

# Cloning and tissue-specific expression of mouse macrophage colony-stimulating factor mRNA

(colony-stimulating factor 1/multiple transcripts/expression library)

TRIPATHI B. RAJAVASHISTH\*, RONALD ENG\*, RICHARD K. SHADDUCK†, ABDUL WAHEED†, CHAD M. BEN-AVRAM‡, JOHN E. SHIVELY‡, AND ALDONS J. LUSIS\*§

\*Departments of Medicine and Microbiology, and Molecular Biology Institute, University of California, Los Angeles, CA 90024; †Department of Medicine, Montefiore Hospital, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213; and ‡Division of Immunology, Beckman Research Institute of the City of Hope, Duarte, CA 91010

Communicated by Eugene P. Cronkite, October 27, 1986

**ABSTRACT** Macrophage colony-stimulating factor (CSF-1) stimulates the production of macrophages from bone marrow progenitor cells. We have identified a cDNA clone for murine CSF-1 by antibody screening of a mouse L-cell cDNA library in the expression vector  $\lambda$ gt11. A screen of about 150,000 recombinant plaques yielded 6 clones that reacted well with an antibody raised against denatured and reduced mouse L-cell CSF-1. These clones were further screened with synthetic oligonucleotides based on the amino-terminal amino acid sequence of CSF-1. One clone, which hybridized to the oligonucleotides, was sequenced and found to contain a single open reading frame. This encompassed 68 amino acids of the mature protein, including the entire amino-terminal sequence we previously reported. This is preceded by what appears to be a 31 amino acid signal peptide. Blot analysis showed that this cDNA hybridizes to a major mRNA species of about 4.5 kilobases (kb) as well as several smaller, less abundant mRNA species (3.8, 2.3, and 1.4 kb) present in mouse L cells. A similar pattern of hybridization was observed with mRNA from a human pancreatic carcinoma cell line that produces CSF-1. Striking differences in the qualitative and quantitative expression of mRNA species for CSF-1 were observed in various mouse tissues. Liver expressed primarily a 1.4-kb species, heart and lung expressed primarily a 4.5-kb species, brain expressed high levels of both the 4.5-kb and 1.4-kb species, and intestine lacked detectable CSF-1 transcripts. Southern blot analysis suggests that the CSF-1 gene is present as a single copy in the mouse haploid genome and that it is not rearranged or amplified in L cells.

Colony-stimulating factors (CSFs) are a family of glycoproteins that regulate the growth and differentiation of committed hematopoietic progenitor cells and affect the function of mature cells (1, 2). CSFs are identified *in vitro* by their ability to stimulate the formation of colonies of differentiated cells in semisolid agar cultures of bone marrow progenitor cells (3). The functional and structural properties of the CSFs have been best characterized in the mouse, where four distinct CSFs have been clearly identified. M-CSF (CSF-1) stimulates primarily the growth of the monocyte-macrophage lineage (4), G-CSF stimulates the production of neutrophilic granulocytes (5), GM-CSF (CSF-2) stimulates the production of both granulocytes and macrophages (6), and multi-CSF (interleukin 3, IL-3) affects the growth of all myeloid precursors, inducing mixed colonies *in vitro* (7). cDNA sequences encoding murine or human GM-CSF, multi-CSF, and G-CSF have been isolated, and the recombinant factors have been expressed in eukaryotic cells (8-11).

The murine CSF-1 has been purified from L-cell conditioned medium. The purified factor is a heterogeneous glycoprotein ( $M_r \approx 70,000$ ) consisting of two identical subunits linked by disulfide bonds (2, 4). Mild reduction of the factor results in subunit dissociation and the loss of all biological activity. Following the removal of carbohydrate, the polypeptide subunit of CSF-1 was estimated to have a molecular weight of about 15,000 (4). Recently, we reported the amino-terminal amino acid sequence of mature mouse L-cell CSF-1 (12, 13).

We now report the molecular cloning and expression of murine CSF-1 cDNA sequences in the expression vector  $\lambda$ gt11. The clone has been used to examine the expression of CSF-1 mRNA and should prove useful for structural and functional studies of the factor.

