Regulation of Expression of the Human \textit{MTH1} Gene Encoding 8-Oxo-dGTPase

\textbf{ALTERNATIVE SPlicing OF TRANSCRIPTION PRODUCTS*}

(Received for publication, February 10, 1997, and in revised form, May 8, 1997)

Hisanobu Oda‡, Yusaku Nakabeppu‡, Masato Furuichi‡, and Mutsuo Sekiguchi‡§§

From the ‡Department of Biochemistry, Medical Institute of Bioregulation, Kyushu University, Fukuoka 812-82, Japan and the §Department of Biology, Fukuoka Dental College, Fukuoka 814-01, Japan

The enzyme 8-oxo-7,8-dihydrodeoxyguanosine triphosphatase (8-oxo-dGTPase) hydrolyzes 8-oxo-dGTP to 8-oxo-dGMP, thereby preventing misincorporation of 8-oxo-dGTP into DNA. We investigated expression of \textit{MTH1} gene encoding 8-oxo-dGTPase. Large amounts of \textit{MTH1} mRNA were present in thymus and testis, embryonic tissues, and certain cell lines. In peripheral blood lymphocytes, the level of \textit{MTH1} mRNA was significantly increased after concomitant treatment with phytohemagglutinin and interleukin-2. Analyses of the 5' regions of the \textit{MTH1} transcripts revealed that 7 types of \textit{MTH1} mRNAs, which may be produced by transcription initiation at different sites and/or alternative splicing. The \textit{MTH1} gene consists of 5 major exons, some of which are composed of differentially processed segments. All types of \textit{MTH1} mRNAs carry the entire coding region, and may be functional. Three ATG initiation codons in-frame were found in the 5' regions of some of the \textit{MTH1} mRNAs. There is a polymorphic alteration at the 5' splicing site (GT to GC) located in exon 2, an event which affects splicing patterns of the \textit{MTH1} transcript. Allele frequency of this polymorphism is about 20% among healthy volunteers.

Oxygend radicals produced during normal cellular metabolism damage chromosomal DNA. An oxidized form of guanine, 8-oxoguanine\textsuperscript{1} (8-oxoG; 8-oxo-7,8-dihydroguanine), seems to be pertinent in terms of mutagenesis as well as carcinogenesis (1–4). 8-Oxoguanine nucleotide can pair with cytosine and adenine nucleotides with an almost equal efficiency during DNA replication and, high frequency of mutation is induced (5, 6).

Spontaneous mutagenesis caused by the oxidation of guanine nucleotides can be separated into two pathways, one initiated from oxidation of guanine in DNA and the other from oxidation of guanine nucleotide in the nucleotide pool. The oxidation of guanine residues in DNA results in formation of the 8-oxoG:C pair, which can lead to a C:G to T:A transversion, unless repair is made (7, 8). Studies on \textit{Escherichia coli} mutator mutants revealed that 8-oxoguanine generated in DNA is removed by DNA glycosylase/AP endonuclease, coded by the \textit{mutM} gene (9–12), and that adenine paired with 8-oxoguanine is excised by a specific adenine-DNA glycosylase, coded by the \textit{mutY} gene (13–15). Oxidation of guanine also proceeds in the form of free nucleotides, and an oxidized form of dGTP, 8-oxo-dGTP, is a potent mutagenic substrate for DNA synthesis (16). The MutT protein of \textit{E. coli} hydrolyzes 8-oxo-dGTP to the corresponding nucleoside monophosphate, and lack of the \textit{mutT} gene increases the occurrence of A:T to C:G transversion a thousand-fold over the wild type level (16–19).

Enzymes similar to those of the MutM, MutY, and MutT proteins were found in mammalian cells (20–22). Among them, the MutT-related protein has been studied most extensively (23). Based on the partial amino acid sequence determined with the purified human 8-oxo-dGTPase protein, cDNA for the human enzyme and the genomic sequence were isolated (24, 25). The human gene for 8-oxo-dGTPase was named \textit{MTH1} for human \textit{mutT} homologue and was found to be located on chromosome 7 at p22. With expression of cDNA for human \textit{MTH1} protein in the \textit{mutT\textsuperscript{−}} cells the elevated frequency of spontaneous mutation reverted to normal (24, 25). This means that the human protein may have the same antimutagenic capacity as the \textit{E. coli} MutT protein. cDNAs for mouse and rat enzymes have also been isolated (26, 27).

To examine roles of mammalian 8-oxo-dGTPase in the control of spontaneous mutagenesis as well as carcinogenesis, it is important to examine expression of the gene in various tissues, at different stages of development and under different physiological conditions, and to elucidate molecular mechanisms involved in the regulation of gene expression. We isolated various types of \textit{MTH1} transcripts, each of which has a unique structure at its 5'-terminal region, and their relations to the genomic sequence were elucidated. The existence of a polymorphic alteration which affects splicing of some of the transcripts is also given attention.

