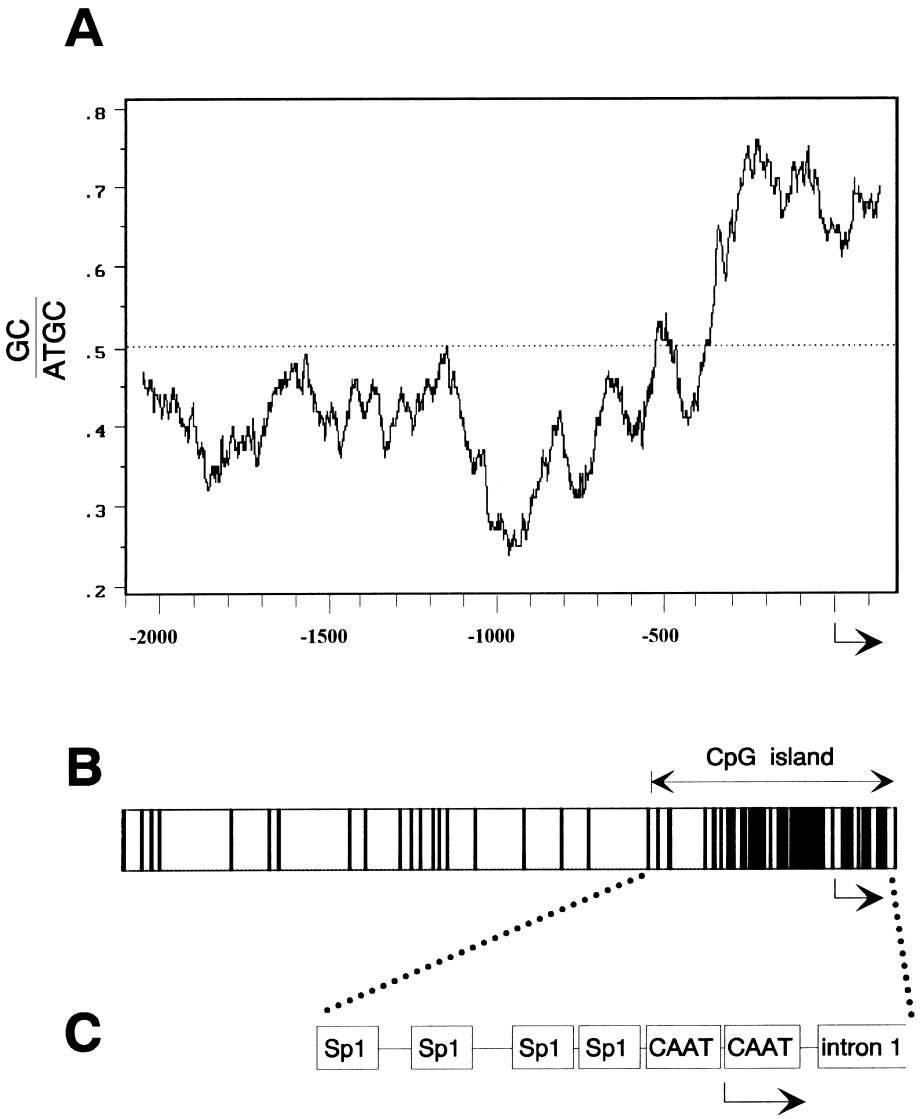

Fig. 5. Properties of the *MTRR* promoter area. (**A**) Proportion of G–C pairs in an interval of 100 nucleotides around each nucleotide of the sequence surrounding the promoter area. Data were generated using the program PBASE from the PCGENE package (Intelligenetics). (**B**) Density of CpG doublets. Each vertical line represents a CpG doublet. (**C**) Potential Sp1 and CAAT elements in the promoter region. The right-angle arrow shows the position of the transcription initiation area.

arrows), with the second G corresponding to the principal start site. The first G in the sequence was designated the upstream transcription start site and assigned the identifier +1. This extends the 5′UTR to 137 bp upstream of the previously reported initiation codon (Leclerc et al., 1998). The transcription initiation sites were confirmed by observing the presence or absence of products in RT-PCR reactions involving antisense primers in exon 2, 3 or 4 paired with sense primers that were positioned downstream or upstream of the identified transcription initiation sites (data not shown). An RT-PCR product bridging the main transcription start site and exon 4 (primers AL528/1103C) was subcloned and sequenced to confirm the identity of the product and absence of intronic sequences.

