BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **245,** 120–126 (1998)
ARTICLE NO. RC988394

Identification of Alternatively Spliced Transcripts Encoding Murine Macrophage Colony-Stimulating Factor

Shinya Suzu, Kiyohiko Hatake,* Jun Ota, Yuji Mishima, Muneo Yamada, Seiichi Shimamura, Fumihiko Kimura,† and Kazuo Motoyoshi†

*Biochemical Research Laboratory, Morinaga Milk Industry Co. Ltd., Higashihara 5-1-83, Zama, Kanagawa 228, Japan; *Division of Hematology, Department of Internal Medicine, Jichi Medical School, Kawachi, Tochigi 329-04, Japan; and †Third Department of Internal Medicine, National Defense Medical College, Tokorozawa, Saitama 359, Japan*

Received February 25, 1998

We have isolated a novel cDNA encoding macrophage colony-stimulating factor (M-CSF) from a murine stromal cell line, ST2. The cDNA included an entire coding sequence of the M-CSF gene but contained an additional sequence of 140 base pairs (bp). Northern blot analysis demonstrated that other murine cell lines such as a fibroblastic cell line (L) and a stromal cell line (PA6) also expressed the transcripts corresponding to the clone. The nucleotide sequence analyses of the cDNA and the cloned M-CSF genome revealed that the 140-bp insertion sequence was part of intron 1 which separated exon 1 and exon 2: the former contained part of the amino acid residues of the signal sequence and the latter the rest of the signal sequence and the first 22 amino acid residues of the mature protein. The insertion of the 140-bp intron sequence not only changed the amino acid sequence of the signal peptide but also generated an in-frame termination codon. However, instead of the dysfunction of the original initiation codon, the 140-bp insertion sequence contained a putative ATG initiation codon that preserved the original open reading frame. Finally, we found that the cDNA directed the expression of a secreted and biologically active M-CSF protein when it was introduced into COS7 cells and M-CSF activity in the culture supernatants was measured using an M-CSF-dependent cell line. These results indicate the presence of an alternatively spliced M-CSF transcript which utilizes an alternate initiation codon in order to specify active M-CSF protein. © 1998 Academic Press

Macrophage colony-stimulating factor (M-CSF) is a regulatory cytokine that stimulates the proliferation, differentiation, and survival of monocytes/macrophages and their committed progenitor cells (1). The expression of M-CSF receptor and the action of M-CSF have been thought to be restricted to monocyte/macrophage lineage in hematopoietic systems (2, 3). However, we and other investigators recently reported that M-CSF had an ability to develop bone marrow fibroblastic stromal cells (4-6). Although it remains to be elucidated whether or not M-CSF acts directly on the progenitor of the stromal cells, this is a novel function of M-CSF (5, 6).

The human M-CSF gene is comprised of 10 exons and 9 introns spanning 20 kilo bases (kb) of the genome (7), which is presumably conserved in murine systems (8). Although M-CSF is transcribed from a single gene (7, 9), it is well known that multiple splice variants of the transcript are produced from the gene (7, 9-11). For example, in the original isolation of human M-CSF cDNA from a pancreatic carcinoma cell line, phorbol myristate acetate was shown to induce at least seven M-CSF transcripts ranging from 1.5 to 4.5 kb (10). At least four different forms of M-CSF cDNA have been isolated and sequenced (7, 9-13). In addition to the complexity in the transcriptional regulation of the M-CSF gene, the translational and posttranslational modification of the primary products derived from the multiple species lead to the production of several different protein forms, e. g., a secreted and a membrane-associated form.

The secreted form is encoded by two transcripts of 2.5 and 4.0 kb, which differ only in their 3′ noncoding region: the 2.5 kb- or 4.0 kb transcript uses exon 9 or 10, respectively, of the M-CSF gene (7, 9, 11-15). They contain the complete exon 6 encoding an amino acid sequence which should be cleaved by proteases, leading to the rapid secretion of the protein from the cell (14, 15). In spite of the presence of a single open reading frame, these transcripts specify at least two protein species, i. e., an M-CSF protein with the Mr of 85,000 and that with the Mr greater than 200,000 (11, 12, 14-16). The latter was uniquely modified by a chondroitin sulfate glycosaminoglycan chain (16-18). Due to this modification, this type of M-CSF can bind to type V collagen, and this might allow the species to associate with extracellular matrix (19, 20). The other two transcripts of 1.6 and 3.1 kb encoded the M-CSF protein

0006-291X/98 $25.00

associated with the cell membrane (7, 10, 12, 21). These transcripts again differ only in the alternate use of the 3′ nonconding region, i. e., exon 9 or 10 (7, 10, 12, 21). In these transcripts, the splicing out of the large segments of exon 6 deletes a region of 298 amino acid residues containing the site at which proteolytic cleavage occurs before the secretion of the active M-CSF described above. Therefore, the product from the transcripts is stably expressed at the cell surface and is also biologically active (14, 22).

