Journal of Cellular Biochemistry 40:91–107 (1989)

Human Macrophage Colony-Stimulating Factor Heterogeneity Results From Alternative mRNA Splicing, Differential Glycosylation, and Proteolytic Processing

P.J. Shadle, L. Aldwin, D.E. Nitecki, and K. Koths

Departments of Protein Chemistry (P.J.S., K.K.) and Chemistry, (L.A., D.E.N.), Cetus Corporation, Emeryville, California 94608

The single gene for human macrophage colony-stimulating factor (M-CSF, or CSF-1) generates multiple mRNA species that diverge within the coding region. We have characterized translation products of these mRNA species from native and recombinant sources. Immunoblots of reduced native M-CSF indicate that multiple glycosylated species ranging from 25 kd to 200 kd are secreted by human monocytes and cell lines. In contrast, CV-1 cells expressing a short M-CSF clone secrete only 24 kd recombinant M-CSF. Synthetic peptide antibodies were developed to distinguish between secreted recombinant M-CSF from long and short mRNA splicing variants. Immunoblot analysis indicates that alternative mRNA splicing generates some M-CSF protein heterogeneity. Most secreted MIA PaCa-2 M-CSF reacts with long-clone-specific antibody. Lectin affinity chromatography shows that variable glycosylation contributes significantly to MIA PaCa-2 M-CSF size heterogeneity. In addition, cell lysates also contain larger M-CSF species that apparently undergo proteolytic processing before secretion. The data indicate that M-CSF protein heterogeneity results from both pre- and post-translational processing.

Key words: synthetic peptide antibodies, HL-60, MIA PaCa-2, monocyte

Macrophage colony-stimulating factor (M-CSF), also known as CSF-1, is a hematopoietic growth factor that induces proliferation and differentiation of cells of the monocyte/macrophage lineage [1,2]. Although evidence for only a single M-CSF gene has been reported [3], considerable heterogeneity in M-CSF mRNA has been detected in certain cell lines [3,4]. Several cDNA clones have been isolated and sequenced. The results indicate that alternative mRNA splicing occurs within the cod-

P.J. Shadle's present address is Process and Product Development, California Biotechnology, Inc., 2450 Bayshore Parkway, Mountain View, CA 94043.

Received July 6, 1988; accepted December 7, 1988.

ing region of exon 6 [5,6]. The degree to which this mRNA variability is reflected in heterogeneous M-CSF translation products is presently unclear. Human urinary M-CSF has been reported to be a glycosylated homodimer composed of 23 to 28 kd subunits [5,7–11]. Both human and mouse M-CSF have also been purified from culture supernatants of cell lines, and the reported size of the dimeric molecule was about 68 kd [12–18]. M-CSF has been estimated to contain 10–50% carbohydrate by weight [16] and some size heterogeneity reported for mouse L cell M-CSF was presumed to reflect glycosylation heterogeneity [19]. Comparison of M-CSF protein and cDNA sequences indicates that M-CSF precursor proteins undergo glycosylation as well as proteolytic processing that removes a leader sequence and a C-terminal domain that contains a membrane-spanning domain [3,5,6,20–22]. Cell-associated M-CSF protein has been detected in detergent lysates of cells by radioimmunoassay (RIA) [2,23], but has not been further characterized.

Although antibodies to native, dimeric human M-CSF have been developed previously [7,16,24], they were not capable of detecting reduced human or mouse M-CSF on immunoblots. Recently, using recombinant human M-CSF (rM-CSF) as the antigen [22] (Kawasaki et al., submitted), high-titer, specific rabbit antisera have been developed that detect reduced as well as nonreduced M-CSF on immunoblots. Such rM-CSF antibodies have allowed detection and characterization of M-CSF species that were not described in previous studies. Using antibodies to rM-CSF, as well as antibodies to synthetic peptides of M-CSF sequences, we have analyzed M-CSF protein made by human cell lines and primary cultures of monocytes. We show here that M-CSF secreted from various sources displays a range of size that is much wider than is commonly seen in products of a single gene. Cell-associated M-CSF was also characterized and shown to differ from secreted M-CSF in size. The data suggest that variability in mRNA splicing, glycosylation, and proteolytic processing generate this unusual protein heterogeneity.

