Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 3101–3105, May 1986
Biochemistry

Genomic cloning, characterization, and multilineage growth-promoting activity of human granulocyte-macrophage colony-stimulating factor

(growth factor/erythropoiesis/megakaryopoiesis/gene sequence/transient expression)

KENNETH KAUSHANSKY*, PATRICK J. O'HARA†, KATHLEEN BERKNER†, GERALD M. SEGAL*, FREDERICK S. HAGEN†, AND JOHN W. ADAMSON*

*Division of Hematology RM-10, University of Washington, Seattle, WA 98195; and †ZymoGenetics, Inc., 2121 North 35th Street, Seattle, WA 98103

Communicated by Earl W. Davie, December 23, 1985

ABSTRACT Through the use of long single-sequence oligonucleotide probes, the complete gene for human granulocyte-macrophage colony-stimulating factor (hGM-CSF) has been cloned from a human genomic library. The gene is 2.5 kilobases in length, contains three introns, and is present as a single copy in the human genome. When subcloned into the mammalian expression vector pD3, the gene directs the synthesis of authentic hGM-CSF. In addition to its stimulation of *in vitro* granulopoiesis and monopoiesis, recombinant hGM-CSF stimulates *in vitro* erythropoiesis and megakaryopoiesis.

Colony-stimulating factors (CSFs) are acidic glycoproteins that are required for the survival, proliferation, and differentiation of hematopoietic colony-forming cells in culture (1). Individual CSFs were originally defined by the cell type(s) in the resulting colonies. Hence, granulocyte-macrophage CSF (GM-CSF) stimulates the production of granulocytes and macrophages from their progenitor cells. There are at least three distinct human CSFs [M-CSF or CSF-1 (2), GM-CSF or CSFα (3), and G-CSF or CSFβ (4)]. A fourth CSF [multi-CSF or interleukin 3 (IL-3)] has been purified from murine sources (5).

Murine GM-CSF is a 23-kDa protein that has been purified to homogeneity from medium conditioned by endotoxin-stimulated mouse lung (6). A partial amino acid sequence has been reported (7), and the cDNA for the molecule has been cloned (8). Human GM-CSF (hGM-CSF) has been partially purified from placental-conditioned medium (3) and purified from Mo cell-conditioned medium (9). The cDNA for GM-CSF has been cloned recently from Mo cells (10), from a helper T-cell line (11), and from normal lymphocytes (12). The sequence predicts a polypeptide of 144 amino acids and codes for a 22-kDa glycoprotein that stimulates *in vitro* GM production. In addition, GM-CSF stimulates mature neutrophils to localize at sites of inflammation (13), enhances their cytotoxicity (14), and stimulates eosinophils to kill parasites (15).

Physiologic studies with GM-CSF have been hampered by inadequate amounts of impure preparations. To better understand the function of CSFs and their mechanism of action in hematopoiesis, we have used long single-sequence oligonucleotide probes to screen a human genomic library and have cloned and expressed the gene for GM-CSF. By analyzing the gene and its product, we hope to better understand its method of regulation, to dissect the functional organization of the molecule, and to examine the role of GM-CSF in normal and in stimulated hematopoiesis.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Oligonucleotide Probes and Screening. Partially complementary pairs of oligonucleotides, 40 bases in length based on the sequence of a murine GM-CSF cDNA, were end-labeled with [γ-^{32}P]ATP, annealed, and filled in using the Klenow fragment of DNA polymerase (5–10 $\times$ 10^8 cpm/μg, ref. 16). Clones (4 $\times$ 10^6) of a once-amplified λ Charon 4A phage library (kindly provided by Tom Maniatis) were plated at a density of 2.5 $\times$ 10^4 per 137-mm plate. Nitrocellulose replica filters were prepared and prehybridized in 0.75 M NaCl/75 mM sodium citrate, pH 7, 5$\times$ concentrated Denhardt's solution, 20% formamide, 50 mM phosphate (pH 7), 1 mM sodium pyrophosphate, 100 μM ATP, and 50 μg of sonicated salmon sperm DNA per ml at 37°C overnight. Boiled probe (1 $\times$ 10^6 cpm/ml) was added and the filters were allowed to hybridize at 37°C overnight. The filters were washed twice in 30 mM NaCl/3 mM sodium citrate/0.1% NaDodSO$_4$ at room temperature for 15 min and at 40°C for 30 min.

Genomic Sequencing. A 3.2-kilobase (kb) *Hin*dIII/*Eco*RI fragment was prepared by digestion, size-fractionation, and electroelution and was subcloned into pUC13 (pUC hGM-CSF). A detailed restriction map was generated. Discrete double-digestion fragments were subcloned into M13 mp18 or mp19 vectors and sequenced by the dideoxynucleotide chain-termination method (17). The entire gene was sequenced in both directions.