3.5. Activity of promoter deletion constructs

To identify regions important for promoter activity, a series of segments from the 5′ end of the gene was cloned into the promoterless *cat* reporter vector pBLCAT3, and the resulting promoter-*cat* constructs were transiently transfected into COS cells (Fig. 7). Three plasmid constructs with sequential 5′-deletions and a fixed 3′-end (restriction site *Pst*I), named pH-P, pAfl-P and pBsa-P were analysed. The highest CAT activity (14-fold the pBLCAT2 positive control) was observed for pH-P, which contains the longest promoter segment. pAfl-P and pBsa-P gave intermediate activities (6.8- and eightfold the pBLCAT2 activity). This decrease in activity suggests the presence of positive elements in

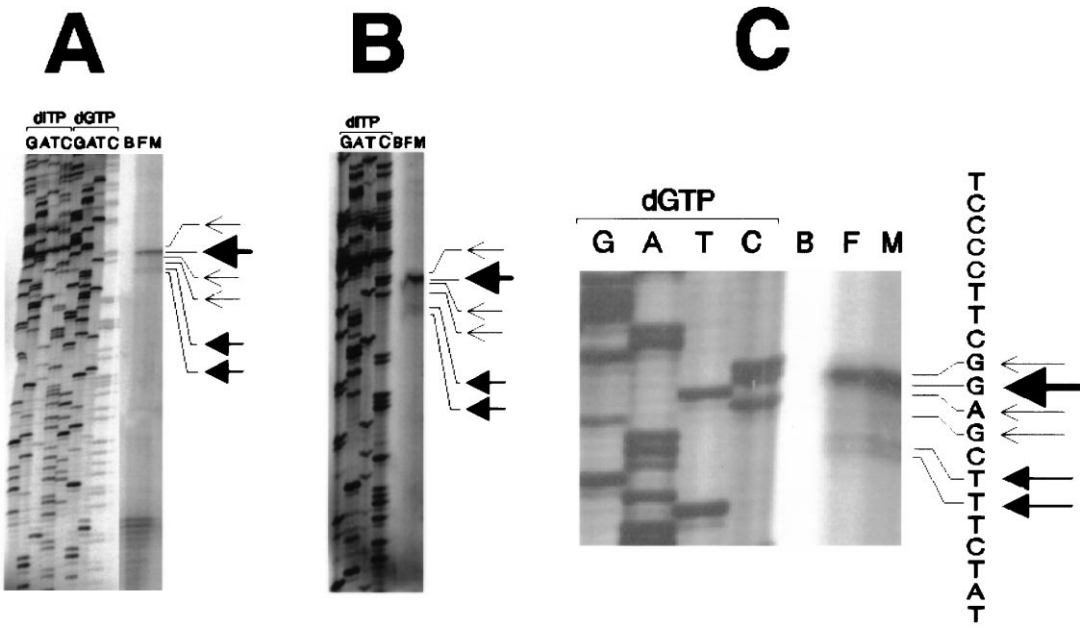

Fig. 6. Primer extension analysis. (**A**) Primer extension was performed with mRNA from human fibroblasts (5 μg, lane F) or myotubes (2 μg, lane M), or 5 μg of tRNA (lane B) using antisense primer AG124, specific for *MTRR* mRNA. Extension products corresponding to major and minor transcription initiation sites are indicated by large and small arrows, respectively. The cDNAs produced were electrophoresed along with a sequencing ladder (GATC) representing the sequence of the 5′ region of *MTRR* determined with the same primer as that used for primer extension. (**B**) Primer extension was performed as in (**A**), with primer AK480. (**C**) Enlarged view of the bands obtained in an independent experiment performed with primer AG124.

