Genomic Organization and Expression of a Human Gene for Myc-associated Zinc Fingerprint Protein (MAZ)*

(Received for publication, March 9, 1998, and in revised form, May 19, 1998)

Jun Song‡§, Hiroo Murakami‡§, Hatsumi Tsutsui‡, Xiaoren Tang‡, Masatoshi Matsumura¶, Keiichi Itakura, Ichiro Kanazawa***, Kailai Sun‡, and Kazunari K. Yokoyama‡‡‡

From the ‡Tsukuba Life Science Center, RIKEN, 3-1-1 Koyadai, Tsukuba, Ibaraki 305, Japan, the §Department of Medical Genetics, China Medical University, Shenyang 110001, China, the *Institute of Applied Biochemistry, University of Tsukuba, Tsukuba, Ibaraki 305, Japan, the †Department of Molecular Genetics, Beckman Research Institute of The City of Hope, Duarte, California 91010, and the **Department of Neurology, Institute of Brain Research, Faculty of Medicine, University of Tokyo, Tokyo, Japan

We have cloned and characterized the genomic structure of the human gene for Myc-associated zinc finger protein (MAZ), which is located on chromosome 16p11.2. This gene is transcribed as an mRNA of 2.7 kilobases (kb) that encodes a 60-kDa MAZ protein. A 40-kb cosmid clone was isolated that includes the promoter, five exons, four introns, and one 3’-untranslated region. All exon-intron junction sequences conform to the GT/AG rule. The promoter region has features typical of a housekeeping gene: a high G + C content (88.4%); a high frequency of CpG dinucleotides, in particular within the region 0.5 kb upstream of the site of initiation of transcription; and the absence of canonical TATA and CAAT boxes. An S1 nuclease protection assay demonstrated the presence of multiple sites for initiation of transcription around a site 174 nucleotides (nt) upstream of the ATG codon and such expression was reflected by the promoter activity of a MAZ promoter/CAT (chloramphenicol acetyltransferase) reporter gene. Cis-acting positive and negative elements controlling basal transcription of the human MAZ gene were found from nucleotides (nt) −383 to −248 and nt −2500 to −948. Moreover, positive and negative autoregulatory elements were also identified in the regions from nt −248 to −189 and from nt −383 to −248 after co-transfection of HeLa cells with plasmids that carried the MAZ promoter/CAT construct and the MAZ-expression vector. Our results indicate that the 5’-end flanking sequences are responsible for the promoter activities of the MAZ gene.

The c-myc protooncogene is a member of a family of genes that encode DNA sequence-specific transcription factors with basic, helix-loop-helix, and leucine zipper domains. The Myc protein binds to DNA as heterodimers with a related polypeptide, Max (1–4). Appropriate regulation of expression of the human c-myc gene is necessary for the proliferation and differentiation of cells and for progression of the cell cycle, and deregulation of the expression of c-myc is associated with tumorigenesis and apoptosis (1, 4). Regulation of the expression of the human c-myc gene occurs at multiple levels, which include the initiation, the termination, and the attenuation of transcription (2, 4). In proliferating cells, the initiation of transcription of the c-myc gene is controlled by two major promoters, P1 and P2, and the RNA initiated from the P2 promoter accounts for 80–90% of the total RNA initiated from the P0, P1, and P2 promoters (4, 5). Initiation of transcription from the P2 promoter requires at least three cis-elements: ME1a2, E2F, and ME1a1 (6, 7). Several transcription factors, including Sp1 (5, 8), the Myc-associated zinc finger protein (MAZ) (9, 10), Pur-1 (11), and E2F (12) bind to these elements in vitro and in vivo.

The MAZ protein was identified as a transcription factor that binds to a GA box (GGGAGGG) at the ME1a1 site, to the attenuator region of P2 within the first exon of the c-myc gene, and to a related site that is involved in the termination of transcription of the gene for complement 2 (C2) (5, 9). Kennedy and Rutter (11) identified the Pur-1 protein as a GAGA box binding factor that binds to rat genes for insulin I and II and to the human gene for islet amyloid polypeptide (11). We recently reported the isolation of a cDNA clone for a member of the family of MAZ proteins in human islet cells (13). MAZ protein plays a role in the control of the initiation of transcription of genes for the adenovirus major late protein (14), CD4 (15), the serotonin receptor (16), and hematopoietic transcription factor (17), as well as in the termination of transcription between the closely spaced human genes for complement (8) and in the termination of transcription of the introns of the mouse gene for IgM-D (8). Therefore, MAZ appears to be a transcription factor with a dual role in the initiation and termination of transcription. We showed previously that MAZ is essential for the ME1a1-mediated expression of the c-myc gene during the neuroectodermal differentiation of P19 cells (18) and for the nucleate-hypersensitive element-mediated transcription of the c-myc gene in islet β-cells (13).

To gain a better understanding of the regulation of expression of MAZ, of the splicing mechanism, of the differential polyadenylation and of the potential interactions of MAZ with other factors, we isolated a human genomic gene for MAZ from cosmid and YAC libraries. We characterized the genomic structure of the gene for MAZ protein and identified regulatory elements in 5’-end flanking sequences that are involved in...
FIG. 1. Nucleotide and predicted amino acid sequences of the human gene for MAZ. The nucleotide sequence has been deposited in GenBank under accession number D89880. The nucleotides are numbered on the right, starting with the first residue shown. Exons are shown in uppercase letters and are shaded. The major site of initiation of transcription, as determined by the 51 nucleotide protection assay, is indicated by an arrow. The stop codon of the coding sequences is indicated by a star. Numbering of amino acids is shown on the right, starting with the site of initiation of translation, M. The polyalanine repeats are double underlined. The zinc finger motifs are boxed. The polyadenylation signal is indicated by dotted line.
basal transcription and in the autoregulation of the gene for MAZ by the MAZ protein itself.