\textbf{EXPERIMENTAL PROCEDURES}

Chemicals—[\textalpha-\textsuperscript{32}P]dCTP and [\textgamma-\textsuperscript{32}P]ATP were purchased from Amersham International plc (Buckinghamshire, United Kingdom). Recombinant \textit{Taq} DNA polymerase, restriction enzymes, T4 DNA polymerase, T4 DNA ligase, and T4 polynucleotide kinase were obtained from Takara Shuzo (Kyoto, Japan) and Toyobo Co. (Osaka, Japan). T4 RNA ligase and calf intestinal alkaline phosphatase were purchased from New England Biolabs, Inc. (Beverly, MA) and Boehringer Mannheim (Mannheim, Germany), respectively. DNA labeling kits were obtained from Nippon Gene (Toyama, Japan). 1 Kb DNA Ladder and 0.24–9.5 Kb RNA Ladder as size standards, 10 × RPMI medium, cesium chloride,
and formamide were purchased from Life Technologies, Inc. (Gaithersburg, MD). Other chemicals were obtained from Wako Pure Chemical Industries Ltd. (Osaka, Japan). Sources of other materials are given in the text.

**Methods and Culture**—Lymphoblastoid cell lines established from healthy volunteers were kindly provided by Drs. T. Tana and T. Sasaki. These cell lines and a Jurkat line, a human T cell leukemia cell line, were cultured in RPMI 1640 medium containing 10% fetal calf serum. HeLa S3 cells were grown in Dulbecco's modified Eagle's medium supplemented with 5% fetal calf serum and 5% horse serum at 37 °C in a humidified atmosphere with 5% CO2. Peripheral blood mononuclear cells were isolated using Ficoll-Paque (Pharmacia Biotech, Uppsala, Sweden). Peripheral blood lymphocytes (PBL) were obtained from peripheral blood mononuclear cells by removing monocytes adhering to the flask wall. PBL were cultured in RPMI 1640 medium containing 10% fetal calf serum, with or without 50 IU/ml interleukin-2 (Shionogi Co., Ltd., Japan) and 10 μM/ml phytohemagglutinin-P (Difco) for 72 h.

**Preparation of RNA—**Total cellular RNA was prepared by the procedure of Chirgwin et al. (28). Briefly, cells in midlog growth phase were harvested, washed with ice-cold phosphate-buffered saline, and lysed with 4 M guanidine thiocyanate (Fluka Chemie AG, Buchs, Switzerland), 25 mM sodium citrate, 7% 2-mercaptoethanol, 0.5% sodium N-lauroyl sarcosinate (pH 7.0). The RNA was precipitated on a 5.7 M cesium chloride cushion by centrifugation at 32,000 rpm for 4 h in a Beckman SW27.5 rotor at 15 °C. The pellets were dissolved in 4 M guanidine thiocyanate, 25 mM sodium citrate, 7% 2-mercaptoethanol, extracted with phenol and chloroform, precipitated with ethanol, then dissolved in diethylpyrocarbonate-treated water. To obtain poly(A)+ RNA, an oligo(dT)-cellulose spun column (Pharmacia Biotech Inc.) was used, according to the manufacturer's manual. Purified poly(A)+ RNA was dissolved in 10 mM Tris-HCl (pH 7.6), 1 mM EDTA. RNA prepared from various human tissues were purchased from CLONTECH Laboratories, Inc. (Palo Alto, CA).

**Northern Blot Analysis—**Northern blotting was done as described (29). RNA (12 to 15 μg) was dissolved in 24 μl of sample buffer (50% formamide, 6.5% formaldehyde, 20 mM MOPS, pH 7.0), denatured at 65 °C for 5 min, and electrophoresed on a 1.2% agarose containing 6.5% formaldehyde. RNA was transferred onto nitrocellulose membrane (BA-85, Schreicher & Schuell Inc., Dassel, Germany) by capillary transfer (30) and the membrane was baked at 80 °C for 2 h in vacuo. After prehybridization in 5 × SSPE (1 × SSPE: 0.15 mM NaCl, 10 mM sodium phosphate, pH 7.0, 1 mM EDTA), 5 × Denhardt's solution, 0.1% sodium dodecyl sulfate (SDS), 200 μg of sonicated salmon sperm DNA/ml, overnight at 42 °C, the membrane was hybridized with 32P-labeled probe (labeled 46S oligo(dT)25-Bam I-Asp I fragment of MTH1) (31) (10 ng/ml: specific radioactivity ≈ 4 × 108 cpm/μg) in 50% formamide, 5 × SSPE, 5 × Denhardt's solution, 0.1% SDS, 200 μg of sonicated salmon sperm DNA per ml, for 16 h at 42 °C. The membrane was washed twice in 2 × SSC (1 × SSC: 0.15 mM NaCl, 0.015 M sodium citrate), 0.2% SDS at room temperature for 15 min, and then twice in 0.1 × SSC, 0.2% SDS for 15 min at 37 °C. Radioactivity on the membrane was measured in a BAS 2000 Bio-image analyzer (Fuji Photo Film Co., Ltd., Japan) and 109 cpm/mg of sonicated salmon sperm DNA/ml, 0.36 mM actinomycin D, 0.36 mM unlabeled dNTP, and 60 units of RNAsin (Promega, Madison, WI).