Analysis of Human Genomic DNA. High molecular weight human genomic DNA was purified from leukocyte lysates and then digested with restriction endonucleases, size-fractionated by 0.6% agarose gel electrophoresis, transferred to nitrocellulose, and prehybridized as above with the addition of 0.5 mg of yeast tRNA per ml. A 251-base-pair (bp) *Sst* I/*Pvu* II fragment was isolated from pUC hGM-CSF, subcloned into pSP65, and transcribed into [^{32}P]RNA by SP6 RNA polymerase (3.3 $\times$ 10^8 cpm/μg, ref. 18). The blot was hybridized with 2 $\times$ 10^6 cpm/ml overnight, washed in 0.3 M NaCl/0.03 M sodium citrate/0.1% NaDodSO$_4$ at 20°C, 40°C, and 65°C for 30 min, dried, and exposed for 4 days.

Gene Expression. A 2.6-kb *Bst*EII/*Eco*RI fragment was subcloned into pD3. Its orientation was confirmed by restriction analysis (pD3 hGM-CSF). COS cells were transfected with 10 μg of pD3 hGM-CSF by calcium phosphate precipitation (19). The cells were incubated in Dulbecco's minimal essential medium supplemented with 10% heat-inactivated fetal calf serum at 37°C under 7% CO$_2$ in air. Incubation medium was collected 3 days after transfection and assayed.

Abbreviations: CSF, colony-stimulating factor; GM, granulocyte-macrophage; hGM-CSF, human GM-CSF; kb, kilobase(s); IL-3, interleukin 3; BFU-E, erythroid burst-forming cell(s); CFU-Meg, megakaryocyte colony-forming cell(s); CFU-GM, EM colony-forming cell(s); bp, base pair(s).

Supernatant (1.5 ml) was concentrated 5-fold by ultrafiltration and applied to a 0.9 cm × 40 cm column of Bio-Gel P-100 equilibrated in Tris-buffered saline with 0.1% Tween 20. Fractions were collected, sterilized by filtration, and assayed.

Hematopoietic Colony Assays. Normal human bone marrow was fractionated on a Ficoll/Hypaque density gradient (1.077 g/ml). Low-density mononuclear cells were depleted of adherent cells by double plastic adsorption and of T cells by erythrocyte-rosetting (20). For GM colony growth, 5×10^4 cells were plated in 1 ml of α medium supplemented with 10% preselected fetal calf serum, antibiotics, and 0.1–10% of the sample to be assayed and made semi-solid with methylcellulose (0.9%). Fifty units of GM-CSF was defined by the dilution that produced half-maximal stimulation of GM colony formation per 10^5 target cells. For the growth of erythroid burst-forming cells (BFU-E), 1.0 unit of recombinant human erythropoietin (Amgen, Thousand Oaks, CA) was added on day 4 to otherwise standard cultures (21). For the growth of megakaryocyte colony-forming cells (CFU-Meg), 25% human plasma was substituted for fetal calf serum (22). All cultures were incubated for 13 days at 37°C under 5% CO_2 in air. A GM colony was defined as 50 or more cells and a megakaryocyte colony was defined as 4 or more cells. Erythroid bursts contained at least 500 hemoglobin-containing cells. The cellular composition of GM colonies was evaluated by culture in 0.3% agar. Whole 1-ml cultures were fixed and stained for myeloperoxidase, counterstained with toluidine blue, and examined by light microscopy.

RESULTS

Genomic Library Screening. Clones (4×10^6) of a once-amplified human genomic library were screened using a single-sequence 70-base oligonucleotide probe derived from the middle of the murine GM-CSF cDNA (Fig. 1). Of 72 plaques giving an initial hybridization signal, 10 were plaque-purified and evaluated by restriction analysis and blotting. Additional probes to the 5′ and 3′ ends of the hGM-CSF cDNA were designed once it became available (10) and were used to confirm authenticity and to determine the extent of homology. Three clones were found to hybridize to the 5′ end probe, 3′ end probe, and original probe from the middle of the cDNA for murine GM-CSF and were found to be identical (λ hGM-CSF).

Structure of the Gene for hGM-CSF. Restriction analysis of λ hGM-CSF revealed a 16.2-kb insert with two internal *Eco*RI sites. All three probes hybridized with a 6.7-kb fragment. Digestion and blotting of this clone with *Eco*RI and *Hin*dIII indicated that the gene for hGM-CSF is <3.2 kb in length. This fragment was subcloned into pUC13 (pUC hGM-CSF). Dideoxynucleotide sequencing was carried out on restriction fragments generated from pUC hGM-CSF. The genomic sequence for hGM-CSF was compared to several hGM-CSF cDNA clones (10–12) and is presented in Fig. 2. Based on this comparison, the gene for hGM-CSF contains three introns of 98, 687, and 808 bp. There were several sequence differences noted in the coding regions of hGM-CSF.

```
   AAG ATA TTC GAG CAG GGT CTA CGG GGC AAT TTC ACC   Murine cDNA
   lys ile phe glu gln gly leu arg gly asn phe thr   Murine Protein
   glu leu tyr lys  -   -   -   -   -  ser leu  -    Human Protein
5' GAG CTG TAC AAG CAG GGT CTA CGG GGC AGC CTC ACC   Probe I
                                        3' GAG TGG   Probe II

   AAA CTC AAG GGC GCC TTG AAC ATG ACA GCC AGC TAC   Murine cDNA
   lys leu lys gly ala leu asn met thr ala ser tyr   Murine Protein
    -   -   -   -  pro  -  thr                       Human Protein
   AAT C 3' Probe I
   TTT GAG TTC CCG GGG AAC TGG TAC TGT CGG TCG A 5'  Probe II
```

