MOLECULAR AND CELLULAR BIOLOGY, June 1993, p. 3191-3201 Vol. 13, No. 6
0270-7306/93/063191-11$02.00/0

Expression of mRNA Encoding the Macrophage Colony-Stimulating Factor Receptor (c-*fms*) Is Controlled by a Constitutive Promoter and Tissue-Specific Transcription Elongation

XIE YUE, PAULA FAVOT, TIMOTHY L. DUNN, A. IAN CASSADY, AND DAVID A. HUME*

Centre for Molecular Biology and Biotechnology, University of Queensland, Brisbane, Queensland 4072, Australia

Received 16 November 1992/Returned for modification 2 February 1992/Accepted 1 March 1993

The gene encoding the receptor for macrophage colony-stimulating factor 1 (CSF-1), the c-*fms* proto-oncogene, is selectively expressed in immature and mature mononuclear phagocytes and trophoblasts. Exon 1 is expressed only in trophoblasts. Isolation and sequencing of genomic DNA flanking exon 2 of the murine c-*fms* gene revealed a TATA-less promoter with significant homology to human c-*fms*. Reverse transcriptase primer extension analysis using exon 2 primers identified multiple clustered transcription initiation sites. Their position was confirmed by RNase protection. The same primer extension products were detected in equal abundance from macrophage or nonmacrophage sources of RNA. c-*fms* mRNA is acutely down-regulated in primary macrophages by CSF-1, bacterial lipopolysaccharide (LPS), and phorbol myristate acetate (PMA). Each of these agents reduced the abundance of c-*fms* RNA detectable by primer extension using an exon 3 primer without altering the abundance of presumptive short c-*fms* transcripts detected with exon 2 primers. Primer extension analysis with an intron 2 primer detected products at greater abundance in nonmacrophages. Templates detected with the intronic primer were induced in macrophages by LPS, PMA, and CSF-1, suggesting that each of the agents caused a shift from full-length c-*fms* mRNA production to production of unspliced, truncated transcripts. The c-*fms* promoter functioned constitutively in the RAW264 macrophage cell line, the B-cell line MOPC.31C, and several nonhematopoietic cell lines. Macrophage-specific expression and responsiveness to selective repression by LPS and PMA was achieved by the incorporation of intron 2 into the c-*fms* promoter-reporter construct. The results suggest that expression of the c-*fms* gene in macrophages is controlled by sequences in intron 2 that act by regulating transcription elongation.

Macrophage colony-stimulating factor 1 (CSF-1) promotes the survival, proliferation, and differentiation of cells of the mononuclear phagocyte series (29). Its biological activities are mediated by binding to a plasma membrane receptor, the product of the c-*fms* proto-oncogene, which is a ligand-dependent protein tyrosine kinase (24). Expression of c-*fms* mRNA and CSF-1 binding activity is largely restricted to mononuclear phagocytes and placental trophoblasts (24). Although there has been considerable progress in understanding tissue-specific gene expression in other hematopoietic cell lineages, macrophage differentiation remains poorly understood at the level of gene transcription. The most extensive studies have been of the lysozyme gene in chickens, which is controlled by a complex interaction between macrophage-specific enhancer elements and silencers that restrict gene expression in nonmacrophages (10). The paucity of information about macrophage-specific gene expression has been attributed to the difficulty of obtaining high transfection efficiencies with primary mammalian macrophages or myelomonocytic cell lines (25), a technical obstacle that we have recently overcome in a study of the constitutive elements controlling the expression of the urokinase plasminogen activator gene in macrophages (3). The limited literature on macrophage-specific gene expression has been largely dependent upon the use of transgenic animals (2, 25), which places limitations on detailed analysis of *cis*-acting sequences. Preliminary characterization of the human gene encoding c-*fms* revealed that in the two cell types that express the gene, trophoblasts and macrophages, transcription is initiated from two separate promoters, separated by a 25-kb intron (30). A limited analysis of transcriptional regulation of the two promoters in human cell lines (20) suggested that a 550-bp segment flanking exon 2 can direct macrophage-specific expression of a reporter gene. In this report, we show that the murine c-*fms* gene contains a constitutive promoter and that production of full-length mRNA is controlled by sequences in intron 2 that regulate transcript elongation.

MATERIALS AND METHODS

Cell lines and cell culture. The cell lines RAW264, Lewis lung carcinoma (LLC1 or LL/2), L929 (fibrosarcoma), BALB/3T12-3 (embryonic fibroblasts), BALB/c 10ME HD.A 5R.1 (methylcholanthrene-transformed fibroblasts), BALB/c C1.7 (embryonic fibroblasts), and MOPC.31C were obtained from the American Type Culture Collection. They were maintained in RPMI 1640 plus 10% fetal bovine serum. Primary bone marrow-derived macrophages (BMDM) were produced by cultivation of murine femoral bone marrow cells in recombinant human CSF-1 (a gift from Chiron Corp.) as described previously (14, 15). Lipopolysaccharide (LPS; Re595 from *Salmonella minnesota*) and phorbol myristate acetate (PMA) were obtained from Sigma.

Mouse c-*fms* genomic clone isolation and plasmid construction. A mouse genomic DNA library from the A17 murine T-cell lymphoma cell line in λEMBL3A (a gift from Mark

* Corresponding author.

Hogarth) was screened with a 5′ restriction fragment of the murine c-*fms* cDNA (21) kindly provided by Larry Rohrschneider. A clone containing a 14-kb *Sal*I insert was isolated, the restriction map of which was consistent with the identity of this clone as the *Sal*I genomic DNA fragment which encompasses the 5′ end of the murine c-*fms* gene (9). A 7-kb *Apa*I fragment bracketing the 5′ end of the gene was subcloned into pBluescript (Stratagene) to yield pXY*fms*1, and subfragments were sequenced by the dideoxy-chain termination method, using a Pharmacia T7 polymerase kit and double-stranded plasmid templates. The 3′ *Apa*I site in this clone is at position 134 in the murine cDNA sequence (21). The restriction map of this fragment is presented in Fig. 7, in the context of the analysis of reporter constructs. Chloramphenicol acetyltransferase (CAT) reporter constructs were generated by subcloning the desired restriction fragments from the *Apa*I genomic DNA fragment in pBluescript into the multiple cloning site of the vector pCAT-Basic (Promega). The positive control vector pHβAPr-CAT, a gift from Peter Gunning, contains a 4.3-kb *Eco*RI-*Alu*I fragment of the human β-actin promoter linked to the CAT gene. An additional set of reporter constructs was constructed in the luciferase reporter gene plasmid pGL2-Basic (Promega). This vector has the advantage of both increased sensitivity and a low nonspecific background due to the presence of a polyadenylation signal, to prevent read-through transcripts from the vector, placed upstream of the multiple cloning site. The desired restriction fragments from the *Apa*I genomic DNA clone were also cloned into the multiple cloning site of this vector. In the case of the intron-containing c-*fms* plasmid, pGL6.7*fms*, the c-*fms* genomic DNA was removed from pXY*fms*1 as an *Spe*I fragment and inserted into the same site in pGL2-Basic. The 5′ *Spe*I site in this case is approximately 50 bp into the c-*fms* promoter sequence, while the 3′ *Spe*I site is in the pBluescript multiple cloning site. Plasmid pGL2-Control, containing the simian virus 40 (SV40) early promoter and enhancer, was used in some experiments. Because this vector displayed some apparent tissue specificity, another control vector was produced by inserting the *Eco*RI-*Alu*I fragment of the human β-actin promoter (see above) into the multiple cloning site of pGL2-Basic. This plasmid is referred to as pGL-HβAPr.

RNA extraction, RNase protection, and primer extension analysis. RNA was isolated by the method of Evans et al. (8). Briefly, cells were lysed by shearing in 8 M guanidine-HCl–0.3 M sodium acetate–1% sarcosyl. After centrifugation, the supernatant was ethanol precipitated, and the pellet was suspended in 8 M guanidine–0.3 M sodium acetate, reprecipitated, washed twice with 70% ethanol, and dissolved in Tris-EDTA buffer. Primer extension was performed exactly as described by Sambrook et al. (22) except that *N*-2-hydroxyethylpiperazine-N′-2-ethanesulfonic acid (HEPES) buffer was substituted for piperazine-N,N′-bis(2-ethanesulfonic acid) (PIPES). Superscript RNase H^- Moloney murine leukemia virus reverse transcriptase was obtained from Bethesda Research Laboratories. The following primers were used (their positions in the c-*fms* gene are described in Results): primer A, 5′-TTGGAGAGTCGGCTGGGCACG GGGCTCCCAGCTGCTAGTTCT; primer B, 5′-ATGCCA AACTGTGGCCAGCAGCAGGACCAGAGGAGGCCCC; primer E, 5′-CCAGTTCTGGGCCACTAGGCTCGATGA CAGGGGCCCCCTGACCATG; and primer F, 5′-GAGGCA TCACATGCAGGACAGCGATGCCCTCTTTGCAACACT CCCCC.