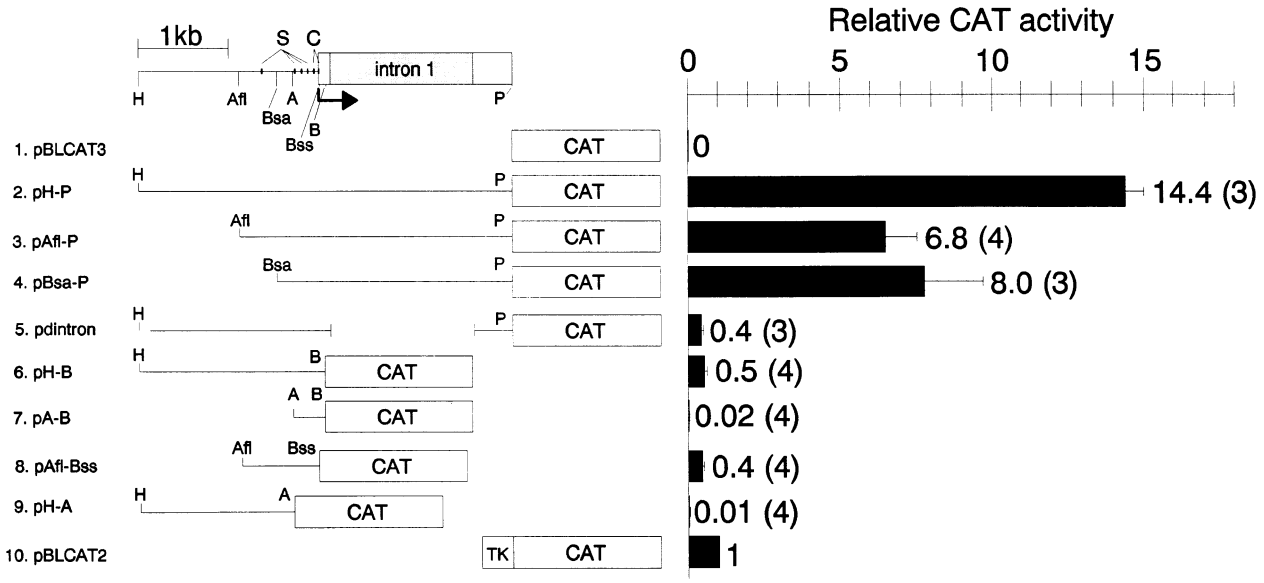

Fig. 7. Human *MTRR* promoter activity in COS cells. Comparison of the various promoter-*cat* fusion constructs, the TK promoter-*cat* fusion construct as positive control and the promoterless vector as negative control. A linear representation of the *MTRR* 5′ region with relevant restriction sites is shown at the top (H, *Hind*III; AfI, *Afl*II; Bsa, *Bsa*AI; A, *Avr*II; Bss, *Bss*HII; B, *Bgl*I; P, *Pst*I). Intron 1 is shown as a grey box, flanked by exons 1 and 2, represented as white boxes. Letter S and C represent potential Sp1 and CCAAT elements. The *MTRR* promoter sequences fused to the *cat* gene are represented by lines flanked by letters corresponding to the identity of the restriction enzymes used for plasmids construction. The *cat* gene and *TK* promoter are shown in boxes. Each construct was cotransfected with pCMVβ into COS cells, and CAT activities were measured 72 h later. The transfection efficiency for each *cat* construct was normalized to the β-galactosidase activity and compared to those obtained for pBLCAT2. The data represent the mean ± SEM from several (number indicated) independent transfections performed with each construct.