Although the 4.0 kb form is predominant in many cells (7, 9-12), all the available evidence suggests to us that there are splice variants which are yet to be fully characterized. The cloning of minor splice variants is therefore necessary to clarify the significance of such species, if any. In this paper, we report the identification and characterization of a novel alternatively spliced M-CSF transcript.

EXPERIMENTAL PROCEDURES

Cell culture. Murine bone marrow stromal cell lines, ST2 (23) and MC3T3-G2/PA6 (PA6; 24), and the murine fibroblastic cell line L were obtained from the Riken Cell Bank (Tsukuba, Japan). They were routinely maintained in DMEM medium (Life Technologies, Grand Island, NY) supplemented with 10% (v/v) heat-inactivated fetal calf serum (FCS). Monkey kidney COS7 cells (Riken Cell Bank) were also routinely cultured with DMEM medium containing 10% FCS. Murine M-NFS-60 cells (25), which are dependent on the presence of M-CSF for their proliferation and were used for the determination of M-CSF activity in this study, were obtained from the American Type Culture Collection (ATCC, Rockville, MD). The cells were routinely maintained in RPMI1640 medium (Life Technologies) containing 10% FCS and recombinant human M-CSF at a concentration of 100 ng/ml (26).

Polymerase chain reaction (PCR). Total RNA was isolated from a subconfluent culture of ST2, PA6, or L cells by using RNA zol B reagent (TEL-TEST, Friendswood, TX) (27). One microgram of total RNA was reverse transcribed by using 2.5 mM random 9 mers, 2.5 U Avian Myeloblastosis Virus reverse transcriptase, and reagents from a commercial RNA PCR kit (RNA LA PCR Kit; TaKaRa Shuzo, Kyoto, Japan) (28). The PCR amplification was performed by using LA Taq polymerase (TaKaRa Shuzo) and according to the recommendations of the manufacturer (28). For the amplification of the entire coding region of M-CSF, two primers (P1 and P2) were generated, which were based on the previously described murine cDNA sequences (12). P1 hybridized to nucleotide sequences 139-163 in exon 1, and its sequence was 5′-TGCCGGGACCCAGCTGCCCGTATGA-3′. P2 was anti-sense to nucleotide sequences 1801-1825 in exon 8, and its sequence was 5′-TCCTTTCTATACTGGCAGTTCCACC-3′. Thus, the primer pair P1/P2 would generate at least two fragments which were derived from the transcript containing the complete exon 6 and that containing the truncated exon 6 (12). The PCR conditions were as follows: 30 sec of denaturation at 94°C, 30 sec of annealing at 60°C and 2 min of extension at 72°C (28). Genomic DNA was also isolated from ST2 cells by using a reagent containing SDS and protease (GENOME DNA Isolation Kit; BIO 101, Vista, CA) (29). The PCR amplification of the region including intron 1 of M-CSF genome was performed by using the DNA as the template, LA Taq polymerase (TaKaRa Shuzo), and two oligonucleotide primers (P1 and P3). The primer P1 hybridized to nucleotides in exon 1 as described above. P3 was anti-sense to nucleotide sequences 226-250 in exon 2; its sequence was 5′-TACTCCTGCTCATGAGGAGACAGAC-3′ (12). The PCR conditions were as follows: 20 sec of denaturation at 98°C, and 3 min of annealing and extension at 68°C (28). The analysis of PCR products was performed by comparing the predicted PCR fragment length with the actual PCR products after the ethidium bromide staining of products separated by electrophoresis on a 1 or 2% Seakem GTG agarose gel (FMC BioProducts, Rockland, ME).

Cloning and sequencing of PCR products. The PCR products were ligated directly into a pCR 2.1 vector (Invitrogen, Carlsbad, CA) (30) and sequenced by the dye terminator cycle sequencing method (Perkin Elmer, Foster City, CA). Sequencing was carried out in both orientations of the cDNAs, using an M13 reverse primer, M13 forward primer (29), and appropriately synthesized primers which hybridized to the sequences of the templates.

Northern blot analysis. Total RNA was prepared as described above. The RNA (10 μg/lane) was fractionated by electrophoresis through a 1.5% agarose gel containing 2.2 M formaldehyde (29). The size-fractionated RNA was transferred from the gel to a nylon membrane (Hybond N+; Amersham, Buckinghamshire, England) and fixed by baking. The resulting filter was probed with cDNAs labeled with [^{32}P]dCTP (ICN Biomedicals, Irvine, CA), using the random oligonucleotide priming reaction (MegaPrime DNA Labelling Systems; Amersham). The hybridization was performed at 65°Cusing rate-enhanced hybridization solution (Rapid-hyb buffer; Amersham). The filters were washed with 2x SSC (1x SSC = 0.15 M NaCl/0.015 M sodium citrate) containing 0.1% SDS at room temperature, and the final washings were done with 0.1× SSC containing 0.1% SDS at 65°C (29). The relative mobilities of different species of M-CSF mRNA were visualized by exposure of the filters to Hyperfilm MP (Amersham).