At the end of the reaction, 20 μl of 1 M NaOH was added to the 30-μl incubation, and the RNA template was hydrolyzed for 15 min at 65°C before neutralization with 20 μl of 1 M HCl plus 20 μl of 1 M Tris-HCl (pH 7.6). The transcribed DNA was recovered by ethanol precipitation and separated by electrophoresis on 6% polyacrylamide–8 M urea sequencing gels in parallel with a sequencing ladder generated by using the same primer and lpxyfms1 as a template.

For RNase protection assays, a 410-bp c-*fms* genomic DNA fragment (*Xba*I [−203] to *Sma*I [+207]) or a 200-bp 5′ cDNA fragment (*Eco*RI in the host vector to *Bst*EII at bp 186 [21]) was cloned into pBluescript. A ^{32}P-labeled antisense RNA probe was synthesized by using T3 RNA polymerase. For dot blot hybridization, a sense probe was also synthesized by using T7 RNA polymerase and treated identically. The reaction mixture contained 2 μg of template DNA in 100 μl of reaction buffer (40 mM Tris-HCl [pH 7.6], 6 mM $MgCl_2$, 2 mM spermidine, 5 mM NaCl, 10 mM dithiothreitol, 200 μg of bovine serum albumin per ml, 500 μM ATP, CTP, and GTP, 25 μM UTP, 10 μl of [α-^{32}P]UTP [10 μCi; 800 Ci/mmol; New England Nuclear]). The reaction was terminated by addition of DNase I, the products were separated on 5% polyacrylamide gels, and the labeled probe band was excised and eluted overnight in 50% deionized formamide. The probe (50,000 cpm) was added to 10 μg of DNase I-treated RNA, precipitated in 70% ethanol, and redissolved in 20 μl of 4 mM HEPES (pH 6.4)–0.1 mM EDTA–40 mM HCl–80% deionized formamide. After heating to 85°C for 15 min, the mixture was cooled to 45°C and allowed to hybridize overnight. Then 350 μl of RNase buffer (10 mM Tris-HCl [pH 7.5], 5 mM EDTA, 300 mM NaCl, 0.4 μg of RNase T_1, 8 per ml, μg of RNase A per ml) was added, and the mixture was incubated for 30 min at 37°C. The reaction was stopped by addition of 20 μl of 10% sodium dodecyl sulfate (SDS) plus proteinase K (final concentration, 12.5 μg/ml) and incubation for 30 min. Following phenol-chloroform extraction and ethanol precipitation in the presence of 20 μg of tRNA, the protected bands were separated on 8% polyacrylamide sequencing gels.

Transient transfection. Transient transfections and CAT assays were carried out as previously described (3). Briefly, 5×10^6 cells in 250 μl were transfected by electroporation at 750 V/cm (0.4-cm cuvettes) and 960-μF capacitance in RPMI 1640–10% fetal calf serum at room temperature and returned immediately to culture at 37°C. After 24 h for luciferase or 48 h for CAT assays, cells were harvested and assayed for reporter gene activity. The luciferase activity was assayed by using reagents supplied by Promega and measured with a Berthold luminometer. Activities were normalized to cell protein assayed by using a Bradford microprotein assay with reagents supplied by Bio-Rad.

RESULTS

c-*fms* promoter sequence. Figure 1 shows the DNA sequence of the mouse and human c-*fms* genes flanking exon 2. Because of the heterogeneous transcription initiation sites identified below, the two sequences are numbered relative to the initiation codon. Between −160 and −110, the alignment is tenuous, as both species have an extended polypurine tract which in the mouse includes 24 consecutive A residues. The alignment between the two species is also interrupted around −250 by the insertion in the mouse of a short GT repeat that is absent from the human gene. Beyond the conserved region shown, there is no obvious homology between the next 250 bp of mouse sequence and 600 bp of human sequence (not shown). In this region, the human gene contains an *Alu* repeat (20). The most conserved parts of the

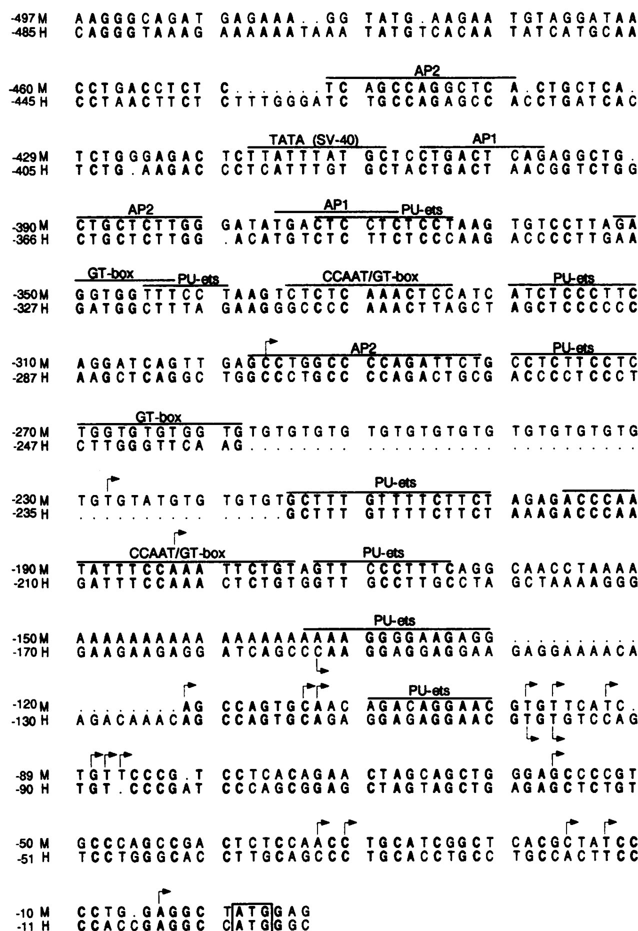


FIG. 1. DNA sequence of the proximal promoter of murine c-*fms*. The sequence of the human gene is derived from reference 20. Alignment between the two species was ascertained by using MacVector software. Arrows mark major transcription start sites in the two species, and numbering is shown relative to the first base of the initiation codon. Putative binding sites for known transcription factors are indicated and discussed in the text.

c-*fms* promoters of the mouse and human genes contain several sites similar to the binding sites of the ubiquitous CCAAT box-binding proteins (Fig. 1), but none fits perfectly the consensus sequence for CBP-1, CBP-2, or NF-1 (6). The c-*fms* promoters in both species contain multiple polypurine and polypyrimidine strings containing the 5′-GGAA-3′ recognition site of the *ets* family of transcription factors (31), including two copies in each species of the PU box (5′-GAGGAA-3′), the binding site for the transcription factor PU.1, which is restricted to B cells and macrophages (17). One other feature of interest is the multiple AP1-like (5′-TGACTCA-3′) and AP2-like (5′-CCCAGGC-3′) sites in both species, both of which might confer phorbol ester responsiveness (5). Finally, the murine c-*fms* promoter contains the TATA-like sequence 5′-TTATTTATGC-3′ at −252 (Fig. 1), but this sequence does not appear to be conserved in the human gene.

Reverse transcriptase primer extension and RNase protection analysis of c-*fms* mRNA in macrophages and nonmacrophages. To determine the transcription start site of the mouse c-*fms* promoter flanking exon 2, we performed reverse transcriptase primer extension analysis using DNase I-treated RNA from murine BMDM (Fig. 2). Primer extension using a 40-bp oligonucleotide (primer A) complementary to the 5′ end of the reported murine cDNA sequence (−75 in Fig. 1) (21) revealed numerous candidate transcription start sites (Fig. 2A). No bands were observed when nucleotides were omitted from the reverse transcriptase reaction (Fig. 2A) or when tRNA was used as a template (not shown). The major candidate transcription start sites are indicated in Fig. 1. While some of them correlate exactly with those identified in human monocyte-like cells (20), others extend further upstream. The same pattern of transcription start sites was observed when RNA from the macrophage cell line RAW264 was used (Fig. 2A). Surprisingly, identical primer extension products at comparable abundance were also produced with use of RNA from nonmacrophage cell lines such as the B-lymphocyte cell line MOPC.31C (Fig. 2A). To eliminate the possibility of artifact, the primer extension was repeated with use of a second 40-bp oligonucleotide (primer B) which hybridized to the end of the second exon, approximately 80 bp 3′ of primer A. Again, the same products at comparable abundance were obtained with use of RNA from primary macrophages, RAW264 cells, and MOPC.31C B cells (Fig. 2B). Evidence for this lack of specificity was extended to nonhematopoietic cells, L929 fibrosarcoma (Fig. 2B) and Lewis lung carcinoma (not shown). Finally, when primer E, complementary to the 5′ end of exon 3, was used, extension products were generated from primary macrophage or RAW264 RNA (Fig. 2C) but not from MOPC.31C, Lewis lung carcinoma (Fig. 2C), or L929 (not shown) RNA. This finding suggested that c-*fms* transcripts are initiated in nonmacrophages but terminate before the beginning of exon 3.