**Primer extension—**A reaction mixture containing 6 pmol of MTH-P primer in 33 μl of reaction mixture, using a first strand cDNA synthesis kit (Pharmacia Biotech), the remaining MTH-P primer was removed using PrimeErase Quik Push Column (Stratagene, La Jolla, CA). 100 μl of the purified ss-cDNA solution was mixed with 75 μl of 2 mM NaOH for RNA hydrolysis and ss-cDNA was precipitated with ethanol. The 5′ end of BML primer was ligated to the 3′ end of synthesized ss-cDNA using T4 RNA ligase. PCR was performed in a reaction mixture containing an appropriate amount of ss-cDNA, 10 pmol each of MTH1–17 and BM1 primers, 200 ng of ss-cDNA, 200 μg/ml of RNase-free DNase, 10 mM Tris-HCl (pH 8.3), 50 mM KC1, 1.5 mM MgCl2, 0.001% gelatin, and 2.5 units of rTaq DNA polymerase in a final volume of 50 μl. The initial denaturation was done at 94 °C for 1 min, then amplification was performed by 40 cycles of denaturation at 94 °C for 1 min, annealing at 55 °C for 2 min, and extension at 72 °C for 1 min, using a DNA thermal cycler (Perkin-Elmer). Following the first PCR, 1 μl of the reaction mixture was added to the second PCR reaction mixture. Cycles of the second PCR were run in the same manner except that MTH2–17 and BM2 primers were used. The PCR products were analyzed in 6% PAGE and cloned into pT7Blue T-vector (Novagen, Madison, WI).

**RT-PCR—**Total RNA (1 μg) was reverse-transcribed into ss-cDNA with MTH-P as a specific primer, using First Strand cDNA Synthesis Kit (Pharmacia Biotech). To amplify a specific 5′ region for each type of MTH1 mRNA, PCR was done in 50 μl of a reaction mixture containing the first strand cDNA, 2.5 units of rTaq DNA polymerase, 10 μM sense primer for each type of MTH1 mRNA (T1 to T4), 10 μM antisense primer MTH2–17, and 200 μM dNTP. The initial denaturation was done at 94 °C for 1 min, then amplification was performed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 60 °C for 1 min, and extension at 72 °C for 2 min, followed by extension at 72 °C for 2 min. The PCR product (10 μl) was digested with NcoI, and analyzed by 6% PAGE. To amplify full-length MTH1 cDNA, oligo(dT)17 was used as a primer for synthesis of the first-strand cDNA. PCR was performed as described above, except that Ex531 was used as an antisense primer. PCR products (10 μl) were separated on 3% agarose gel electrophoresis (NuSieve 3:1 Agarose, FMC BioProducts, Rockland, ME). DNA was transferred onto nylon membrane (Hybond N+, Amersham) and the filter was processed for hybridization with32P-labeled 461-bp NcoI-Asp I fragment of MTH1 cDNA as a probe. Radioactivity on the membrane was measured in BAS 2000 Bio-image analyzer.

**DNA sequencing—**Nucleotide sequence was determined using Dye Terminator or Dye Primer Cycle Sequencing FS Ready Reaction Kits and model 373A automated DNA sequencer (Perkin Elmer), according to the manufacturer's instructions.

**Genomic PCR and Direct Sequencing—**Genomic DNA was extracted from various types of human cell lines using ISOGEN kits (Nippon Gene). To amplify the entire region for MTH1 exon 2, PCR was done in 100 μl of reaction mixture containing 100 ng of genomic DNA, 10 μM of each sense (668-5) or antisense primer (491-3), 200 μM dNTP, and 2.5 units of rTaq DNA polymerase. The initial denaturation was performed at 95 °C for 1 min, then amplification was done by 35 cycles of denaturation at 95 °C for 45 s, annealing at 65 °C for 45 s, and extension at 72 °C for 45 s, followed by extension at 72 °C for 3 min. The PCR product (254 bp) was purified using Microcon 100 (Amicon Inc., Beverly, MA) and subjected to direct sequencing using Dye Terminator Cycle Sequencing FS Ready Reaction Kits and oligonucleotides GT-GC1 and GT-GC2 as primers.