the segment between sites *Hin*dIII and *Afl*II. An additional construct, pdintron, with the same 5′ and 3′ ends as pH-P but lacking intron 1, gave a detectable activity, but it was at a far lower level, corresponding to about half the activity obtained with pBLCAT2 and 30-fold below the activity obtained with pH-P. This result

underscores the critical presence of intron 1 for obtaining such a high level of expression. Two other plasmid constructs with a distinct 3′-end (restriction site *Bgl*I), named pH-B and pA-B, were evaluated. The pH-B construct showed positive CAT activity, but it was at about half the pBLCAT2 control and similar to the activity exhibited by pdintron. This indicates the presence of positive elements between the *Hin*dIII and *Bgl*I sites. However, there was near background activity with either the pH-A or pA-B constructs, indicating that disruption of the sequence at the *Avr*II site interrupts a segment necessary for promoter function. The intronless constructs pH-B, pdintron and pAfl-Bss all gave a significant, but relatively low, level of CAT activity, suggesting that their sequence in common, pAfl-Bss, contains positive promoter elements. Taken together, these results suggest that three regions contribute to expression activity. Region −1503 to −630 (*Hin*dIII to *Afl*II) stimulated activity in the intron-containing constructs. Region −629 to 52 (*Afl*II to *Bss*HII), including both sides of the *Avr*II site at position −286, contains promoter elements to provide a minimal, albeit suboptimal, promoter function. The presence of a third region, spanning positions 53–1792 (*Bss*HII to *Pst*I) and containing intron 1, produced a dramatic stimulation of activity.

3.6. Other transcripts in the neighbourhood of the MTRR gene

The sequence immediately surrounding the *MTRR* gene does not match any known gene. However, three overlapping EST sequences (GenBank Accession Nos AA314900, AA316122 and R28412) are found within the region at −997 to −2408, upstream of the *MTRR* exon 1. Clone 133158 (GenBank Accession No R28412) was sequenced to assess the nature of its 3′ end. The location of the poly(A) tail, which is preceded by a AATAAA polyadenylation signal (data not shown), is compatible with the existence of a transcription unit that is antiparallel to *MTRR*. Another EST sequence (GenBank Accession No AA772426) corresponds to positions −2489 to −2227. It overlaps with the three EST contigs, but is oriented in parallel to the *MTRR* transcript and contains a poly(A) tail, preceded by a AATAAA sequence. Finally, another cDNA (GenBank Accession No AI016694) shows a match of 432 bp in the area −4000. It has a poly(A) tail, but no AATAAA consensus signal. Its polarity is the same as *MTRR*.

3.7. Physical mapping of human MTRR gene

We previously assigned the *MTRR* gene to chromosome 5, but regional localization was limited (Leclerc et al., 1998). Radiation hybrid analysis positioned the gene between markers D5S406 and D5S478, which is consistent with FISH results that allowed us to determine the gene's cytogenetic position as 5p15.2–15.3 (Leclerc et al., 1998). Three positive YACs (941a2, as well as the overlapping clones, 789d2 and 757d2), were shown to contain the whole of the *MTRR* gene (data not shown). These YACs are part of a larger contig, partially represented in Fig. 8. Their localization on the Whitehead map shows that they share a common region that includes STS markers WI4410 and WI9255. YAC clones 827a7 and 904h7 are positive for the flanking markers WI1755 and D5S1957, respectively, but not for *MTRR*, WI4410 and WI9255, while 789d2 and 941a2 are positive for the latter three markers as well as WI1755 and D5S1957, respectively. These data place the *MTRR* gene between WI1755 and D5S1957. To confirm this location, we tested the four STS markers (WI1755, WI4410, WI9255 and D5S1957) by PCR with PAC clone 104K2 and four BAC clones, 138P20, 259D10, 189I20 and 82N18 containing the full *MTRR* gene (data not shown). Three of the BAC clones, 138P20, 259D10 and 189I20, were positive for WI4410 and WI9255. In sum, a comparison of all genomic clones analysed in this study restricts the location of the *MTRR* gene to the interval between markers WI1755 and D5S1957.