c-*fms* mRNA is down-regulated in macrophages by the ligand, CSF-1, and by bacterial LPS and phorbol esters (PMA) (11, 23). The actions of LPS are associated with cessation of c-*fms*-specific transcription in nuclear run-on transcription assays (11). We confirmed that each of these agonists down-regulated c-*fms* mRNA in BMDM and RAW264 cells by Northern (RNA) analysis (data not shown, but see also Fig. 6). Figure 3 shows the result of primer extension analysis with RNA from BMDM starved of CSF-1 (which maximally up-regulates c-*fms* mRNA) and then treated with CSF-1, LPS, or PMA. Using the exon 2 primer (primer B), no effect of CSF-1, LPS, or PMA on the yield of primer extension products was observed. Conversely, when primer E (at the beginning of exon 3) was used, there was a clear decrease in the yield of all of the extension products in response to each of the three agonists. The magnitude of this decrease is indicated by quantitation of the total radioactivity incorporated into extension products with use of an AMBIS Radioanalyser (Fig. 3). The same pattern of results was obtained with use of the three stimuli on RAW264 cells (not shown).

If there is a block to transcription elongation prior to the beginning of exon 3 in nonmacrophages and in macrophages treated with agonists, it should be possible to detect transcripts containing intron 2. To investigate this possibility, the sequence of the mouse gene early in intron 2 was ascertained. Sequencing of this region was particularly difficult because of extensive secondary structure which was not resolved by using *Taq* polymerase at higher temperature. Readable sequence required the use of deaza-dGTP and 8 M urea–30% formaldehyde denaturing gels. Alignment of the mouse and human sequences in this region reveals low-level homology (Fig. 4). Regions of transcription attenuation in eukaryotic genes have commonly been associated with GC-

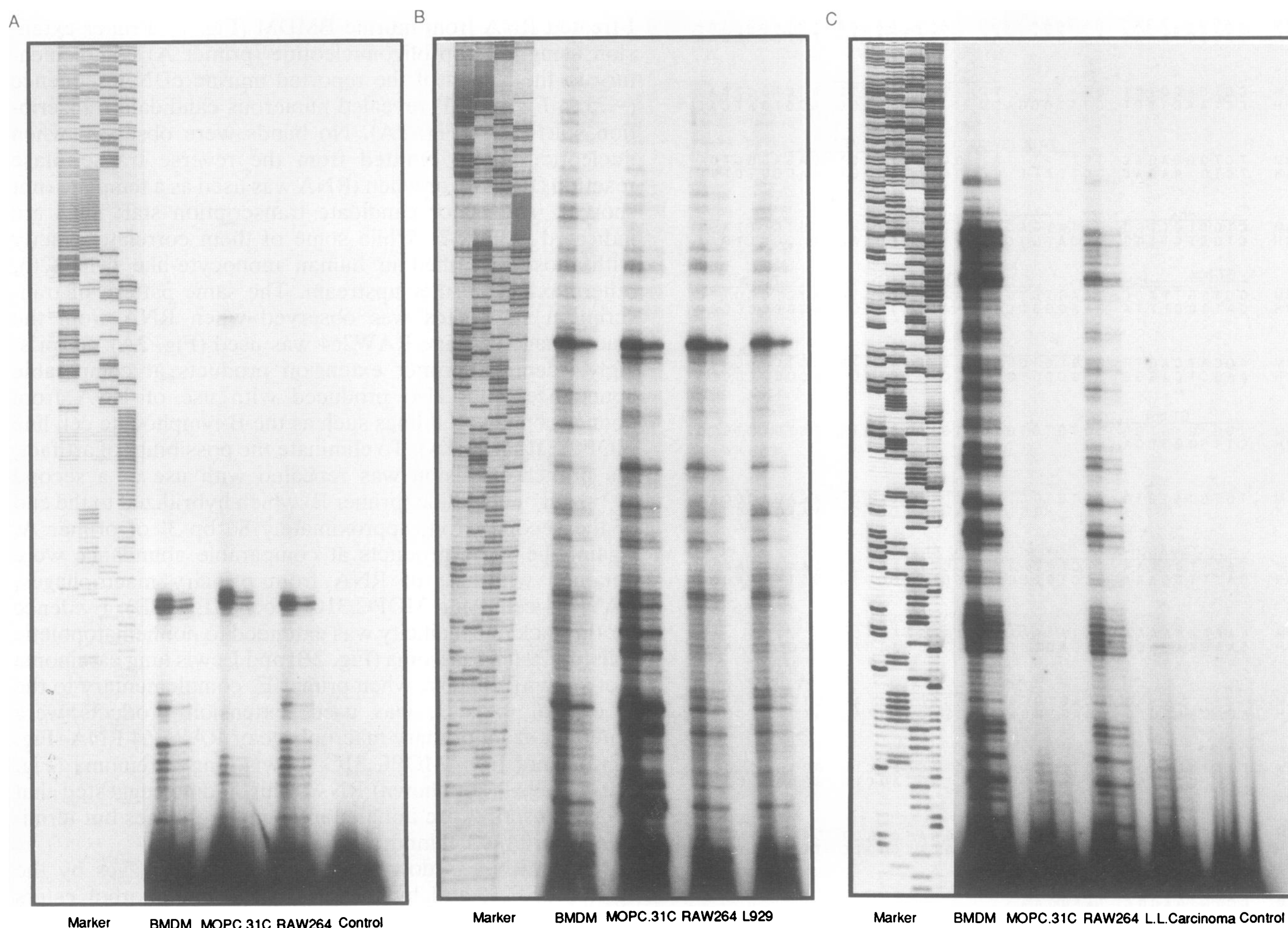


FIG. 2. Primer extension analysis of sites of c-*fms* transcription initiation. RNA from the cell lines indicated was annealed with radiolabeled primer A (which extends to the end of the published murine c-*fms* cDNA sequence) (A), primer B (which hybridizes to the end of exon 2) (B), or primer E (which hybridizes to the beginning of exon 3) (C), incubated with reverse transcriptase and deoxynucleotides, and separated on a polyacrylamide gel in parallel with a sequencing ladder derived by using the same primer and double-stranded c-*fms* genomic DNA in pBluescript as a template. In control reactions, deoxynucleotides were omitted.

rich elements containing dyad symmetry which might potentially form stable stem-loops in transcribed RNA or single-stranded DNA (27). The intronic sequences of both the mouse and human c-*fms* genes contain extended GC-rich repeats. Potential stem-loops formed by these repeats are shown in Fig. 4 and compared with the TAR region of the human immunodeficiency virus, a known stable secondary structure (27).

Primer extension was repeated by using a primer (primer F) based upon the intron 2 sequence around the predicted stem loop in Fig. 4. Identical primer extension products were detected in all cell types examined, except BMDM starved of CSF-1 to maximally up-regulate c-*fms* mRNA (Fig. 5). In these cells, the level of intron-containing RNA was much lower but was increased by each of the agonists, LPS, PMA, and CSF-1, shown to down-regulate full length c-*fms* mRNA (Fig. 3). The data suggest that the level of intron-containing transcripts varies inversely with the level of full-length c-*fms* transcripts and are consistent with the existence of a site of selective transcription elongation within intron 2.

In an attempt to identify the sites of transcription attenuation in intron 2 and to independently validate the primer extension results, we performed RNase protection studies. c-*fms* genomic DNA extending from −203 to +210 was cloned into pBluescript, and a radiolabeled RNA probe was generated by using T3 RNA polymerase. Both BMDM and RAW264 RNA protected the labeled probe against RNase digestion, producing multiple protected bands (Fig. 6A). The position of these bands relative to the splice donor site at the end of exon 2 correlated precisely with the clustered transcription start sites identified by primer extension in Fig. 2. RNA obtained from BMDM treated with CSF-1, LPS, or PMA was less effective in producing all of the protected bands, indicating that the abundance of each of the bands correlates with the abundance of full-length c-*fms* mRNA. When RNA from L929, Lewis lung carcinoma, or MOPC.31C cells was used, no protected bands were detected even when the amount of RNA added was increased to 50 μg and the autoradiographs were exposed for a prolonged period (data not shown).