Transfection and analysis of M-CSF. The cDNA fragments containing a putative open reading frame of M-CSF were subcloned into a mammalian expression vector, pCEF. The vector is a derivative of pRc/CMV (Invitrogen, San Diego, CA), which we made by the inserting HindIII/XbaI fragment of a vector pEF-BOS containing the promotor region of polypeptide chain elongation factor-1 (31) into the pRc/CMV vector. The resulting plasmid DNAs were transfected into COS7 cells by using LipofectAMINE reagent (GIBCO BRL, Gaithersburg, MD) according to the manufacturer's instructions (32). The media were removed 72 hrs later and assayed for the concentration of M-CSF by measuring the proliferation of M-CSF-dependent M-NFS-60 cells (25, 26). In brief, cells were seeded into 96-well culture plates at a concentration of 1×10^4 cells per well and incubated for 44 hrs in the presence or absence of experimental samples. The recombinant murine M-CSF (R&D Products, Minneapolis, MN) was also included in the assay. Following the addition of [^{3}H]thymidine (ICN Biomedicals), its incorporation into DNA was measured after a 6-h period (26).

RESULTS

Identification and cloning of alternatively spliced M-CSF transcript. We performed a reverse transcription PCR (RT-PCR) analysis using total RNA obtained from murine bone marrow stromal cells (ST2), which were shown to constitutively produce M-CSF (6). As shown in Fig. 1, the RT-PCR reaction using primers P1 and P2 resulted in the amplification of four different products, whose sizes were approximately 1850, 1700, 950, and 800 base pairs (bp). The primers were chosen so as to amplify the entire coding region of M-CSF, and so that the primer pair would generate fragments with sizes of 1686 or 800 bp. The former would be derived from the transcript containing the complete exon 6, and the latter from the transcripts containing the truncated exon 6 (12). Thus, the fragments of approximately 1850

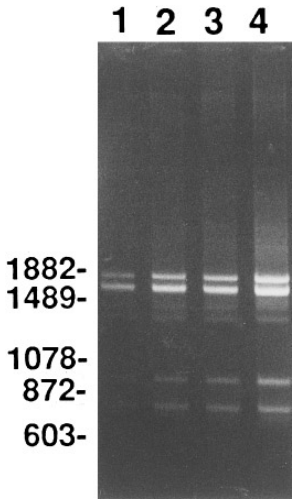


FIG. 1. RT-PCR amplification of fragments derived from M-CSF transcripts. An RT-PCR was performed by using total RNA from ST2 cells and the primer pair P1/P2. An aliquot was removed from the reaction mixture at the ends of 28 cycles (lane 1), 30 cycles (lane 2), 32 cycles (lane 3) and 34 cycles (lane 4) of the PCR amplification. The amplified products were then analyzed by electrophoresis on 2% agarose gel and visualized after ethidium bromide staining of the gel. The molecular mass standards are shown at the *left* in bp.

and 950 bp were unexpected amplification products in the RT-PCR reaction. The Southern blot analysis with M-CSF cDNA as a probe revealed that these additional fragments actually contained an M-CSF-related sequence (data not shown). We therefore cloned these two fragments together with the expected amplification products and analyzed their nucleotide sequences.

The DNA sequence analysis of the cDNAs derived from the two additional fragments revealed that both possessed an insertion of a 140-bp sequence in the coding region (indicated by the underline in Fig. 2A). In addition, the difference in length between the additional two fragments was demonstrated to be due to the result of an alternate use of exon 6 that had been reported (12): the longer fragment contained the complete exon 6, whereas the shorter contained the truncated form of exon 6. We also found that the 140-bp sequence was inserted at the junction of the exon 1-derived sequence and the exon 2-derived sequence (7, 12, 33). The published sequence of murine M-CSF demonstrated that exon 1 encoded the 5′ untranslated region and the first 13 amino acid residues of the signal sequence, and exon 2 encoded the remaining 19 amino acid residues of the signal sequence and the first 22 amino acid residues of the mature protein (7, 12, 13, 21, 33). The insertion of the 140-bp sequence resulted in the generation of TGA stop codon 31 bp downstream from the beginning of the inserted sequence (Fig. 2A). However, there were repetitive ATG codons in the inserted sequence, which overlapped the TGA stop codon (indicated by the boldface in Fig. 2A). If the ATG codon is used as a translational initiation codon, the reading frame of M-CSF is preserved in these cDNAs, though the first 13 amino acid residues of a 32-amino acid signal peptide will be replaced with an unrelated sequence of 37 amino acid residues in the cDNAs (indicated by the underline in Fig. 2B).