MATERIALS AND METHODS

Synthetic Peptides

In order to develop antibodies specific for the protein products of alternative M-CSF mRNA splicing, two predicted amino acid sequences were chosen from the long and short coding region cDNA clones. Peptide I spanned predicted amino acids 146–159 of the mature coding region of the short cDNA clone and had the sequence: Cys-Ser-Ser-Gln-Gly-His-Glu-Arg-Gln-Ser-Glu-Gly-Ser-Ser. Peptide II spanned predicted amino acids 171–183 of the long M-CSF cDNA clone with a Cys residue added at the C-terminus, and had the sequence Ala-Ser-Val-Ser-Pro-His-Gln-Pro-Leu-Ala-Pro-Ser-Met-Cys. Figure 3 indicates the positions of peptides I and II in the coding region of each M-CSF cDNA clone.

Peptides were synthesized by the solid-phase method [25] on a Biosearch 9500 automated peptide machine, cleaved with hydrogen fluoride, and purified by preparative HPLC using a Waters Delta Prep 3000 instrument, on a 15–20 μm Vydac C4 PrepPAK column. Cysteine, either in the natural sequence (peptide I) or added at the C-terminal end of the sequence (peptide II), was used to crosslink the peptides to protein carriers or to activated Sepharose for affinity purification of antisera. For immunization, keyhole limpet hemocyanin (KLH, Calbiochem, La Jolla, CA) was conjugated to purified peptides with the heterobifunctional agent mal-sac-HNSA (the

N-maleimido-6-aminocaproyl ester of 1-hydroxy-2-nitro-4-benzenesulfonic acid) as previously described [26]. Bovine serum albumin (BSA) peptide conjugates were prepared similarly and used to assay immunoglobulin production and specificity.

Peptide Columns for Affinity Purification

Omega-aminohexyl-Sepharose 4B containing 2–10 μmol amino group/ml (Sigma, St. Louis, MO) was swelled in 0.1 M phosphate buffer, pH 7.5. A threefold molar excess of mal-sac-HNSA over amino groups was added and anion release monitored as described for 30 min or until reaction ceased [26]. The gel was washed with 0.1 M phosphate, pH 6.0, until the yellow color eluted. Peptide was added at a twofold molar excess over amino groups on the gel, and the gel was incubated on a gently rocking shaker at 4°C for 24 h to react free sulfhydryl groups on the peptide with maleimido groups on the resin. The gel was washed with phosphate-buffered saline, pH 7.2 (PBS), followed by 0.5 M NaCl in PBS to remove uncoupled peptide, and was stored at 4°C in PBS containing 0.02% azide. The final packed gel had 0.1–2.0 μmol per ml of covalently bound peptide.

Development of Rabbit Antibodies

M-CSF antibodies. A cDNA clone of M-CSF [3], truncated in the coding region at predicted amino acid 150 of the mature protein, was expressed in *Escherichia coli* (Kawasaki et al., submitted). Two rabbits were immunized with 250 μg of purified rM-CSF in complete Freund's adjuvant and boosted after 28 days and at 14 day intervals thereafter with 100 μg of rM-CSF in incomplete Freund's adjuvant. Antisera were collected after 3 boosts. IgG was purified by affinity chromatography on protein A agarose (Pharmacia, Piscataway, NJ), following manufacturer's instructions. This antibody is referred to as rM-CSF antibody.