**EXPERIMENTAL PROCEDURES**

**Plasmids**—pCMV-MAZ was constructed as described previously (13). MAZ-CAT reporter plasmids, namely pMAZCAT0, pMAZCAT1, pMAZCAT2, pMAZCAT3, pMAZCAT4, pMAZCAT5, and pMAZCAT6 were constructed by digesting the DNA of the promoter region with HindIII/XmnI, StuI/XmnI, MscI/XmnI, DpnI/PvuII, BssHII/XmnI, and HgaI/XmnI, respectively, and then subcloning, via the pHindIII linker, into the HindIII site of pSV00CAT (19). Similarly, deletion mutants, namely, pMAZCAT2-d and pMAZCAT3-d, were constructed by digesting the insert DNA of pMAZCAT2 or pMAZCAT3 with DpnI/PvuII or PvuII/BssHII, respectively, and subcloning into the HindIII site of the pSV00CAT vector.

**Screening of a Library of Human Genomic DNA**—A cosmid library (20) was constructed from the genomic DNA of the HL-A-homologous B-lymphoblastoid cell line AKIBA (A24, Bw52, Dw12, DQw1, and Cw63), which had been partially digested with Sau3AI, with subsequent ligation of fragments to the cosmid vector pWE15 (Stratagene, La Jolla, CA). The library was screened by colony hybridization with 1.8- and 0.7-kb EcoRI fragments of MAZ cDNA from pCMVMAZ (13) as probes (21). Filters were prehybridized at 68 °C for 30 min in a solution that contained 6 × SSC (1 × SSC: 150 mM sodium chloride, 15 mM sodium citrate, pH 7.0), 5 × Denhardt’s solution (0.1% Ficoll (type 400; Phar- macia LKB, Uppsala, Sweden), 0.1% polyvinylpyrrolidone, 0.1% bovine serum albumin (fraction V; Sigma)), 0.1% SDS, and 0.1 mg/ml denatured calf-thymus DNA and then hybridized in the same buffer supplemented with 1 × 10^6 cpm/ml of radioactive probe at 68 °C for 20 h. The hybridization probe for MAZ was prepared by the random primed labeling method (22) using a Random Primed DNA Labeling Kit (Boehringer Mannheim, Mannheim, Germany). Filters were washed three times at room temperature for 20 min in 2 × SSC, 0.1% SDS and once at 65 °C for 30 min in 0.2 × SSC, 0.1% SDS and then exposed to XAR-5 film (Eastman Kodak Co., Rochester, NY) with an intensifying screen. A 40-kb DNA insert of a cosmid clone was digested with EcoRI and then subcloned into the pBluescriptII SK+ vector (Stratagene) for further studies. The human MAZ-yeast artificial chromosome (YAC) recombinant clones, y645D4 and y976H4 were isolated from a library of CGM1 DNA (CHEF, Paris, France) by PCR-mediated methods according to the protocol provided by CHEF. The primers for YAC screening were described elsewhere (13, 18).
Nucleotide Sequencing—DNA sequencing was carried out by the dideoxy chain termination method (23) with an automated DNA se-
quencer (ABI 373A; Applied Biosystems Inc., Foster City, CA) using a DNA Sequencing Kit (Ref. 24; Dye Terminator Cycle Sequencing Ready
Reaction; Applied Biosystems Inc.). In the case of DNA fragments with an
unusual high G + C content, cycle sequencing was performed
according to the protocols provided by the manufacturer (Applied Bio-
systems Inc.). Nucleotide and deduced amino acid sequences were an-
alyzed with the GCG program (25).

Southern and Northern Blotting Analysis—High-molecular weight
DNA from human cells was extracted, digested, fractionated on a 1% gel,
transferred onto a nylon membrane, and described as described
(21). The DNA on the membrane was allowed to hybridize with the probe,
and the membrane was washed and exposed to Kodak XAR-5 film with an intensifying screen as described elsewhere (21). Multiple
tissue Northern blots were obtained from CLONTECH (human MTN II,
7759–1, and human MTN III, 7767–1; CLONTECH, Palo Alto, CA). The
blots were hybridized and washed before autoradiography as described
elsewhere (21). A 1.6-kb EcoRI fragment of cDNA for MAZ and a 0.6-kb
EcoRI/BamHI fragment of the human DNA (~1.0 kb to ~0.4 kb relative
to the major site of initiation of transcription) were radiolabeled for use
as DNA probes for further hybridization.

5′ Nuclease Assay—Total RNA from HeLa cells or human peripheral
blood lymphocytes (PBL) was prepared by the guanidine thiocyanate
method as described elsewhere (24). S1 nuclease protection was con-
ducted essentially as described previously (26) using a 208-nt probe (nt
95°C for 40 s, 60°C for 40 s and 72°C for 40 s. Sequences of primers for
MAZ and glyceraldehyde-3-phosphate dehydrogenase were as follows:
sense primer, 5′-TTCTTGTCAGCCGTCGTTG-3′; antisense primer, 5′-
CCTGGAATTTGACCTGTTG-3′ and sense primer, 5′-TCCACACCTCT-
GCTGGCTA-3′; antisense primer, 5′-ACACATCGTACCATCAC-
CACCA-3′.

Reverse Transcription (RT)-Polymerase Chain Reaction—The reac-
tion mixture for RT-PCR containing cDNA (equivalent to 20 or 25 ng of
total RNA), specific primer sets (5 pmol each), 50
was carried out in a total volume of 20 μl that contained 10 ng of each
primer, 10 ng of DNA, 200 μM each dNTP, ΔTh polymerase buffer
(Toyo, Kyoto, Japan), and 2.5 units of ΔTh polymerase (Toyo,
and then subjected to 30 cycles of 95°C for 40 s, 60°C for 30 s, and 72°C
for 30 s in a DNA thermal cycler (9600 model; Perkin-Elmer, Foster
City, CA). The products of PCR were separated by electrophoresis on a
9% polyacrylamide gel.