**RESULTS—**

**Expression of the MTH1 Gene in Human Cells and Tissues—**Northern blot analysis was performed to determine levels of MTH1 mRNA in various tissues and cultured cell lines.
shown in Fig. 1, a band corresponding to 0.8-kilobase MTH1 mRNA was detected in almost all of the samples examined, although intensities of the bands varied considerably. Radioactivities hybridized to MTH1 cDNA were quantified using an image analyzer, and standardized for radioactivities hybridized to 18 S ribosomal RNA, as shown in Table I. Relatively large amounts of mRNA were found in thymus and testis. Fetal tissues (brain and liver) also contained large amounts of MTH1 mRNA.

The content of MTH1 mRNA in Jurkat cells, a human T cell leukemia cell line, was exceedingly high, as compared with findings in adult and fetal human tissues, evidence in accord with the observation that Jurkat cells contain a large amount of MTH1 protein with 8-oxo-dGTPase activity (20). HeLa cells, which also have a high proliferating capacity, had a lower but significantly large amount of MTH1 mRNA.

**Induction of MTH1 mRNA Synthesis**—To assess the correlation between the cellular content of MTH1 mRNA and the proliferating capacity of cells, we measured mRNA levels in resting and actively proliferating cells. Lymphocytes in peripheral blood are mostly in the G0 state, and when stimulated with phytohemagglutinin and interleukin-2, they enter the proliferating cell cycles within 72 h, as evidenced by flow-cytometric analysis (data not shown). Total RNAs were prepared from the two phases of cells and levels of MTH1 mRNA determined by Northern blot analysis. As shown in Fig. 2, a large amount of MTH1 mRNA was found in the RNA sample from the proliferating cells while a much lesser amount of mRNA was seen in the sample from the resting cells. About 7 times a larger amount of mRNA were found in the proliferating phase, as compared with findings in the resting one. Thus, expression of the MTH1 gene is regulated so as to reflect the proliferative state of cell.

**Primer Extension Analysis of Transcription Products**—To determine the transcription initiation site(s) for the MTH1 gene, a primer extension experiment was done, the result of which is shown in Fig. 3. Poly(A)+ RNA was prepared from Jurkat cells, which exhibit an extremely high level of expression of the gene, and used for the analysis. Major extension products were detected around the 70-base region, and multiple extension products larger than 100 bases were also seen, at a relatively lower level. We obtained an essentially similar
found various types of cDNAs according to their composition of the segments. These cDNAs were designated type 2, 3, and 4. MTH1 extended products from type 1 mRNA. Consequently, we found that the primer extension of Jurkat poly(A) fragments derived from HinfI-digested pBR322; lane 2, 2.0 μg of HeLa S3 poly(A)+ RNA; lane 3, 1.5 μg of Jurkat poly(A)+ RNA; lane 4, 4.5 μg of Jurkat poly(A)+ RNA; lane 5, 6.0 μg of yeast tRNA; lane 6, 6.0 μg of Jurkat poly(A)+ RNA. An arrow indicates the major extension product.

result with poly(A)+ RNA prepared from HeLa S3 cells, which produce about a quarter of the amount of mRNA, compared with Jurkat cells. Almost the same band patterns were obtained when the reaction was performed at either 37 or 42 °C (lanes 4 and 6 in Fig. 3), suggesting that these bands represent authentic transcription products, not artificially terminated ones. These results indicate that there are multiple 5′ ends of the MTH1 transcripts.

Sequence Analysis of the Multiple Transcription Products—To determine the nucleotide sequences of multiple 5′ ends of the MTH1 transcripts, SLIC-PCR was applied to amplify 5′ parts of the MTH1 cDNA species (Fig. 4A). If the MTH1 transcript has a N-base extended sequence beyond the MTH-Pr primer, a SLIC-PCR product of (75 + N) bp length, which is longer than the corresponding primer-extension product by 45 bases should be amplified. Indeed, various lengths of amplified products as seen in the primer extension were obtained. These products have poly(A)+ RNA and primer extension product from MTH1 mRNA. In SLIC-PCR, MTH1 mRNA was reverse-transcribed to produce ss-cDNA, using MTH-P as a primer. Oligonucleotide BML was ligated to the ss-cDNA. To amplify the 5′ region of MTH1 cDNA, nested PCR was performed using 2 sets of primers (1st, BM1 and MTH1–17; 2nd, BM2 and MTH2–17). The hatched boxes represent the known MTH1 cDNA sequence (24) while the open boxes represent the upstream sequence revealed by the present analyses. "N"b indicates number of bases extended beyond the MTH-Pr primer. Solid line represents the sequence derived from the oligonucleotide BML. B, SLIC-PCR products obtained from Jurkat cells analyzed by 6% PAGE. An arrowhead indicates the 112-bp PCR product.

cDNA has an insertion between the first and the second exon of type 1 cDNA (Fig. 5). In type 2 cDNA, there are two types of insertions and, according to types of insertion, it can be divided into 2A and 2B. Type 2A carries a 51-bp insertion while type 2B a 124-bp insertion, the latter being composed of the 51-bp sequence and an additional 73-bp sequence. The 51-bp and the 73-bp sequences were designated exon 2a and 2c, respectively, since they were found to exist contiguously in the genomic sequence (see Fig. 6).

