Differential Transcription of Exon 1 of the Human c-fms Gene in Placental Trophoblasts and Monocytes

JANE VISVADER AND INDER M. VERMA*

Molecular Biology and Virology Laboratory, The Salk Institute, P.O. Box 85800, San Diego, California 92138

Received 17 October 1988/Accepted 9 December 1988

Structural analysis of the 5' end of the human c-*fms* gene revealed that a large intron of about 25 kilobases separates an upstream noncoding exon (exon 1) from the signal peptide-containing exon (exon 2). Northern (RNA) blot analysis, S1 nuclease mapping, and primer extensions showed that exon 1 is transcribed in placenta but not in cells of the monocytic lineage. This is due to the differential usage of promoters, separated by approximately 25 kilobases, in a cell-specific manner. One major c-*fms* transcript was observed in U-937 cells, whereas multiple initiation sites for transcription appeared to be utilized in placental cells. Nucleotide sequence comparisons showed that the 3' end of the human platelet-derived growth factor receptor gene lies approximately 350 base pairs upstream of the major initiation sites for c-*fms* transcription in placental trophoblasts.

The *fms* proto-oncogene is the cellular homolog of the transforming gene (v-*fms*) of feline sarcoma virus and encodes a transmembrane glycoprotein with tyrosine protein kinase activity (24, 25, 27). The *c-fms* product is closely related and probably identical to the receptor for the macrophage colony-stimulating factor, CSF-1 (29). This growth factor is required for the survival, proliferation, and differentiation of mononuclear phagocytes and their committed bone marrow progenitors (30). Furthermore, expression of the CSF-1 receptor represents one of the earliest and most definitive markers of monocytic differentiation. Recent studies have demonstrated that monocytes also produce CSF-1 (12, 23), suggesting that the growth factor functions in an autocrine fashion to regulate the growth and differentiation of these cells.

The expression of c-*fms* is restricted to cells of the monocyte-macrophage lineage (19, 34), placental tissue (17), and cells of trophoblastic origin (13, 18). Uterine CSF-1 levels in the mouse have been shown to be regulated by the synergistic action of female sex steroids and to be elevated 10,000-fold during pregnancy (21). This increase in CSF-1 and the observations that placenta and trophoblast cell lines express the *c*-*fms* product imply that CSF-1 may have an additional role of regulating placental trophoblast proliferation and differentiation in pregnancy.

A 25-kilobase (kb) intron separates exons 1 and 2. A human placental cosmid library, constructed by inserting partial Mbol restriction fragments of genomic DNA into the pWE15 vector (32), was screened with both 32 P-labeled fms cDNA fragments and oligonucleotides complementary to the 5'untranslated region of the human fms cDNA sequence (7). A restriction map of 54 kb of chromosomal DNA was derived from three overlapping cosmid clones and is shown in Fig. 1A. The exon containing the signal peptide sequence was designated exon 2 and found to lie approximately 25 kb downstream from the first noncoding exon, which was designated exon 1 on the basis of results from S1 nuclease mapping and primer extension experiments (see later). The precise location of exon 3 has not yet been determined. The nucleotide sequences of exons 1 and 2 (229 base pairs [bp]) are shown in Fig. 1B together with partial sequences spanning the intron-exon junctions and were deduced by comparison of the cDNA and genomic sequences. The human *c-fms* cDNA sequence reported by Coussens et al. (7) was derived from a human placental cDNA library. We have confirmed the 3'-most 59 nucleotides of the designated exon 1 region, found in six separate cDNA clones after screening with an exon 1-specific oligonucleotide. The 5'-most eight nucleotides of the published cDNA sequence differ from those in the genomic sequence, possibly due to the occurrence of an aberration during cDNA synthesis. The exact size of the upstream noncoding exon (exon 1) is not yet known; however, data indicating the possible location of transcription initiation sites are described in a later section (see Fig. 3).