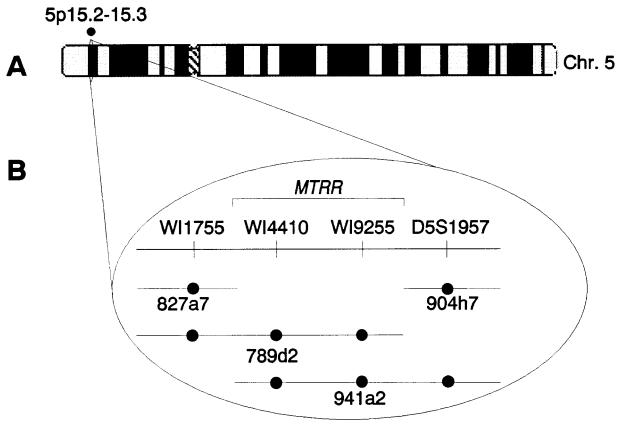

Fig. 8. Mapping of *MTRR* gene. (**A**) The *MTRR* gene was assigned to chromosome 5 by somatic cell hybrids analysis. The black dot above the chromosome shows the region 5p15.2–15.3, identified by FISH. (**B**) Partial physical map of chromosome 5 showing the fine location of *MTRR*. The horizontal line at the top represents the region of chromosome 5 harbouring the *MTRR* gene, with pter to the left and the centromere to the right. Markers are indicated above the line and relevant oligonucleotide primers and PCR conditions were as described (Leclerc et al., 1998). YACs are represented below the horizontal line, with each corresponding address. YAC clone 757d2, which is not shown, contains the same markers as 789d2. *MTRR* gene is located inside the interval marked with a bracket.

4. Discussion

We have characterized the entire transcription unit of the *MTRR* gene. The gene spans about 34 kb on

chromosome 5, region 5p15.2–15.3, between closely related STS markers WI1755 and D5S1957. The exon sequences encompass about 10% of the gene. The validity of our proposed genomic organization is supported by several lines of evidence. A primer extension analysis was performed with two different RNA sources and yielded a strong nucleotide band with weaker bands for a few adjacent nucleotides. These start sites are in agreement with results obtained by RT-PCR and are further supported by the proximity of potential binding sites for ubiquitous transcription factors, including a nearby upstream CAAT box as well as Sp1, AP-1 and AP-2 elements. A second CAAT element is found downstream of the cap site, a property observed in other genes (Fan et al., 1997). Analysis of the *MTRR* gene structure shows conservation of the position and type of several introns relative to those observed in the genes coding for its *C. elegans* ortholog and for rat CPR, supporting the proposal that these genes share a common origin (Leclerc et al., 1998). The size of the mRNA species predicted by the sum of the 15 exons (3302 bp) and a 200 base poly(A) tail agrees with the species of approximately 3.6 kb observed by Northern blot analysis (Leclerc et al., 1998).

In our initial studies on *MTRR* cDNA sequences (Leclerc et al., 1998), we reported alternative splicing involving two segments of 26 and 154 bp. Knowledge of the gene structure clarifies the mechanisms leading to these splicing products. The 26 bp deletion arises from the use of a potent 5′ splice site located close to the end of exon 1. The deletion of 154 bp results from complete omission of exon 3 and yields a truncated protein missing a part of the FMN binding site and completely deficient in FAD and NADPH binding sites.

We observed a minisatellite in intron 6, that, although showing a limited number of alleles, could be a useful marker for various genetic studies. The number of repetitive units observed in intron 6 of *cblE* patients is no different from those found in the general population, indicating that variability in length of the minisatellite is not one of the causes of *cblE disease*, at least for the 11 patients that were analysed in this study.

We showed that *cat* constructs containing the 5′ area of the gene, including the promoter region, exon 1, intron 1 and a portion of exon 2 gave a very potent activity. However, several constructs containing fragments of this large sequence were relatively poorly active or inactive by comparison with the parent sequence, suggesting that the transcriptional regulation of the *MTRR* gene is complex. We can distinguish three regions involved in transcriptional activity: a core promoter and two flanking regions required for optimal activity.