Type 3 cDNA carries, in place of exon 1a, exon 1b sequence that locates downstream of the exon 1a sequence in the genomic DNA. This type of cDNA also carries two types of insertions corresponding to those of type 2A and 2B, and they were named type 3A and 3B, accordingly. Type 4 cDNA, which carries exon 2a in place of exon 1a or 1b, is classified into 4A and 4B. They carry insertions also corresponding to those of 2A and 2B, respectively. Constitutions of these various types of cDNAs are shown in Fig. 5.

Seven Types of MTH1 mRNAs—To show that these various types of MTH1 mRNAs are indeed present in cell, RT-PCR analysis was performed using total RNA prepared from Jurkat cells. Four types of PCR primers specific for the 5′ end of each MTH1 mRNA species (T1 to T4, shown in Fig. 5) were prepared, and in combinations with the common 3′-primer (MTH2–17), specific regions of the mRNA were amplified as cDNA fragments. Fig. 7 shows patterns of PAGE analysis. Except for the case of type 1 MTH1 mRNA that exists as a single form, two types each of cDNA fragments, corresponding to those having 51- and 124-bp insertions, were detected. When the amplified fragments were digested by NcoI restriction en-

![Fig. 3. Primer extension analysis of MTH1 mRNA.](image-url)
zyme, which recognizes the CCATGG sequence present at the site for the MTH1 initiation codon, expected sizes of smaller fragments were produced. An additional small band present in lane 4 was not affected by NcoI digestion, and was disregarded.

To determine if all of the multitypes of MTH1 mRNAs carry the coding sequence for MTH1 protein, we set up an experiment to amplify the entire coding region for MTH1 protein with the use of 5' primers specific for each type of MTH1 mRNA. Total RNA was prepared from the human thymus, and the first strand cDNA was reverse-transcribed using an oligo(dT)18 primer to select functional poly(A)+ mRNAs. As shown in Fig. 8, expected sizes of PCR products for 7 types of MTH1 mRNA were obtained. Since there are some additional bands beside those with expected sizes, the PCR products were subjected to Southern hybridization with 32P-labeled MTH1 cDNA as a probe. Only bands with expected sizes for the MTH1 mRNA were obtained. Since there are some additional bands beside those with expected sizes, the PCR products were subjected to Southern hybridization with 32P-labeled MTH1 cDNA as a probe. Only bands with expected sizes for the MTH1 mRNA were obtained.
Fig. 8. Existence of the entire coding sequence in all types of MTH1 mRNAs. As-cDNA was synthesized from total RNA, prepared from thymus, using oligo(dT)$_{18}$ primer. The full-length MTH1 coding sequence was amplified using a 5' primer specific for each type of cDNA and a 3' primer complementary to the 3'-untranslated region of MTH1 mRNA. Upper panel, PCR products were analyzed by 3% agarose gel electrophoresis. The presence or absence of the primers and the template in the reaction mixture are shown by + or −. Lower panel, Southern blot analysis of RT-PCR products. DNA was transferred onto nylon membrane. The filter was processed for hybridization with $^{32}$P-labeled 461-bp Neo-I$\alpha$Sp1 fragment of MTH1 cDNA as a probe.

Species were detected (Fig. 8, lower panel). Thus, it seems clear that human cells contain functional multiforms of MTH1 mRNAs with various 5' configurations.

Formation of 7 Types of MTH1 mRNAs by Alternative Splicing—We reported earlier that the human MTH1 gene is composed of four exons (25). The present study has revealed that an additional exon must exist between the previously characterized first and second exons and, moreover, these exons are composed of more than two segments, which are transcribed together but spliced in a different manner. Based on alignment of the cDNA sequences and the genomic sequence, a new feature of the MTH1 gene, composed of 5 exons, now appears, for which exons are renamed according to their arrangements. The overall structure of the gene thus defined is shown in the upper part of Fig. 9.

Various patterns of splicing, which would produce different forms of mRNAs, are shown in the lower part of Fig. 9. Alternative splicing of a single primary transcript would produce all types of mRNAs, although there remains the possibility that three types of primary transcripts, one for type 1 and 2, and the others for types 3 and 4, might be transcribed from different initiation sites. It should be noted that sequences that fit the GT/AG rule exist around each of exon-intron junctions.