Reverse Transcription (RT)-Polymerase Chain Reaction—The reac-
tion mixture for RT-PCR containing cDNA (equivalent to 20 or 25 ng of
total RNA), specific primer sets (5 pmol each), 50 μM dNTPs, 1× buffer
for KOD Dash and 0.5 units of KOD Dash (Toyo, Osaka, Japan) was
subjected to a DNA thermal cycler. Samples were heated at 95°C for 2 min
to denature the template DNA and then subjected to 25 cycles of

RESULTS

Cloning and Determination of the Genomic Structure of the
Human Gene for MAZ—Screening of a cosmids library of the
generic DNA from human B-lymphoblastoid AKIBA cells with the
1.8- and 0.7-kb EcoRI fragment of the coding region and the
untranslated region of pMAZi cDNA, respectively (13), as probe
yielded seven clones. DNA from the seven cosmid clones was
digested with EcoRI and then allowed to hybridize with the
same 1.8-kb DNA probe. A single 9.0-kb EcoRI fragment was
detected in all seven cosmid clones and subcloned into Blue-
scriptII SK−. The complete restriction maps of the seven sub-
clones were identical. From one of the seven subclones, pJSMAZ9.0E, we further subcloned the 1.3-kb PstI-PstI fragment and the
3.5-kb PstI-EcoRI fragment into pBluescriptII SK− to generate pJSMAZ1.3P/P and pJSMAZ3.5P/E, respectively. A 1.1-kb EcoRI fragment adjacent to the inserted DNA of
pJSMAZ9.0E was also isolated and further subcloned into pBluescriptII SK− to generate pJSMAZ1.1E. These plasmids, pJSMAZ1.3P/P, pJSMAZ3.5P/E, and pJSMAZ1.1E were fully
sequenced and characterized (Fig. 1).

Intron-exon boundaries were determined from the divergence
of the genomic sequence from the sequence of the cDNA
for the human protein and the presence of consensus splice
donor-acceptor sequences was confirmed (32, 33). The locations
of five exons and four introns within the 6.0 kb of human
genomic DNA cloned in this study are shown in Fig. 1. Both the
5′- and the 3′-splice junctions and the sizes of exons and introns
are shown in Table I. The exons were small, with the exception of
exon 5, which extended over 1028 nucleotides. Similarly, the
introns were relatively small, ranging from 84 to 277 bp, with
the exception of intron 4, which extended over 1.5 kb of DNA.
The region that encoded the zinc finger motif of the MAZ
protein extended from exon 2 to exon 4 and was interrupted by
two introns. The 3′-untranslated region of the MAZ gene that
we isolated was about 1.1 kb in length (Figs. 1 and 2 and Table
1). We also isolated two human MAZ-YAC recombinant clones,
y654D4 and y976H4, from a YAC library of EB virus-transformed
human peripheral lymphocytes (CGM1). The two
cloned YACs spanned 560 kb of DNA that included the MAZ gene (Fig. 2). The NotI DNA fragment covered the entire genomic region of the MAZ gene and was subcloned into the pWE15 vector for generation of a physical map of the inserted DNA. The nucleotide sequence of the MAZ gene in the YAC clones was identical to that of the cloned cosmid DNA (data not shown).

**Sequence of the 5’-Flanking Region of the MAZ Gene**—We compared the genomic MAZ sequence with corresponding sequences of cDNAs for members of the MAZ family that had been reported previously (9–11). We found relatively limited homology in the 5’-flanking region even though the sequences of the coding regions of these genomic and cDNA clones were identical (Fig. 3).

To examine whether the human genome contains several MAZ genes, we designed 5’-end primers on the basis of the 5’-end upstream sequences of human genomic MAZ DNA and two representative cDNAs (see Fig. 3b; 13, 15). The 3’-end primers corresponded to the conserved sequences in the genomic MAZ gene. Then human total genomic DNA and mixtures of cDNA were used as templates for PCR. The primer combinations corresponding to the genomic MAZ sequences resulted in successful amplification, while with MAZi and clone 33813 cDNAs, we failed to amplify any DNA fragments (Fig. 3c). Thus, the differences in the 5’-flanking regions of both MAZi and clone 33819 cDNAs might have been due to artifacts in cloning due to the high G + C content of this region. Alternatively, differential splicing is also a possible explanation in the case of MAZi. Detailed characterization of the physical map and the nucleotide sequences of YAC-MAZ clones supported this conclusion (data not shown).

A particular striking feature of the MAZ gene was its high G + C content. The average G + C content of the 6.0-kb genomic MAZ DNA was 68.9%; that of the upstream 2.5-kb part of the 6.0-kb genomic MAZ DNA was 76.2% and that of the downstream 2.5-kb part of the 6.0-kb genomic MAZ DNA was 61.2%. The 500-bp region upstream of the ATG initiation codon had a G + C content of 88.4%. It is noteworthy that two regions of 71 and 77 bp (nt –103 to –33 and nt –306 to –230 relative to the ATG codon, respectively), had extremely high G + C contents of 98% and 93%, respectively.
The GC-rich region contained restriction sites for some rare “CG cutters,” such as NarI, BssHII, SrfI, and EcoI (Fig. 2b), which are characteristic of “CpG islands” (34–36). The 1.3-kb region at the 5′-end of the MAZ gene contained 171 copies of CpG and 196 copies of GpC dimers, whereas the 2-kb region at the 3′-end of the MAZ gene contained 165 GpC and only 70 CpG dinucleotides. A total of 45 HpaII sites were found in the MAZ gene (Fig. 2c). Thirty-four of them were found within the 3-kb region that encompassed the promoter and the first two exons of the MAZ gene, whereas only eight HpaII sites were found within a 3-kb region of the 3′-end (Fig. 2c). An abundance of HpaII sites, also referred to as HTF islands, in addition to the CpG islands, is characteristic of housekeeping genes (34–36). The presence of a CpG island immediately upstream of the ATG codon strongly suggests that this region contains the promoter sequence of the gene for MAZ.