FIG. 1. Synthetic oligonucleotide probes for screening λ Charon 4A human genomic library. Two partially complementary oligonucleotides were synthesized based on the region of the murine cDNA sequence containing two potential glycosylation sites. The amino acids known to differ in the human polypeptide are noted.* The sequence of the probe was altered based on the frequency of codon usage for human proteins (23).

*Golde, D. W., Annual Meeting of the American Society of Hematology, December 1–4, 1984, Miami, FL.

The cDNA for hGM-CSF is reported to contain a sequence heterogeneity at amino acid position 117 (ACT/Thr or ATT/Ile, ref. 10). We have found ATT/Ile at this position as have others (11, 12). The cDNA sequence reported by Cantrell *et al.* (12) differs from the genomic sequence reported here at two positions: in the codons for amino acid 27 (ACA/Thr) and for amino acid 102 (AAA/Lys). We have found ACG/Thr and AAG/Lys for these positions, in agreement with others (10, 11). Given the relatively high transcription error rate of reverse transcriptase in generating cDNA molecules (32), sequence differences are not unexpected.

Gene Copy Number. Human genomic DNA was prepared, digested with various restriction endonucleases that do not cut the hGM-CSF gene, fractionated by agarose gel electrophoresis, and transferred to nitrocellulose. An RNA probe representing a 251-bp *Sst* I/*Pvu* II fragment containing exon 1 and part of intron 1 was prepared and used to probe the genomic digests (Fig. 3). A single band hybridizes with probe in all lanes under washing conditions of moderate stringency.

Gene Expression. pD3 is a simian virus 40 (SV40) ori-based plasmid vector and was used to express biologically active GM-CSF from COS cells. It contains a unique *Bcl* I cloning site separating an adenovirus late promoter and tripartite leader sequence from a poly(A) signal. The vector also contains the enhancer sequence from SV40 and a set of splice donor/acceptor signals. A 2.6-kb *Bst*EII/*Eco*RI fragment of λ hGM-CSF from the TATA box to the poly(A) site was subcloned into pD3. Plasmid DNA was transfected into COS cells and incubation medium was collected. GM-CSF was detected in the culture medium by bioassay. In five separate transfection experiments, the crude conditioned medium contained $1.9–2.9 \times 10^4$ units of GM-CSF per ml. Approximately 40% of the colonies contained eosinophils. Neutrophil, monocyte, and mixed granulocyte/monocyte colonies were nearly equally represented. In addition, recombinant hGM-CSF was able to enhance the growth of BFU-E and CFU-Meg to form colonies *in vitro*. At a concentration of crude COS cell-conditioned medium that stimulated half-maximal GM colony growth, 50% of a maximal number of erythroid bursts and 17% of a maximal number of megakaryocyte colonies developed. Occasional mixed erythroid-myeloid colonies were present.

The COS cell supernatant was concentrated and size-fractionated by gel filtration. The GM-CSF produced by the transfected COS cells eluted in a broad-based peak with a relative molecular mass of 20–23 kDa (Fig. 4). The activity stimulating CFU-GM, BFU-E, and CFU-Meg cofractionated during this partial purification. The most active fractions were pooled (as indicated in Fig. 4) and were evaluated by dose–response analysis (Fig. 5). A concentration-dependent increase in the growth of all colony types was observed with inhibition seen at higher concentrations.