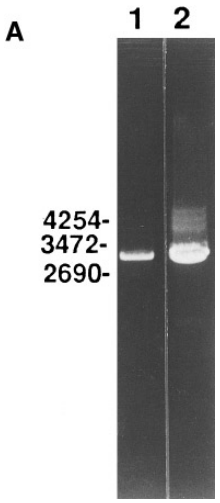
We next examined the origin of 140-bp insertion sequence. For this purpose, we performed a PCR amplification of the M-CSF genome containing intron 1 that separated two exons, i.e., exons 1 and 2. An approximately 3.1 kb fragment was detected (Fig. 3A), which would be expected based on the reported intron-exon structure of the human M-CSF gene (7, 33). The fragment was therefore cloned and the genomic DNA was analyzed by sequencing. As a result, we found that the 140-bp insertion sequence was part of intron 1. It was located near the 3′ end of the intron (indicated by the parentheses in Fig. 3B). We therefore tentatively designated the 140-bp sequence exon 1.1. The 5′ end of exon 1.1 was preceded by a possible acceptor sequence (indicated by the boldface in Fig. 3B), and the 3′ end was followed by a possible donor sequence (indicated by the italics in Fig. 3B) as discussed below.

The cDNA containing the entire coding region of M-CSF (12) and the cDNA fragment isolated from the region within exon 1.1 were each used as a probe to examine the expression of M-CSF transcript. The Northern blot analysis of ST2 cells and two additional two cell lines (i. e., the murine bone marrow stromal cell line PA6 and the murine fibroblastic cell line L) with the wild-type cDNA revealed a major hybridizing species of approximately 4.0 kb as well as a minor species of approximately 2.5 kb (Fig. 4, lanes 1-3). The longer exposure of the blot revealed the presence of at least four less- abundant species ranging from 1.5 to 4.5 kb (data not shown). Hybridization of another blot with the cDNA fragment specific to the sequence of exon 1.1 established that all three cell lines expressed the M-CSF transcript possessing exon 1.1 (Fig. 4, lanes 4-6). The size of the transcript was indistinguishable from that of a major species detected on the blot probed with the wild-type cDNA. The Northern blot analysis was performed under conditions of high stringency, as described in Experimental Procedures (29). Taken together with the finding that exon 1.1 was only part of intron 1 (Fig. 3B), the M-CSF transcript containing the exon 1.1 sequence was demonstrated to be due to an alternative splicing event.

Expression of M-CSF cDNA in COS7 cells. Finally, we investigated whether the M-CSF transcript containing exon 1.1 could direct the production of biologically active M-CSF protein. Initially, we made four plasmids which carried different constructions of the cDNA insert, i. e., the cDNA containing exon 1.1 and the complete exon 6 ("CSF-A" in Fig. 5), that containing the complete exon 6 but not exon 1.1 ("CSF-B"), that

FIG. 2. DNA sequence and deduced amino acid sequence of PCR products. The amplified fragments of approximately 1850 and 950 bp shown in Fig. 1 were cloned and the cDNA inserts were analyzed by sequencing. (A) The 140-bp insertion sequence not found in the reported cDNAs is underlined. The initiation codon which is thought to be utilized in the reported cDNAs is indicated by italics. The in-frame ATG codons within the 140-bp insertion sequence are indicated by boldface. The nucleotide difference between this sequence and the sequence reported by Ladner et al. (12) is shown above the nucleotide sequence (T → C at nucleotide position 310). The change was also observed in the cDNA clones derived from the amplified fragments of 1700 and 800 bp shown in Fig. 1, which represented the reported structure of M-CSF transcript. This change alters the amino acid (Val → Ala) as shown in Fig. 2B. In any case, the single amino acid change was not critical to the function of the protein, as shown in Fig. 5. The cDNAs derived from the fragments of approximately 1850 and 950 bp differed in the structure of exon 6, as reported with the previously isolated cDNAs (12). Since the sequence downstream exon 6 of the isolated cDNAs agrees with that reported by Ladner et al. (12), we omitted the nucleotide sequence as well as the deduced amino acid sequence for that portion. (B) The presumed amino acid sequence by the isolated cDNA is depicted. The amino acid sequence not found in the reported cDNAs is underlined. The sequence deduced by the reported cDNAs is shown above this sequence (12). The first amino acid in the mature M-CSF protein (Lys) in indicated by the boldface.

containing exon 1.1 and the truncated exon 6 ("CSF-C"), and that containing the truncated exon 6 but not exon 1.1 ("CSF-D"). These plasmids were introduced into COS7 cells, and the media conditioned by the transfected cells were assayed for the presence of biologically active M-CSF using an M-CSF-dependent M-NFS-60 cell line. As shown in Fig. 5, the plasmids containing exon 1.1 ("CSF-A" and "CSF-C") directed the