Several consensus motifs were found in the promoter region of the MAZ gene. Numerous CT-tract motifs, namely, one (TCCT)2, one (TCC)3, one (TCC)2, and four (TCCC)3, were also detected at positions 1, 7, 9, and 92 bp were also detected (Fig. 5). These results indicated that transcription of the MAZ gene has features of the family of “housekeeping” genes (39). Moreover, the expression of the MAZ gene appeared to be ubiquitous and independent of cell type (see Fig. 8).

Characterization of the Site of Initiation of Transcription—The 5′-boundary of exon 1 was determined by S1 nuclease protection analysis of the SmaI/SmaI fragment (208 nt) of the promoter region of the MAZ genomic clone (Fig. 5). A strong signal corresponding to 106 bp was detected by the S1 protection analysis in the presence of total RNA from HeLa cells. In addition, five weak signals corresponding to 100, 98, 97, 93, and 92 bp were also detected (Fig. 5). These results indicated that transcription of the MAZ gene started at multiple sites, namely, at positions +1, +7, +9, +10, +14, and +15. All of these sites are located within a CT repeated sequence (see Fig. 4).

MAZ is a Single-Copy Gene Located on Chromosome 16p11.2—Southern hybridization analysis using DNA from HeLa cells and human PBL showed that a single 9.0-kb EcoRI fragment hybridized with the MAZ-specific cDNA probe (Fig. 6). When DNA was digested with BamHI, we detected three major MAZ fragments, of 4.4, 2.1, and 1.6 kb, respectively, using human PBL and HeLa cells DNAs. These results suggested that MAZ might be encoded by a single, unique gene. Digestion with BamHI revealed common 4.4-, 2.1-, and 1.6-kb fragments in WI-38, MKN7, MKN28, HGC27, GCY1, MKN45, PBL, and HeLa cells. The weak band of the EcoRI-generated 7.2-kb fragment might represent cross-hybridization with the DP-1 gene, which exhibits weak sequence homology to the gene for MAZ (Ref. 40; data not shown). The reciprocal hybridization of human PBL DNAs with the genomic BamHI fragments of MAZ DNA as probes yielded distinct fragments of 4.4, 2.1, and 1.6 kb (Fig. 6c). These results suggest that MAZ is encoded by a single, unique gene. The physical mapping of human-YAC recombinant clones confirmed the presence of a single, unique gene for MAZ (data not shown).

We next attempted to determine the chromosomal location of the human MAZ gene by in situ hybridization with the MAZ cosmid clone on spreads of replicated prometaphase chromosomes that had been prepared from phytohemagglutinin-stimulated normal human male lymphocytes. As shown in Fig. 7, fluorescent spots were observed on chromosome 16 and this result was reproducible. To map the MAZ gene with greater precision, we performed a series of FISH experiments using various human-YAC clones that encompassed the MAZ region. These results indicated that MAZ is located within a 3-kb region of the 3′-end of the MAZ gene (data not shown).
FIG. 4. Putative promoter elements in the 5′-untranslated region of the human MAZ gene. The nucleotide sequence (GenBank accession number D89880) is numbered with the major site of initiation of transcription as +1 (several detected sites are indicated by bent arrows). The GC repeats ((GC)₄, (GC)₅, (CCG)₃, and (CCG)₈) and CT tracts are shown by heavily and lightly outlined boxes, respectively. Sp1- and AP-2-binding sites are underlined by dotted lines and uninterrupted lines, respectively. The two G + C stretches are indicated by wavy lines above the sequence. The ATG codon for initiation of translation is indicated in boldface.

accuracy, we compared the immunofluorescence micrographs and the Q-banding patterns of the same cells. This analysis clearly indicated that the human MAZ gene mapped to band 16p11.2. To confirm this result, we performed a similar experiment using the immunogold detection method, with biotinylated DNA derived from the MAZ gene as the probe (41). The silver grains were concentrated on the same region of chromosome 16 (data not shown).

Characterization of the Three Major Transcripts—We examined the tissue distribution of MAZ transcripts using human multitissue Northern blots (Fig. 8). Transcripts were identified in all the tissues examined, albeit at different levels. We detected three mRNAs, of 1.6, 2.7, and 4.6 kb, respectively, with the latter two transcripts being major species. The transcripts of 2.7 and 4.6 kb were present in all tissues examined but in the liver, in particular, the level of the 2.7-kb transcript was very low. The levels of the 2.7- and 4.6-kb transcripts in the heart, placenta, pancreas, spleen, prostate, colon, peripheral blood leukocytes, thyroid, and adrenal gland were higher than those in other tissues. The 1.6-kb transcript was detected mainly in the heart, placenta, pancreas, spleen, prostate, colon, thyroid, spinal cord, trachea, and adrenal gland.