Differential expression of exon 1 in placental and monocytic cells. Total RNA (5) was extracted from placenta, the placental trophoblast cell line JEG-3, human peripheral blood monocytes, and the myelomonocytic cell lines U-937 and HL-60 and used for Northern (RNA) blot analysis (14, 16) with c-fms probes. The U-937 and HL-60 cell lines were treated with the phorbol ester TPA for 48 h to induce the expression of c-fms mRNA (28) prior to extraction of RNA. A probe encompassing exon 2 plus partial intron 1 and 2 sequences hybridized to fms mRNA in all cell types and tissue (Fig. 2A, lanes 1 to 5). However, an exon 1 probe only detected fms mRNA in placenta (lane 6) and the JEG-3 cell line (lane 7). No signal could be detected in the other cell lines, even after a fivefold-longer exposure (lanes 8 to 10). Thus, there is cell-specific expression of exon 1 in placenta and cells of the monocytic lineage. A clear size difference between fms mRNA from placenta and TPA-treated U-937 cells can be observed on Northern blot analysis as shown in Fig. 2B; this could be due to the lack of exon 1 in U-937 fms mRNA. Southern blot analysis has demonstrated that the fms gene in the U-937 cell line is intact and contains exon 1 (C. Van Beveren and J. Visvader, unpublished data). Thus, the differential hybridization of exon 1 probe is not due to gene rearrangement. The lack of expression of exon 1 in normal blood monocytes indicates that this is not associated with the abnormal phenotype of U-937 and HL-60 cells.

Run-on transcription experiments (8) with nuclei from JEG-3 cells and TPA-treated U-937 and HL-60 cells confirmed that exon 2 is transcribed in all cell lines but that exon

^{*} Corresponding author.



FIG. 1. Structure of the 5' end of the human c-*fms* gene. (A) Restriction map of the area encompassing exons 1 and 2 which contains the signal peptide sequence. The locations of the overlapping cosmid clones and the EcoRI, KpnI, and BamHI restriction endonuclease cleavage sites are shown. There are multiple, closely spaced BamHI sites located upstream of the most 5' BamHI site depicted; these have yet to be mapped precisely and are not shown. (B) The sequences spanning the intron-exon junctions of the 5'-flanking region of the human c-*fms* gene. A 25-kb intron separates exons 1 and 2. The predicted signal peptide sequence is indicated. Potential transcription initiation sites, identified by S1 nuclease and primer extension experiments (described in the legend to Fig. 3), are depicted by the thick arrows. The 5' end of the published cDNA sequence is indicated by a thin arrow. The nucleotide sequence of the 3' untranslated (UT) region of the human platelet-derived growth factor receptor (6, 9) is underlined.

1 is only expressed detectably in the placental trophoblast cell line (Fig. 2C). The locations of exons 1 and 2 are shown in a schematic diagram (Fig. 2D); restriction fragments containing these exons were cloned in pUC-based vectors. A

 $10-\mu g$ sample of each plasmid was immobilized onto nitrocellulose filters and hybridized with labeled nuclear RNA. The relative intensities of the bands varied between experiments, probably due to the efficiency of hybridization. All



FIG. 2. Analysis of c-*fms* mRNA in U-937, HL-60, and JEG-3 cell lines, peripheral blood monocytes, and placenta. (A) Northern blot analysis with exon 1- and exon 2-derived probes. The locations of these fragments are indicated in panel D. The random primer technique with $[\alpha^{32}P]dCTP$ was used to prepare probes. Total RNA (10 µg) from placenta (lanes 1 and 6), the placental trophoblast cell line JEG-3 (lanes 2 and 7), peripheral blood monocytes (lanes 3 and 8), the monomyelocytic cell line HL-60 (lanes 4 and 9), and the monocytic cell line U-937 (lanes 5 and 10) was loaded onto 1.7% agarose–formaldehyde gels. The latter two cell lines were treated for 48 h with TPA (20 ng/ml) to induce *fms* mRNA before isolation of RNA. (B) Size difference between *fms* mRNA in U-937 cells and placenta detected by Northern blot analysis. U-937 cells were treated with TPA (20 ng/ml) for either 24 or 48 h. A full-length *fms* cDNA clone was used to prepare ³²P-labeled probe. (C) Nuclear run-on analysis with ³²P-labeled nuclear RNA from monocytic and placental trophoblast cell lines. Nitrocellulose filters were prepared with 10 µg of DNA per slot. The arrows indicate the less-intense exon 1-specific bands. (D) Restriction sites surrounding exons 1 and 2, used for the preparation of exon 1- and 2-specific probes. The solid blocks indicate exons, which are separated by a 25-kb intron. B, P, E, D, and N represent *Bam*HI, *Pst*I, *Eco*RI, *Dra*I, and *Nco*I restriction sites, respectively.