DISCUSSION

Human genomic libraries containing inserts of 7–20 kb have been constructed in λ cloning vectors and have been used successfully to obtain full-length genomic clones. Recently,

```
AAGCTTGCTGAGAGTGGCTGCAGTCTCGCTGCTGGATGTGCACATGGTGGTCATTCCCTCTGCTCACAGGGGCAGGGGTCCCCCTTACTGGACTCAGGTTGCCCCCTGCTCCAGGTCCT
                                                                                                                    100

GGGTGGGAGCCCATGTGAACTGTCAGTGGGGCAGGTCTGTGAGAGCTCCCCTCACACTCAAGTCTCTCACAGTGGCCAGAGAAGAGGAAGGCTGGAGTCAGAATGAGGCACCAGGGCGGG
                                                                                                     200

CATAGCCTGCCCAAAGGCCCCTGGGATTACAGGCAGGATGGGGAGCCCTATCTAAGTGTCTCCCACGCCCCACCCCAGCCATTCCAGGCCAGGAAGTCCAAACTGTGCCCCTCAGAGGGA
                                                                                 300

GGGGGCAGCCTCAGGCCCATTCAGACTGCCCAGGGAGGGCTGGAGAGCCCTCAGGAAGGCGGGTGGGTGGGCTGTCGGTTCTTGGAAAGGTTCATTAATGAAAACCCCCAAGCCTGACCA
                                                            400

CCTAGGGAAAAGGCTCACCGTTCCCATGTGTGGCTGATAAGGGCCAGGAGATTCCACAGTTCAGGTAGTTCCCCCGCCTCCCTGGCATTTTGTGGTCACCATTAATCATTTCCTCTGTGT
                  500                                                                                                     600
                                                                              1                                       10
                                                                             MET Trp Leu Gln Ser Leu Leu Leu Leu Gly Thr Val Ala Cys
ATTTAAGAGCTCTTTTGCCAGTGAGCCCAGTACACAGAGAGAAAGGCTAAAGTTCTCTGGAGG ATG TGG CTG CAG AGC CTG CTG CTC TTG GGC ACT GTG GCC TGC
                                                                                                                700
                    20                                      30                                      40
Ser Ile Ser Ala Pro Ala Arg Ser Pro Ser Pro Ser Thr Gln Pro Trp Glu His Val Asn Ala Ile Gln Glu Ala Arg Arg Leu Leu Asn
AGC ATC TCT GCA CCC GCC CGC TGC CCC AGC CCC AGC ACG CAG CCC TGG GAG CAT GTG AAT GCC ATC CAG GAG GCC CGG CGT CTC CTG AAC

                    50
Leu Ser Arg Asp Thr Ala Ala Glu Met
CTG AGT AGA GAC ACT GCT GCT GAG ATG GTAAGTGAGAGAATGTGGGCCTGTGCCTAGGCCACCCAGCTGGCCCCTGACTGGCCACGCCTGTCAGCTTGATAACATGACATT
   800                                                                                                                  900
                                             60
                Asn Glu Thr Val Glu Val Ile Ser Glu Met Phe Asp Leu Gln
TTCCTTTTCTACAG AAT GAA ACA GTA GAA GTC ATC TCA GAA ATG TTT GAC CTC CAG GTAAGATGCTTCTCTCTGACATAGCTTTCCAGAAGCCCCTGCCCTGGGG
                                                                                                                    1000

TGGAGGTGGGGACTCCATTTAGATGGCACCACACAGGGTTGTCCACTTTCTCTCCAGTCAGCTGGCTGCAGGAGGAGGGGTAGCAACTGGGTGCTCAAGAGGCTGCTGGCCGTGCCCC
                                                                                                        1100

TATGGCAGTCACATGAGCTCCTTTATCAGCTGAGCGGCCATGGGCAGACCTAGCATTCAATGGCCAGGAGCTACCAGGGGACAGGTGGTAAAGTGGGGGTCACTTCATGAGACAGGAGCT
                                                                                           1200

GTGGGTTTGGGGCGCTCACTGTGCCCCGAGACCAAGTCCTGTTGAGACAGTGCTGACTACAGAGAGGCACAGAGGGGTTTCAGGAACAACCCTTGCCCACCCAGCAGGTCCAGGTGAGGC
                                                                1300

CCCACCCCCCTCTCCCTGAATGATGGGGTGAGAGTCACCTCCTTCCCTAAGGCTGGGCTCCTCTCCAGGTGCCGCTGAGGGTGGCCTGGGCGGGGCAGTGAGAAGGGCAGGTTCGTGCCT
                                        1400

GCCATGGACAGGGCAGGGTCTATGACTGGACCCACGCTGTGCCCCTCCCAAGCCCTACTCCTGGGGGCTGGGGGCAGCAGCAAAAAGGAGTGGTGGAGAGTTCTTGTACCACTGTGGGCA
     1500                                                                                                     1600