## MATERIALS AND METHODS

**Construction of L-Cell cDNA Library in  $\lambda$ gt11.** Poly(A)<sup>+</sup> RNA was used as template for the synthesis of single-stranded DNA with an oligo(dT) primer and avian myeloblastosis virus reverse transcriptase. Double-stranded cDNA was prepared essentially according to the method of Gubler and Hoffman (14), in which second-strand synthesis is mediated by RNase H, DNA polymerase I, and bacteriophage T4 DNA ligase. The double-stranded cDNA was methylated at internal *Eco*RI restriction sites, blunt-ended with T4 DNA polymerase, and ligated to *Eco*RI linkers according to Young and Davis (15). Excess linkers were removed from cDNA by *Eco*RI digestion followed by Sephadex G-75 gel filtration and two rounds of ethanol precipitation in 2.0 M ammonium acetate. To decrease the background of phage lacking inserts, the *Eco*RI-cut  $\lambda$ gt11 DNA was treated with calf intestinal alkaline phosphatase prior to ligation to cDNA. The ligated DNA was packaged using an extract purchased from Amersham and amplified on *Escherichia coli* Y1088 (15).

**Antibody Screening.** Antibodies specific for L-cell CSF-1 were raised in rabbits as described (12). Antibody screening of the L-cell cDNA library was performed according to the method of Young and Davis (15). In brief, the phage library was plated on *E. coli* Y1090 at a density of  $2 \times 10^4$  phage per 150-mm plate and isopropyl  $\beta$ -D-thiogalactopyranoside-induced plaques were transferred to nitrocellulose filters. Filters were then incubated with a 1:1000 dilution of antibody for 20 hr and then <sup>125</sup>I-labeled protein A to visualize the bound antibody. Positive plaques were selected after two or more rounds of plaque purification.

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.

Abbreviation: CSF, colony-stimulating factor.

§To whom requests for reprints should be addressed.