Peptide antibodies. Peptide/KLH conjugates were dissolved in PBS, emulsified in Freund's complete adjuvant, and 500 μg of each were injected into the axial lymph nodes of rabbits. Boosting was done similarly with 250 μg of antigen in incomplete Freund's adjuvant, and sera were collected 7 days after each boost. Sera were screened for reactivity with peptide-BSA conjugates on dot-blots. After 5–6 boosts, positively reacting IgG was purified on protein A-Sepharose and then affinity-purified on peptide-Sepharose columns. Briefly, IgG was applied to a 10 ml peptide-Sepharose column equilibrated in PBS, pH 7.2. The column was washed with PBS until the absorbance of the effluent returned to baseline, then with 1 M NaCl to remove nonspecifically absorbed material. Antipeptide IgG was eluted with 0.1 M glycine-HCl buffer, pH 2.5, neutralized with 1 M Tris-HCl, pH 8.3, and dialyzed into Tris-buffered saline (TBS), pH 7.5. Aliquots were stored at −20°C at 2 mg/ml until use.

M-CSF Bioassay and RIA

M-CSF bioactivity was determined in a colony-forming assay performed on mouse bone marrow cells [3]. Colonies were counted after 7 days of culture, and 1 colony was defined as 1 unit of M-CSF activity. M-CSF was also detected by RIA, performed as described previously [14,27], using purified MIA PaCa-2 M-CSF as a standard and a rabbit antibody to human urinary M-CSF [2,4,14]. RIA results are reported in equivalent units/ml, calculated from calibration of the MIA PaCa-2 M-CSF standard in the colony-forming bioassay [2,3].

Cell Culture

Pancreatic carcinoma cell lines MIA PaCa-2 and PANC-1 were obtained from the American Type Culture Collection (Rockville, MD) and cultured as described previously [14,27]. M-CSF production was induced in MIA PaCa-2 cells cultured in roller bottles in serum-free Dulbecco's MEM (DMEM) containing phorbol myristate acetate (PMA, 10 ng/ml) and retinoic acid (10 μmol/L) as described [14]. Culture medium containing M-CSF was collected after 3 days. PANC-1 cells were cultured in T-flasks in DMEM containing 10% fetal calf serum (FCS), and culture medium containing M-CSF was collected after 3–4 days. The HL-60 human leukemia cell line was propagated in T-flasks in DMEM containing 10% FCS; M-CSF production was stimulated in serum-free medium using PMA (10 ng/ml) [28]. Culture medium containing M-CSF was collected after 2–3 days.

Human monocytes were purified from peripheral blood of normal donors by sequential density gradient centrifugation in Ficoll and Percoll followed by adherence [28]. Contamination of purified monocytes with other cells was typically 5% or less. Monocytes were cultured in T-flasks in MEM containing PMA (10^{-10} M) for 24 h, and culture medium containing M-CSF was collected.

The CV-1 cell line was transformed with a plasmid containing an M-CSF coding sequence under control of the SV40 T promoter [29]. rM-CSF expression was initiated by infection with SV40 virus as described, and cell supernatants were harvested 4 days later. rM-CSF was purified as described previously, using batch adsorption to DEAE-Sepharose, affinity chromatography on lentil lectin agarose, and hydrophobic interaction HPLC [22].

Cell extracts were made by washing adherent cell layers gently with PBS and extracting in situ with cold 50 mM Tris-HCl, pH 8.0, containing 50 mM NaCl, 10 mM EDTA and 0.5% Triton X-100. Fresh phenylmethylsulfonylfluoride (0.2 mM) (PMSF) and leupeptin (5 μg/ml) were added to inhibit proteolysis, and extraction was performed on ice for 10 min. One milliliter of extraction buffer contained about $2–4 \times 10^7$ cells. Extracts were clarified by centrifugation at 12,000*g* and immediately prepared for RIA by dilution into assay buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.1% NP-40, 1 mg/ml BSA), or for SDS-PAGE by boiling in SDS gel sample buffer containing 1% mercaptoethanol. The cell extraction buffer did not affect the RIA. All samples except those from CV-1 cells were concentrated by trichloroacetic acid (TCA) precipitation prior to electrophoresis.