We define region −629 to 52 as the core promoter. All the candidate Sp1, AP-1, AP-2 and CCAAT sites fall within the minimal promoter-active sequence, which also contains the transcription initiation area. A TATA-less promoter with a cluster of Sp1 boxes in close proximity to the transcription start site(s), is typical of housekeeping genes. Many such promoters display multiple start sites, while others have a unique major transcription start site. In the *CAD* gene, Kollmar et al. (1994) found that an Sp1 site at −49 substituted for the missing TATA box and plays a major role in start site selection, as reported for other activator binding sites, in other TATA-less promoters at similar distances (Kollmar et al., 1994). As found for *MTRR*, with an Sp1 site at position −53, transcription of the *CAD* gene starts at a single major site. Other elements may be involved in start site selection, such as the sequences at positions 6–9 and 29–31, corresponding to the sequence TTTCN(1–30)GCA observed in several TATA-less promoters (Means and Farnham, 1990). This sequence was proposed to bind to the Housekeeping Initiator Protein 1 (HIP1), which directs initiation of transcription in the *DHFR* gene when it is coupled to an Sp1 site.

The region 5′ to the core promoter (positions −1503 to −630) also contributes to optimal activity. Several potential promoter elements are present in this segment, and the presence of a myogenin element (CAAGGGCA) at position −1150 could account for the high expression observed in skeletal muscle (Leclerc et al., 1998). The CpG island starts at position −546, and the presence of this portion of the CpG island may contribute to optimal activity.

We also conclude that the presence of a third region (pos 53–1792), located immediately downstream of the core promoter and encompassing intron 1, has an important impact on promoter activity. It is well known that certain promoters function more efficiently when accompanied by an intron. In some cases, 10- to 100-fold more mRNA is produced from intron-containing constructs (Kriegler,1990). We observed that the presence of intron 1 increases promoter activity some 30-fold. Analysis of the 5′end of *MTRR* cDNAs in cDNA libraries, as well as *cat* constructs containing parts of intron 1 failed to provide any evidence for the existence of an autonomous promoter in intron 1 (data not shown). The CpG island extends into that region until position 782. Since more than half of the CpG island is intronic, it is possible that the presence of additional CpG island sequences is responsible for increased level of expression by intron 1. The high activity associated with the presence of intron 1 may also be due to the existence of positive intronic elements (GenBank Accession No AF121202; McEwen and Ornitz, 1988; Tone et al., 1999) or to a stabilizing effect related to the splicing event (Kriegler, 1990). Promoters and cap sites are the most difficult portions of genes to define and understand, and it is a challenge to assess how disparate promoter sequences and the proteins they bind contribute to the initiation of transcription. Our work constitutes a first step in defining and characterizing the *MTRR* promoter.

Several segments of chromosome 5 have been characterized, including 3.5 Mb of region 5p15.2 (Church et al., 1997), but genomic segments encompassing the *MTRR* gene had not been reported previously. Furthermore, our analysis of the region upstream of the *MTRR* gene revealed the existence of corresponding EST sequences containing poly(A) tracts. Potential genes occur on both DNA strands of this region of the chromosome, a concept reported for a growing number of sequenced genes.

We previously reported a cDNA-derived 698 amino acid sequence for human methionine synthase reductase (Leclerc et al., 1998). The N-terminal sequence deduced from the largest cDNA initially reported suggested a cytosolic location for the enzyme, consistent with its role in the reductive activation of methionine synthase, a cytosolic enzyme. The sequence presented here for exon 1 extends the previously reported cDNA by 58 bp. The alternative splicing event involving 26 bp in exon 1 allows for two potential translation products, giving a second, predicted translation product of 725 amino acids in addition to the 698-amino-acid protein previously described. Both ATG codons are in a good context for initiation of translation (Kozak, 1991). The inclusion of the 26 bp yields the 698-amino-acid protein previously predicted (Leclerc et al., 1998), with an N-terminus consistent with cytosolic targeting. The mRNA form lacking the 26 bp segment can initiate at the upstream ATG, making the protein 27 amino acids longer (GenBank Accession No AF121214). When this predicted protein sequence was evaluated for a potential mitochondrial presequence with the program TRANSPEP, the amino acid sequence MGAASVRAGARLVEVALCSFTVTCLEVMRRF was identified. A mitochondrial location for methionine synthase reductase has to be confirmed by biochemical and/or immunological studies. If the alternative splicing at the 5′ end of the *MTRR* mRNA generates a mitochondrial presequence, then a role for methionine synthase reductase in mitochondria will need to be identified and would raise the possibility that this enzyme could act on substrates other than methionine synthase, as is the case for the corresponding flavodoxin/flavodoxin reductase system in bacteria (Osborne et al., 1991; Bianchi et al., 1993).