Because of the extensive GC-rich inverted repeats in exon 2 and the beginning of intron 2, we reasoned that internal stem loops in the intron-containing truncated transcripts could be interfering with efficient hybridization. To eliminate these potential structures from the probe, we prepared a new probe template by subcloning the 5′ end of murine c-*fms* cDNA (21) to the *Bst*EII site at bp 186 into pBluescript. With labeled antisense RNA transcribed from this segment used

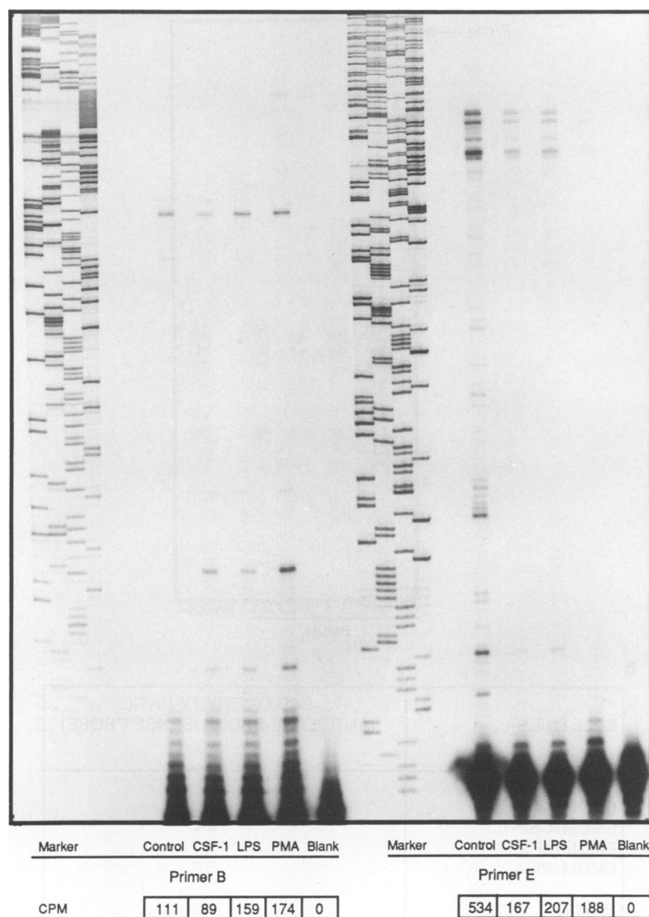


FIG. 3. Effects of CSF-1, LPS, and PMA on c-*fms* transcription attenuation in BMDM. BMDM were washed and incubated overnight in the absence of CSF-1 to maximally induced c-*fms* mRNA. They were then incubated for 4 h with CSF-1 (1,000 U/ml), LPS (100 ng/ml), or PMA (10^{-7} M) prior to RNA isolation. RNA (50 μg) from each treatment was incubated with radiolabeled primer B (end of exon 2) or primer E (beginning of exon 3) and extended with reverse transcriptase. The extension products were separated on a sequencing gel in parallel with a sequencing reaction. Quantitation was achieved by scanning the entire gel with an AMBIS Radioanalyser and counting the total radioactivity in the extension products in each lane. The results expressed as total counts per minute are displayed below each lane.

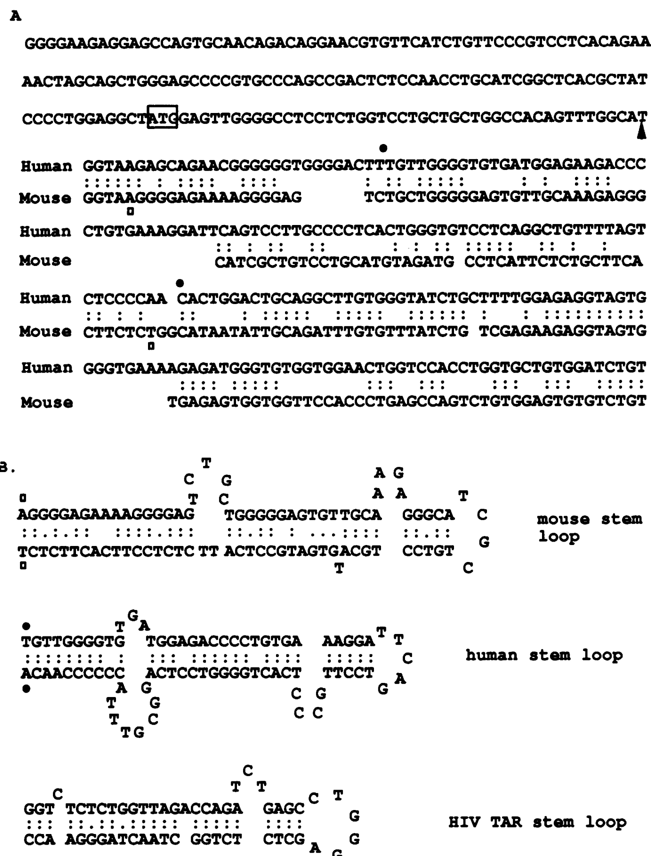


FIG. 4. Sequence analysis of the second intron of murine c-*fms*. (A) Sequence of murine c-*fms*. New sequence starts at the splice site indicated by an arrow, which is in the identical position of the human gene. Significant alignment with the human c-*fms* genomic sequence obtained from EMBL (accession number X14720) was determined by using MacVector software. Both species contain a long inverted repeat between the bases indicated by the squares (mouse) and circles (human). These sequences have the potential to form stable stem-loops in transcribed RNA or single-stranded DNA as indicated in panel B. The stem-loop structure formed by the human immunodeficiency virus (HIV) TAR region is shown for comparison (see Discussion).

as a probe, RNA from BMDM or RAW264 cells generated the expected 186-base protected band, but it was still impossible to detect any protected bands when nonmacrophage RNA was used (data not shown).

The only way to eliminating secondary structures from the transcripts is to denature and immobilize them. Accordingly, the 5′ c-*fms* antisense RNA probe was used in a dot blot hybridization on macrophage and nonmacrophage RNA bound to nitrocellulose membranes. The result (Fig. 6B) indicates that macrophages and nonmacrophages do indeed contain approximately equal levels of transcripts that hybridize to the 5′ end of c-*fms*. Furthermore, as suggested by the primer extension data in Fig. 3, CSF-1, LPS, and PMA had no significant effect on the abundance of transcripts detected with the 5′ c-*fms* probe. The reasons why short transcripts cannot be detected in nonmacrophages by RNase protection are considered in Discussion.

Transient transfection analysis. The expression of similar levels of presumptive short c-*fms* transcripts in macrophages and nonmacrophages implies that the c-*fms* promoter is not tissue specific. Promoter specificity was assayed directly by transient transfection using reporter gene constructs linking the CAT gene to the c-*fms* promoter. Because of the heterogeneity of transcription start sites identified in Fig. 2, most of the constructs were designed to include the entire 5′ untranslated region of c-*fms* by retaining the c-*fms* initiation codon. In the first experiment, a series of CAT reporter construct deletions was transfected into RAW264 cells. The most extensive construct contained ca. 3.5 kb of 5′ flanking sequence. Sequential deletion of the 5′ flanking sequence produced little alteration of the level of reporter gene expression (Fig. 7a). Even deletion to −200 reduced but did not abolish the constitutive activity in RAW264 cells. 3′ truncation of the promoter to −63 (Fig. 1) (which means that the CAT rather than the c-*fms* initiation codon is used) also had little effect. This construct is comparable to the human c-*fms* promoter constructs used by Roberts et al. (20) in which the

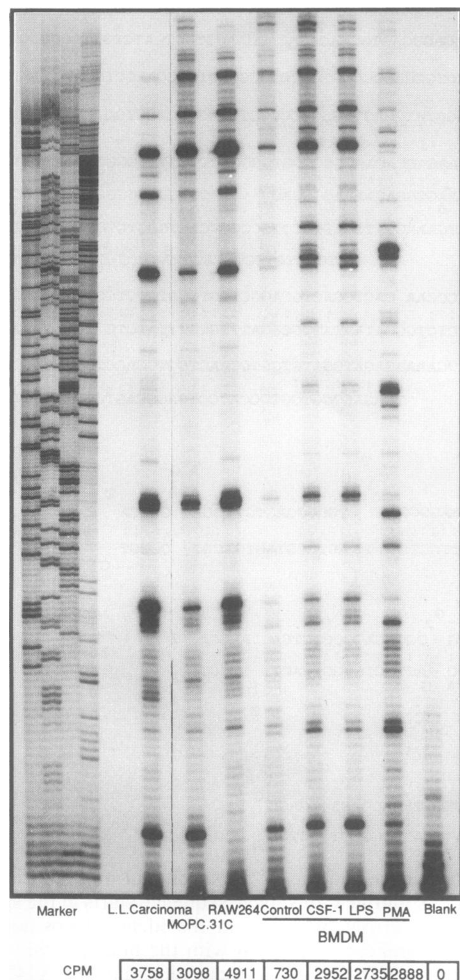
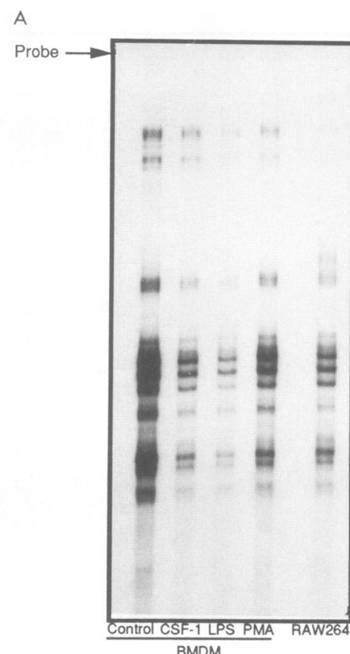