B

exon 1

TGCCGGGACCCAGCTGCCCGTATGACCGCGCGGGGCGCCGCGGGGCGCTGCCCTTCTTCG]

GTAAGCTGCAACCGTGGCGCGCGGGGCCCGGGCCGGGCTGGGGCAGGAGCTCTGCAGCAA

GCAGCAGGCGGCTCTGCGGGGCCACTTGGAGCGGACAGCCCCTTCTCGCCAGCTGCCCAG

GCTTCTCGGCCCTGGGTTCTGGCTTCCTTACTGCTTCTAAGCTACGGCTGAGGCCCGCGG

TATTTCAATAGCTGCTCCTGGGGCTGCAGCGCTTTGCGAGGTAAAGAGAAGGCTGCTCAT

CCCATTGCACGGAGGAGAACACTGAGCTACCCACAGGCTGGA----------TGTAA

exon 1.1

GATGGAAAGGTGGCTCATCATTTAAAAACACTTGTTG**CTCTTGCAG** [TGGATCCAGGTT

CAAATCCCAGAATCTAGATGATGGTTCACAACCATTCATACCTCCAGGTCCAGGAGATTC

CAAGAAGGAAGCACATAGAGTGTGTGCATGCAAGCAGGCAAAATATTCATAAACACAAAA

CAAAAACA]**GTAAGC**CCTTTTAAAATTGTCTTTTAAAGTACTCTGTATTGTTGAGCTGCT

CTGGGTTAACCTCAGAAGAGTTCTTGGAGAAGATGGCATTTGGGCTATTACCAAGGGAAG

AGTGAGGTGGTGCCCTGGCTGATAAGAATCTAATGGTGCTGCAGCATGGGAATGGGGCGT

TGTTTCTTTGCCCTTAGTGGGACATGGGGGTAGCTAGGGAGAGGAAGTGCTGGAAAGACA

exon 2

TGACTCCCTTCTGTTATAG [ACATGGCTGGGCTCCCGGCTGCTGCTGGTCTGTCTCCT

CATGAGCAGGAGTA

FIG. 3. PCR amplification of the genomic region containing exon 1 of M-CSF gene and partial characterization of the genomic clone. (A) The PCR amplification was performed by using genomic DNA from ST2 cells and the primer pair P1/P3. An aliquot was removed from the reaction mixture at the ends of 25 cycles (lane 1) and 30 cycles (lane 2) of the PCR amplification. The amplified products were then analyzed by electrophoresis on 1% agarose gel and visualized after ethidium bromide staining of the gel. The molecular mass standards are shown at the *left* in bp. (B) The amplified fragment shown in Fig. 3A was cloned and the cDNA inserts were analyzed by sequencing. The 140-bp insertion sequence, which we designated exon 1.1, is indicated by parentheses. The possible acceptor sequence and donor sequence for exon 1.1 are indicated by the boldface.

production of biologically active M-CSF into the media, as did those without exon 1.1 ("CSF-B" and "CSF-D"), although the activities in the media conditioned by COS7 cells transfected with the plasmids containing exon 1.1 were somewhat lower than those obtained by using the plasmids without the exon. As mentioned above (Fig. 2A), the original ATG codon did not specify the M-CSF protein in the cDNAs containing exon 1.1. Therefore, the result shown in Fig. 5 indicated that an alternate initiation codon was functionally utilized in the cDNAs. When we removed the complete sequence of exon 1 which included the original ATG codon from the cDNAs with exon 1.1 and introduced the resulting plasmids ("dCSF-A" and "dCSF-C" in Fig. 5) into COS7 cells, the result of the bioassay showed that these truncated cDNAs also directed the expression of active M-CSF, supporting our conclusion described above.

DISCUSSION

We have shown that murine ST2 cells express an M-CSF transcript that has not been previously described. The transcript contained the sequence within intron 1 (Figs. 2 and 3). The presence of this transcript was confirmed by Northern blot analysis (Fig. 4B). Since only part of the intron sequence was found, the transcript seemed to be due to an alternative splicing event and not simply to incomplete splicing. An examination of the sequences surrounding the inserted sequence, which we designated exon 1.1, might support this hypothesis. The 3′ end of exon 1.1 was followed by a possible donor sequence (GTAAGC, Fig. 3B) (34), which was similar to the sequence at the intron/exon boundaries of the human M-CSF gene (7). The 5′ end of exon 1.1 was preceded by a possible acceptor sequence (CTCTTGCAG, Fig. 3) (34). However, the site had a short pyrimidine tract, a critical determinant of the strength of splice acceptors (35), and

