Nuclear proteins interacting with the promoter region of the human granulocyte/macrophage colony-stimulating factor gene

(DNA-binding proteins/gene expression/tissue specificity/growth factors/hemopoiesis)

M. Frances Shannon, Jennifer R. Gamble, and Mathew A. Vadás

Division of Human Immunology, Institute of Medical and Veterinary Science, Frome Road, Adelaide, South Australia 5000

Communicated by J. F. A. P. Miller, September 18, 1987

ABSTRACT The gene for human granulocyte/macrophage colony-stimulating factor (GM-CSF) is expressed in a tissue-specific as well as an activation-dependent manner. The interaction of nuclear proteins with the promoter region of the GM-CSF gene that is likely to be responsible for this pattern of GM-CSF expression was investigated. We show that nuclear proteins interact with DNA fragments from the GM-CSF promoter in a cell-specific manner. A region spanning two cytokine-specific sequences, cytokine 1 (CK-1, 5' GAGAITC- CAC 3') and cytokine 2 (CK-2, 5' TACCTTA 3') bound two nuclear proteins [nuclear factor (NF)-GMA and NF-GMb] from GM-CSF-expressing cells in gel retardation assays. NF-GMb was inducible with phorbol 12-myristate 13-acetate and accompanied induction of GM-CSF message. NF-GMb was absent in cell lines not producing GM-CSF, some of which had other distinct binding proteins. NF-GMA and NF-GMb eluted from a heparin-Sepharose column at 0.3 and 0.6 M KCI, respectively. We hypothesize that the sequences CK-1 and CK-2 bind specific proteins and regulate GM-CSF transcription.

Human granulocyte/macrophage colony-stimulating factor (GM-CSF) is a 22-kDa glycoprotein that stimulates the formation of granulocyte, macrophage, granulocyte/macrophage, and eosinophil colonies from normal bone marrow progenitor cells in vitro (1). GM-CSF has direct action also on the function of mature peripheral blood granulocytes (2, 3). The gene encoding human GM-CSF is 2.5-3 kilobases long (4, 5) and maps to the long arm of chromosome 5 (6).

Antigen- or mitogen-activated T cells and T-cell lines produce relatively high levels of GM-CSF (7-9). Other cytokines such as interleukin 1 and tumor necrosis factor activate the expression of the GM-CSF gene in endothelial cells (10). Primary human stromal cells can also be induced to produce GM-CSF (11).

It appears that there may be multiple forms of control of GM-CSF production, both at a transcriptional and posttranscriptional level. mRNA stability has been shown to be involved in controlling the induction of mouse GM-CSF in macrophages (12), and a sequence in the 3'-untranslated region is responsible for the instability of GM-CSF mRNA (13). It has been reported (14) that a 650-base-pair (bp) fragment from the promoter region of the human GM-CSF gene was involved in T-cell-specific expression of the gene and was only active in phytohemagglutinin/phorbol 12-myristate 13-acetate (PMA)-stimulated T cells.

Computer analysis of the mouse and human GM-CSF genes has shown that the most highly conserved sequences are in the promoter region and in the 3'-noncoding sequence of the mRNA (5), indicating the potential importance of these regions in the regulation of expression of these genes.

Sequences shared with other cytokines can also be found especially in the promoter region of the genes (5, 15, 16). Sequences within this promoter region, especially those sequences shared with other cytokine genes, may be the binding sites for nuclear proteins that confer cell specificity or inducibility on the GM-CSF gene.

In this study, we have examined the interaction of the GM-CSF promoter with nuclear proteins from cells that express GM-CSF and from cells where GM-CSF is not produced. We have also used synthetic oligonucleotides to examine the interaction of nuclear proteins with conserved sequences in the GM-CSF promoter. We report cell-specific interactions with promoter regions of the GM-CSF gene.

MATERIALS AND METHODS

DNA Probes. A 650-bp Pst I fragment from the promoter region of the human GM-CSF gene (Fig. 1a) was cloned into pUC18 and was the source of all fragments used in gel retardation assays (the GM-CSF genomic clone was a gift from J. Gasson, School of Medicine, University of California, Los Angeles). The fragments routinely used were a 445-bp Sac I fragment that was end-labeled with T4 DNA polymerase (Pharmacia, Uppsala, Sweden) (17) and a 199-bp Dde I–Sac I fragment derived from the 5' end of the larger Sac I fragment (Fig. 1a) that was radiolabeled by end-filling with the Klenow fragment of DNA polymerase 1 (Bresa, Adelaide, Australia) (17).