                                                    70                                      80
                                          Glu Pro Thr Cys Leu Gln Thr Arg Leu Glu Leu Tyr Lys Gln Gly Leu Arg Gly Ser Leu
CTTGGCCACTGCTCACCGACGAACGACATTTTCCACAG GAG CCG ACC TGC CTA CAG ACC CGC CTG GAG CTG TAC AAG CAG GGC CTG CGG GGC AGC CTC
                                                                                       1700
              90                                     100
Thr Lys Leu Lys Gly Pro Leu Thr Met Met Ala Ser His Tyr Lys Gln His Cys Pro Pro Thr Pro
ACC AAG CTC AAG GGC CCC TTG ACC ATG ATG GCC AGC CAC TAC AAG CAG CAC TGC CCT CCA ACC CCG GTGAGTGCCTACGGCAGGGCCTCCAGCAGGAA
                                                                                                                 1800

TGTCTTAATCTAGGGGGTGGGGTCGACATGGGGAGAGATCTATGGCTGTGGCTGTTCAGGACCCCAGGGGGTTTCTGTGCCAACAGTTATGTAATGATTAGCCCTCCAGAGAGGAGGCAG
                                                                                                          1900

ACAGCCCATTTCATCCCAAGGAGTCAGAGCCACAGAGCGCTGAAGCCCACAGTGCTCCCCAGCAGGAGCTGCTCCTATCCTGGTCATTATTGTCATTACGGTTAATGAGGTCAGAGGTGA
                                                                                      2000

GGGCAAACCCAAGGAAACTTGGGGCCTGCCCAAGGCCCAGAGGAAGTGCCCAGGCCCAAGTGCCACCTTCTGGCAGGACTTTCCTCTGGCCCCACATGGGGTGCTTGAATTGCAGAGGAT
                                                           2100

CAAGGAAGGGAGGCTACTTGGAATGGACAAGGACCTCAGGCACTCCTTCTGCGGGAAGGGAGCAAAGTTTGTGGCCTTGACTCCACTCCTTCTGGGTGCCCAGAGACGACCTCAGCCCAG
                                      2200

CTGCCCTGCTCTGCCCTGGGACCAAAAAGGCAGGCGTTTGACTGCCCAGAAGGCCAACCTCAGGCTGGCACTTAAGTCAGGCCCTTGACTCTGGCTGCCACTGGCAGAGCTATGCACTCC
        2300                                                                                                    2400

TTGGGGAACACGTGGGTGGCAGCAGCGTCACCTGACCCAGGTCAGTGGGTGTGTCCTGGAGTGGGCCTCCTGGCCTCTGAGTTCTAAGAGGCAGTAGAGAAACATGCTGGTGCTTCCTTC
                                                                                                        2500
                                                                  110                                     120
                                                                 Glu Thr Ser Cys Ala Thr Gln Ile Ile Thr Phe Glu Ser Phe Lys Glu
CCCCACGTTACCCACTTGCCTGGACTCAAGTGTTTTTTATTTTTCTTTTTTTAAAG GAA ACT TCC TGT GCA ACC CAG ATT ATC ACC TTT GAA AGT TTC AAA GAG
                                                                                                          2600
                        130                                     140             144
Asn Leu Lys Asp Phe Leu Leu Val Ile Pro Phe Asp Cys Trp Glu Pro Val Gln Glu ***
AAC CTG AAG GAC TTT CTG CTT GTC ATC CCC TTT GAC TGC TGG GAG CCA GTC CAG GAG TGAGACCGGCCAGATGAGGCTGGCCAAGCCGGGGAGCTGCTCTC
                                                                                                     2700

TCATGAAACAAGAGCTAGAAACTCAGGATGGTCATCTTGGAGGGACCAAGGGGTGGGCCACAGCCATGGTGGGAGTGGCCTGGACCTGCCCTGGGCCACACTGACCCTGATACAGGCATG
                                                                                           2800

GCAGAAGAATGGGAATATTTTATACTGACAGAAATCAGTAATATTTATATATTTATATTTTTAAAATATTTATTTATTTATTTATTTAAGTTCATATTCCATATTTATTCAAGATGTTTT
                                                                   2900

ACCGTAATAATTATTATTAAAAATATGCTTCTACTTGTCCAGTGTTCTAGTTTGTTTTTAACCATGAGCAAATGCCAGTGGGTGCCTGCCTTCCCATGAGGCAGGGGAGGAGGAAACGG
                                       3000

GGAGGTGGAGAGGGGGCGGGGGCCTCCCAGGCGTTGGGCACTATCCAAGGGCCAACACTGTCAGAGCAGAGGGGAGGTGAGAGCCGGGCATAGTCGGAATTC
      3100                                                                                            3194
```

