Human Methionine Synthase

cDNA CLONING, GENE LOCALIZATION, AND EXPRESSION*

Linda H. Chen‡, Mei-Lan Liu‡, Hye-Yeol Hwang‡, Li-Sheng Chen‡, Julie Korenberg§, and Barry Shane¶

From the ‡Department of Nutritional Sciences, University of California, Berkeley, California 94720-3104 and the §Ahmanson Department of Pediatrics, Division of Genetics, Cedars-Sinai Research Institute, University of California, Los Angeles, California 90048

Human cDNAs for methionine synthase (5-methyltetrahydrofolate-homocysteine S-transmethylase; EC 2.1.1.13) have been isolated from fetal and adult liver and HepG2 libraries. The cDNAs span 7.2 kilobases (kb) and consist of a 394-base pair upstream untranslated region, a 3795-base pair open reading frame encoding a 1265-residue 140.3-kDa protein, and about 3 kb of 3’ region. The deduced protein sequence shares 53 and 63% identity with the Escherichia coli and the presumptive Caenorhabditis elegans proteins, respectively, and contains all residues implicated in B12 binding to the E. coli protein. Several potential polymorphisms and a cryptic splice deletion were detected in the coding region of the cDNAs. A polymorphism that results in a D919G modification in the protein is fairly common in human DNA samples. Northern analyses of poly(A) mRNA indicated two major species of about 8 and 10 kb in human tissues and some minor, partially spliced species. mRNA levels were highest in the pancreas, skeletal muscle, and heart of the adult and in the kidney in the fetus and were low in adult liver. Genomic clones were isolated and the 5’ region was analyzed. Exon 1 is preceded by a number of potential promoter sites, including an E box, CAAT boxes, and a GC box, but this region lacks a TATA element. The human methionine synthase gene was localized to chromosome region 1q42.3–43 by in situ hybridization.

Methionine synthase, one of two B12-dependent mammalian enzymes, catalyzes the remethylation of homocysteine to methionine and the concurrent demethylation of 5-methyltetrahydrofolate to tetrahydrofolate (1). Under conditions of B12-depletion, such as pernicious anemia, loss of methionine synthase activity leads to a “methyl folate trap.” The depletion of other folate coenzymes results in defective DNA synthesis and the potential role of polymorphisms in the enzymes involved in this cycle in disturbances of one-carbon metabolism. As a prelude we have isolated and characterized various human methionine synthase cDNAs. In this report, we describe the molecular cloning of human methionine synthase cDNAs and the localization, expression, and partial characterization of its gene.

EXPERIMENTAL PROCEDURES

Materials—α-32P-dATP (1000 Ci/mmol), [α-32P]dCTP (6000 Ci/ mmol), [γ-32P]ATP (6000 Ci/mmol), and [γ-32P]ATP (2000 Ci/mmol) were obtained from DuPont NEN. DNA restriction and modifying enzymes and RNase A were obtained from Boehringer Mannheim, Promega, or New England Biolabs. AmpliTaq and rTh DNA polymerases were from Perkin-Elmer. Nitran membranes were obtained from Schleicher and Schuell. Oligonucleotide primers were synthesized by the Micro-Chemical Facility (University of California, Berkeley). Multiple Tissue Northern blots for human adult and fetal poly(A) mRNA and human adult liver, kidney, and placenta total RNA were obtained from Clontech or were isolated from HepG2 and MCF-7 cells. All other materials were obtained from commercial vendors.

Isolation of cDNA and Genomic Clones—Total RNA was isolated from HepG2 cells using Trizol (Life Technologies, Inc.). Two μg were reversely transcribed at 48 °C for 45 min and amplified using Access reverse transcription-PCR (Promega) using a degenerate reverse primer (5’-TTNGGNTNCCNGCRTTNG-3’) corresponding to amino acids 281–275 of the putative C. elegans protein (11) and a forward primer (5’-ACTGGNACATGATHCAR-3’) corresponding to amino acids 25–30 to generate a 767-bp product. PCR conditions were 94 °C for 1 min, 60 °C for 1 min, and 72 °C for 2 min for 45 cycles. Human adult and fetal liver cDNA libraries in Agt10 (Clontech) and a HepG2 cDNA library in λZap (Stratagene) were screened (106 plaques/library) with [32P]dCTP labeled primers generated using the Random Primed DNA labeling kit (Boehringer Mannheim) and the PCR-generated probe as the template. Following a second round of screening and plaque purification, 5–6 positive clones were obtained from each library. Phage...