Acknowledgements

This work was supported by grants from the Medical Research Council of Canada Group in Medical Genetics (GR-13297) and NIH-National Heart, Blood and Lung Institute (HL58955). S.W.S. is a scholar of the Medical Research Council of Canada. We thank Timothy Johns and Eric Shoubridge for preparation of human myotubes. We are grateful to the I.M.A.G.E. consortium (Livermore, CA) for providing the human cDNA clones 704947 and 133158.

References

An, G., Hidaka, M., Siminovitch, L., 1982. Expression of bacterial β-galactosidase in animal cells. Mol. Cell Biol. 2, 1628–1632.

Bianchi, V., Reichard, P., Eliasson, R., Pontis, E., Krook, M., Jörnvall, H., Haggård-Ljungquist, E., 1993. *Escherichia coli* ferredoxin NADP+ reductase: activation of *E. coli* anaerobic ribonucleotide reduction, cloning of the gene (*fpr*) and overexpression of the protein. J. Bacteriol. 175, 1590–1595.

Chartrain, N.A., Geller, D.A., Koty, P.P., Sitrin, N.F., Nussler, A.K., Hoffman, E.P., Billiar, T.R., Hutchinson, N.I., Mudgett, J.S., 1994. Molecular cloning, structure, and chromosomal localization of the human inducible nitric oxide synthase gene. J. Biol. Chem. 269, 6765–6772.

Church, D.M., Yang, J., Bocian, M., Shiang, R., Wasmuth, J.J., 1997. A high-resolution physical and transcript map of the cri du chat region of human chromosome 5p. Genome Res. 7, 787–801.

Fan, W., Christensen, M., Eichler, E., Zhang, X., Lennon, G., 1997. Cloning, sequencing, gene organization, and localization of the human ribosomal protein RPL23A gene. Genomics 46, 234–239.

Gardiner-Garden, M., Frommer, M., 1987. CpG islands in vertebrate genomes. J. Mol. Biol. 196, 261–276.

Gulati, S., Chen, Z.Q., Brody, L.C., Rosenblatt, D.S., Banerjee, R., 1997. Defects in auxiliary redox proteins lead to functional methionine synthase deficiency. J. Biol. Chem. 272, 19 171–19 175.

Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A.E., Kel, O.V., Ignatieva, E.V., Ananko, E.A., Podkolodnaya, O.A., Kolpakov, F.A., Podkolodny, N.L., Kolchanov, N.A., 1998. Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. Nucleic Acids Res. 26, 362–367.

Jarrett, J.T., Hoover, D.M., Ludwig, M.L., Matthews, R.G., 1998. The mechanism of adenosylmethionine-dependent activation of methionine synthase: a rapid kinetic analysis of intermediates in reductive methylationof Cob(II)alamin enzyme. Biochemistry 37, 12 649–12 658.

Kingston, R.E., Sheen, J., 1998. Isotopic assays for reporter gene activity, in: Current Protocols in Molecular Biology. Wiley, New York, pp. 9.7.1–9.7.11.

Kollmar, R., Sukow, K.A., Sponagle, S.K., Farnham, P.J., 1994. Start site selection at the TATA-less carbamoyl-phosphate synthase (glutamine-hydrolyzing)/aspartate carbamoyltransferase/dihydroorotase promoter. J. Biol. Chem. 269, 2252–2257.