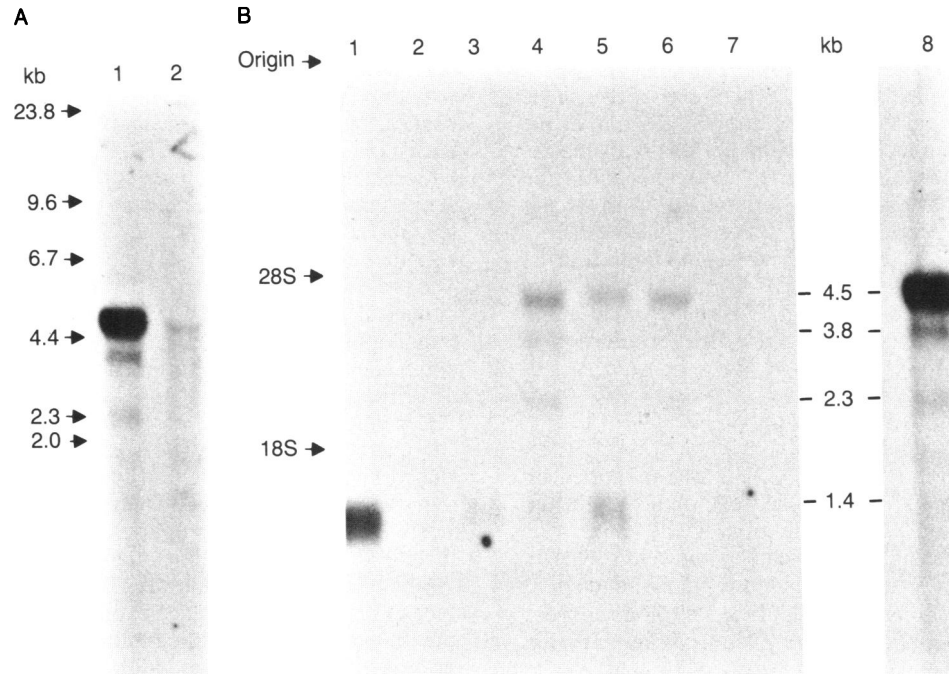


FIG. 3. Blot-hybridization analysis of CSF-1 mRNA. Poly(A)<sup>+</sup> RNA from mouse L cells and from human MIA-PaCa-2 cells (A, lanes 1 and 2, respectively) or from mouse liver, intestine, kidney, lung, brain, heart, and spleen (B, lanes 1-7, respectively; L-cell RNA in lane 8) was denatured and electrophoresed in 1.0% agarose gel with formaldehyde and then blotted to nitrocellulose filters. The filters were probed with <sup>32</sup>P-labeled mouse CSF-1 cDNA. The filter in A was washed at low stringency (0.15 M NaCl/15 mM sodium citrate at 45°C), while the filter in B was washed at high stringency (15 mM NaCl/1.5 mM sodium citrate at 65°C). Filters were then autoradiographed overnight at -70°C with an intensifying screen. Size markers (*Hind*III-digested phage λ DNA) are indicated in A. The positions of large (28S) and small (18S) ribosomal RNA and the sizes of RNA bands that hybridized to the probe are shown in B.

GenBank<sup>¶</sup> and the National Protein Information Resource data bases<sup>||</sup>, using the programs SRCHN\*\* and SEARCH<sup>††</sup>. The mouse L-cell CSF-1 cDNA sequence did not show significant homology to any other protein in these data bases. In addition, the mouse CSF-1 sequence did not show significant homology to GM-CSF (8), multi-CSF (10), G-GSF (11), erythropoietin (23), or interleukin 1 (24).

**Tissue-Specific Expression of Multiple CSF-1 Transcripts.** The entire mouse CSF-1 cDNA insert was isolated and used as a probe to examine the expression of CSF-1 gene transcripts in cell lines and mouse tissues. Blot-hybridization analysis of mouse L-cell poly(A)<sup>+</sup> RNA revealed a major hybridizing species of about 4.5 kilobases (kb) as well as several less abundant species of about 3.8, 2.3, and 1.4 kb (Fig. 3A). The pattern of transcripts in the human pancreatic carcinoma cell line MIA-PaCa-2 which produces CSF-1, examined under conditions of reduced hybridization stringency, was similar to that of mouse L cells (Fig. 3A).

Numerous tissues are capable of elaborating CSF-1 (25, 26); therefore, we examined poly(A)<sup>+</sup> RNA from several mouse tissues (Fig. 3B). Lung and heart contained a pattern of transcripts similar to that observed in L cells, although at considerably reduced levels. Liver, on the other hand, exhibited high levels of the 1.4-kb species but lacked the larger species. The predominant RNA species in brain and

kidney was 1.4 kb long, although these tissues also expressed the larger species. Spleen had very low levels of CSF-1 transcripts, and intestine lacked detectable transcripts (Fig. 3B). Thus, the pattern of CSF-1 transcripts present in various tissues is complex and highly tissue-specific.

The fact that the CSF-1 probe hybridizes to the 1.4-kb mRNA species observed in liver and certain other tissues under conditions of high stringency suggests that the transcript corresponds to CSF-1. Moreover, the transcript is similar in size to the putative bioactive mRNA present in human MIA-PaCa-2 cells (22). To further test whether the transcript represents a distinct but homologous mRNA species, we performed primer-extension studies (27). A 15-base-long oligonucleotide (CAG GTG TCC ATT CCC) complementary to nucleotides 127-141 of the CSF-1 coding region was synthesized, labeled with <sup>32</sup>P by kinase, and used to prime synthesis of cDNA from liver, intestine, and L-cell mRNA. A unique extension product of about 200 nucleotides was observed using mRNA from liver and L cells, which contain CSF-1 transcripts as judged by RNA blot hybridization, whereas no extension product was observed using mRNA from intestine, which lacks hybridizing CSF-1 transcripts (data not shown). Thus, the 1.4-kb transcript present in liver appears to correspond to CSF-1.

The presence of high levels of a CSF-1 transcript in liver raises the possibility that liver is a major source of the protein. This was tested by assaying mouse liver conditioned medium for CSF-1 activity in a mouse bone marrow system (28). No activity was observed, consistent with the earlier observations of Bradley *et al.* (26). Thus, it appears either that the hepatic 1.4-kb transcript does not encode bioactive CSF-1 or that the protein is not properly processed in liver.