FIG. 2. Genomic sequence of hGM-CSF. The locations of consensus sequences are underlined, and include the "minus 100" signal (24), "CAAT" box (25), "TATA" box (26), cap site (26), ribosomal binding site (27), initiation codon, intron–exon boundaries (28), internal intron splice signals (29), poly(A) signals (30), and a transcription termination signal (31). Bases matching these consensus sequences are noted by Ṅ. Potential N-linked glycosylation sites (Asn-Xaa-Thr/Ser) are noted by ◆. The putative cap site and poly(A) site [as deduced from data of Wong *et al.* (10)] are noted by 🡇. The signal peptide cleavage site is indicated by ⬇ (10).

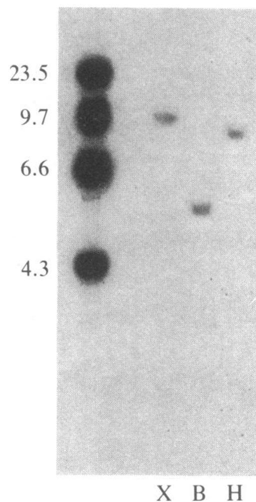

FIG. 3. Southern blot analysis of human genomic DNA. Human genomic DNA purified from leukocyte lysates was digested with *Xba* I (X), *Bam*HI (B), and *Hin*dIII (H). Sizes are shown in kb.

several genes have been cloned by using long single-sequence oligonucleotide probes (33). Assuming that homology would exist between murine GM-CSF and hGM-CSF, we designed

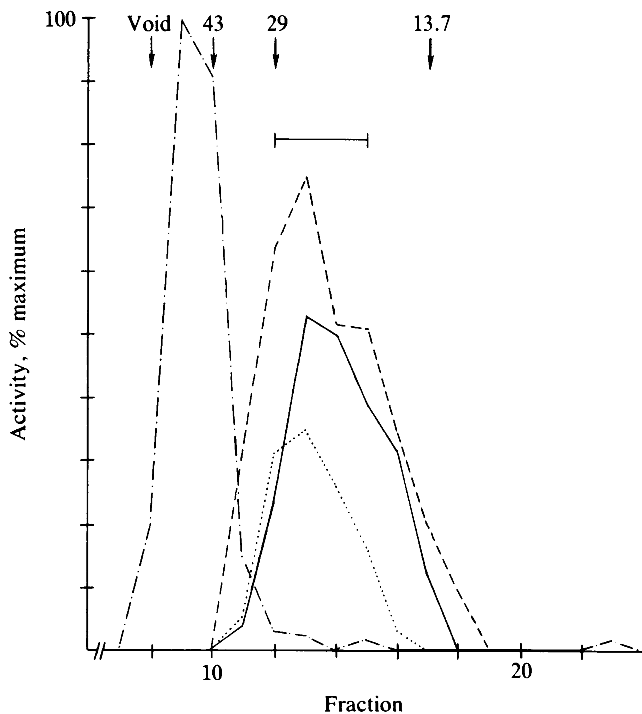

FIG. 4. Size fractionation of recombinant hGM-CSF. Concentrated medium from COS cells transiently expressing recombinant hGM-CSF was size-fractionated by gel filtration. The OD_{280} (–·–·–) was monitored and 1.25-ml fractions were collected and assayed for stimulation of GM colony-forming cells (CFU-GM) (—), BFU-E (– – –), and CFU-Meg (·····). Activity is expressed as the % of the numbers of colonies that formed in the presence of 2.5% phytohemagglutinin-stimulated lymphocyte-conditioned medium. This stimulation resulted in 192 ± 14 (mean ± SD) GM colonies, 63 ± 6 erythroid bursts, and 13 ± 2 megakaryocyte colonies per 10^5 low-density, nonadherent, T-cell-depleted marrow cells. Negative controls included α medium and the supernatant of COS cells transfected with an irrelevant gene, chloramphenicol acetyltransferase (pDCAT). Data points represent the mean of triplicate platings in three separate experiments. Molecular mass standards included blue dextran (2×10^6 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa), and RNase (13.7 kDa).

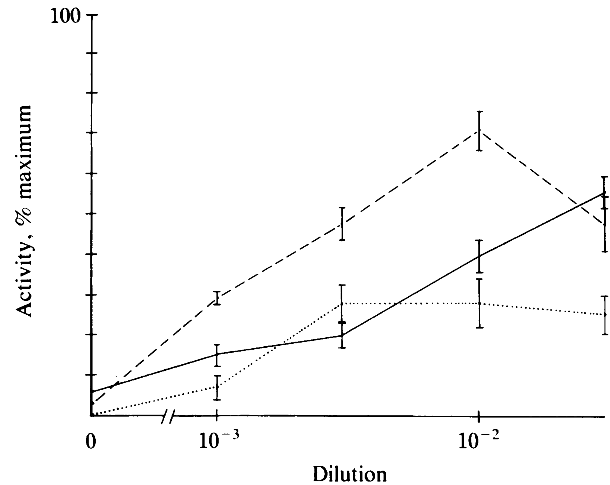

FIG. 5. Dose–response analysis of recombinant hGM-CSF. Partially purified recombinant hGM-CSF was diluted in α medium and assayed in bone marrow cultures for stimulation of CFU-GM (—), BFU-E (– – –), and CFU-Meg (·····). Activity is expressed as stated in the legend to Fig. 4.

oligonucleotide probes from the murine cDNA sequence to probe a human genomic library constructed in λ Charon 4A (34).

Based on this strategy, a full-length genomic clone for hGM-CSF was isolated. The gene contains three intervening sequences. Comparison of this sequence to consensus sequences for eukaryotic regulatory elements reveals several regions of significant homology. Beginning at base 597 is an 11-base sequence matching the canonical TATA box sequence ($G^{G}_{T}GTATA^{A}_{T}A^{TA}_{AG}$) at all but a single position (26). Thirty bases downstream is a sequence similar to a consensus cap site sequence [YYCAYYYYY (where Y = pyrimidine), ref. 26], suggesting that transcription initiates at base 629, 35 bases upstream from the initiation codon.