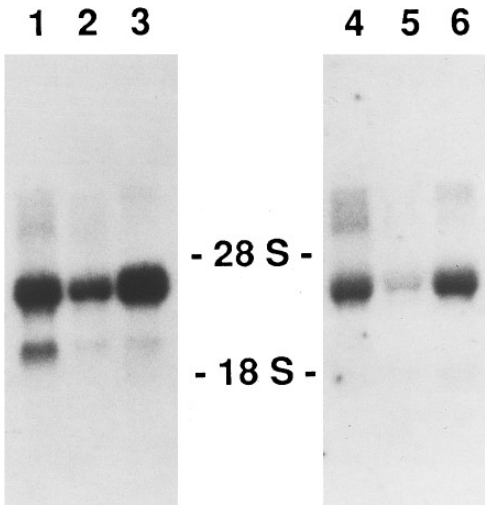


FIG. 4. Northern blot analysis of M-CSF transcripts. Total RNA from ST2 cells (lanes 1 and 4), PA6 cells (lanes 2 and 5), and L cells (lanes 3 and 6) was analyzed for the presence of M-CSF transcript by Northern blot analysis. The hybridization was performed with the cDNA insert obtained by cloning the fragment of approximately 1700 bp shown in Fig. 1, which represented the reported structure of M-CSF transcript (lanes 1-3), or the cDNA fragment whose sequence was specific to exon 1.1 (lanes 4-6), as a probe. The latter was prepared from the cDNA with exon 1.1 by using the restriction enzymes BamHI and SspI. Markers are mammalian ribosomal RNAs.

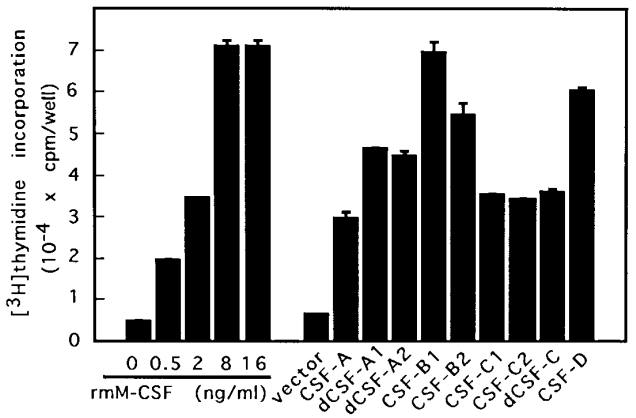


FIG. 5. M-CSF activities in the media conditioned by COS7 cells transfected with M-CSF cDNAs. The cDNA inserts were constructed as follows. The four different fragments amplified by RT-PCR shown in Fig. 1 were initially cloned into pCR2.1 vector. The cDNA inserts were then excised from the vector and recloned into the vector pCEF. Therefore, there were four types of the cDNA insert, i.e., the cDNA containing exon 1.1 and the complete exon 6 (CSF-A), that containing the complete exon 6 but not exon 1.1 (CSF-B), that containing exon 1.1 and the truncated exon 6 (CSF-C), and that containing the truncated exon 6 but not exon 1.1 (CSF-D). We also made two types of deletion mutants (dCSF-A and dCSF-C). This was accomplished by the removal of the 5′ portion of CSF-A or CSF-C cDNA using a restriction enzyme XbaI, which deleted the first 89 nucleotides shown in Fig. 2A. The pCEF vector and the resulting plasmids were introduced into CoS7 cells. The number of the cDNA represents the independent isolation of the plasmids. The media conditioned by COS7 cells were assayed for M-CSF activity after a 500-fold dilution using M-NFS-60 cells. The murine recombinant M-CSF was also included in the assay at the indicated concentration.

its sequence was very similar to the upstream acceptor site of exon 6 of the human gene (7), which is bypassed in at least some transcripts (9-13). It is therefore likely that the splice acceptor flanking exon 1.1 is relatively weak. Since we detected the transcript which had exon 1.1 and the truncated exon 6 as well as the transcript which had exon 1.1 and the complete exon 6 (Figs. 1 and 2), the choice of exon 1.1 and the utilization of two acceptor sites of exon 6 appeared to be independent phenomena. The examination of the sequence of the ends of two introns that was generated by the insertion of exon 1.1 revealed that both introns conformed to the GT-AG rule (Fig. 3B) (35).

The expression of the transcript with exon 1.1 was not specific to ST2 cells. As shown in Fig. 4B, other murine cell lines including PA6 and L also expressed the M-CSF transcript possessing the exon. We confirmed this result by performing an RT-PCR with total RNA obtained from these cells using the primer pair P1/P2 (data not shown). We did not obtain an evidence for the presence of human equivalents for the murine transcript. In an experiment I which we performed an RT-PCR with RNA obtained from a human stromal cell line, KM102, human umbilical vein endothelial cells, human peripheral monocytes or a human osteoblastic cell line, MG63 (26,36). we did not detect the amplified product corresponding to the murine transcript (data not shown). However, the presence of a human transcript which carried an intron sequence was reported. In the original isolation of human M-CSF cDNA from a pancreatic carcinoma cell line, nine of ten cDNA clones isolated contained a 115 bp of part of intron 2, which includes an in-frame stop codon (10). It was also demonstrated that this form did not direct the production of an active protein (10).