Polymorphic Alteration Affecting Splicing Pattern—When total RNA prepared from human thymus was subjected to RT-PCR using type-specific 5’ primers, cDNA fragments corresponding to 7 different types of MTH1 mRNAs were detected, as expected from the foregoing experiments. However, a quite different pattern was obtained with RNA prepared from PBL, provided by a healthy volunteer. In this case, only four types of RNAs, corresponding to type 1, 2B, 3B, and 4B, were found (Fig. 10). Sequence analysis of the cDNA and the genomic DNA obtained from the same person revealed that it has a homozygous T to C base substitution at the beginning of exon 2c segment (see Fig. 5). This site serves as the 5’ splicing site during maturation of type 1, 2A, 3A, and 4A mRNA and, thus, the base change should abolish proper splicing at this site, leaving segments 2b and 2c connected.

To examine distribution of this particular base change among the human population, genomic DNAs were prepared from tumor cell lines (Jurkat and HeLa S3) and lymphoblastoid cell lines established from 10 healthy Japanese volunteers, and the corresponding region was amplified and sequenced. In most cases, including DNAs from the two tumor cell lines, this site is homozygous for GT, in two cases it is heterozygous as GT/GC, and only one case shows homozygocity for GC. Thus, this site is polymorphic in the human population.

DISCUSSION

To understand regulatory mechanisms of expression of the MTH1 gene, determination of the transcription initiation site is essential, to be followed by identification of the promoter region and characterization of cis-regulatory sequences. In the process of this approach, we found 7 types of MTH1 mRNA with different 5’ sequences in vivo. Comparisons of the sequences for these mRNAs and the MTH1 genomic sequence revealed that the MTH1 gene consists of 5 major exons, some of which consist of two or three segments differentially processed.

In most or all of human cells and tissues, type 1 mRNA which lacks the exon 2 sequence is present as a major form, as judged from primer extension analysis and the frequency of isolated clones as SLIC-PCR products. The other 6 forms of MTH1 mRNAs carry at least one segment of exon 2, implying the...
existence of some regulation for exon selection in the course of processing of pre-\textit{MTH1} mRNAs. It is of interest to note that for type 1 mRNA, the major transcript lacks any segment of exon 2. On the other hand, type 4A and 4B mRNA, which lack the exon 1 sequence, carry the exon 2a sequence, which is not found in other types of mRNAs.

For efficient splicing reactions of pre-mRNA in mammalian cells, it has been pointed out that at least four types of cis-elements present within intron segments, are required (33). A 5’ splice site which directs the beginning of an intron has a consensus sequence of (CA)AG:GU(A/G)AGU, and GU is the 5’ end of the intron. On the other hand, a 3’ splice site which determines the 3’ end of an intron has a much poorer consensus sequence, (NC/U)AG:(G/A), in which AG is its 3’ end. In addition, two sequences present near the 3’ splice site within an intron are important and are called a branch point sequence (UNCURAC) and a polypyrimidine-rich region. The former is necessary for a lariat formation with the 5’ splice site and is located in 18 to 38 nucleotides upstream of the 5’ splice site, while the latter is found between the former and the 3’ splice site. At each exon-intron junction of the \textit{MTH1} gene, the consensus sequences for the 5’ splice sites are well conserved, as follows: exon 1a (CAG:GAU), exon 1b (GU:GUGAc), 2b (Aga:GUGA), and 2c (AAG:GUGAu). The 3’ splice site preceding exon 3 has a well conserved feature for the consensus sequence, namely, it has a 23-nucleotide length polypyrimidine tract and a plausible branch point sequence (gGCUGGAC). On the other hand, in the 3’ splice site preceding exon 2b, a much shorter polypyrimidine tract with a poorly conserved branch point sequence (aUGGAC) was found. These features of the cis-elements for splicing may explain why exon 1a is more efficiently joined with exon 3 rather than with exon 2b, thus resulting in a predominant formation of type 1 \textit{MTH1} mRNA in cell.

Type 4A and 4B mRNAs lack a sequence corresponding to exon 1 and the most 5’ region of these mRNAs corresponds to 2a segment. No 2a segment was found in other mRNA species that carry one of the exon 1 segments. It should be noted that no sequence similar to the 3’ splice site is present around the 5’ end of 2a segment, thus it is likely that 2a segment serves as part of the first exon for type 4A and 4B \textit{MTH1} mRNAs.

In \textit{MTH1} transcripts, there are alternative selections of one or more of three exon 2 segments. A polymorphic alteration, Aga:GUGA to Aga:GeGAGU, was found in the 5’ splice site sequence, located at the beginning of the exon 2c segment. In peripheral blood lymphocytes prepared from an individual who is homozygous for this polymorphic alteration, type 2A, 3A, and 4A \textit{MTH1} transcripts were not detected. Frequency of this type of polymorphic alteration in the \textit{MTH1} allele is about 20%, implying that about 4 out of 100 individuals may have the homozygous alterations regarding this particular site. Such a homozygote was found in a group of young healthy volunteers, but to date there is no apparent phenotype. Such variants might provide a useful means to examine molecular mechanisms of alternative splicing as well as for estimation of the biological significance of various forms of \textit{MTH1} protein, as discussed below.