Expression of MAZ during the Cell Cycle—As shown in Fig. 9a, the level of expression of MAZ protein was modulated by the cell cycle. The 60-kDa MAZ protein was produced during the G₀ and G₂ phases, but it was not detected during early S phase (Fig. 9a). When serum-starved normal diploid WI-38 cells were stimulated by addition to the medium of a high concentration of serum, the level of expression of MAZ appeared to be reduced at 12 h; the protein disappeared at 30 h and then it reappeared at 36 to 48 h (Fig. 9a). The level of expression of human HLA-associated β₂-microglobulin was not changed significantly during the cell cycle (Fig. 9a). These results indicated that the level of MAZ protein was modulated in a cell cycle-dependent manner. At early S phase, in particular, we were unable to detect the expression of MAZ protein. We next examined the level of MAZ mRNA by RT-PCR during cell cycle and found the similar changes as that of MAZ protein. By contrast, the level of expression of glyceraldehyde-3-phosphate dehydrogenase was unaltered (Fig. 9b). The CAT activities of WI-38 cells transformed with the pMAZCAT0 reporter construct reflected the variations in the level of the MAZ protein during the cell cycle (see Fig. 9c). Furthermore, the introduction of a MAZ expression vector into HeLa cells significantly and dose dependently enhanced the promoter activity of the pMAZCAT0 reporter plasmid (Fig. 9d). Thus, expression of the MAZ gene appears to be controlled during the cell cycle and to be regulated by the MAZ protein itself.

Characterization of the Promoter—To examine whether the 5′-flanking region of the human MAZ gene contained a functional promoter, we generated a series of chimeric MAZ promoter/CAT gene reporter constructs, as depicted schematically in Fig. 10. These constructs were used to transfect HeLa cells to characterize the regulatory elements of the human MAZ pro-
Promoter activity was retained after deletion to position −383 relative to the major site of initiation of transcription (pMAZCAT3; Fig. 10a). Further deletion to positions −248, −189, and −40 (pMAZCAT4, pMAZCAT5, and pMAZCAT6) resulted in significant decreases in promoter activity. Moreover, the internal deletion mutant pMAZCAT2-d (with a deletion from nt −383 to −248) had diminished promoter activity. Therefore, DNA sequences between nt −383 and −248 appeared to be required for high basal promoter activity in HeLa cells. Negative elements might be present between nt −948 and −2500 because the activity due to pMAZCAT0 was significantly repressed as compared with that due to pMAZCAT1. These results clearly demonstrated that the 5′-end flanking region contained the functional promoter of the human gene for MAZ.

As shown in Fig. 9, a and b, the levels of MAZ protein and MAZ mRNA and expression of the MAZ gene varied in a cell-cycle-dependent manner. Moreover, increased concentrations of the MAZ expression vector resulted in significant induction of the CAT reporter activity of the MAZ-CAT gene (Fig. 9d). Therefore, we examined whether the MAZ protein might be able to regulate the expression of its own gene. We co-transfected cells with a MAZ-CAT reporter construct and a MAZ expression plasmid (pCMV-MAZ) and monitored expression of the MAZ gene. As shown in Figs. 10c, the CAT activity due to the pMAZCAT4 construct was increased significantly by cotransfection with pCMV-MAZ, while that of the deletion construct pMAZCAT3-d was not, suggesting that an element between nt 248 and 189 might be involved in positive enhancement of the regulation of expression of the MAZ gene. The CAT activity due to pMAZCAT2 was lower than that due to the deletion mutant pMAZCAT2-d in the presence of pCMV-MAZ, suggesting that the deleted region (nt −383 to −248) might play a role in down-regulation of the MAZ gene by the MAZ protein. The regulatory sequences seemed to overlap one another.

The CAT activity due to pMAZCAT0 was enhanced significantly in the presence of MAZ protein, whereas that of pMAZCAT1 was not. Thus, the regions from nt −248 to −189 and from nt −2500 to −948 appear to contain putative positive regulatory elements, in contrast to the region from nt −383 to −248 that seems to contain negative regulatory elements.

**DISCUSSION**

The present study revealed the exon-intron structure of the human gene for MAZ that spans approximately 6.0 kb and consists of promoters, five exons, four introns, and a 3′-untranslated region. In addition, physical mapping studies of MAZ-YAC clones demonstrated that the MAZ gene is a single and unique gene. All exon-intron boundaries begin with GT at the 5′-end and terminate with AG at the 3′-end, conforming to the GT-AG rule (42). When compared with three reported cDNAs for human MAZ, the insert in our clone was most similar to clone 33819 from HeLa cells (the coding region was the same as that of clone 33819). The major differences were found in the 5′-end promoter region (see Fig. 3). The corresponding region of the gene for MAZ from human islets is missing about 35 nucleotides (G + C; from nt +69 to +128). From the results of Southern blotting and RFL-PCR with cosmids and YAC recombinant clones, it appeared that MAZ might be encoded by a single gene (Figs. 3 and 6). The heterogeneity of the 5′-end promoter regions of cDNA sequences for MAZ might be due to artifacts that arose because of the high G + C content. Alternatively, differential splicing might have occurred in the case of the MAZ gene from human islets.

Multiple sites for initiation of transcription were found within 174 bp upstream of the site for initiation of translation by the S1 nuclease protection assay (Fig. 5). These sites are located in a putative initiator sequence near the major site for initiation of transcription that matches the 5′-YYC(A/T)YYYYY-3′ (Y, pyrimidine) consensus sequence (43). Furthermore, the results of a transcription experiment in vitro showed that the sequence in the vicinity of the defined site of initiation of transcription was sufficient to promote faithful transcription (data not shown). We tried to confirm the site of initiation of transcription in a primer extension assay. However, we were unsuccessful since the reverse transcriptase did not read through the 5′-end regions with a high G + C content. Furthermore, the MAZ promoter-CAT construct also demonstrated the significant activity of the promoter (Figs. 9 and 10). Taken together, our results demonstrate that transcription of the MAZ gene originated from a TATA-less promoter in vivo and in vitro and that the 5′-end region really contained the promoter region of the MAZ gene required to generate a 2.7-kb mRNA. The exact nature of the transcription factors and the specificity of expression of the MAZ gene from this promoter region remain to be determined.