Kozak, M., 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. 266, 19 867–19 870.

Kriegler, M., 1990. Gene Transfer and Expression. Stockton Press, New York.

Leclerc, D., Wilson, A., Dumas, R., Gafuik, C., Song, D., Watkins, D., Heng, H.H.Q., Rommens, J.M., Scherer, S.W., Rosenblatt, D.S., Gravel, R.A., 1998. Cloning and mapping of a cDNA for methionine synthase reductase, a flavoprotein defective in patients with homocystinuria. Proc. Natl. Acad. Sci. USA 95, 3059–3064.

Luckow, B., Schütz, G., tz. CAT constructions with multiple unique restriction sites for the functional analysis of eukaryotic promoters and regulatory elements. Nucleic Acids Res. 15, 5490.

MacGregor, G.R., Caskey, C.T., 1989. Construction of plasmids that express *E. coli* beta-galactosidase in mammalian cells. Nucleic Acids Res. 17, 2365.

McEwen, D.G., Ornitz, D.M., 1988. Regulation of the fibroblast growth factor receptor 3 promoter and intron 1 enhancer by SP1 family transcription factors. J. Biol. Chem. 273, 5349–5357.

Marsden, P.A., Heng, H.H.Q., Scherer, S.W., Stewart, R., Hall, A.V., Shi, W.M., Tsui, L.C., Schappert, K.T., 1993. Structure and chro-

mosomal localization of the human constitutive endothelial nitric oxide synthase gene. J. Biol. Chem. 268, 17 478–17 488.
Means, A.L., Farnham, P.J., 1990. Transcription initiation from the dihydrofolate reductase promoter is positioned by HIP1 binding at the initiation site. Mol. Cell. Biol. 10, 653–661.
Osborne, C., Chen, L.M., Matthews, R.G., 1991. Isolation, cloning, mapping, and nucleotide sequencing of the gene encoding flavodoxin in *Escherichia coli*. J. Bacteriol. 173, 1729–1737.
Porter, T.D., Beck, T.W., Kasper, C.B., 1990. NADPH-cytochrome P-450 oxidoreductase gene organization correlates with structural domains of the protein. Biochemistry 29, 9814–9818.
Proudfoot, N.J., Whitelaw, E., 1988. Termination and 3′end processing of eukaryotic RNA in: Transcription and Splicing. IRL Press, Oxford, p. 97.
Rosenblatt, D.S., Cooper, B.A., Pottier, A., Lue-Shing, H., Matiaszuk, N., Grauer, K., 1984. J. Clin. Invest. 74, 2149–2156.
Rozen, R., 1996. Molecular genetic aspects of hyperhomocysteinemia and its relation to folic acid. Clin. Invest. Med. 19, 171–178.
Schug, J., Overton, C., 1997. TESS: Transcription elements search software on the WWW. Technical Report CBIL-TR-1997-1001-V0.0 of the Computational Biology and Informatics Laboratory, School of Medicine, University of Pennsylvania.
Shoubridge, E.A., Johns, T., Boulet, L., 1996. Use of myoblast cultures to study mitochondrial myopathies. Meth. Enzymol. 264, 465–475.
The *C. elegans* sequencing consortium 1998. Genome sequence of the nematode *C. elegans*: a platform for investing biology. Science 282, 2012–2018.
Tone, M., Diamond, L.E., Walsh, L.A., Tone, Y., Thompson, S.A.J., Shanahan, E.M., Logan, J.S., Waldmann, H., 1999. High level transcription of the complement regulatory protein CD59 requires an enhancer located in intron 1. J. Biol. Chem. 274, 710–716.
Wilson, A., Platt, R., Wu, Q., Leclerc, D., Christensen, B., Yang, H., Gravel, R.A., Rozen, R., 1999. A common variant in methionine synthase reductase combined with low cobalamin (vitamin B_{12}) increases risk for Spina Bifida. Mol. Genet. Metab. 67, 317–323.