**Analysis of CSF-1 Genomic Sequences.** Mouse genomic DNA isolated from spleen cells and L-cell DNA were examined by Southern blotting and hybridization with the CSF-1 cDNA probe. Digestion of DNA with *Eco*RI, *Hae* III,

<sup>¶</sup>National Institutes of Health (1983) Genetic Sequence Databank: GenBank (Research Systems Div., Bolt, Beranek, and Newman, Cambridge, MA), Tape Release 40.0.

<sup>||</sup>Protein Identification Resource (1985) Protein Sequence Database (Natl. Biomed. Res. Found., Washington, DC), Release 8.0.

\*\*Lipman, D. J. & Wilbur, W. J. (1982) RBSA-Program for Bioscience Similarity Analysis (National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, Mathematical Research Branch, National Institutes of Health, Bethesda, MD).

<sup>††</sup>Orcutt, B. C. & Dayhoff, M. O. (1982) PIR-Reports, NVC 0484, and Pro-0484 (Natl. Biomed. Res. Found., Washington, DC).

and *Hind*III yielded a similar pattern for mouse genomic and L-cell DNA. *Eco*RI gave two bands at 10.0 and 12.0 kb, *Hae* III gave a single band at about 1.0 kb, and *Hind*III gave a single band at about 10.0 kb (Fig. 4). The presence of a single band after *Hae* III and *Hind*III digestion suggests that the CSF-1 gene is present as a single copy per haploid genome. Also, the fact that the pattern and intensity of the bands from L-cell DNA were not discernibly different from those observed with mouse genomic DNA suggests that the CSF-1 gene is not amplified or grossly rearranged in mouse L cells, despite the much higher levels of CSF-1 transcripts in L cells than in the mouse tissues examined. Obviously, point mutations or other small mutations of the CSF-1 would not be detected by our analysis.

## DISCUSSION

A partial cDNA clone for mouse CSF-1 was identified by antibody screening of a mouse L-cell cDNA library in the expression vector  $\lambda$ gt11. The identity of the clone was established by comparing the predicted amino acid sequence encoded by the cDNA with the amino-terminal amino acid sequence of mature mouse L-cell CSF-1. The clone appears to contain 99 nucleotides of the 5' untranslated region followed by a region encoding a 31 amino acid signal peptide and 68 amino acids of the mature protein. The CSF-1 sequence is not homologous to sequences for various other hematopoietic growth factors, including GM-CSF (8), multi-CSF (10), and G-CSF (11). The lack of homology of members of the CSF family is surprising, considering their overlap in functional control of granulocyte and macrophage differentiation (1-3). The CSF-1 cDNA was used as a probe to

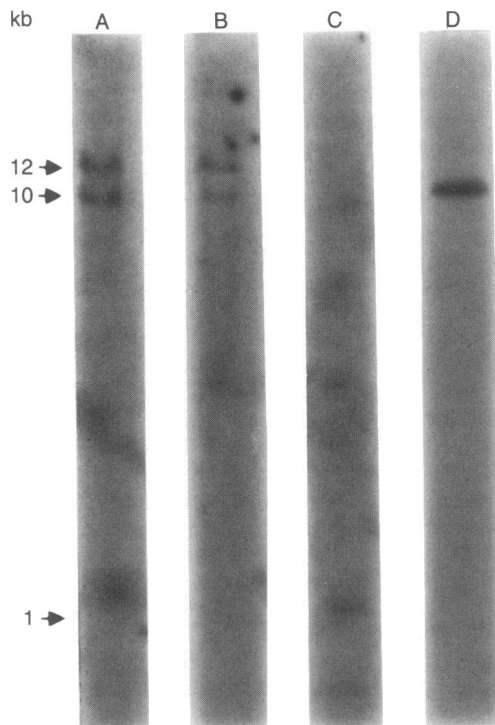


FIG. 4. Southern blot analysis of mouse genomic DNA and L-cell DNA. Mouse genomic DNA or DNA isolated from L cells was digested with various restriction enzymes, electrophoresed in 0.7% agarose gels, blotted onto nitrocellulose, and hybridized to  $^{32}$ P-labeled CSF-1. Lane A: L-cell DNA digested with *Eco*RI. Lanes B-D: mouse genomic DNA digested with *Eco*RI, *Hae* III, or *Hind*III, respectively. L-cell DNA gave a pattern similar in intensity and size to genomic DNA after digestion with *Hae* III or *Hind*III (data not shown). The sizes of DNA bands are indicated on the basis of DNA size markers run in a parallel lane (not shown).