FIG. 5. Detection of RNA containing sequences homologous to c-*fms* intron 2 by reverse transcriptase primer extension. DNase-treated RNA was prepared from each of the cell lines indicated and from BMDM treated with CSF-1, LPS, or PMA as indicated in the legend to Fig. 3. RNA (50 μg) was hybridized to radiolabeled primer F, which is based upon the sequence in the vicinity of the potential stem-loop early in murine c-*fms* intron 2 (Fig. 4), and extended by using reverse transcriptase. The extension products were separated on a sequencing gel in parallel with a sequencing ladder generated by using the same primer. Quantitation was achieved by scanning the entire gel with an AMBIS Radioanalyser and counting the total radioactivity in the extension products in each lane. The results (expressed as total counts per minute) are displayed below each lane.

CAT gene was fused at −57. Treatment of the RAW264 cells transfected with the reporter constructs with PMA had no effect on CAT activity (Fig. 7a). The activities of long (pFMS3.5CAT) and short (pFMS0.3CAT) c-*fms* reporter constructs were also tested in nonmacrophages, using a β-actin promoter–CAT construct as a transfection control. The relative activities of the two c-*fms* constructs did not distinguish RAW264 cells from the B-cell line MOPC.31C or any of the nonhematopoietic murine tumor cell lines tested (Fig. 8).

To address the role of intron 2 in generating tissue specificity, the c-*fms* genomic DNA fragments were transferred to the luciferase reporter construct, pGL2-Basic. Apart from the availability of a more convenient set of



B

CELL LINES	RADIOACTIVITY RATIO (ANTISENSE PROBE /SENSE PROBE)
BMDM	5.3
BMDM+CSF-1	4.4
BMDM+LPS	3.7
BMDM+PMA	5.6
RAW264	4.1
RAW264+CSF-1	6.1
RAW264+LPS	4.1
RAW264+PMA	11.4
Lewis Lung Carcinoma	5.3
L929	8.6
MOPC.31C	4.1

FIG. 6. (A) RNase protection analysis of the 5′ ends of c-*fms* transcripts expressed in macrophages. RNase protection was carried out as described in Materials and Methods. The RNA samples derived from BMDM treated with various agonists or from RAW264 cells are the same as those shown in Fig. 3. Shown are the protected bands separated on an 8% polyacrylamide sequencing gel, with the position of the undigested 420-base probe (run on a separate lane on the gel) indicated by the arrow. The size of each protected band was ascertained by comparison with a sequencing ladder run in parallel. Under precisely the same conditions, RNA from MOPC.31C, L929, or Lewis lung carcinoma cells failed to protect the probe from digestion (not shown). (B) Detection of short c-*fms* transcripts by RNA dot blot hybridization. Total RNA (10 μg) from the cell lines indicated was dissolved in 50% formamide–8% formaldehyde–1 × SSC (0.15 M NaCl plus 0.015 M sodium citrate), heated to 68°C for 15 min, transferred to Hybond N (Amersham) membranes under vacuum, washed with 10 × SSC, and then baked overnight at 80°C. The treatments of the BMDM and RAW264 cells are the same as in Fig. 3 and 5. Duplicate blots were hybridized with either ^{32}P-labeled antisense or sense RNA homologous to the 5′ end of c-*fms* to cDNA bp 186 prepared as for RNase protection (see Materials and Methods). The hybridization was carried out in 50% formamide–5 × SSC–5× Denhardt's solution–0.1% SDS at 42°C. The blots were washed in 0.1 × SSC–0.1% SDS at 65°C and then scanned with an AMBIS Radioanalyser. The radioactivity (counts per minute) present in each dot was counted, and the ratios are presented as the counts bound with the antisense probe/counts bound with the sense probe (corrected for the length of the probe).

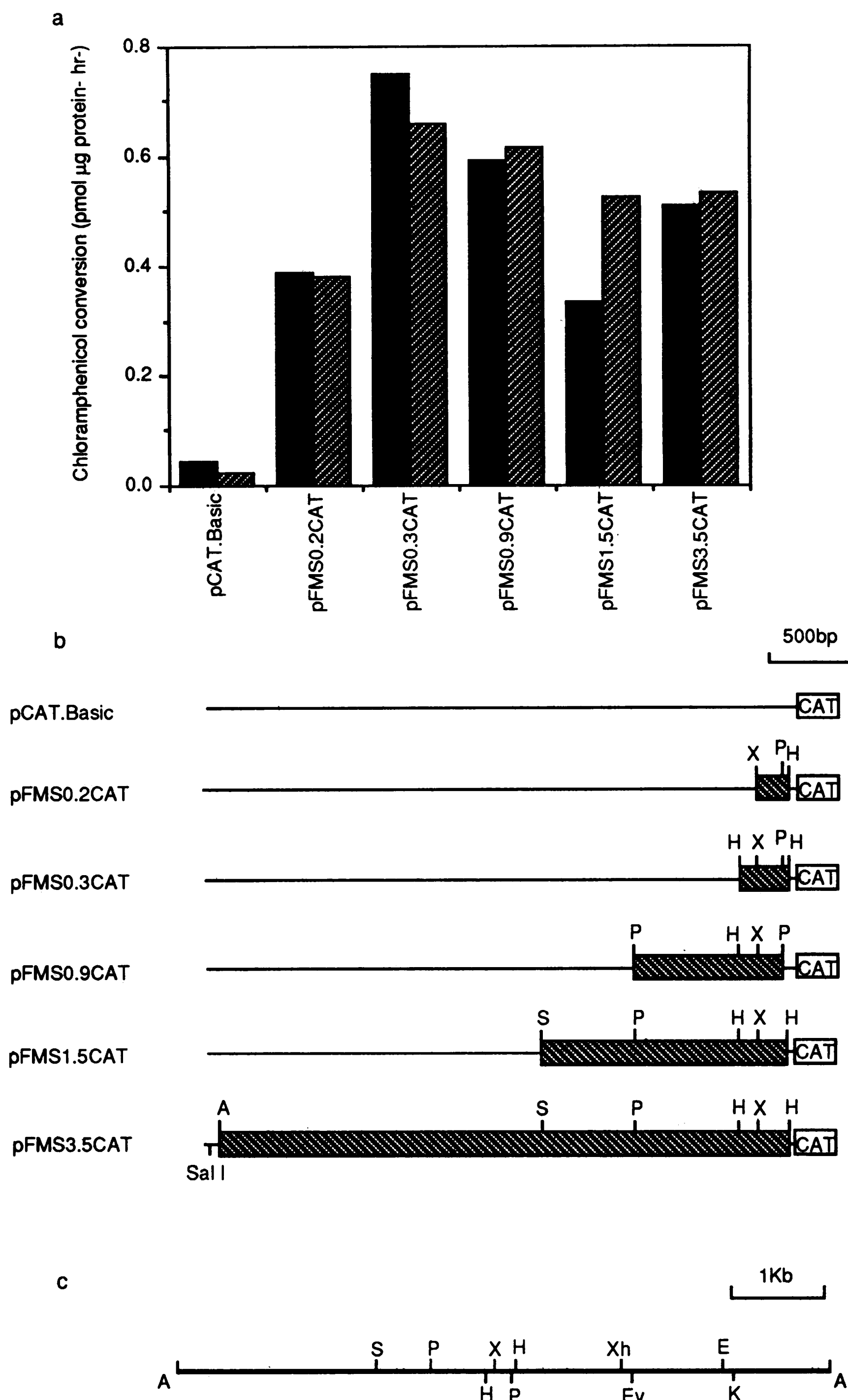


FIG. 7. (a) Transient transfection analysis of the activity of the c-*fms* promoter, using CAT reporter gene constructs. Cells were transfected with 10 µg of reporter plasmid by electroporation (see Materials and Methods), and CAT expression was assayed after 48 h. The map of the reporter constructs is shown in panel b, and the restriction map of the 7-kb *Apa*I genomic DNA fragment encompassing the 5' end of the murine c-*fms* gene is shown in panel c. Restriction enzyme sites: A, *Apa*I; S, *Sac*I; P, *Pvu*II; H, *Hae*III; X, *Xba*I; Xh, *Xho*I; Ev, *Eco*RV. In panel a, RAW264 cells were transfected with each of the constructs; then 10^{-7} M PMA was added to half the cells (shaded bars), while the remainder were untreated (solid bars). The results are averages of two separate experiments, each of which involved two separate transfections and separate preparations of plasmid.