Sequence analysis of the first exon and the 5′-upstream region suggested that this region had an unusually high G + C content. In particular, the average G + C content of the 500-bp region upstream of the ATG codon was 88.4%. It is noteworthy that two regions, extending over 71 and 77 bp (nt −103 to −33 and nt −306 to −230 relative to the ATG codon, respectively), had extremely high G + C contents of 98.6 and 97.4%, respectively (Fig. 4). The 1.3-kb fragment of the 5′-end boundary of the MAZ gene contained 171 copies of CpG and 196 copies of GpC islands. Furthermore, G + C-rich sequences were located before and after the cap site (+1) of the MAZ gene, within 280 bp. These two long G + C-rich sequences might contribute to changes in DNA conformation and might be modified, for example, by methylation (34–36). Such changes might explain the strength of the promoter of the MAZ gene under certain conditions and in different types of cells.
CT tracts were present in the region upstream of the small 77-bp region with high G + C content. Such CT tracts are strongly reminiscent of structures described in the gene for the receptor of epidermal growth factor (EGFR) (44) and in other promoters, such as those of the ets-2 (45), c-Ki-ras (46), and c-myc (4, 13) genes. Such repeats are thought to form triplex or H-DNA structures that are sensitive to a variety of nucleases (47). In an examination of CT tracts in the gene for epidermal growth factor receptor, Johnson et al. (44) showed that a specific factor (CTF) binds these sequences and that deletion of the CT tracts significantly down-regulates transcription. We reported similarly that a MAZ protein binds the CT tracts of the c-myc promoter region (13).

(CCG)$_n$ repeats were present in the 71-bp GC-rich stretch and several (CCG)$_n$, (GC)$_m$, (GC)$_n$ repeats were also present in the 500-bp long GC-rich region. GC-rich promoters and the absence of a TATA sequence are characteristic of housekeeping genes (34–36, 43). In fact, Northern blotting analysis revealed the expression of MAZ transcripts in a large variety of human tissues. Examination of the DNA sequence upstream of exon 1 failed to reveal any TATA box or CAAT box. Similar findings have been obtained for a variety of oncogenes, such as human ets-1 (48), human ets-2 (45), human fgr (49), human c-src (50) and murine c-Ki-ras (46); for genes for growth factors and their receptors, such as human epidermal growth factor receptor (44) and insulin-like growth factor receptor (51); and for housekeeping genes, such as the gene for adenosine deaminase (52). The promoters of such genes have a number of common characteristics, such as the presence of multiple sites for initiation of transcription.

Fig. 6. Southern blotting analysis of human genomic DNA. a, Southern blots of EcoRI digests of human genomic DNAs from PBL and HeLa cells. b, Southern blots of the BamHI digests of human genomic DNAs isolated from PBL, HeLa, WI-38 (lane 1), MKN7 (lane 2), MKN28 (lane 3), MKN45 (lane 4), HGC27 (lane 5), and GCV1 (lane 6) cells. c, hybridization of a BamHI digest of human PBL genomic DNA with the various BamHI-generated DNA fragments from MKN7 cells.

d, schematic representation of the BamHI-generated DNA fragments used for the hybridization in c. The sizes of fragments are shown on the right in c. The DNA probes used for the hybridization are indicated.

Fig. 7. Fluorescence in situ hybridization mapping with a cosm ide clone of the MAZ gene at chromosome 16p11.2. a, the fluorescein isothiocyanate-labeled cosm id MAZ clone was used for fluorescence in situ hybridization of metaphase chromosomes, prepared from phytohemagglutinin-stimulated peripheral blood lymphocytes. b, Q-banding pattern of the same chromosome as shown in a. Arrows indicate the positive signal at chromosome 16p11.2. c, schematic representation of the loci of the MAZ gene at chromosome 16p11.2. A large and a small dot indicate 10 and 1 spot, respectively.

Fig. 8. Northern blotting analysis of expression of MAZ mRNA. Commercially purchased multiple-tissue Northern blot filters, with approximately 2.0 µg/lane of poly(A)$^+$ RNA from each indicated tissue, were allowed to hybridize with a 1.8-kb $^{32}$P-radiolabeled DNA fragment of a cDNA clone for human MAZ (13).
transcription, presumably because of the absence of a TATA box and a CAAT box, and they often have an unusually high G + C content. We found 16 consensus binding sites for Sp1 and 26 binding sites for AP-2 that were located close together or as partially overlapping sites in the 5'-end promoter region. There were also several potential binding sites for the epidermal growth factor receptor-specific transcription factor (ETF) in the promoter region. This protein is thought to replace TATA-FIG. 9. Expression of MAZ protein and promoter activity of the MAZ gene during the cell cycle. a, MAZ protein was detected by Western blotting analysis with a lysate of synchronized WI-38 cells and antibodies specific for MAZ or human β2-microglobulin. Cells were synchronized as described under “Experimental Procedures.” Cells reinitiated the cell cycle after G0 arrest and were harvested after a further 1, 6, 12, 18, 24, 30, 36, and 48 h. Upper panel, MAZ protein; lower panel, human β2-microglobulin. b, equivalent of 25 μg of total RNA obtained from synchronized WI-38 cells were examined by RT-PCR. Upper panel, MAZ mRNA; lower panel, glyceraldehyde-3-phosphate dehydrogenase mRNA.

cat, expression of the CAT reporter gene in WI-38 cells stably transformed with the pMAZCAT0 plasmid (10 μg) at indicated times after addition of a high concentration of serum. Percent conversion of acetyltransferase is indicated. d, the promoter activity of the pMAZCAT0 reporter gene (5 μg) was enhanced by addition of increasing amounts of the pCMV-MAZ expression vector. Lane 1, none; lane 2, 0.1 μg; lane 3, 1.0 μg; lane 4, 5.0 μg; lane 5, 10 μg of pCMV-MAZ gene. Percent conversion of acetyltransferase is indicated.