* This study was supported in part by Public Health Service Grant DK-42033 from the Department of Health and Human Services. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed. Tel.: 510-642-5202; Fax: 510-642-0535; E-mail: bandie@violet.berkeley.edu.
DNA from these clones was purified and EcoRI or NotI fragments were subcloned into pBluescript KS+ (Strategene) for further analysis. As sequences became available, library screening was repeated to obtain additional clones extending the 3’ region of the cDNA. Sequences were also extended initially by PCR extension of total cDNA library using methionine synthase-specific primers.

A human genomic library, generated from the lung fibroblast cell line WI38 and cloned in the Lambda FIX II vector (Strategene), was screened (104 plaques) with 32P(dCTP-labeled DNA fragments generated by random priming the PCR-derived genomic DNA and cDNA. The phage DNA from these clones was purified and characterized by restriction mapping and Southern hybridization.

**DNA Sequencing and Intron Size Determination**—DNA was sequenced by the method of Sanger et al. (13) using Sequenase, Version 2.0 (U.S. Biochemical Corp.) or using an Applied Biosciences Model 373A automated DNA sequence analyzer located at the Microsequencing Facility, University of California, Berkeley. The cDNA sequence was verified by sequencing both DNA strands. Exon-intron junctions were determined by direct sequencing across the junctions using oligonucleotide primers based on the cDNA sequence. Intron sizes were determined by sequencing the region by or PCR using flanking primers. Both strands of the nucleotide sequence in the 5’ region of the gene were evaluated for known consensus sequences that have been reported as potential transcriptional regulators using a transcription factor data base (Genetics Computer Group, Madison, WI).

**Nucleotide Sequence**—Total RNA was isolated from HepG2 and MCF-7 cells and 5’-RACE analysis using nested primers for various regions of the methionine synthase cDNA was performed as described previously (14).

**Northern Analysis**—Twenty-five µg of total RNA or 2 µg of poly(A) mRNA from human liver, kidney, muscle, pancreas, or placenta or from HepG2 or MCF-7 cells were fractionated under denaturing conditions on a 1% agarose gel containing 1% formaldehyde and transferred to a Zeta Probe GT membrane (Bio-Rad) by capillary transfer. The membranes were hybridized for 16 h at 42°C in formamide-containing hybridization buffer using 2.0 × 106 cpm of 32P-random-labeled methionine synthase or G3PDH cDNAs as probes. Similar hybridization conditions were used to probe human Multiple Tissue blots prepared with 2 µg of RNA. Blots were quantitated using a phosphorimager (Bio-Rad).

**Chromosomal Localization of Methionine Synthase Gene**—A methionine synthase DNA probe was labeled with biotin-14-dATP (Life Technologies, Inc.) by nick translation and hybridized to metaphase chromosomes prepared from normal male peripheral blood using the bromodeoxyuridine synchronization method (14). Fluorescence in situ hybridization was performed as described previously (15). Two amplifications were carried out using bismutylated antiavidin. To generate clear reverse bands, metaphase chromosomes were counterstained with Chromomycin A2, followed by Diastatycin A (15). The image was captured using a Photometrics cooled CCD camera and a BIDS image analysis system (Oncor).

**RESULTS AND DISCUSSION**

**Cloning of Human Methionine Synthase**—We have previously isolated a number of human cDNA encoding folate-dependent enzymes by functional complementation of *E. coli* mutants (16, 17). Multiple attempts at cloning a human methionine synthase cDNA using various expression vectors and a metE metH *E. coli* mutant, which lacks B12-dependent and -independent methionine synthases, were unsuccessful. This may reflect an inability of *E. coli* to insert or reduce the B9 coenzyme on the human protein, but more likely reflects a lack of full-length cDNA sequences in the expression libraries used. Various degenerate primer pairs, based on the *E. coli* and putative *C. elegans* methionine synthase sequences, were then used to amplify regions of HepG2 RNA by RT-PCR. One pair, from the 5’ region, produced the expected size product (767 bp), and sequence analysis indicated a high degree of identity with the *C. elegans* DNA and deduced protein sequences. The amplified product was used as a probe to isolate 5–6 clones from each of the human fetal and adult liver and HepG2 libraries. None of these clones extended to the 3’ end of the open reading frame. More clones encompassing the 3’ region of the cDNA were obtained by rescreening the libraries with restriction fragments encompassing the 3’ ends of the originally isolated clones. Initial sequences and some probes were also obtained by PCR extension of total library cDNA.