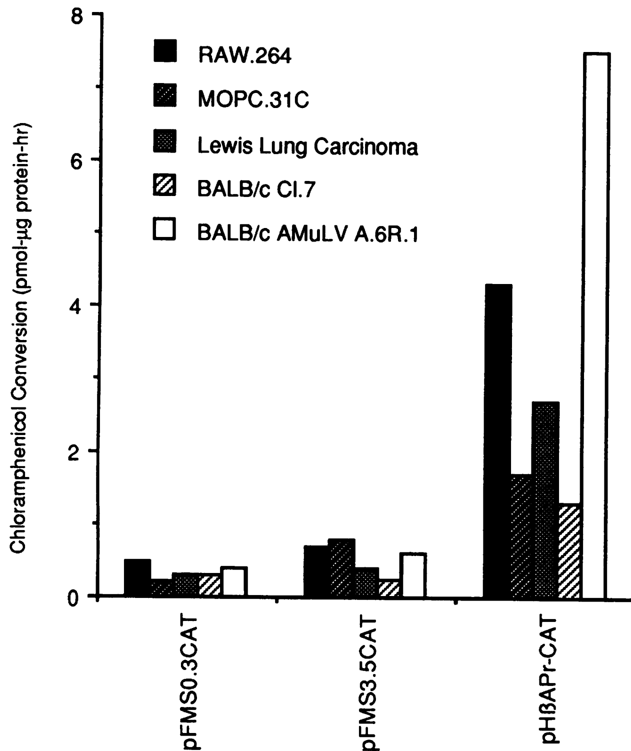


FIG. 8. Comparison of the activities of the short (pFMS0.3CAT) and long (pFMS3.5CAT) c-*fms* promoter constructs in macrophages (RAW264 cells) and a range of nonmacrophage lines with that of the positive control plasmid pHβAPr-CAT. In each case, the cells were transfected by electroporation with 10 μg of plasmid, and CAT activity was assayed 48 h later. The results are averages of duplicate transfections and are representative for each of the cell lines.

restriction sites in the multiple cloning site, pGL-Basic provided a much lower negative control than did pCAT-Basic. Figure 9A compares the activities of three constructs, a 0.3-kb c-*fms* promoter, a 3.5-kb c-*fms* promoter, and a minigene construct (pGL6.7*fms*) containing the 3.5-kb 5′ flanking sequence plus the whole of intron 2 and the beginning of exon 3. The results obtained with the first two constructs confirm those obtained with the CAT vectors; the activity of the c-*fms* promoter relative to the β-actin promoter did not clearly distinguish macrophages from nonmacrophages. The inclusion of the intron 2 sequences reduced reporter gene expression in RAW264 cells but almost completely abolished activity in MOPC.31C, L929, and Lewis lung carcinoma cells. Also shown in Fig. 9A is the comparative activity of the SV40 early promoter and enhancer in the positive control vector pGL-Control. The β-actin promoter was slightly less active in MOPC.31C cells than in RAW264 cells and less again than in L929 and Lewis lung carcinoma cells. We have confirmed that the levels of β-actin mRNA and the rate of transcription of the β-actin gene in run-on transcription assays are similar in each cell population (not shown), so the expression of this construct provides some indication of transfection efficiency in each case. By contrast, pGL-Control (the SV40 early promoter) was 20- to 50-fold more active in MOPC.31C cells than in RAW264 cells, and in L929 and Lewis lung carcinoma cells, it was a further 10-fold less active. These data suggest that the SV40 enhancer contains tissue-specific *cis*-acting elements.

The results of the primer extension experiments in Fig. 3

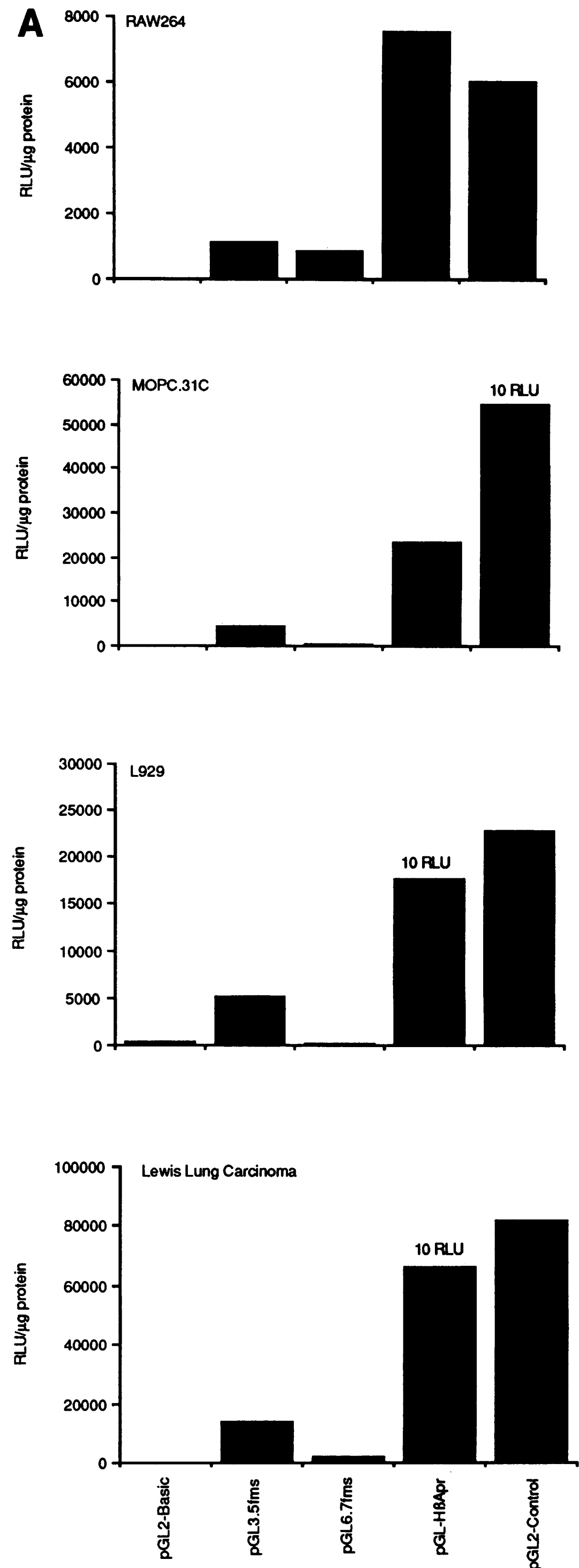


FIG. 9.