The intriguing finding in the present study was that the cDNA with exon 1.1 encoded a secreted- and biologically active M-CSF (Fig. 5). As described above, exon 1.1 included an in-frame stop codon (Fig. 2A). Therefore, these results indicated that the translational initiation codon used by the cDNA differed from that used by the previously isolated cDNAs. The examination of previously reported sequences revealed that there were two in-frame ATG codons between the transcriptional starting point (33) and the codon corresponding to the amino terminus of mature M-CSF protein (37, 38). The upstream ATG is an initiation codon, because the downstream ATG was not found in the human gene (7, 10, 11). Furthermore, the utilization of the downstream ATG caused a significant loss of the signal peptide, especially a complete loss of the region being rich in hydrophobic amino acids (ThrTrpLeuGlySerArgLeuLeuLeuValCysLeuLeu, in Fig. 2B), which is a characteristic of known signal peptides (39). If the ATG codons within exon 1.1, which preserves the reading frame of M-CSF as described above (Fig. 2B) is utilized as an initiation codon, the first 13 amino acid residues of a 32-amino acid signal sequence are replaced with an unrelated sequence comprised of 37 amino acid residues (Fig. 2B). However, this change still preserves the hydrophobic core sequence mentioned above, which might allow the translational product by the cDNA with exon 1.1 to be secreted (Fig. 5). The M-CSF activities in the media conditioned by COS7 cells transfected with the plasmids containing exon 1.1 were lower than those obtained by using the plasmids containing the previously defined structure of M-CSF cDNA. This might be due to the difference in the sequence flanking the ATG initiation codon among the two types of cDNA, which could affect the translational efficiency (40). In any case, this is the first report demonstrating the functional utilization of the alternate initiation codon in M-CSF gene.

At least five groups of investigators have reported the sequences for a portion or a full coding region of murine M-CSF (9, 12, 13, 21, 33). However, the reported sequences differ slightly in exon 1. The sequences of the coding region within exon 1 are

```
ATGACCGCGCGGGGCCGCGCGGGGCCGTGCCCTTCTTCG (21)
                          **          **
ATGACCGCGCGGGGCGCCGCGGGGCGCTGCCCTTCTTCG (12, 13, 33)
                       **
ATGACC - CGCGGGCGGCCGC - - GGCGCTGCCCTTCTTCG (9)
```

The diversity of the sequences within this region might be due to an artifact of the sequencing analysis, because the region contains a stretch of GC residues which might be difficult to analyze. However, it is also possible that the region is the site at which mutations frequently occur. The presence of the M-CSF transcript which utilizes the alternate ATG codon might be a mechanism of escape from a possible inactivation of the M-CSF gene by such mutations.