Fig. 10. Functional activity of the human MAZ promoter. a, summary of fusions of the CAT gene with the gene for human MAZ and transient expression of CAT under control of the MAZ promoter. Open boxes represent the CAT gene. The names of the constructs are shown in the middle. Numbering is relative to the major site of initiation of transcription. Plasmid DNAs (10 μg) were used to transfect HeLa cells and CAT activity was measured as described under “Experimental Procedures.” b, CAT fusions with the human MAZ gene and transient expression of CAT driven by the MAZ promoter (5 μg) after co-transfection with pCMV-MAZ (10 μg). Plasmid DNA was used to transfect HeLa cells and CAT activity was measured as described under “Experimental Procedures.” Percent conversion of acetyltransferase is indicated. c, promoter activities of MAZ-CAT fusion genes are expressed relative to the activity of pSV00CAT (5 μg) in the absence of pCMV-MAZ, which was taken arbitrarily as 1.0. All values are the averages of results from at least three experiments and the standard deviation for each value is indicated.
binding proteins in the control of expression of genes that lack TATA boxes (53).

We identified mRNAs of 1.6, 2.7, and 4.6 kb, respectively. It is of interest that only the 4.6-kb mRNA, was detected when the EcoRI/BamHI DNA fragment that corresponded to the promoter region of the MAZ gene (−1.0 to −0.4 kb relative to the site of initiation of transcription) was used as DNA probe (data not shown). Thus, it is likely that the transcription of the 4.6-kb mRNA is initiated from the far upstream promoter. Studies of transcription and promoter activities in vitro and nucleotide sequencing indicated that each transcript, namely, the 4.6, 2.7, and 1.6 kb transcripts, might have been generated by corresponding independent promoters, which the possibility of differential splicing or differential polyadenylation seems less likely. However, we failed to isolate a full-length cDNA clone that corresponded to the 4.6-kb mRNA by molecular cloning, probably because of the unusual high G + C content of this region, as described above.

The 60-kDa MAZ protein was detected during the G₀ and G₂ phases but not during the early S phase. When the cells were stimulated to reinitiate the cell cycle from the G₀ phase by addition of a high concentration of serum, the expression of MAZ was down-regulated at 12 h, was lowest at 30 h and increased from 36 to 48 h (Fig. 9c). The similar changes in the levels of MAZ mRNA were obtained during the cell cycle (Fig. 9b). These data indicated that expression of the MAZ gene was controlled by the cell cycle, as was the activity of the MAZ promoter (Fig. 9c).

In order to identify the elements that regulate the basal transcription of the MAZ gene, we constructed a series of deletion mutants of a MAZ-CAT fusion gene and transfected HeLa cells with them (Fig. 10). We identified a positive control clone that corresponded to the 4.6-kb mRNA by molecular cloning, probably because of the unusual high G + C content of this region, as described above.

The region for negative control of basal transcription was localized in the far upstream region between nt −2500 and −948. We also showed that the expression of the MAZ gene is controlled by its own product, the MAZ protein. Positive elements for autoregulation by the MAZ protein were putatively identified in the proximal region from nt −248 to −189 and in the distal region from nt −2500 to −948. The former proximal region contains two consensus (TCCC) elements and Sp1 and AP-2-binding sites. Negative autoregulatory elements were found in the region between nt −383 and −248 that is adjacent to the positive element. The regulatory elements for autoregulation by the MAZ protein had a reciprocal relationship to the regulatory elements for basal transcription of the MAZ gene: the enhancer region for basal transcription was “shut off” and the region for negative control of basal transcription was “turned on” by the product of the MAZ gene. This scenario is supported by the observation that forced expression of the MAZ gene resulted in significant and dose-dependent enhancement of the promoter-CAT activities of the pMAZCAT0 construct (Fig. 9d). We do not yet know the functional significance of these elements in autoregulation. We are currently trying to identify the transcriptional factors that are involved in autoregulation of the MAZ gene.

Acknowledgments—We thank Drs. C. Geltinger, T. Murata, N. Adachi, T. Koga, and H. Ugaı for many helpful discussions and critical reading of the manuscript.

APPENDIX

The figure summarizes the CAT fusion genes with the gene for human MAZ (Fig. s1).