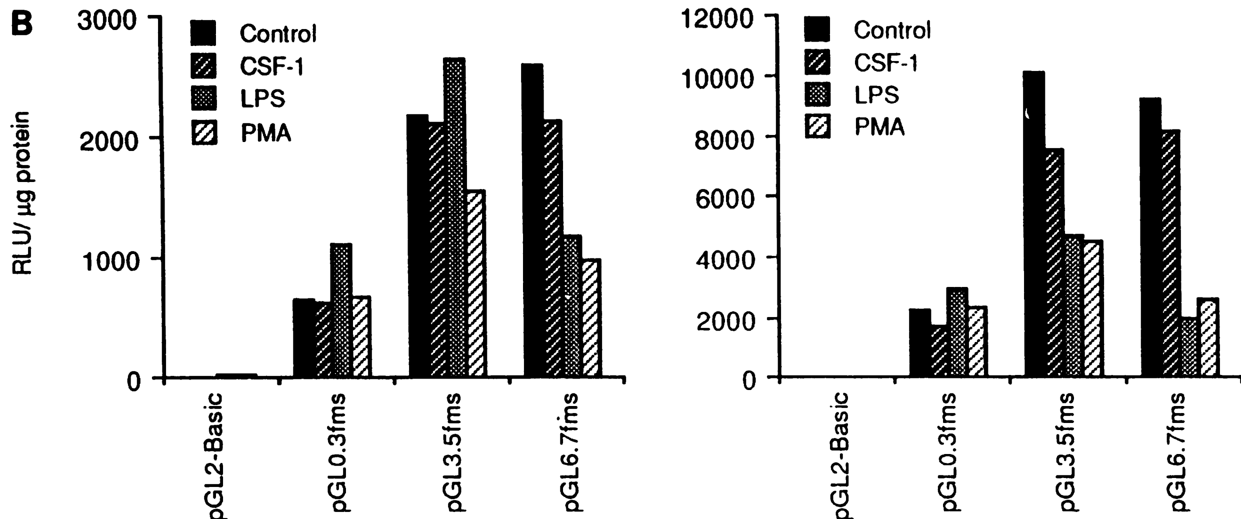


FIG. 9. Transient transfection analysis of the activity of the murine c-*fms* promoter, using luciferase promoter constructs. Cells were transfected by electroporation with 10 μg of plasmid as described in Materials and Methods. pGL0.3*fms* and pGL3.5*fms* contain the same promoter fragments as do the corresponding CAT vectors in Fig. 7 and 8. pGL6.7*fms* contains the entire *Apa*I genomic DNA fragment from pXT*fms*1, so that the reporter gene is fused into exon 3. It is therefore identical to pGL3.5*fms* except for the inclusion of intron 2 and exon 3. In each case, cells were returned to medium at 37°C immediately posttransfection and harvested after 24 h for the assay of luciferase activity. (A) Comparison of the activity of each of the c-*fms* constructs with the promoterless luciferase parent vector (pGL2-Basic), the SV40 promoter (pGL2-Control), and the β-actin promoter (pGL-HβAPr). In some panels, the results from control plasmids are off scale and are presented in units of 10 RLU as indicated. Note that the activity of pGL6.7*fms* in the nonmacrophages is not significantly different from the activity of the promoterless plasmid, pGL2-Basic. The results are averages of duplicate transfections and are representative of three experiments. (B) Effects of LPS (500 ng/ml), PMA (1 μg/ml), and CSF-1 (10^4 U/ml) added immediately posttransfection on the activity of the three *fms* promoter plasmids. In each case, cells from a single transfection were divided into four wells prior to addition of agonist; results are averages of two separate transfections. In both of the experiments shown, PMA and LPS induced the activity of the SV40 promoter construct (pGL2-Control) two- to fivefold but had no effect on the activity of the β-actin promoter, while CSF-1 had no effect on the activity of any promoter. RLU, relative light units.

and 5 suggest that CSF-1, LPS, and PMA reduce c-*fms* mRNA levels in macrophages by blocking transcription elongation within intron 2. Figure 9B shows the effects of the three agonists on c-*fms* promoter activity, using the luciferase reporter constructs above. CSF-1 had no effect in transient transfections with any of the reporter plasmids, which may reflect an artifact of transient transfection. In another study, we showed that the introduction of DNA into macrophages prevents them from responding to CSF-1 with increased growth (28) and prevents induction of the CSF-1-responsive urokinase plasminogen activator promoter (27a). By contrast to the data obtained with the CAT reporter constructs, PMA exerted a slight inhibition of the activity of the longer (3.5-kb) c-*fms* promoter but had a far more marked effect on the activity of the construct containing intron 2. LPS was more selective; it caused no inhibition of the activity of the 3.5- or 0.3-kb c-*fms* promoter but reduced the activity of the intron-containing construct three- to fivefold. The effects of the two agents were selective. In transfections parallel to those shown in Fig. 9, the SV40 promoter in pGL-Control was actually *trans* activated by LPS and PMA, whereas the β-actin promoter was relatively unaffected (not shown).

DISCUSSION

The c-*fms* gene in humans is closely linked to the gene encoding the platelet-derived growth factor (PDGF) receptor on chromosome 5 (20). Exon 1, located less than 500 bp downstream of the PDGF receptor gene, is transcribed only in trophoblasts and choriocarcinomas (20, 30). In human monocyte cell lines, transcription initiates immediately 5′ of exon 2, which contains the initiation codon. Figures 1 and 2 show that the sequences of the mouse and human c-*fms* genes adjacent to exon 2 are conserved and that murine primary macrophages and the macrophage line RAW264 produce transcripts initiated at a cluster of transcription start sites in this region. These data are consistent with analysis of the human c-*fms* gene in macrophages (20), which identified a similar diversity of transcription start sites. The presence of multiple clustered transcription starts is common in TATA-less promoters (1, 26). The other element typically involved in specifying accurate transcription initiation is the CCAAT box, which is usually placed 50 to 90 bp upstream of the transcription start site (6). The repeated CCAAT-like motif around 60 to 100 bp from the multiple transcription start sites of c-*fms*, although it does not fit the CCAAT box consensus sequences, does bind multiple DNA-binding proteins in all cell lines tested, including at least one nuclear protein that also binds the CCAAT box sequence of the ubiquitously active *H2-K* promoter (unpublished data).

Taken together, the transient transfection and reverse transcriptase primer extension analyses in Fig. 2, 3, 5, 7, 8, and 9 suggest that expression of full-length c-*fms* mRNA in macrophages is controlled primarily by sequences in intron 2 that probably mediate tissue-specific transcription elongation. This conclusion contrasts with the claim that the homologous human c-*fms* promoter directs monocyte-specific expression of a CAT reporter gene in transient transfection analysis (20). Given the substantial homology between the promoters from the two species (Fig. 1) and the conserved patterns of expression, it seems unlikely that this

contrast reflects a genuine species difference. One possibility is that the SV40 enhancer included at the 3′ end of the CAT gene in the human c-*fms* reporter constructs (20) confers tissue specificity as well as detectability. This possibility is favored by the data in Fig. 9A, which suggest that the SV40 promoter plasmid is tissue specific. An alternative view of the regulation of c-*fms* during human monocyte/macrophage differentiation derives from the work of Weber et al. (32), who suggested that the gene is controlled via differential mRNA stability. The primer extension data do not completely exclude this as an explanation for tissue specificity and the actions of LPS, PMA, and CSF-1 but argue strongly against it. c-*fms* transcripts were detected in nonmacrophages with exon 2 primers (A and B) and the intronic primer (F) but not with the exon 3 primer (E). Primers B and E are contiguous in the spliced c-*fms* mRNA, and it is difficult to envisage a degradative pathway that spares the 5′ end of c-*fms* so rigorously. In nuclear run-on transcription assays using a full-length cDNA, we could not detect c-*fms* transcripts in L929 or MOPC.31C cells, whereas they were of comparable abundance to β-actin transcripts in RAW264 cells (data not shown). Furthermore, Gusella et al. (11) provided clear evidence that the down-regulation of c-*fms* in a murine macrophage line by LPS occurred at the level of transcript synthesis rather than mRNA stability. Hence, we strongly favor controlled transcription elongation as the explanation of the apparent presence of short c-*fms* transcripts in nonmacrophages and macrophages treated with LPS, PMA, and CSF-1.

Control of gene expression by at the level transcription elongation has been observed in many cellular genes, including the adenosine deaminase, c-*myc*, c-*myb*, c-*fos*, and epidermal growth factor receptor (EGF-R) genes (see references 4, 7, 12, 19, and 27 and references therein). The EGF-R is particularly interesting since it is another member of the protein tyrosine kinase family of receptors and directs tissue-specific differentiation of a number of epithelial cell lineages. The EGF-R gene, like c-*fms*, is expressed from a TATA-less promoter and initiates from multiple clustered sites (12). Promoter activity does not correlate with patterns of expression, and a major determinant of full-length mRNA production is a transcription termination site in the first intron. As with this and most other examples of transcription attenuation, we have not been able to define exactly where the attenuated transcripts end in the case of c-*fms*. Although the primer extension data (Fig. 2 and 3) and dot blot hybridization (Fig. 6B) indicate that short transcripts are as abundant in nonmacrophages as is the full-length mRNA in macrophages, they could not be detected by Northern blotting with a cDNA or a genomic DNA probe (not shown), and no bands protected from S1 nuclease were detected in nonmacrophages in the published studies of the human c-*fms* gene (20, 30). Similarly, we could not detect the short transcripts in nonmacrophages by RNase protection (Fig. 6A). In every case in which the 3′ ends of attenuated transcripts have been identified, they are extremely heterogeneous (27) even when analyzed in cell-free transcription assays (16). Since the 5′ ends of the c-*fms* transcripts are also heterogeneous, part of the reason they cannot be detected in Northern blots or nuclease protection may be that they do not form a definable band on a gel. An additional problem influencing detection of the short transcripts by nuclease protection is the abundance of extended GC-rich inverted repeats in exon 2 and the beginning of intron 2. In fact, the mouse sequence from −70 to +110 contains 63% GC residues. GC-rich stem-loops in both DNA and RNA are remarkably stable (13). Furthermore, there is the potential for interchain hybridization between short c-*fms* transcripts because of the inverted repeats. The existence of duplexes and stem-loops in short c-*fms* transcripts would probably reduce hybridization in nuclease protection assays, and any regions of mismatch or loop formation would not be protected from cleavage.