three of the putative exon 1-containing plasmids (designated 1a, 1b, and 1c) hybridized to ³²P-labeled RNA from JEG-3 nuclei. No signal was detected with U-937 or HL-60 nuclear RNA and exon 1-containing plasmids after fivefold-longer exposures and in three separate experiments. The major c-*fms* transcripts in placenta and JEG-3 cells appeared to initiate at the 5' boundary of exon 1 indicated in Fig. 2D, as determined by S1 mapping. However, there are at least four other *fms*-specific transcripts (Fig. 3B, arrowheads) occurring in lower abundance that extend into the region encompassed by exon 1c.

Multiple initiation sites of transcription in placental cells. The S1 nuclease mapping technique (2) with single-stranded M13 phage-derived probes (4) was used to localize possible initiation sites within the c-fms gene. Since exon 1 was not transcribed in the monocytic cells, we used probes representing either exon 1 or exon 2. Two protected bands of approximately 190 and 185 bases (Fig. 3A, lanes 2 and 3) were observed with total RNA (50 µg) from placental and JEG-3 cells and a uniformly labeled exon 2b probe (Fig. 2D). This probe contains 540 nucleotides of *fms* genomic sequence, extending from the NcoI site in exon 2 to a DraI site located within the 25-kb intron. Only one band, corresponding to 190 bases, was expected on the basis of the proposed intron-exon junction. The lower band may be due to secondary structure formation, as the intensity varied between experiments. One specific band of approximately 155 nucleotides was observed in the case of U-937 RNA (Fig. 3, lane 4). Hybridization of the exon 2 probe to U-937 RNA was consistently lower than to placental RNA, even though the levels of c-fms mRNA per 10 μ g were similar according to Northern blot analysis. This weaker hybridization is probably related to the secondary structure at the 5' end of the U-937 fms mRNA. The other lower-molecular-weight bands observed in placental and U-937 RNA samples were not specific, as determined in several experiments.

As expected, no protected bands were observed with U-937 cellular RNA and the exon 1b probe (Fig. 3B, lane 4; Fig. 2D). In contrast, two major protected bands of approximately 200 and 185 bases were observed with RNA from placental and JEG-3 cells (Fig. 3B, lanes 2 and 3). These bands correspond to potential initiation sites located 80 and 65 bases further upstream of the full-length c-fms cDNA sequence (7). The putative initiation sites are indicated in Fig. 1B and 3D. The noncoding exon 1 would therefore be approximately 185 base pairs long. In addition, at least four less-intense bands in the range of 230 to 240 and 250 to 260 nucleotides were detected reproducibly in six independent experiments by S1 nuclease and RNase protection mapping (data not shown). These results may account for the hybridization of ³²P-labeled RNA from JEG-3 nuclei to the exon 1c probe, located upstream of the major initiation sites (Fig. 2C and D). Thus, transcription of the human c-fms gene appears to initiate at multiple sites in placental cells.



FIG. 3. c-fms promoters are apparently utilized in a cell-specific manner. (A and B) S1 nuclease protection of genomic exon 1 and exon 2 probes with total RNA (50 μ g) from placenta, JEG-3 cells, and U-937 cells. Single-stranded DNA probes were prepared from recombinant phage M13 clones. The exon 2b probe extends from the *DraI* site in the first intron to the *NcoI* site in exon 2, containing 540 bases of *fms* genomic sequence (A). The exon 1b probe contains a *PstI* genomic fragment shown in Fig. 2D (B). tRNA was used as a control (lane 1). The markers represent *MspI*-digested pBR322. The arrows indicate the *fms*-specific protected bands. The arrowheads depict minor protected products observed in placenta and JEG-3 cells. (C) Primer extensions with a 165-nucleotide fragment derived from a phage M13 clone and poly(A)⁺ RNA (5 μ g) from U-937 cells and placenta. This probe contains 135 nucleotides of complementary *fms* cDNA sequence, with the 5' and 3' ends corresponding to position 26 of exon 3 and position 123 of exon 2, respectively. Lane 1 represents a control with RNA from uninduced human erythroleukemia cells. Different preparations of placental RNA were used in lanes 2 and 3. The c-*fms*-specific extension products consistent with the S1 mapping data are indicated by arrows. (D) Schematic representation of the putative c-*fms* transcriptional start sites in placental trophoblasts and U-937 cells. The ATG initiation codon is shown.