Partial Purification of MIA PaCa-2 M-CSF

MIA PaCa-2 M-CSF was purified by a modification of the method described previously [14]. Briefly, culture supernatants were treated with PMSF (0.2 mM) and leupeptin (1 μg/ml) to inhibit protease activity. M-CSF was purified by adsorption on calcium phosphate gel, followed by lectin affinity chromatography and immunoaffinity chromatography. Lectin affinity chromatography was performed on lentil lectin agarose (LcH; E-Y Labs, Burlingame, CA), on wheat germ agglutinin agarose (WGA; E-Y Labs), or on a mixed-bed column of both lectins. The mixed-bed column was made by mixing the two lectin agarose preparations (25 ml of each gel) prior to pouring the column. Lectin columns were equilibrated in Tris-buffered saline (TBS, 50 mM Tris-HCl, pH 7.5, 200 mM NaCl). The glycoprotein fraction containing M-CSF was eluted from LcH using 0.1 M alpha-methylmannoside, and from WGA

using 0.1 M N-acetyl-D-glucosamine in TBS. The mixed-bed column was eluted with TBS containing 0.1 M of each sugar. Eluates were monitored for protein by absorbance at 280 nm, and for M-CSF using RIA. Lectin eluates were dialyzed to remove alpha-methylmannoside, which was toxic in the colony-forming assay. M-CSF was further purified by immunoaffinity chromatography using a rat monoclonal antibody to mouse M-CSF coupled to CNBr-activated agarose [14].

Immunoblotting

Dilute samples were enriched for M-CSF by 10% TCA precipitation or by selective batch precipitation with lectin-agarose gels (E-Y Labs). Both concanavalin A and wheat germ agaroses were used. The TCA precipitate or glycoprotein fraction, containing M-CSF, was eluted using 1% SDS in 0.2 M Tris-HCl, pH 6.8, and applied directly to SDS gels after boiling and reduction with 1% mercaptoethanol. Control samples containing lectin agaroses and media were treated identically and electrophoresed in parallel. SDS-PAGE was done by the method of Laemmli [30].

Immunoblotting was done by the method of Burnette [31]. M-CSF antigen was visualized by autoradiography after either 1) using rabbit antiserum to *E. coli* recombinant M-CSF followed by [^{125}I]-labeled goat antirabbit IgG (New England Nuclear, Boston, MA); 2) using anti-rM-CSF IgG or anti-peptide IgG labeled with [^{125}I] using the chloramine-T method as described below. Preimmune serum was used in the two-antibody development protocol as a control and showed that binding was specific for M-CSF. The single-step method was used on partially purified samples expected to contain IgG that might give a false positive signal.

Rabbit anti-M-CSF IgG and antipeptide IgG were purified as described and radiolabeled with 1 mCi of carrier-free [^{125}I] in 100 μl of 0.1 M $NaPO_4$ buffer, pH 7.5 [32]. Chloramine-T (Sigma, 5 mg/ml, St. Louis, MO) was added in three 10-μl portions over 3 min, and the reaction proceeded on ice for 5 min. Reaction was stopped with 30 μl of sodium metabisulfite and potassium iodide (10 mg/ml) followed by addition of 1.9 ml of buffer containing 0.2% heat-inactivated BSA. Free [^{125}I] was removed by overnight dialysis against 4 × 2 liters of PBS. Using this protocol, IgG was labeled to a specific activity of over 10^7 cpm/μg and the dialyzed pool was 95% TCA-precipitable. About 100 units of M-CSF were detectable in a single band on an immunoblot that was developed as described, using either the one-step or the two-step method.