Those attenuators that have been characterized in prokaryotic and eukaryotic systems often contain GC-rich motifs with dyad symmetry (27). The c-*fms* sequences in exon 2, early in intron 2 (Fig. 4), and in exon 3 (mouse cDNA bp 127 to 139; 5′-CAGGGGGCCCCTG-3′) clearly resemble other attenuator sequences, including those of the c-*fos* gene (5′-TCCCCGGCCGGGGA-3′) (18), the c-*myb* gene (5′-GC CCCCTGTGGGC-3′) (19), and the TAR region of human immunodeficiency virus type 1 (as shown in Fig. 4). Homologies between c-*fms* sequences and the c-*fos* and c-*myb* terminators may be of functional significance. The c-*fos* attenuator is apparently involved in the regulation of c-*fos* transcription elongation in macrophages (7). c-*myb* is expressed in myeloid progenitor cells, and transcriptional elongation is down-regulated during myeloid differentiation (19). There might conceivably be a link between the control of transcription elongation of c-*myb* and c-*fms*.

In summary, we have provided evidence that tissue-specific expression of the murine c-*fms* gene in macrophages and its regulation by specific agonists involve sequences in intron 2 that probably control transcription elongation. Because c-*fms* is the receptor for the major macrophage growth and differentiation factor CSF-1, selective elongation of c-*fms* transcripts may be viewed as the rate-limiting event in macrophage differentiation.

REFERENCES

1. **Blake, M., R. Jambou, A. Swick, J. Kahn, and J. Azikhan.** 1990. Transcription initiation is controlled by upstream GC-box interactions in a TATAA-less promoter. Mol. Cell. Biol. **10:**6632–6641.
2. **Bonnifer, C., M. Vidal, F. Grosveld, and A. Sippel.** 1990. Tissue-specific and position independent expression of the complete gene for chicken lysozyme in transgenic mice. EMBO J. **9:**2835–2842.
3. **Cassady, A. I., K. J. Stacey, K. A. Nimmo, K. M. Murphy, D. vonderAhe, D. Pearson, F. M. Botteri, Y. Nagamine, and D. A. Hume.** 1991. Constitutive expression of the urokinase plasminogen activator gene in RAW264 macrophages involves distal and 5′ non-coding sequences that are conserved between mouse and pig. Nucleic Acids Res. **19:**6839–6847.
4. **Chen, Z., J. W. Innis, M. Sun, D. A. Wright, and R. E. Kellems.** 1991. Sequence requirements for transcriptional arrest in exon 1 of the human adenosine deaminase gene. Mol. Cell. Biol. **11:**6248–6256.
5. **Chiu, R., M. Imagawa, R. Imbra, J. Bockoven, and M. Karin.** 1987. Multiple *cis*- and *trans*-acting elements mediate the transcriptional response to phorbol esters. Nature (London) **329:**648–651.
6. **Chodosh, L., A. Baldwin, R. Carthew, and P. Sharp.** 1988. Human CCAAT-binding proteins have heterologous subunits. Cell **53:**11–24.
7. **Collart, M., N. Tourkine, D. Belin, P. Vassalli, P. Jeanteur, and J.-M. Blanchard.** 1991. c-*fos* gene transcription in murine macrophages is modulated by a calcium-dependent block to elongation in intron 1. Mol. Cell. Biol. **11:**2826–2831.
8. **Evans, R., S. J. Kamdar, and T. M. Duffy.** 1991. Tumor-derived products induce IL-1α, IL-1β, TNF-α and IL-6 gene expression in murine macrophages. Distinctions between tumor induced and bacterial endotoxin-induced gene expression. J. Leukocyte Biol. **49:**474–482.
9. **Gisselbrecht, S., S. Fichelson, B. Sola, D. Bordereaux, A. Hampe, C. Andre, F. Galibert, and P. Tambourin.** 1987. Fre-

quent c-*fms* activation by proviral insertion in mouse myeloblastic leukaemias. Nature **329**:259–261.

10. **Grewal, T., M. Theisen, U. Borgmeyer, T. Grussenmeyer, R. Rupp, A. Stief, F. Qian, A. Hecht, and A. Sippel.** 1992. The −6.1-kilobase chicken lysozyme enhancer is a multifactorial complex containing several cell-type-specific elements. Mol. Cell. Biol. **12**:2339–2350.
11. **Gusella, G., E. Ayroldi, I. Espinoza-Delgado, and L. Varesio.** 1990. Lipopolysaccharide, but not IFN-γ, down-regulates c-fms mRNA protooncogene expression in murine macrophages. J. Immunol. **144**:3574–3580.
12. **Haley, J. D., and M. D. Waterfield.** 1991. Contributory effects of *de novo* transcription and premature transcript termination in the regulation of human epidermal growth factor receptor protooncogene RNA synthesis. J. Biol. Chem. **266**:1746–1753.
13. **Hirao, I., Y. Nishimura, Y.-I. Tagawa, K. Watanabe, and K.-I. Miura.** 1992. Extraordinarily stable mini-hairpins: electrophoretical and thermal properties of the various sequence variants of d(GCGAAGC) and their effect on DNA sequencing. Nucleic Acids Res. **20**:3891–3896.
14. **Hume, D. A., and Y. Denkins.** 1989. Activation of macrophages to express cytocidal activity correlates with inhibition of their responsiveness to macrophage colony-stimulating factor (CSF-1): involvement of a pertussis toxin-sensitive reaction. Immunol. Cell Biol. **67**:243–249.
15. **Hume, D. A., and R. Nayar.** 1989. Encapsulation is not involved in the activities of recombinant gamma interferon associated with multilamellar phospholipid liposomes on murine bone marrow-derived macrophages. Lymphokine Res. **8**:415–427.
16. **Innis, J. W., and R. E. Kellems.** 1991. A heat-labile factor promotes premature 3′ end formation in exon 1 of the murine adenosine deaminase gene in a cell-free transcription system. Mol. Cell. Biol. **11**:5398–5409.
17. **Klemsz, M., S. McKercher, A. Celada, C. vanBeveren, and R. Maki.** 1990. The macrophage and B cell-specific transcription factor PU.1 is related to the *ets* oncogene. Cell **61**:113–124.
18. **Lamb, N., A. Fernandez, N. Tourkine, P. Jeanteur, and J.-M. Blanchard.** 1990. Demonstration in living cells of an intragenic negative regulatory element within the rodent c-*fos* gene. Cell **61**:485–496.
19. **Reddy, C., and E. Reddy.** 1989. Differential binding of nuclear factors to the intron 1 sequences containing the transcriptional pause site correlates with c-*myb* expression. Proc. Natl. Acad. Sci. USA **86**:7326–7330.
20. **Roberts, W., L. Shapiro, R. Ashmun, and A. Look.** 1992. Transcription of the human colony-stimulating factor-1 receptor gene is regulated by separate tissue-specific promoters. Blood **79**:586–593.
21. **Rothwell, V. M., and L. R. Rohrschneider.** 1987. Murine c-*fms* cDNA: cloning, sequence analysis and retroviral expression. Oncogene Res. **1**:311–324.
22. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
23. **Sariban, E., K. Imamura, M. Sherman, V. Rothwell, P. Pantazis, and D. Kufe.** 1989. Downregulation of c-*fms* gene expression in human monocytes treated with phorbol esters and colony-stimulating factor 1. Blood **74**:123–129.
24. **Sherr, C. J.** 1990. Colony-stimulating factor-1 receptor. Blood **75**:1–12.
25. **Skalnik, D., D. Dorfman, A. Perkins, N. Jenkins, N. Copeland, and S. Orkin.** 1991. Targetting of transgene expression to monocyte-macrophages by the gp91-phox promoter and consequent histiocytic malignancies. Proc. Natl. Acad. Sci. USA **88**:8505–8509.
26. **Smale, S., and D. Baltimore.** 1989. The "initiator" as a transcriptional control element. Cell **57**:103–113.
27. **Spencer, C., and M. Groudine.** 1990. Transcription elongation and eukaryotic gene regulation. Oncogene **5**:777–785.

27a. **Stacey, K. J., and D. A. Hume.** Unpublished data.

28. **Stacey, K. J., I. L. Ross, and D. A. Hume.** Electroporation and DNA-dependent cell death in murine macrophages. Immunol. Cell Biol., in press.
29. **Stanley, E. R.** 1986. Action of colony-stimulating factor (CSF-1). CIBA Found. Symp. **118**:29–41.
30. **Visvader, J., and I. M. Verma.** 1989. Differential transcription of exon 1 of the human c-*fms* gene in placental trophoblasts and monocytes. Mol. Cell. Biol. **9**:1336–1341.
31. **Wasylyk, B., C. Wasylyk, P. Flores, A. Begue, D. Leprince, and D. Stehelin.** 1990. The c-*ets* protooncogenes encode transcription factors that cooperate with c-Fos and c-Jun for transcription activation. Nature (London) **346**:191–194.
32. **Weber, B., J. Horiguchi, R. Luebbers, M. Sherman, and D. Kufe.** 1989. Posttranscriptional stabilization of c-*fms* mRNA by a labile protein during human monocytic differentiation. Mol. Cell. Biol. **9**:769–775.