Two complementary 41-bp oligonucleotides were synthesized with EcoRI ends (Fig. 1b). Each oligonucleotide was end-labeled with [γ-32P]ATP (Bresa) and polynucleotide kinase (Pharmacia). The radiolabeled oligonucleotides were annealed by heating to 100°C for 5 min in 25 mM Tris-HCl, pH 7.6/150 mM NaCl and then cooling at room temperature for 15 min. Unlabeled oligonucleotides were also annealed as described above to give a final concentration of 10 ng/μl and used as specific competitors in the binding reactions.

Cell Lines. U5637 is a human bladder carcinoma cell line that constitutively produces GM-CSF and granulocyte colony-stimulating factor (G-CSF) (18). HUT78 is a T-lymphoblastoid cell line derived from a patient with Sezary syndrome (19). SP2/10-Ag14 is a mouse myeloma cell line (20) and LiBr is a human melanoma cell line (21). GM-CSF mRNA can be detected in U5637 and HUT78 cells but not in LiBr and SP2 cells (M.F.S., unpublished results).

All cell lines were routinely grown in RPMI medium supplemented with 10% (vol/vol) fetal calf serum. Cells were harvested at 106 cells per ml or at 80% confluence for nonadherent and adherent cell lines, respectively, following treatment for 6 hr with PMA at 10 ng/ml. Untreated cells were grown for the same period of time but without PMA.

Abbreviations: GM-CSF, granulocyte/macrophage colony-stimulating factor; PMA, phorbol 12-myristate 13-acetate; G-CSF, granulocyte colony-stimulating factor; IL-3, interleukin 3; NF, nuclear factor.
Preparation of Nuclear Extracts. Nuclei were prepared as described by Dignam et al. (22). Nuclear proteins were extracted by constant agitation for 30 min with 0.5 M KCl at 4°C. Following centrifugation at 20,000 rpm for 30 min (Beckman rotor TL100.3), the supernatant was dialyzed against three changes of TM buffer containing 100 mM KCl for 12–16 hr. (TM buffer is 50 mM Tris-HCl (pH 7.6), 12.5 mM MgCl₂, 1 mM EDTA, 1 mM dithiothreitol, 20% (vol/vol) glycerol; ref. 23). The protein extracts were stored at −70°C. Protein concentration was determined using the Bio-Rad protein assay (Bio-Rad).

Gel Retardation Assay. For binding reactions, 0.1–1.0 ng of radiolabeled fragment (5–10,000 cpm) was mixed with 1–3 μg of nuclear extract in a final volume of 20 μl containing 25 mM Tris-HCl (pH 7.6), 6.25 mM MgCl₂, 0.5 mM EDTA, 0.5 mM dithiothreitol, 10% (vol/vol) glycerol, and 80–200 mM KCl. Poly(dl-dC) (0.5–3 μg) was used in all reactions as nonspecific competitor. Specific competitors were added to each reaction mixture as described in individual experiments. The reaction mixtures were analyzed on 5% polyacrylamide gels in low-ionic-strength buffer (24). The gels were preelectrophoresed at 20 V/cm for 2 hr and electrophoresed at the same voltage for 1–2 hr. Following electrophoresis the gels were dried and autoradiographed either overnight or for 1–2 days.

Binding reactions with synthetic oligonucleotides were as described above. The retardation patterns were analyzed on 11% polyacrylamide gels in 0.5 × TBE (1.0 × TBE = 50 mM Tris borate, pH 8.3/1 mM EDTA).

Heparin-Sepharose Column Chromatography. A 1-ml heparin-Sepharose (Pharmacia) column was equilibrated with TM buffer containing 100 mM KCl (TM.1). Six milligrams of crude nuclear extract from PMA-treated U5637 cells was loaded onto the column in the same buffer. Following extensive washing with TM.1, the bound protein was eluted in a stepwise fashion with 3 ml of 0.2 M, 3 ml of 0.3 M, and 3 ml of 0.6 M KCl in TM buffer. The eluates from each salt concentration were collected and dialyzed into TM.1, and the protein concentration was estimated with the Bio-Rad assay. The fractions were tested for binding activity to the synthetic oligonucleotides.