---

**REFERENCES**

1. Cole, M. D. (1986) Annu. Rev. Genet. 20, 561–571
2. Spencer, C. A., and Groudine, M. (1991) Adv. Cancer Res. 56, 1–48
3. Blackwood, E. M., and Eisenman, R. N. (1991) Science 251, 1211–1217
4. Bossone, S. A., Asselin, C., Patel, A. J., and Marcu, K. B. (1992) Annu. Rev. Biochem. 61, 809–860
5. Desjardins, E., and Hay, N. (1993) Mol. Cell. Biol. 13, 5710–5724
6. Moberg, K. H., Logan, T. J., Tyndall, W. A., and Hall, D. J. (1992) Oncogene 7, 411–421
7. Moberg, K. H., Tyndall, W. A., Pyrc, J., and Hall, D. J. (1991) J. Cell. Physiol. 148, 75–84
8. Ashfield, R., Patel, A. J., Bossone, S. A., Brown, H., Campbell, R. D., Marcu, K. B., and Proudfit, N. J. (1994) EMBO J. 13, 5656–5667
9. Bossone, S. A., Asselin, C., Patel, A. J., and Marcu, K. B. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 7452–7456
10. Pyrc, J. J., Moberg, K. H., and Hall, D. J. (1992) Biochemistry 31, 4102–4110
11. Kennedy, G. C., and Rutter, W. J. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 11498–11502
12. Kovesdi, I., Reichel, R., and Nevin, J. B. (1988) Cell 55, 219–228
13. Tsutsui, H., Sakatsume, O., Itakura, K., and Yokoyama, K. K. (1996) Biochem. Biophys. Res. Commun. 226, 801–809
14. Parks, C. L., and Shenk, T. (1997) J. Virol. 71, 9600–9607
15. Duncan, D. D., Stupakoff, A., Hedrick, S. M., Marcu, K. B., and Siu, G. (1995) Mol. Cell. Biol. 15, 3179–3186
16. Parks, C. L., and Shenk, T. (1996) J. Biol. Chem. 271, 4417–4430
17. Bockamp, E.-O., McLaughlin, F., Myrrell, A. M., Goettgen, B., Robb, L., Begley, C. G., and Green, A. R. (1995) Blood 86, 1502–1514
18. Komatsu, M., Li, H., Tsutsui, H., Itakura, K., Matsumura, M., and Yokoyama, K. K. (1997) Oncogene 15, 1123–1131
19. Araki, E., Shimada, F., Shichiri, M., and Ehina, Y. (1988) Nucleic Acids Res. 16, 1627
20. Kawai, J., Ando, A., Sato, T., Nakatsui, T., Tsuchi, K., and Inoko, H. (1989) ImmunoL 142, 312–317
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
22. Feinberg, A. P., and Vogelstein, B. (1983) Anal. Biochem. 133, 6–13
23. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
24. Gou, L. H., and Wu, R. (1983) Methods Enzymol. 100, 60–96
25. Altschul, S. F., Gish, W., Miller, W., and Lipman, D. J. (1990) J. Mol. Biol. 215, 403–410
26. Ancelet, P. M., Brent, R., Kington, R. B., Moore, D. D., Smith, J. A., Seidman, J. G., and Struhl, K. (1987) Current Protocols in Molecular Biology, Wiley, New York
27. Ozawa, K., Murakami, Y., Eki, Y., Soeda, E., and Yokoyama, K. (1992) Genomics 12, 214–220
28. Lichter, P., Chang-Tang, C. J., Call, K., Hennansson, G., Evans, G. A., Housman, D., and Ward, D. C. (1990) Science 247, 64–69
29. Goldman, C. M., Moffat, L. F., and Howard, B. H. (1982) Mol. Cell. Biol. 2, 1044–1051
30. Kitabayashi, I., Kawakami, Z., Chin, R., Ozawa, K., Matsuoka, T., Toyoshima, S., Umesono, K., Evans, R. M., Gachelin, G., and Yokoyama, K. (1992) EMBO J. 11, 167–175
31. Kabot, S., Ozawa, K., Kondoh, S., Soeda, E., Israel, A., Shiriski, K., Fujinaga, K., Ikura, K., Gachelin, G., and Yokoyama K. (1990) EMBO J. 9, 127–135
32. Sharp, P. A. (1981) Cell 23, 443–454
33. Breathnach, R., and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349–383
34. Bird, A. P. (1986) Nature 321, 209–213
35. Lindsay, S., and Bird, A. P. (1987) Nature 327, 336–338
36. Lavia, P., Macleod, D., and Bird, A. (1987) EMBO J. 6, 2773–2779
37. Courey, A. J., and Tjian, R. (1988) Cell 55, 877–896
38. Faist, S., and Meyer, S. (1992) Nucleic Acids Res. 20, 5–26
39. Sehgal, A., Patel, N., and Chaos, M. (1988) Mol. Cell. Biol. 8, 3160–3167
40. Koyano-Nakagawa, N., Nishida, J., Baldwin, D., Arai, K., and Yokota, T. (1994) Mol. Cell. Biol. 14, 5099–5107
41. Bhattacharyya, B., Burns, J., Plannery, D., and McGel, D. O. (1988) Nucleic Acids Res.
Stephen, R. M., and Schneider, T. D. (1992) J. Mol. Biol. 228, 1124–1136
43. Azizkhan, J. C., Jensen, D. E., Piem, A. J., and Wade, M. (1993) Crit. Rev. Eukaryotic Gene Exp. 3, 229–254
44. Johnson, A. C., Jinno, Y., and Merlino, G. T. (1988) Mol. Cell. Biol. 8, 4174–4184
45. Mavrothalassitis, G. J., Watson, D. K., and Papis, T. S. (1990) Oncogene 5, 1337–1242
46. Hoffman, E. K., Trusko, S. P., Morphy, M., and George, D. L. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 2705–2709
47. Pestov, D. G., Dayn, A. Y. Siyanova, E., George, D. L., and Mirkin, S. M. (1991) Nucleic Acids Res. 19, 6527–6532
48. Jorcyk, C. L., Watson, D. K., Mavrothalassitis, G. J., and Papis, T. S. (1990) Oncogene 6, 323–332
49. Patel, M., Leevers, S. J., and Brickell, P. M. (1990) Oncogene 5, 201–206
50. Bonham, K., and Fujita, D. J. (1993) Oncogene 8, 1973–1981
51. Evans, T., DeChiara, T., and Efstatiadis, A. (1988) J. Mol. Biol. 199, 61–81
52. Aronow, B., Lattier, D., Silliger, R., Dasing, M., Hutton, J., Jones, G., Stock, J., McNeish, J., Potter, S., Witte, D., and Wigington, D. (1989) Genes Dev. 3, 1384–1400
53. Kageyama, R., Merlino, G. T., and Pastan, I. (1989) J. Biol. Chem. 264, 15508–15514