RESULTS

Immunoblot Analysis of Secreted Human M-CSF

Multiple species of M-CSF protein were observed when secreted M-CSF from different human cell types was examined on immunoblots (Fig. 1). PANC-1, a pancreatic carcinoma cell line, secreted species which migrated at about 68 kd and 100 kd to 200 kd on nonreducing SDS-PAGE (Fig. 1, lane A). Both species were altered by reduction (lane B). The data suggest that the 68 kd band was reduced to an apparent subunit molecular weight of 35 kd, while the 100 kd to 200 kd smear ran at about 100 kd. The 68 kd species was apparently dimeric, as expected for M-CSF, while the larger species behaved as if it were heterogeneous. Because the rM-CSF antibody used here reacts differently with dimeric and reduced monomeric M-CSF (Kawasaki et al., submitted), a quantitative comparison of the bands in lanes A and B was not possible.

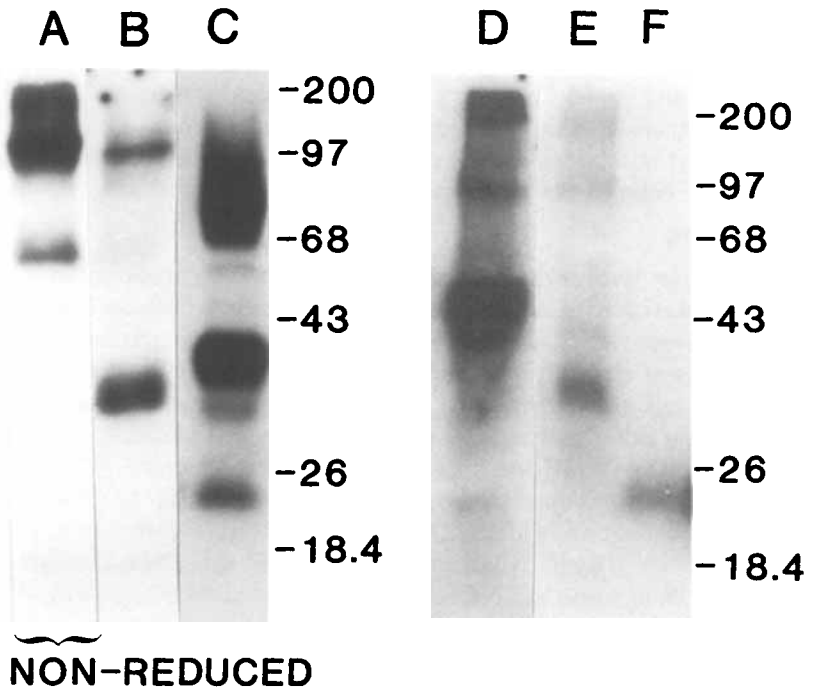

Fig. 1. Comparison of M-CSF protein secreted by different cell types in vitro. Immunoblot of two 10% SDS gels, both developed with [^{125}I]-labeled anti-rM-CSF IgG. **Lane A:** PANC-1 secreted M-CSF (2,000 units), nonreduced sample. **Lane B:** PANC-1 M-CSF, reduced (2,000 units). Iodoacetamide (50 mM) was added to both samples A and B before electrophoresis. **Lane C:** Reduced MIA PaCa-2 secreted M-CSF, purified by calcium phosphate, mixed-bed lectin affinity chromatography, and immunoaffinity chromatography as described in Materials and Methods. **Lane D:** Reduced monocyte M-CSF. **Lane E:** Reduced HL-60 M-CSF. **Lane F:** Reduced, purified CV-1 cell rM-CSF. Cells were cultured as described, and samples in lanes A, B, D, and E were enriched for M-CSF by batch elution from a mixture of lentil lectin and wheat germ agglutinin agaroses.