examine the expression of CSF-1 mRNA in cell lines and various mouse tissues. A complex pattern consisting of at least four distinct poly(A)<sup>+</sup> transcripts, ranging from 4.5 kb to 1.4 kb in size, was observed in L cells and several mouse tissues. Moreover, large qualitative and quantitative differences in the pattern of CSF-1 transcripts were observed between tissues. Mouse liver contains a single 1.4-kb transcript similar in size to the bioactive CSF-1 mRNA observed in human cell line MIA-PaCa-2 (22). Bioassay of mouse liver conditioned medium indicates the absence of CSF-1 activity. The presence of a single transcript and the absence of CSF-1 biological activity are intriguing. It is possible that the transcript in liver is not translated or that there is an inhibitor in the conditioned medium that affects the biological activity of CSF-1.

Recently, Kawasaki *et al.* (22) reported the identification and characterization of cDNA clones corresponding to human urinary CSF. The extensive homology of the human urinary CSF sequence and that of mouse CSF-1 (Fig. 2) indicates that the factors are analogous. Sequencing of several human cDNA clones derived from the human pancreatic carcinoma cell line MIA-PaCa-2 indicated the presence of intervening sequences and of heterogeneity of the 3' end (22), suggesting that the multiple CSF-1 transcripts arise as a result of alternative mRNA splicing. Our Southern blotting studies, suggesting that the CSF-1 gene occurs as a single copy in the haploid genome, are consistent with this hypothesis. Our present findings suggest that the multiple transcripts are not the results of abnormal expression in a transformed cell line, since they occur in normal tissues and since the relative levels of the various RNA species are regulated in a tissue-specific manner.

The functional and regulatory significance of the multiple CSF-1 transcripts is unclear. Unlike most other hematopoietic growth factors, CSF-1 activity is produced by a wide variety of tissues, including brain, heart, and lung (25, 26, 29). Although GM-CSF is also produced by a variety of tissues (30), it appears to be derived exclusively from lymphocytes. Thus, some of the transcripts observed in this study presumably give rise to the active factor. All of the major RNA species are considerably longer than a sequence required to encode the  $M_r$  15,000 mature mouse CSF-1 polypeptide. However, the primary CSF-1 translation product has not yet been identified, and it is possible that CSF-1 undergoes significant proteolytic processing before secretion. We have been unsuccessful in attempts to identify bioactive or immunoreactive CSF-1 following *in vitro* translation or frog oocyte microinjection of mouse L-cell mRNA (A.J.L., K. A. O'Donnell, R.E., and T.B.R., unpublished data). Recently, Burgess *et al.* (31) separated two forms of CSF-1 from mouse L-cell conditioned medium by reversed-phase high-performance liquid chromatography. The forms differed in subunit size, hydrophobicity, and specificity for the production of granulocyte as compared to macrophage colonies in bone marrow cultures (31).

The availability of the mouse CSF-1 cDNA should facilitate studies of the regulation of the factor and its role in normal or abnormal cell proliferation. Of particular importance is the isolation and expression of sequences corresponding to the CSF-1 transcripts, in order to examine their protein products and functions.

This work was supported by National Institutes of Health Grants AM27008, CA15237, and CA32737 and the University of California, Los Angeles Jonsson Comprehensive Cancer Center. T.B.R. is a fellow of the Lady Tata Memorial Trust. A.J.L. is an established investigator of the American Heart Association.