**Nucleotide Sequence**—Overlapping clones from the various cDNA libraries were completely sequenced in both orientations. The cDNA sequence of the 5’ UTR and the coding region of human fetal liver methionine synthase and the deduced protein sequence are shown in Fig. 1. This sequence, plus an additional 3 kb of 3’ UTR (7224 bp), has been deposited in GenBank (accession number U73338). The most 5’ fetal liver cDNA clone started at +15 (Fig. 1), and the first 15 nucleotides of the sequence were obtained by 5’-RACE analysis (see below) of cDNA derived from total HepG2 and MCF-7 RNA. With some exceptions (described below), identical sequences were found for the adult liver and HepG2 clones. The sequence contains a consensus polyadenylation signal and a short poly(A) tail. The 3’ UTR also contains an AUu sequence.

The major open reading frame codes for a 140.3-kilodalton protein of 1265 amino acids. This is similar to the mass of the purified monomeric pig liver methionine synthase (150 kDa) and the human placental protein (160 kDa), although the placental protein was reported to be a heterotrimer of subunit size 95, 45, and 35 kDa. The deduced human protein shares 63 and 53% sequence identity with the *C. elegans* and *E. coli* proteins, respectively, and many of the other residues are conservative substitutions (Fig. 2). Amino acid residues implicated in B12 cofactor binding in the *E. coli* protein (9, 10) are conserved in the human protein (His-785, Asp-783, Ser-836, Phe-723, Phe-729, and Leu-730 in the human sequence), and the regions around these residues show a high degree of identity (Fig. 2). There is little homology to other proteins in the data bases. The N-terminal third of the protein does show limited regions of identity with the recently described sequence for betaine:homocysteine methyltransferase (19), a B12- and folate-dependent methionine synthase present in some mammalian tissues. These regions may be involved in homocysteine binding.

The human gene contains a 4-kb intron located between residues 428 and 429 at the start of the long open reading frame. The similarity observed between the human and *C. elegans* nucleotide and protein sequences starts with exon 2, and exon 1 and intron 1 show no similarity with the *C. elegans* sequence. The exon 1 sequence was found in all three cDNA libraries by PCR analysis. RT-PCR of total RNA from human liver, kidney, and placenta, using three different primer pairs that spanned exons 1 and 2, gave the expected size bands in each case, indicating that exon 1 was not an artifact of the cDNA libraries.

The cDNA contains an upstream open reading frame (bases 12–434, Fig. 1) that potentially encodes a very basic 14.9-kDa protein (pl 12.3) of 141 amino acids that has no similarity with any protein in the data bases. The sequence around the start ATG is suboptimal for translation initiation, and the ATG is also positioned too close to the putative CAP site for efficient translation (20). The sequence around a second ATG (bases 35–37) is also suboptimal for translation initiation, whereas the third ATG (395–397) is in an optimal context for transla-
FIG. 1. Nucleotide sequence of human fetal liver methionine synthase cDNA. The 5' UTR, open reading frame, and derived amino acid sequence are shown. The first 15 nucleotides were obtained by 5'-RACE. The position of the region amplified using degenerate primers and upstream ATG codons are underlined.

Human Methionine Synthase cDNA

3630
tion initiation and encodes the start of the methionine synthase protein.

Polymorphisms, Deletions and/or Library Anomalies—Because defects in methionine synthase may potentially play a role in hyperhomocysteinemia, the cDNAs were screened for variants. The A at position 3150 in the fetal clones and an adult liver clone was changed to a G in another adult liver clone and in HepG2. This results in a D919G modification in the protein in the region believed to be involved in the binding of accessory proteins involved in cofactor reduction. This appears to be a fairly common polymorphism. In a preliminary analysis involving 44 human DNA samples, 16 were heterozygous and 1 homozygous for this mutation. The functional consequences, if any, of this polymorphism remain to be investigated.