RESULTS

Specific Interaction of Nuclear Proteins with the Promoter Region of the GM-CSF Gene. A 445-bp Sac I fragment from the GM-CSF promoter region (Fig. 1a) was radiolabeled and mixed with nuclear extracts prepared from the U5637 bladder carcinoma cell line that had been treated for 6 hr with PMA at 10 ng/ml. Gel retardation assays resulted in a single retarded band (Fig. 2A), which was specifically blocked by competition with 100 ng of unlabeled Sac I fragment but not with 500 ng of poly (dl-dC) (Fig. 2A, compare lanes 2 and 4).

To define more precisely the promoter region responsible for the protein–DNA interaction, competition experiments were carried out with a number of smaller fragments spanning the 445-bp region. The binding could be successfully blocked by competition with a 199-bp Dde I–Sac I fragment (Fig. 2A, lane 6) comprising the region closest to the transcription initiation point (Fig. 1a). Fragments further upstream did not compete successfully in the retardation reactions. Fig. 2A (lanes 8 and 9) shows the inability of a 130-bp Pst I–Sac I fragment (positions −603 to −457) to block the binding by competition.

When binding reactions were carried out with radiolabeled 199-bp Dde I–Sac I fragment and nuclear extract from PMA-stimulated U5637 cells, one main retarded complex was observed (Fig. 2B, lane 1). A fainter complex, migrating closer to the free DNA was seen on longer exposure of the gels to x-ray film (Fig. 2B). The formation of these complexes could be specifically blocked by competition with 100 ng of unlabeled fragment (Fig. 2B, lane 2).

Comparison of Promoter DNA–Protein Interactions from Various Cell Lines. To determine if any proteins interacting

Fig. 1. (a) Restriction enzyme map of the promoter region of the human GM-CSF gene. The sequence is numbered from the start of transcription (+1). The sequence between the Sac I site at position −11 and position −96 is expanded to show both the sequence conserved between the cytokine genes and the TATA box sequence. P, Pst I; S, Sac I; D, Dde I. (b) Sequences of the 41-bp synthetic oligonucleotides (GM) that span the conserved cytokine sequence (GAGATTCAC). Conserved regions are underlined. Sequence of the unrelated 32-bp oligonucleotide X is also shown.

\[ \text{GM} \]
\[ 5'\text{AATCTGTCAGCCGCAATCCGCATGGTCATGGC} \]
\[ 3'\text{GCTATCGTTCGTCAGGCATCCCAATG} \]

\[ X \]
\[ 5'\text{CTGAGAGCTCCGCGTCAGGCAGACCTTC} \]
\[ 3'\text{CAGAGGAGGCGAGCTGCTGTTCCAGAGAGCT} \]
with the 445-bp or 199-bp GM-CSF promoter fragments are cell-type-specific, we made nuclear protein extracts from U5637, HUT78, and SP2 cell lines treated for 6 hr with PMA at 10 ng/ml. Fig. 3A shows the retarded bands obtained with extracts from each cell line following binding to the 445-bp Sac I fragment. U5637 and HUT78 extracts show the same retarded bands, with the major band corresponding to the U5637 band described above (band 3). A weaker band migrating higher on the gel is also visible in both extracts (band 2). The SP2 extract, however, shows a retarded band (band 1) that migrates higher on the gel than either retarded complex from the other cell lines.

We also examined the interaction between these extracts and the smaller 199-bp fragment. The three cell lines examined showed different retardation patterns (Fig. 3B). The main retarded band observed with U5637 cell extracts (band 1) was also present in extracts from the other two cell lines but at reduced levels. HUT78 cell extracts generated a different band (band 2) as did SP2 cell extracts (band 3). Band 2 was seen in some U5637 cell extracts. These results show that each of these cell lines contains a distinct set of proteins capable of interacting with the GM-CSF promoter.

A Conserved Cytokine-Specific Sequence Binds Nuclear Proteins. Two complementary oligonucleotides (each 41 bp long) spanning the sequence 5' GAGATTCCAC 3' (Fig. 1b), were synthesized to investigate the interaction of nuclear proteins with this sequence. Two specific retarded complexes, labeled a and b in Fig. 4A, were generated with extracts from PMA-stimulated U5637 cells. These complexes will be referred to as nuclear factor (NF)-GMa and NF-GMb.