1. Burgess, A. W. & Metcalf, D. (1980) *Blood* 56, 947-958.
2. Wing, E. J. & Shaduck, R. K. (1985) in *Biological Response*

- Modifiers*, ed. Torrence, P. F. (Academic, New York), pp. 219–243.
3. Metcalf, D. (1984) *The Hematopoietic Colony Stimulating Factors* (Elsevier, Amsterdam).
  4. Das, S. K. & Stanley, E. R. (1982) *J. Biol. Chem.* **257**, 13679–13684.
  5. Nicola, N. A., Metcalf, D., Matsumoto, M. & Johnson, G. R. (1983) *J. Biol. Chem.* **258**, 9017–9021.
  6. Burgess, A. W., Camakaris, J. & Metcalf, D. (1977) *J. Biol. Chem.* **252**, 1998–2003.
  7. Bazill, G. W., Haynes, M., Garland, J. & Dexter, T. M. (1983) *Biochem. J.* **210**, 747–759.
  8. Gough, N. M., Gough, J., Metcalf, D., Kelson, 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. O., 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.
  10. Yokota, T., Lee, F., Rennick, D., Hall, C., Arai, N., Mosmann, T., Nabel, G., Canter, H. & Arai, K.-i. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1070–1074.
  11. Nagata, S., Tsuchiga, M., Asano, S., Kaziro, Y., Yamazaki, T., Yamamoto, O., Hirata, Y., Kubota, N., Oheda, M., Nomura, H. & Ono, M. (1986) *Nature (London)* **319**, 415–418.
  12. Ben-Avram, C. M., Shively, J. E., Shadduck, R. K., Waheed, A., Rajavashisth, T. & Lusic, A. J. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4486–4489.
  13. Ben-Avram, C. M., Shively, J. F., Shadduck, R. K., Waheed, A., Rajavashisth, T. & Lusic, A. J. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7801.
  14. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25**, 264–269.
  15. Young, R. A. & Davis, R. W. (1983) *Science* **222**, 778–782.
  16. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
  17. Suggs, S. V., Wallace, R. B., Hirose, T., Kawashima, E. H. & Itakura, K. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 6613–6617.
  18. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
  19. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
  20. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
  21. Kreil, G. (1981) *Annu. Rev. Biochem.* **50**, 317–348.
  22. Kawasaki, E. S., Ladner, M. B., Wang, A. M., Arsdell, J. V., Warren, M. K., Coyne, M. Y., Schweickart, V. L., Lee, M. T., Wilson, K. J., Boosman, A., Stanley, E. R., Ralph, P. & Mark, D. F. (1985) *Science* **230**, 291–296.
  23. Jacobs, K., Shoemaker, C., Rudersdorf, R., Neill, S. D., Kaufman, R. J., Mufson, A., Seehra, J., Jones, S. S., Hewick, R., Fritsch, E. F., Kawakita, M., Shimizu, T. & Miyake, T. (1985) *Nature (London)* **313**, 806–810.
  24. Gasson, J. C., Golde, D. W., Kaufman, S. E., Westbrook, C. A., Hewick, R. M., Kaufman, R. J., Wong, G. G., Temple, P. A., Leary, A. C., Brown, E. L., Orr, E. C. & Clark, S. C. (1985) *Nature (London)* **315**, 768–771.
  25. Shadduck, R. K. & Metcalf, D. (1975) *J. Cell Physiol.* **86**, 247–252.
  26. Bradley, T. R., Stanley, E. R. & Sumner, M. A. (1971) *Aust. J. Exp. Biol. Med. Sci.* **49**, 595–603.
  27. Rajavashisth, T. B., Kaptein, J. S., Reue, K. L. & Lusic, A. J. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8085–8089.
  28. Waheed, A. & Shadduck, R. K. (1979) *J. Lab. Clin. Med.* **94**, 180–194.
  29. Sheridan, J. W. & Stanley, E. R. (1971) *J. Cell Physiol.* **78**, 451–460.
  30. Nicola, N. A., Burgess, A. W. & Metcalf, D. (1979) *J. Biol. Chem.* **254**, 5290–5299.
  31. Burgess, A. W., Metcalf, D., Kozka, I. J., Simpson, R. J., Vairo, G., Hamilton, J. A. & Nice, E. C. (1985) *J. Biol. Chem.* **260**, 16004–16011.