HepG2 cDNA contained a G1158A polymorphism, which would result in a C255Y modification of the protein. This polymorphism was not observed in other clones or in 11 randomly screened human lymphocyte DNA samples. A T3970C modification was also observed in some clones, but this does not affect the protein sequence. Several other potential polymorphisms were observed in the 3' UTR.

A potential cryptic splice variant was observed in an adult liver cDNA clone with a 113-bp deletion of bases 1470–1582, inclusive. Bases 1470, 1471, 1582, and 1583 in the cDNA encode consensus GT-AG splice signals, but the deletion is 1 base shorter and could arise from GT-AA (or TC-AG) missplicing. The deleted region encodes a region of methionine synthase that shares considerable identity with the C. elegans and E. coli proteins, so it is unlikely that the cDNA shown in Fig. 1 contains a small intronic region. In addition, the deletion results in two slightly overlapping open reading frames that would encode 380-residue (40.9 kDa) and 858-residue (95.9 kDa) proteins. Because of the similarity of these sizes to the sizes of the subunits reported for the placental enzyme (18), we are investigating whether two proteins can be translated from this variant. PCR analysis using primers flanking the deletion indicates that 5–10% of the methionine synthase cDNAs in the different libraries contain this deletion. Preliminary studies, however, have not demonstrated the deletion in mRNA from human tissues.

One adult human liver clone contained the first 2460 bases of methionine synthase coupled to the complete cDNA for fibrinogen. Although this most likely represents a library construction anomaly and the fibrinogen gene is located on a different chromosome (4q28) from methionine synthase, the possibility exists that this reflects a chromosomal translocation. Although most of the cDNA clones contained large inserts (5 kb), some also contained intronic sequences from intron 3 (between cDNA bases 733 and 734), intron 4 (between bases 803 and 804) or intron 5 (between bases 896 and 897).

mRNA Distribution—Northern analysis of mRNA from human adult and fetal tissues indicated two main species of about 8 and 10 kb and a minor band at 4.4 kb (Fig. 3). Small amounts of larger species were also detected. Only slight variations in the relative proportions of the two main species were noted between tissues (Table I). These patterns were observed in multiple Northern analyses of commercially available human mRNA samples, as well as with total poly(A) mRNA extracted from HepG2 and MCF-7 cells. In all cases, methionine synthase mRNAs appear to be of very low abundance, and exposure times of greater than 24 h for detection with a phosphorimager and over 1 week with film were required to give reasonable signals with 2 μg of total poly(A) mRNA.

The reason for the larger major species is not totally clear. It
seems unlikely that the cDNA is lacking an additional 2 kb of sequence at the 5' end, and multiple 5'-RACE analyses, using nested primers to exon 1 (see below) or to exon 2, and 3'-RACE PCR analyses have failed to provide any evidence for an alternate exon 1 or 3' sequence or for alternate splicing in the region of the cDNA reported in this paper.

Northern analyses using probes complementary to intron 3 (2.4 kb) or 4 (1.7 kb) gave clear hybridization signals in the region of the higher molecular size bands shown in Fig. 3 (above 10 kb), and it is clear that the poly(A)-selected mRNA contains partially spliced methionine synthase RNA. A probe complimentary to intron 5 (2 kb) also hybridized to a higher size species but also gave a 10-kb signal. The 10-kb band obtained with the intron 5 probe was less intense than the higher size band, suggesting that retention of intron 5 can only account for part of the 10-kb mRNA species.

We are in the process of characterizing the intron/exon structure of the methionine synthase gene. The high incidence of intron retention in the cDNA libraries and the presence of partially spliced forms of poly(A) mRNA suggests that retention of additional introns may explain the abundance of the 10-kb mRNA species, but this remains to be investigated using additional intron probes.