The formation of these two complexes could be completely blocked by competition with 50 ng of unlabeled annealed oligonucleotides (Fig. 4A, lane 5) but not with the same concentration of an unrelated oligonucleotide (Fig. 4A, lane 10) or with 3 μg of poly(dI-dC) (data not shown). The apparent enhancement of NF-GMa and NF-GMb seen here with increasing concentrations of the nonspecific competitors is not reproducible between experiments. Other retarded complexes migrating higher on the gel or closer to the free DNA are not consistently observed, and their formation cannot be blocked by competition with increasing concentrations of specific competitor (Fig. 4A). Identical results were obtained with PMA-stimulated HUT78 extracts (data not shown). Increasing the salt concentration in the binding reactions from 80 mM to 200 mM greatly reduces this nonspecific interaction and enhances the specific interactions by ≈3-fold (data not shown). Subsequent binding reactions were, therefore, carried out at 200 mM KCl.

The two retarded complexes could result either from interaction with multimers of the same protein or with two distinct proteins. The nuclear extract from U5637 cells was fractionated on a heparin-Sepharose column. The proteins eluted from the column with 0.1 M, 0.2 M, 0.3 M, and 0.6 M KCl were tested in retardation assays (Fig. 4B). The protein(s) responsible for the NF-GMa complex eluted from the column with 0.3 M KCl and that responsible for NF-GMb eluted with 0.6 M KCl (Fig. 4B, lanes 5 and 6 and lanes 7 and 8, respectively), suggesting the involvement of two distinct proteins in these complexes.

Cell-Specific Interactions with the Conserved Cytokine Sequence. We have compared the retardation band patterns obtained with extracts prepared from U5637, HUT78, LiBr, and SP2 cell lines. The extracts were prepared from cells treated for 6 hr with PMA at 10 ng/ml. Both of the specific retarded bands NF-GMa and NF-GMb were obtained using the radiolabeled oligonucleotides and nuclear protein was used for PMA-treated U5637 and HUT78 cells (Fig. 5A, lanes 1 and 2), although HUT78 cell extracts always yielded a 3- to 4-fold higher concentration of the proteins involved in both these complexes. Extracts from the SP2 cell line did not result in either of the specific complexes but gave a diffuse retarded band migrating above NF-GMa (Fig. 5A, lane 3). This interaction is blocked by competition with increasing concentrations of poly(dI-dC) (data not shown). Extracts from PMA-treated LiBr cells bound to the GM-CSF-specific oligonucleotide giving rise to NF-GMa but not NF-GMb. The amount of NF-GMa formed with extracts from LiBr...
cells was always lower than that seen in U5637 cells by a factor of ≈3 (Fig. 5A, lane 4).

Effect of PMA Stimulation on the Formation of NF-GMa and NF-GMb DNA–Protein Complexes. To determine if PMA treatment induced the production of the proteins involved in the two cytokine-specific complexes, extracts were prepared from the U5637 and LiBr cell lines before and after PMA treatment. The retarded patterns obtained with these extracts indicate that in U5637 cells the protein involved in the NF-GMb complex was induced 5- to 10-fold (Fig. 5B, compare lanes 1 and 2). The induction seen here in the level of NF-GMa is not consistent between experiments. In LiBr cells this induction was not observed, and no change in the level of NF-GMa was seen (Fig. 5B, lanes 3 and 4).

DISCUSSION

The experiments described show that there are cell-specific as well as inducible nuclear proteins that interact with DNA fragments from the GM-CSF promoter.

Nuclear proteins from U5637 cells, a cell line that expresses GM-CSF mRNA, specifically bound to both a 445-bp fragment, spanning a large area of the promoter, and a 199-bp fragment in a region close to the transcription start site. The SP2 cell line, where GM-CSF mRNA is not detectable, has proteins that bound to the promoter region of the GM-CSF gene but formed a complex distinct from that found in U5637 or HUT78 cells, where mRNA for GM-CSF is detectable. Since the human and mouse GM-CSF promoters share >80% of the sequences in the first 350 bp, it seems valid to compare human and mouse cell lines. The presence of cell-specific DNA binding proteins has now been shown for a number of gene promoters where it has been hypothesized that these interactions are involved in cell-specific expression (24–28). It is possible then that the binding proteins from SP2 cells may be interacting with common promoter regions such as the TATA box, not involved in tissue-specific transcription, or with negative regulator regions of the promoter, such as those DNA-binding proteins shown to be involved in the negative regulation of the interferon β gene (29, 30). Some of the factors identified here that bind to the GM-CSF promoter may be involved in the inducible or tissue-specific nature of GM-CSF expression.