Two lines of evidence support the assignment of transcriptional initiation to this site. The cDNA cloned by Lee *et al.* (11) begins two bases downstream from this position. As their cDNA library was prepared by the method of Okiyama and Berg, the 5′ terminus of the clone is likely to remain intact. More importantly, primer-extension experiments using mitogen-stimulated lymphocyte RNA presented by Cantrell *et al.* (12) places the cap site only three bases downstream from the position we have selected based on nucleic acid homology. As simultaneous genomic sequencing data were not available in this report (12), the data represent only an approximation of the precise cap site.

Although the TATA box is thought to be the most critical transcription regulatory sequence (24), other sequences also modulate levels of transcription. The "CAAT box" and "minus 100 region" have been shown to regulate the level of transcription of globin genes (24). A 9-base sequence beginning 86 bp upstream from the putative cap site matches the consensus CAAT box sequence ($GG^{C}_{X}TCAATCT$, ref. 25) in five positions, and a 9-base sequence beginning 137 bases from the cap site matches the consensus minus 100 region sequence [RRCYNCACCC (where R = purine), ref. 24] in eight positions. An infrequently seen ribosomal initiation site (27) is present immediately upstream from the initiation codon. There are donor/acceptor consensus splice junction sites bounding all three introns (28, 29). Finally, two sites similar to the poly(A) signal (AATAAA, ref. 30) are present ≈300 bp from the termination codon and are followed by a 3′

termination signal (PYGTGTTYY, ref. 31) ≈20 bases downstream.

High molecular weight genomic DNA was obtained from normal blood leukocytes and digested with restriction endonucleases known not to cut the gene for hGM-CSF. Under conditions of moderate stringency, a single band, each larger than the gene for hGM-CSF, is seen on Southern blot analysis. These data suggest that the gene for hGM-CSF is present in a single copy.

The gene for hGM-CSF directs the synthesis of a biologically active hematopoietic growth factor. Using medium conditioned by transiently expressing COS cells, we were able to demonstrate that, at least, granulocyte, macrophage, eosinophil, erythroid, and megakaryocyte progenitors were stimulated to proliferate and differentiate into mature cells. To further investigate the progenitor cell range, we partially purified the recombinant hGM-CSF by gel-filtration chromatography. The activity that stimulated CFU-GM, BFU-E, and CFU-Meg was enhanced, cofractionated, and eluted with a molecular mass of ≈21 kDa. Half-maximal stimulation of the growth of the various colony types was elicited by similar concentrations of recombinant hGM-CSF.

The multilineage growth-promoting activity of recombinant hGM-CSF demonstrated in this report likely represents the direct stimulation of various hematopoietic progenitor cells. The very low background of GM, erythroid burst, and megakaryocyte colony growth seen in our cultures suggests that most of the accessory cells have been removed and are unlikely to mediate the hGM-CSF effect. Consistent with this interpretation, Metcalf *et al.* (35) in earlier studies have shown that murine GM-CSF could stimulate the first four to five cell divisions of single BFU-E and CFU-Meg. More recently, these same workers and others have reported that murine GM-CSF can stimulate megakaryocyte colony growth as well (36, 39). Thus, hGM-CSF, like its murine counterpart, appears to act on a range of progenitors to promote their growth and differentiation.

Note. Since this manuscript was initially submitted, the sequence of the hGM-CSF gene (37) as well as the multilineage effect of the growth factor (38) have been reported.

We thank Nancy Lin for performing the BFU-E and CFU-GM assays, Frank Grant and Charles Gray for technical assistance, Maggie Insley for preparation of the oligonucleotides, Ila McCullough and Margo Rogers for assistance in preparation of the manuscript, and Ko Kurachi and Earl W. Davie for many insightful discussions. K.K. is the recipient of a Physician-Scientist award (AM-01304) of the National Institutes of Health. The work was supported in part by National Institutes of Health Grants CA-31615 and AI 22772.