In adult human tissues, methionine synthase mRNA levels were expressed at the highest levels in pancreas, skeletal muscle, and heart, and hepatic levels were much lower (Fig. 3; Table I). Although high levels might be expected in the pancreas, which is a very active organ for methylation reactions, other distributions are unexpected. About 90% of the whole body methyl group requirement is for creatine synthesis (21), which occurs primarily in the pancreas, kidney, and liver and to a lesser extent in other tissues (22). Creatine in muscle tissue is derived from other tissues. Methionine synthase is widely distributed in tissues and, unlike enzymes of the transsulfuration pathway, is expressed early in fetal development with the highest levels in nonhepatic tissue (23). The distribution of methionine synthase mRNA in mid-gestational fetal tissues was highest in the kidney (Table I). When the data were normalized to G3PDH mRNA levels, higher levels of expression were still observed in adult pancreas, heart, brain, and placenta but not in skeletal muscle, and fetal liver was still lower than other fetal tissues (Table I). It is not clear that using G3PDH mRNA as an internal standard gives a true reflection of mRNA abundance, especially for muscle. G3PDH mRNA is elevated up to 5-fold in muscle tissue (24), and its relative expression in fetal tissues is unknown. Approximately equal amounts of mRNA were used for these Northern analyses, and the nonnormalized values shown in Table I may be a better reflection of relative mRNA abundance.

Methionine synthase levels in mammalian cells are influenced by folate status, and the ability of cells to grow in the absence of methionine is partly rate-limited by synthase activ-

![Fig. 3. Methionine synthase mRNA levels in human adult and fetal tissues.](image)

**Fig. 3. Methionine synthase mRNA levels in human adult and fetal tissues.** Northern analysis of multiple tissue adult and fetal mRNA blots were conducted as described under “Experimental Procedures” using methionine synthase (A) and G3PDH (B) probes.

![Fig. 4. Nucleotide sequence of 5' region of human methionine synthase gene.](image)

**Fig. 4. Nucleotide sequence of 5’ region of human methionine synthase gene.** The DNA sequence is numbered from the +1 position of exon 1 (shown in lowercase). Potential regulatory sites are boldfaced. Potential transcription initiation sites suggested by primer extension analysis are indicated (●).

| Tissue       | Methionine synthase | G3PDH | Methionine synthase/G3PDH |
|--------------|---------------------|-------|--------------------------|
|              | 8 kb    | 10 kb | 8 kb/10 kb | 8 kb | G3PDH | 10 kb/G3PDH |
| Adult        |         |       |            |       |       |            |
| Heart        | 8.6     | 7.8   | 1.1        | 3.1   | 2.6   | 2.4        |
| Brain        | 2.1     | 2.0   | 1.0        | 0.6   | 3.5   | 3.2        |
| Placenta     | 4.0     | 2.8   | 1.4        | 1.1   | 3.5   | 2.5        |
| Lung         | 1.0     | 0.9   | 1.1        | 0.6   | 1.6   | 1.5        |
| Liver        | 1.0     | 0.8   | 1.2        | 1.0   | 1.0   | 0.8        |
| Skeletal muscle | 5.3 | 4.1   | 1.3        | 7.6   | 0.7   | 0.5        |
| Kidney       | 1.0     | 0.8   | 1.2        | 1.2   | 1.1   | 0.9        |
| Pancreas     | 6.5     | 5.8   | 1.1        | 0.5   | 12.9  | 11.2       |
| Fetus        |         |       |            |       |       |            |
| Brain        | 2.0     | 3.1   | 0.6        | 1.6   | 1.2   | 1.8        |
| Lung         | 2.3     | 2.8   | 0.8        | 1.9   | 1.1   | 1.3        |
| Liver        | 3.1     | 2.7   | 1.2        | 4.0   | 0.7   | 0.6        |
| Kidney       | 6.0     | 6.7   | 0.9        | 4.1   | 1.4   | 1.5        |
Figure 5. Fluorescence in situ localization of the methionine synthase gene to human chromosomal region 1q42.3–43. A human chromosomal preparation was hybridized with a 5-kb methionine synthase DNA probe labeled with biotin-14-dATP. The FTIC signals are clearly shown at the chromomycin and distamycin reverse-banded chromosomal region of 1q42.3–43. The human chromosome 1 ideogram (courtesy of U. Francke) shows the location of the methionine synthase gene.

ity levels. The availability of a human cDNA and gene for methionine synthase will allow studies on the regulation of expression of the protein and on the role of this enzyme in regulating homocysteine remethylation.