The sequence of type 1 \textit{MTH1} mRNA is essentially the same as that of the originally isolated cDNA (24), which directs formation of a 18-kDa \textit{MTH1} protein. Type 2A, 3A, and 4A mRNAs would code for the same polypeptide as a sole product, although they have different 5’-untranslated regions, which may serve as elements for modulating their translation efficiencies (34). It is noteworthy that individuals with the homozygous alterations at the beginning of exon 2c do not produce type 2A, 3A, and 4A mRNAs. The finding that these mRNA species encode the same 18-kDa \textit{MTH1} polypeptide may provide an explanation.

In three other types of \textit{MTH1} mRNAs with connected exon 2b-2c segments (type 2B, 3B, and 4B mRNA), there are three additional in-frame initiation codons (ATG1, ATG2, and ATG3) in their 5’-region (see Fig. 5). Interestingly, translation from ATG1 is interrupted by a stop codon, TGA, located at the beginning of exon 2c. In individuals in whom T at this site is substituted by C, the open reading frame from the ATG1 can get read-through into that for the 18-kDa \textit{MTH1} polypeptide, forming a larger polypeptide \textit{MTH1} with a molecular mass of 22,505. If translation is initiated from ATG2 and ATG3, polypeptides with molecular masses of 20,282 and 19,454 would be produced. Thus, the polymorphic alteration in exon 2c must affect expression patterns for various forms of \textit{MTH1} protein. Physiological effects of such variation might be small, but attention to long-term effects on spontaneous mutagenesis is needed. The sequence around ATG4 codon (ACCATGGGCC) fits well the consensus sequence for the translation initiation site, but others not so well, suggesting that the frequency of initiation from these sites, if any, may be low. By using an \textit{in vitro} transcription/translation system, we detected multiple forms of translation products derived from 2B, 3B, and 4B cDNA. We also detected multiple polypeptides that specifically react with anti-hMTH1 antibody, in crude extracts from various human cell lines. One must take into account that about 5% of \textit{MTH1} protein exists in the mitochondrial matrix while the remaining \textit{MTH1} protein is present in the cytoplasm (35). There is the possibility that the \textit{MTH1} polypeptides with extended N-terminal sequences might be preferentially translocated to the mitochondria. Alternatively, the N terminally extended portion of the \textit{MTH1} polypeptide may modulate its stability and/or function \textit{in vivo}.

Around the 5’ ends of exon 1a, 1b, and 2a no consensus sequence for the 3’ splice site was found, further strengthening the idea that transcription of the \textit{MTH1} gene may be initiated from just upstream from these exons. We found that the sequence TCACCTCC, located right upstream of exon 1a, is identical to the sequence for the initiator of murine TuT gene (36). In addition, a stretch of 13-bp sequence just upstream of exon 1a is well conserved in the promoter region for the mouse.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig10.png}
\caption{Altered patterns of \textit{MTH1} mRNAs in a splicing variant. ss-cDNA was synthesized from total RNA using \textit{MTH1} specific primer (MTH-P). cDNA was amplified using a 5’ primer specific for each type of \textit{MTH1} mRNA (T1 to T4) and a 3’ primer complementary to the \textit{MTH1} coding sequence (MTH2–17). RT-PCR was done using total RNAs from thymus and from PBL prepared from a healthy volunteer. PCR products were analyzed by 6% PAGE.}
\end{figure}
**MTH1** (37). Assuming that the adenine residue located 6 nucleotides upstream of the 5' end of exon 1a is the putative transcription initiation site for human **MTH1** mRNA, the 13-bp sequence is located –17 to –29 in the human **MTH1** gene (–25 to –37 in the mouse genome). This 13-bp sequence contains a potential binding site for Ets family proteins, and two additional potential binding sites for Ets family proteins were found near the putative initiation site and within exon 1a. Ets family proteins are known to regulate transcription in response to multiple developmental and mitogenic signals (38–41), and especially, expression of Ets-1 is regulated during both thymocyte development and T-cell activation (42, 43). These sites may be involved in the induction of **MTH1** gene expression by phytohemagglutinin and interleukin-2. There is no TATA and CCAAT-like sequence in the 5' upstream region of the human **MTH1** gene, but potential Sp1-binding sites are located 50 and 160 bases upstream of the initiator sequence. This region is highly GC-rich, consistent with the notion that the **MTH1** gene is a housekeeping gene.

**Acknowledgments**—We extend special thanks to Drs. T. Tana and T. Sasaizuki for providing lymphoblastoid cell lines, Drs. S. Oda and S. Mizuno for pertinent advice, and M. Ohara for comments on the manuscript.