REFERENCES

1. Clark, S. C., and Kamen, R. (1987) *Science* **236,** 1229–1237.
2. Stanley, E. R., Guilbert, L. J., Tushinski, R. J., and Bartelmez, S. H. (1983) *J. Cell. Biochem.* **21,** 151–159.
3. Rettenmier, C. W., Roussel, M. F., and Sherr, C. J. (1988) *J. Cell. Sci.* **9,** 27–44.
4. Deryugina, E. I., Ratnikov, B. I., Bourdon, M. A., and Muller-Sieburg, C. E. (1994) *Exp. Hematol.* **22,** 920–918.
5. Deryugina, E. I., Ratnikov, B. I., Bourdon, M. A., Gilmore, G. L., Shadduck, R. K., and Muller-Sieburg, C. E. (1995) *Blood* **86,** 2568–2578.
6. Yamada, M., Suzu, S., Akaiwa, E., Wakimoto, N., Hatake, K., Motoyoshi, K., and Shimamura, S. (1997) *J. Cell. Physiol.,* in press.
7. Ladner, M. B., Martin, G. A., Noble, J. A., Nikoloff, D. M., Tal, R., Kawasaki, E. S., and White, T. J. (1987) *EMBO J.* **6,** 2693–2698.
8. Roth, P., and Stanley, E. R. (1992) *Curr. Top. Microbiol. Immunol.* **181,** 141–167.
9. Rajavashisth, T. B., Eng, R., Shadduck, R. K., Waheed, A., Ben-Avram, C. M., Shively, J. E., and Lusis, A. J. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 1157–1161.
10. Kawasaki, E. S., Ladner, M. B., Wang, A. M., Arsdell, J. V., Warren, M. K., Coyne, M. Y., Schweickart, V. L., Lee, M. T., Wilson, K. J., Boosman, A., Stanley, E. R., Ralph, P., and Mark, D. F. (1985) *Science* **230,** 291–296.
11. Wong, G. G., Temple, P. A., Leary, A. C., Witek-Giannotti, J. S., Yang, Y. C., Ciarletta, A. B., Chung, M., Murtha, P., Kriz, R., Kaufman, R. J., Ferenz, C. R., Sibley, B. S., Turner, K. J., Hewick, R. M., Clark, S. C., Yanai, N., Yokota, H., Yamada, M., Saito, M., Motoyoshi, K., and Takaku, F. (1987) *Science* **235,** 1504–1508.
12. Ladner, M. B., Martin, G. A., Noble, J. A., Wittman, V. P., Warren, M. K., McGrogan, M., and Stanley, E. R. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 6706–6710.
13. Delamarter, J. F., Hession, C., Semon, D., Gough, N. M., Rothenbuhler, R., and Mermod, J. J. (1987) *Nucleic Acids Res.* **15,** 2389–2390.
14. Rettenmier, C. W., and Roussel, M. F. (1988) *Mol. Cell. Biol.* **8,** 5026–5034.
15. Manos, M. M. (1988) *Mol. Cell. Biol.* **8,** 5035–5039.
16. Suzu, S., Ohtsuki, T., Yanai, N., Takatsu, Z., Kawashima, T., Nagata, N., and Motoyoshi, K. (1992) *J. Biol. Chem.* **267,** 4345–4348.
17. Price, L. K. H., Choi, H. U., Rosenberg, L., and Stanley, E. R. (1992) *J. Biol. Chem.* **267,** 2190–2199.
18. Kimura, F., Suzu, S., Nakamura, Y., Wakimoto, N, Kanatani, Y., Yanai, N., Nagata, N., and Motoyoshi, K. (1994) *J. Biol. Chem.* **269,** 19751–19756.
19. Suzu, S., Ohtsuki, T., Makishima, M., Yanai, N., Kawashima, T., Nagata, N., and Motoyoshi, K. (1992) *J. Biol. Chem.* **267,** 16812–16815.
20. Ohtsuki, T., Suzu, S., Hatake, K., Nagata, N., Miura, Y., and Motoyoshi, K. (1993) *Biochem. Biophys. Res. Commun.* **190,** 215–222.
21. Pogue-Geile, K., Sakakeeny, M. A., Panza, J. L., Sell, S. L., and Greenberger, J. S. (1995) *Blood* **85,** 3478–3486.
22. Stein, J., Borzillo, G. V., and Rettenmier, C. W. (1990) *Blood* **76,** 1308–1314.
23. Nishikawa, S.-I., Ogawa, M., Nishikawa, S., Kunisada, T., and Kodama, H. (1988) *Eur. J. Immunol.* **18,** 1767–1772.
24. Kodama, H., Amagai, Y., Koyama, H., and Kasai, S. (1982) *J. Cell. Physiol.* **112,** 89–95.
25. Nakoinz, I., Lee, M. T., Weaver, J. F., and Ralph, P. (1990) *J. Immunol.* **145,** 860–864.
26. Suzu, S., Kimura, F., Ota, J., Motoyoshi, K., Itoh, T., Mishima, Y., Yamada, M., and Shimamura, S. (1997) *J. Immunol.* **159,** 1860–1867.
27. Chomczynski, P., and Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159.
28. Mullis, K. B., Faloona, F., Scharf, S. J., Saiki, R. K., Horn, G. T., and Erlich, H. A. (1986) *Cold Spring Harbor Symp. Quant. Biol.* **51,** 263–273.
29. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *in* Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
30. Mead, D. A., Pey, N. K., Herrnstadt, C., Marcial, R. A., and Smith, L. M. (1991) *Bio/Technology* **9,** 657–663.
31. Mizushima, S., and Nagata, S. (1990) *Nucleic Acids Res.* **18,** 5322.
32. Hawley-Nelson, P., Ciccarone, V., Gebeyehu, G., Jessee, J., and Felgner, P. L. (1993) *Focus* **15,** 73–79.
33. Harrington, M. A., Edenberg, H. J., Saxman, S., Pedigo, L. M., Daub, R., and Broxmeyer, H. E. (1991) *Gene* **102,** 165–170.
34. Mount, S. M. (1982) *Nucleic Acids Res.* **10,** 459–472.
35. Jackson, I. J. (1991) *Nucleic Acids Res.* **19,** 3795–3799.
36. Ohtsuki, T., Suzu, S., Nagata, N., and Motoyoshi, K. (1992) *Biochim. Biophys. Acta.* **1136,** 297–301.
37. Ben-Avram, C. M., Shively, J. E., Shadduck, R. K., Waheed, A., Rajavashisth, T., and Lusis, A. J. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 4486–4489.
38. Ben-Avram, C. M., Shively, J. E., Shadduck, R. K., Waheed, A., Rajavashisth, T., and Lusis, A. J. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 7801.
39. Kreil, G. (1981) *Annu. Rev. Biochem.* **50,** 317–348.
40. Kozak, M. (1986) *Cell* **44,** 283–292.