Organization of the 5′ Region of the Human Methionine Synthase Gene—A Lambda Fix II library was screened as described under “Experimental Procedures.” Nine clones were obtained after screening, eight of which were shown to be different by restriction mapping and Southern analysis (not shown). The 5′ region of the gene was also obtained by anchor PCR using a human PromoterFinder library (Clontech). A clone encompassing about 4 kb of the upstream sequence and exon 1 was further characterized. The region immediately 5′ to the transcription start site of exon 1 (+1 as defined by the longest 5′-RACE product; Fig. 4) contains a number of potential promoter sites, including an E box at −125 to −120, two CAAT boxes (CTF/NF-1 sites (25)) at −103 to −97 and −72 to −66, and a GC box (Sp1 binding site (26)) at −55 to −46, but lacks a TATA sequence. These characteristics are often attributed to “housekeeping genes” because the first group of genes found to have TATA-less promoters encoded proteins required for cellular metabolism (27). No obvious promoter characteristics could be ascribed to the 1-kb DNA region 5′ to this region.

Because we can not currently completely account for the 10-kb methionine synthase mRNA species, we have carried out extensive 5′-RACE analyses using primers complimentary to exon 1 and exon 2 regions. We have not obtained any evidence for an alternate exon 1, and all RACE products were consistent with the transcription start site region indicated in Fig. 4. The sequences of most of the RACE products terminated in the region +1 to +20, although some terminated later in this sequence.

Primer extension analysis has proved to be very difficult with methionine synthase mRNA due to its extremely low abundance. Using a primer starting at position +63, faint positive signals were obtained of sizes centering around 56 and 65 bp, as indicated in Fig. 4. This was observed with mRNA from a variety of cell lines and tissues. This is consistent with the transcription start site indicated in Fig. 4. An additional faint signal of 102 bp was also observed, which would suggest a start site of −40 (Fig. 4), which is downstream of the potential transcription factor binding sites. However, we have not detected any 5′-RACE product equivalent to this species.

Studies are currently underway to determine trans-acting factors that bind to regulatory elements and that control expression of the human methionine synthase gene.

Chromosomal Localization of Methionine Synthase—Human methionine synthase was previously localized to chromosome 1 by somatic cell hybridization (28). A 4.9-kb adult liver cDNA clone, covering the cDNA region 734–3399 and containing part of intron 3 and all of intron 4, was used to localize the gene to human chromosome band 1q42.3–43 (Fig. 5). Two independent experiments were performed, and over 50 metaphase cells were evaluated. Signals were clearly seen on two chromatids of at least one chromosome band (1q42.3–43) in over 70% of cells at no other sites in greater than 1% of cell.

Acknowledgments—We thank Sophia Stewart and Xiao-Ning Chen (Cedars-Sinai Research Institute) for assistance in the FISH analysis.

REFERENCES
1. Matthews, R. G. (1984) in Folate and Pterins (Blakley, R. L., and Benkovik, S. J., eds) Vol. 1, pp. 497–553, Wiley Interscience, New York
2. Shane, B., and Stokstad, E. L. R. (1985) Annu. Rev. Nutr. 5, 115–141
3. Shane, B. (1989) Vitam. Horm. 45, 263–335
4. Brattstrom, L., Lindgren, A., Israelsson, B., Malinow, M. R., Norrving, B., Upson, B., and Hamfelt, A. (1992) Eur. J. Clin. Invest. 22, 214–221
5. Selhub, J., Jacques, P. F., Bostom, A. G., D’Agostino, R. B., Wilson, P. W., Belanger, A. J., O’Leary, D. H., Wolf, P. A., Schaefer, E. J., and Rosenberg, I. H. (1995) New Engl. J. Med. 332, 286–291
6. Frost, P., Blom, H. J., Miles, R., Guyette, P., Sheppard, C. A., Matthews, R. G., Boer, G. J. H., den Heijer, M., Kruitman, L. A. J., van den Heuval, L. P., and Rozen, R. (1996) Nat. Genet. 10, 111–113
7. Jacques, P. F., Bostom, A. G., Williams, R. R., Ellison, R. C., Eckfeldt, J. H., Rosenberg, I. H., Selhub, J., and Rozen, R. (1996) Circulation 93, 7–9

2 J.-C. Hsu and B. Shane, unpublished data.