**REFERENCES**

1. Kasai, H., and Nishimura, S. (1984) *Nucleic Acids Res.* 12, 2137–2145
2. Ames, B. N., and Gold, L. S. (1991) *Mutat. Res.* 250, 3–16
3. Ames, B. N., Shigenaga, M. K., and Hagen, T. M. (1993) *Proc. Natl. Acad. Sci. U. S. A.* 90, 7915–7922
4. Janssen, Y. M. W., Houten, B. V., Borm, P. J. A., and Mossman, B. T. (1993) *Lab. Invest.* 69, 261–274
5. Wood, M. L., Dizdaroglu, M., Gajewski, E., and Essigmann, J. M. (1990) *Biochemistry* 29, 7024–7032
6. Shibutani, S., Takeshita, M., and Grollman, A. P. (1991) *Nature* 349, 431–434
7. Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S., and Loeb, L. A. (1992) *J. Biol. Chem.* 267, 166–172
8. Mertio, M. (1993) *Proc. Natl. Acad. Sci. U. S. A.* 90, 1122–1126
9. Cabrera, M., Nghiem, Y., and Miller, J. H. (1988) *J. Bacteriol.* 170, 5405–5407
10. Chung, M. H., Kasai, H., Jones, D. S., Inoue, H., Ishikawa, H., Ohtsuka, E., and Nishimura, S. (1991) *Proc. Natl. Acad. Sci. U. S. A.* 88, 4690–4694
11. Michaels, M. L., Pham, L., Cruz, C., and Miller, J. H. (1991) *Nucleic Acids Res.* 19, 3629–3632
12. Nghiem, Y., Cabrera, M., Cupples, C. G., and Miller, J. H. (1988) *Proc. Natl. Acad. Sci. U. S. A.* 85, 2709–2713
13. Au, K. G., Clark, S., Miller, J. H., and Modrich, P. (1989) *Proc. Natl. Acad. Sci. U. S. A.* 86, 8877–8881
14. Michaels, M. L., Cruz, C., Grollman, A. P., and Miller, J. H. (1992) *Proc. Natl. Acad. Sci. U. S. A.* 89, 7022–7025
15. Makl, H., and Sekiguchi, M. (1992) *Nature* 355, 273–275
16. Yanofsky, C., Cox, E. C., and Hryn, K. (1966) *Proc. Natl. Acad. Sci. U. S. A.* 55, 274–281
17. Igarashi, H., Tsuzuki, T., Kakuma, T., Takimoto, K., Nishida, J., Tsukui, T., and Sekiguchi, M. (1993) *J. Biol. Chem.* 268, 14659–14665
18. Sekiguchi, M., and Takeshige, K. (1995) *Genomics* 24, 485–490
19. Kakuma, T., Nishida, J., Tsuzuki, T., and Sekiguchi, M. (1995) *J. Biol. Chem.* 270, 23942–23948
20. Cai, L. K., Meis, T., Tsuzuki, T., and Sekiguchi, M. (1996) *Carcinogenesis* 18, 2343–2350
21. Akiyama, M., Maki, H., Sekiguchi, M., and Horiuchi, T. (1989) *J. Biol. Chem.* 264, 2721–2726
22. McGoldrick, J. P., Yeh, Y.-C., Solomon, M., Essigmann, J. M., and Lu, A.-L. (1985) *Mut. Res.* 159, 899–916
23. Macleod, K., Leprince, D., and Stehelin, D. (1992) *Proc. Natl. Acad. Sci. U. S. A.* 89, 11021–11025
24. Au, K. G., Clark, S., Miller, J. H., and Modrich, P. (1989) *Proc. Natl. Acad. Sci. U. S. A.* 86, 3949–3952
25. Tajiri, T., Makl, H., and Sekiguchi, M. (1995) *Mutat. Res.* 336, 257–267
26. Mo, J.-Y., Makl, H., and Sekiguchi, M. (1992) *Proc. Natl. Acad. Sci. U. S. A.* 89, 11024–11028
27. Sekiguchi, M., and Takeshige, K. (1995) *J. Biol. Chem.* 270, 23942–23948
28. Baptiste, J., Edwards, D. M., Delort, J., and Mallet J. (1993) *Current Methods and Applications* (White, B. A., ed) pp. 365–385, Humana Press, Totowa, NJ
29. Sharp, P. A. (1985) *Science* 225, 766–771
30. Sharp, P. A. (1993) *Science* 260, 176–179
31. Sharp, P. A. (1985) *Science* 225, 766–771
32. Sharp, P. A. (1993) *Science* 260, 176–179
33. Sharp, P. A. (1985) *Science* 225, 766–771
34. Sharp, P. A. (1993) *Science* 260, 176–179
35. Sharp, P. A. (1985) *Science* 225, 766–771
36. Sharp, P. A. (1993) *Science* 260, 176–179