1. Burgess, A. W. & Metcalf, D. (1980) *Blood* **56**, 947–958.
2. Das, S. K. & Stanley, E. R. (1982) *J. Biol. Chem.* **257**, 13679–13684.
3. Nicola, N. A., Metcalf, D., Johnson, G. R. & Burgess, A. W. (1979) *Blood* **54**, 614–627.
4. Nicola, N. A., Begley, C. G. & Metcalf, D. (1985) *Nature (London)* **314**, 625–628.
5. Ihle, J. N., Keller, J., Oroszlan, S., Henderson, L. E., Copeland, T. D., Fitch, F., Prystowsky, M. B., Goldwasser, E., Schrader, J. W., Palaszynski, E., Dy, M. & Lebel, B. (1983) *J. Immunol.* **131**, 282–288.
6. Burgess, A. W., Camakaris, J. & Metcalf, D. (1977) *J. Biol. Chem.* **252**, 1998–2003.
7. Sparrow, L. G., Metcalf, D., Hunkapiller, M. W., Hood, L. E. & Burgess, A. W. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 292–296.
8. Gough, N. M., Gough, J., Metcalf, D., Kelso, A., Grail, D., Nicola, N. A., Burgess, A. W. & Dunn, A. R. (1984) *Nature (London)* **309**, 763–767.
9. Wong, G. G., Witek, J. S., Temple, P. A., Wilkens, K. M., Leary, A. C., Luxenberg, D. P., Jones, S. S., Brown, E. L., Kay, R. M., Orr, E. C., Shoemaker, C., Golde, D. W., Kaufman, R. J., Hewick, R. M., Clark, S. C. & Wang, E. C. (1985) in *Growth Factors and Transformation*, eds. Feramisco, J., Ozanne, B. & Stiles, C. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), Vol. 3, pp. 235–242.
10. Wong, G. G., Witek, J. S., Temple, P. A., Wilkens, K. M., Leary, A. C., Luxenberg, D. P., Jones, S. S., Brown, E. L., Kay, R. M., Orr, E. C., Shoemaker, C., Golde, D. W., Kaufman, R. J., Hewick, R. M., Wang, E. A. & Clark, S. C. (1985) *Science* **228**, 810–815.
11. Lee, F., Yokota, T., Otsuka, T., Gemmell, L., Larson, N., Lujh, J., Arai, K. & Rennick, D. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4360–4364.
12. Cantrell, M. A., Anderson, D., Cerretti, D. P., Price, V., McKereghan, K., Tushinski, R. J., Mochizuki, D. Y., Larsen, A., Grabstein, K., Gillis, S. & Cosman, D. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6250–6254.
13. Gasson, J. C., Weisbart, R. H., Kaufman, S. E., Clark, S. C., Hewick, R. M. & Wong, G. G. (1984) *Science* **226**, 1339–1342.
14. Vadas, M. A., Nicola, N. A. & Metcalf, D. (1983) *J. Immunol.* **130**, 795–799.
15. Dessein, A. J., Vadas, M. A., Nicola, N. A., Metcalf, D. & David, J. R. (1982) *J. Exp. Med.* **156**, 90–103.
16. Studencki, A. B. & Wallace, R. B. (1984) *DNA* **3**, 7–15.
17. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
18. Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K. & Green, M. R. (1984) *Nucleic Acids Res.* **12**, 7035–7056.
19. Graham, F. L. & van der Eb, A. J. (1973) *Virology* **52**, 456–467.
20. Bagby, G. C., Vasiliki, D., Bennett, R. M., Vandenbark, A. A. & Garewal, H. S. (1981) *J. Clin. Invest.* **68**, 56–63.
21. Powell, J. S., Fialkow, P. J. & Adamson, J. A. (1982) *J. Cell. Physiol.*, Suppl. 1, 79–85.
22. Kimura, H., Burstein, S. A., Thorning, D., Powell, J. S., Harker, L. A., Fialkow, P. J. & Adamson, J. W. (1984) *J. Cell. Physiol.* **118**, 87–96.
23. Grantham, R., Gautier, C., Gouy, M., Jacobzone, M. & Mercier, R. (1981) *Nucleic Acids Res.* **9**, r43–r73.
24. Dierks, P., van Ooyen, A., Cochran, M. D., Dobkin, C., Reiser, J. & Weissmann, C. (1983) *Cell* **32**, 695–706.
25. Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. (1980) *Nucleic Acids Res.* **8**, 127–142.
26. Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. & Chambon, P. (1980) *Science* **209**, 1406–1414.
27. Kozak, M. (1984) *Nucleic Acids Res.* **12**, 857–872.
28. Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459–472.
29. Keller, E. B. & Noon, W. A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7417–7420.
30. Fitzgerald, M. & Shenk, T. (1981) *Cell* **24**, 251–260.
31. McLauchlan, J., Gaffney, D., Whitton, J. L. & Clements, J. B. (1985) *Nucleic Acids Res.* **13**, 1347–1368.
32. Gopinathan, K. P., Weymouth, L. A., Kunkel, T. A. & Loeb, L. A. (1979) *Nature (London)* **278**, 857–859.
33. Ullrich, A., Berman, C. H., Dull, T. J., Gray, A. & Lee, J. M. (1984) *EMBO J.* **3**, 3361–3364.
34. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15**, 1157–1174.
35. Metcalf, D., Johnson, G. R. & Burgess, A. W. (1980) *Blood* **55**, 138–147.
36. Quesenberry, P. J., Ihle, J. N. & McGrath, E. (1985) *Blood* **65**, 214–217.
37. Miyatake, S., Otsuka, T., Yokota, T., Lee, F. & Arai, K. (1985) *EMBO J.* **4**, 2561–2568.
38. Emerson, S. G., Sieff, C. A., Wang, E. A., Wong, G. G., Clark, S. C. & Nathan, D. G. (1985) *J. Clin. Invest.* **76**, 1286–1290.
39. Metcalf, D. (1986) *Blood* **67**, 257–267.