Primer extension analysis (Fig. 3C) (15) was performed by using a uniformly ³²P-labeled single-stranded phage M13derived primer of 165 nucleotides containing fms sequences from exons 2 and 3, with the 3' terminus at position 123 of exon 2. Lane 1 shows a control in which total RNA extracted from uninduced human erythroleukemia cells lacking fmsmRNA was used for primer extension. A nonspecific band of about 230 nucleotides was observed. In the case of TPAinduced U-937 cells, a major band of approximately 255 to 260 nucleotides was observed (lane 4, arrow). The location of the 5' end of c-fms mRNA in U-937 cells as determined by primer extension correlated with the size of the S1 nucleaseprotected fragment (Fig. 3A, lane 4), indicating that the monocyte-specific transcript is not alternatively spliced from a separate exon 1. Thus, the major U-937 c-fms transcript initiates within the noncoding region of exon 2 (Fig. 1B and 3D). When the same primer was used with placental RNA,

multiple bands of various intensities were observed (Fig. 3C, lanes 2 and 3). However, only the upper two bands marked by arrows were consistent with the size predicted by S1 mapping analysis (Fig. 3B, lanes 2 and 3). It is possible that the lower-molecular-weight primer-extended products, which terminated within exon 1, represent prematurely terminated products due to the high G+C content of this region.

Conclusion. The *c-fms* gene has previously been shown to be distributed over a relatively large region (>30 kb) of human genomic DNA and to contain multiple regions of intervening sequence (11). Our structural analysis of the 5' end of this gene has extended the map a further 40 kb upstream of the exon containing the signal peptide sequence (exon 2). A 25-kb intron lies between this exon and the first exon, which appears to contain noncoding sequence only. Interestingly, the 3' untranslated region of the human plate-

let-derived growth factor receptor gene (6, 9) lies only approximately 350 base pairs upstream of the two major transcription initiation sites utilized by the c-fms gene in placental trophoblasts (Fig. 1B), consistent with the recent finding of Roberts et al. (26). It has yet to be determined whether a functional promoter is located in this intergenic region or lies further upstream. The c-fms gene shares common structural features with two other genes possessing intrinsic tyrosine kinase activity: the proto-oncogene c-abl and the epidermal growth factor receptor. Both of these genes are very large and are composed of multiple exons. The c-abl gene covers 230 kb, with a first intron of at least 200 kb (3), while the epidermal growth factor receptor spans a 110-kb locus and contains an 18-kb intron separating exons 1 and 2 (10).

Multiple transcription start sites are apparently utilized in the synthesis of placental c-fms mRNA, as detected by S1 nuclease analysis. This is a common feature of several other oncogenes, including c-abl (1), c-myb (33), c-mos (22), and N-myc (31), as well as other growth factor receptor genes, such as the epidermal growth factor receptor (10). The sequence upstream of the apparent initiation sites in the c-fms gene lacks well-defined promoter elements such as TATA and CAAT boxes or SP1-binding sites. Furthermore, these features are also absent in the region upstream of the intron 1-exon 2 junction, where there should be another functional promoter utilized by cells of the monocyte-macrophage lineage. It will be of interest to delineate the elements and regulatory mechanisms which confer cellspecific expression of the c-fms receptor in placental trophoblasts and monocyte-macrophage cells.

The c-fms oncogene product together with CSF-1 appears to play a crucial role in hemopoiesis and during pregnancy. Interestingly, the first exon of the c-fms gene is expressed in placenta but not in cells of the monocytic lineage. The data presented in this paper indicate that this is due to differential promoter usage in the two cell types. However, the significance of differential expression of an apparent noncoding exon remains to be determined. It is noteworthy that the human α_1 -antitrypsin gene utilizes different promoters 5' of noncoding exons in a cell-specific manner. Transcription from the two α_1 -antitrypsin promoters was found to be mutually exclusive in macrophages and hepatocytes (20).