Comparison of the promoter regions from a number of cytokine genes has revealed some sequences that are conserved between these genes and across species (5, 14, 15). One decanucleotide sequence [cytokine 1 (CK-1)], 5’ GGRRTTYCCAY 3’ (where R = A or G and Y = C or T) is found in both human and mouse interleukin 2, interleukin 3 (IL-3), GM-CSF, and G-CSF genes (Fig. 6). In addition a second sequence [cytokine 2 (CK-2)], 5’ TCAGRTA 3’, lying on the 3’ side of the decanucleotide, is conserved in both human and murine GM-CSF and IL-3 genes (Fig. 6). This sequence is not found in human or murine G-CSF or interleukin 2. The CK-1 sequence is also repeated further upstream in the human GM-CSF (14) and murine IL-3 (32) genes but without the extra flanking conserved sequence (CK-2). We designed our 41-bp oligonucleotide probe spanning the CK-1 and CK-2 sequences to elucidate their role in nuclear protein binding. These two sequences are flanked by GM-CSF sequences that are not conserved in the other genes.

Two DNA–protein complexes that are specific for the 41-bp oligonucleotide spanning these conserved sequences were identified. It would appear that these two complexes are generated by two or more different proteins since the proteins involved in the NF-GMa and NF-GMb complexes are eluted from a heparin-Sepharose column at 0.3 M and 0.6 M KCl, respectively. It has yet to be determined which specific nucleotides are involved in the binding of these proteins.
but have ever, transfection assays. The binding of these cells to gene expressed protein-DNA is likely, therefore, to coordinate the activation of cytokine genes. U5637 (25, 26). We have investigated the role of cytokine genes in the NK1- and GM-CSF-mediated activation of U5637 cells, which express the GM-CSF receptor (27, 28). It is possible that the GM-CSF receptor is involved in the activation of U5637 cells, and that the GM-CSF receptor may be involved in the activation of cytokine genes. This is supported by the finding that the GM-CSF receptor is expressed on the surface of activated U5637 cells. In addition, the expression of the GM-CSF receptor and the GM-CSF receptor-associated protein is induced by GM-CSF in U5637 cells. However, the role of the GM-CSF receptor in the activation of cytokine genes remains to be determined.

Fig. 6. Conserved sequences found in the promoter region of cytokine genes. CK-1 and CK-2 sequences are underlined. All the sequences have been compared to the human GM-CSF sequence. Nonconserved bases are shown in lowercase letters. Numbering is relative to the transcription start site of each gene (+ 1). h: Human; m: murine; IL-2: interleukin 2. Numbers in parentheses after gene name refer to reference number in list of references.

The CK-1 sequence has been postulated to account for the coordinate expression of some of these genes in activated T cells (14, 15). However, G-CSF is not expressed in activated T cells (13) despite the fact that its promoter contains a copy of CK-1. It seems unlikely, therefore, that this sequence alone is responsible for T-cell expression of GM-CSF or IL-3. However, the extended conserved sequence of CK-1 and CK-2 noted above between GM-CSF and IL-3 genes may be involved in T-cell expression of these genes. It is noteworthy that IL-3 and GM-CSF are often coordinately expressed following Con A stimulation of murine T-cell clones (36).

Treatment of U5637 cells with PMA results in an increase in GM-CSF mRNA and an increase in the level of the NF-GMb complex but not NF-GMa. The ability to induce the NF-GMb complex in PMA-treated U5637 cells suggests that the protein(s) involved in the NF-GM complex may be responsible for the inducibility of the GM-CSF gene. However, unstimulated cells also transcribe the GM-CSF gene but have undetectable levels of NF-GMb. Therefore, NF-GMb may not be essential for basal level transcription of the gene in this tumor cell line but could be involved in increasing transcriptional efficiency following stimulation. The situation observed here parallels that found with proteins binding to the octamer motif of the immunoglobulin genes (25, 26). Two proteins, one of which is lymphoid specific, bind to this motif. The lymphoid-specific protein is inducible with lipopolysaccharide (26), as is NF-GMb with PMA in U5637 cells. Extracts from the human melanoma cell line LiBr contain protein(s) that may be equivalent to NF-GMa, but NF-GMb could not be induced by PMA treatment of these cells. Nuclear proteins prepared from the mouse splenic cell line did not bind specifically to the GM-CSF conserved sequence. It would appear, therefore, that the protein(s) that bind to this GM-CSF sequence are limited in their cellular distribution. The result is consistent with results (14) that show that a GM-CSF-chromatophenol acetyltransferase fusion gene is expressed in activated T cells but not in a human B-lymphoblastoid cell line. The elucidation of the role of these protein–DNA complexes in GM-CSF transcription awaits analysis of the DNA fragments in vivo by using transfection assays.