We are grateful to G. Wahl and G. Evans for the human placental cosmid library and to S. Gregory for human monocyte RNA samples. We also thank W. Lamph for assistance with the computer analysis, B. Sefton, R. Mitchell, C. Van Beveren, and P. Sassone-Corsi for discussions, and P. McClintock for typing the manuscript.

J.V. is supported by the Jane Coffin Childs Memorial Fund. This work was supported by funds from the National Institutes of Health and the American Cancer Society to I.M.V.

LITERATURE CITED

- 1. Ben-Neriah, T., A. Bernards, M. Paskind, G. Q. Daley, and D. Baltimore. 1986. Alternative 5' exons in c-abl mRNA. Cell 44:577-586.
- 2. Berk, A. J., and P. A. Sharp. 1978. Structure of the adenovirus 2 early mRNAs. Cell 14:695-711.
- Bernards, A., C. M. Rubin, C. A. Westbrook, M. Paskind, and D. Baltimore. 1987. The first intron in the human c-abl gene is at least 200 kilobases long and is a target for translocations in chronic myelogenous leukemia. Mol. Cell. Biol. 7:3231-3236.
- 4. Bruening, G., A. R. Gould, P. J. Murphy, and R. H. Symons. 1982. Oligomers of avocado sunblotch viroid are found in infected avocado leaves. FEBS Lett. 148:71–78.
- Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. 1979. Isolation of biologically active ribonucleic acid

from sources enriched in ribonuclease. Biochemistry 18:5294-5299.

- Claesson-Welsh, L., A. Eriksson, A. Moren, L. Severinsson, B. Ek, A. Ostman, C. Betsholtz, and C.-H. Heldin. 1988. cDNA cloning and expression of a human platelet-derived growth factor (PDGF) receptor specific for B-chain-containing PDGF molecules. Mol. Cell. Biol. 8:3476–3486.
- Coussens, L., C. Van Beveren, D. Smith, E. Chen, R. L. Mitchell, C. M. Isacke, I. M. Verma, and A. Ullrich. 1986. Structural alteration of viral homologue of receptor protooncogene *fms* at carboxyl terminus. Nature (London) 320: 277-280.
- 8. Greenberg, M. E., and E. B. Ziff. 1984. Stimulation of 3T3 cells induces transcription of the c-*fos* proto-oncogene. Nature (London) **311**:433–438.
- Gronwald, R. G. K., F. J. Grant, B. A. Haldeman, C. E. Hart, P. J. O'Hara, F. S. Hagen, R. Ross, and D. F. Bowen-Pope. 1988. Cloning and expression of a cDNA coding for the human platelet-derived growth factor receptor: evidence for more than one receptor class. Proc. Natl. Acad. Sci. USA 85:3435–3439.
- Haley, J., N. Whittle, P. Bennett, D. Kinchington, A. Ullrich, and M. Waterfield. 1987. The human EGF receptor gene: structure of the 110 kb locus and identification of sequences regulating its transcription. Oncogene Res. 1:375–396.
- Heisterkamp, N., J. Groffen, and J. R. Stephenson. 1983. Isolation of v-fms and its human cellular homology. Virology 126: 248–258.
- Horiguchi, J., M. K. Warren, P. Ralph, and D. Kufe. 1986. Expression of the macrophage colony-stimulating factor (CSF-1) during human monocytic differentiation. Biochem. Biophys. Res. Commun. 141:924–930.
- Izhar, M., P. D. Siebert, R. G. Oshima, W. C. DeWolf, and M. N. Fukuda. 1986. Trophoblastic differentiation of human teratocarcinoma cell line HT-H. Dev. Biol. 116:510-518.
- Lehrach, H., D. Diamond, J. M. Wozney, and H. Boedtker. 1977. RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical re-examination. Biochemistry 16:4743–4751.
- McKnight, S. L., E. R. Gavis, R. Kingsbury, and R. Axel. 1981. Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. Cell 25:385-398.
- Mitchell, R. L., L. Zokas, R. D. Schreiber, and I. M. Verma. 1985. Rapid induction of the expression of proto-oncogene *fos* during human monocytic differentiation. Cell 40:209–217.
- Muller, R., D. J. Slamon, E. D. Adamson, J. M. Tremblay, D. Muller, M. J. Cline, and I. M. Verma. 1983. Transcription of c-onc genes c-ras^{Ki} and c-fins during mouse development. Mol. Cell. Biol. 3:1062–1069.
- Muller, R., J. M. Tremblay, E. D. Adamson, and I. M. Verma. 1983. Tissue and cell type specific expression of two human c-onc genes. Nature (London) 304:454–456.
- Nienhuis, A. W., H. F. Bunn, P. H. Turner, T. Gopal Venkat, W. G. Nash, S. J. O'Brien, and C. J. Sherr. 1985. Expression of the human c-fms proto-oncogene in haemopoietic cells and its deletion in the 5q⁻ syndrome. Cell 42:421–428.
- 20. Perlino, E., R. Cortese, and G. Ciliberto. 1987. The human α_1 -antitrypsin gene is transcribed from two different promoters in macrophages and hepatocytes. EMBO J. 6:2767–2771.
- Pollard, J. W., A. Bartocci, R. Arcecis, A. Orlofsky, M. D. Ladner, and E. R. Stanley. 1987. Apparent role of the macrophage growth factor, CSF-1, in placental development. Nature (London) 330:484–486.
- Propst, F., M. P. Rosenberg, A. Iyer, K. Kaul, and G. F. Vande Woude. 1987. c-mos proto-oncogene RNA transcripts in mouse tissues: structural features, developmental regulation, and localization in specific cell types. Mol. Cell. Biol. 7:1629–1637.
- Rambaldi, A., D. C. Young, and J. D. Griffin. 1987. Expression of the M-CSF gene by human monocytes. Blood 69:1409–1413.
- Rettenmeir, C. W., J. H. Chen, M. F. Roussel, and C. J. Sherr. 1985. The product of the c-fms proto-oncogene: a glycoprotein with associated tyrosine kinase activity. Science 228:320–322.
- 25. Rettenmeir, C. W., M. F. Roussel, C. O. Quinn, G. R. Kitch-

ingman, A. T. Look, and C. J. Sherr. 1985. Transmembrane orientation of glycoproteins encoded by the v-fms oncogene. Cell 40:971–981.

- Roberts, W. M., A. T. Look, M. F. Roussel, and C. J. Sherr. 1988. Tandem linkage of human CSF-1 receptor (c-fms) and PDGF receptor genes. Cell 55:655-661.
- Roussel, M. F., C. W. Rettenmeir, A. T. Look, and C. J. Sherr. 1984. Cell surface expression of v-fms-coded glycoproteins is required for transformation. Mol. Cell. Biol. 4:1999–2009.
- Sariban, E., T. Mitchell, and D. Kufe. 1985. Expression of the c-fms proto-oncogene during human monocytic differentiation. Nature (London) 316:64-66.
- Sherr, C. J., C. W. Rettenmeir, R. Sacca, M. F. Roussel, A. T. Look, and E. R. Stanley. 1985. The c-fms proto-oncogene product is related to the receptor for the mononuclear phagocyte growth factor, CSF-1. Cell 41:665–676.
- 30. Stanley, E. R., L. J. Guilbert, R. J. Tushinski, and S. H.

Bartelmez. 1983. CSF-1: a mononuclear phagocyte lineagespecific haemopoietic growth factor. J. Cell. Biochem. 21: 151–159.

- Stanton, L. W., and J. M. Bishop. 1987. Alternative processing of RNA transcribed from NMYC. Mol. Cell. Biol. 7:4266–4272.
- 32. Wahl, G. M., K. A. Lewis, J. C. Ruiz, B. Rothenberg, J. Zhao, and G. A. Evans. 1987. Cosmid vectors for rapid genomic walking, restriction mapping and gene transfer. Proc. Natl. Acad. Sci. USA 84:2160-2164.
- Watson, R. J., P. J. Dyson, and J. McMahon. 1987. Multiple c-myb transcript cap sites are variously utilized in cells of mouse haemopoietic origin. EMBO J. 6:1643-1651.
- Woolford, J., V. Rothwell, and L. Rohrschneider. 1985. Characterization of the human c-fms gene product and its expression in cells of the monocyte-macrophage lineage. Mol. Cell. Biol. 5:3458–3466.