MIA PaCa-2 M-CSF was also heterogeneous, with several distinct species ranging in size from about 24 kd to over 100 kd under reducing conditions (Fig. 1, lane C). All of the bands were identified as M-CSF because 1) blot strips of the same sample developed in parallel with preimmune IgG did not show any reaction with antibody, and 2) absorption of the rM-CSF antibody with purified rM-CSF before blot development removed the signal. M-CSF species secreted by cultured peripheral blood monocytes and HL-60 cells were also analyzed under reducing conditions and gave complex patterns (Fig. 1, lanes D,E, respectively). In contrast to the multiple bands of nonrecombinant M-CSF protein, rM-CSF expressed from a cDNA with a short coding region [3] in CV-1 cells had an apparent Mr of 24 kd (Fig. 1, lane F) [22].

MIA PaCa-2 M-CSF is Heterogeneous in Glycosylation

Previously, M-CSF was reported to be a single 24 kd species on reducing SDS-PAGE when purified from cells induced by the same method used here [14]. However, the use of a lectin affinity column in this and several other similar purifications

apparently resulted in single-step losses of 50% or more [2,14,16,33]. To improve the yield of M-CSF activity and to determine whether glycosylation heterogeneity explained the consistently low yields, we investigated the behavior of MIA PaCa-2 M-CSF on different lectin affinity columns. Lentil lectin (LcH), which recognizes alpha-mannose residues, binds many glycoproteins having N-linked carbohydrate moieties. LcH may, however, fail to bind N-linked carbohydrate that is highly sialylated. Such N-linked species are generally able to bind to wheat germ agglutinin (WGA), which recognizes N-acetylglucosamine (GlcNAc) and sialic acid residues. WGA also is able to recognize O-linked carbohydrate moieties. These lectins were chosen as most likely to bind glycoproteins having the N-linked, complex carbohydrate moieties which have previously been suggested to be present in M-CSF [16,34].

Table I shows that a significant fraction of M-CSF activity bound specifically to either LcH agarose (row 1) or to WGA agarose (row 2). The WGA-binding fraction was distinct from the LcH-binding fraction because all detectable M-CSF present in the LcH-unbound fraction bound to WGA agarose (Table I, row 3). In contrast, neither unbound fraction bound to the same column when reapplied (rows 4 and 5). Further, M-CSF activity bound quantitatively to a mixed-bed column of both lectin agaroses, doubling the yield of this purification step (Table I, row 6). Both WGA and LcH-binding fractions contained bioactive M-CSF, as determined by activity in a mouse bone marrow colony-forming assay [27].

Since differences in terminal decoration of carbohydrate chains of glycoproteins affect their charge and apparent size, as well as lectin-binding, we performed immunoblot analyses of MIA PaCa-2 M-CSF eluted from the lectin columns described above. Figure 2 shows that M-CSF in the mixed-bed lectin column eluate (lane A) was more heterogeneous than M-CSF eluted from wheat germ lectin (lane B) or M-CSF eluted from lentil lectin (lane C). M-CSF in the mixed-bed lectin column eluate closely resembled M-CSF in the original supernatant that was purified by an alternate method employing controlled pore glass [47], (not shown). Thus, variable

TABLE I. MIA PaCa-2 M-CSF Glycosylation Heterogeneity Analysis by Lectin Affinity Chromatography*

			% M-CSF (RIA) Recovered	
	Sample	Applied to—	Unbound	Bound & eluted
1	$CaPO_4$ eluate	Lentil lectin[a]	44	40
2	$CaPO_4$ eluate	Wheat germ[b]	36	52
3	LcH unbound	Wheat germ[b]	0	90
4	LcH unbound	Lentil lectin[a]	96	1
5	WGA unbound	Wheat germ[b]	94	2
6	$CaPO_4$ eluate	Mixed bed LcH + WGA	0	85

*MIA PaCa-2 cells were cultured in serum-free medium as described in Materials and Methods, and conditioned medium (6 liters) was partially purified on calcium phosphate gel [14]. Lectin affinity chromatography was performed on aliquots of the calcium phosphate pool as described. M-CSF activity in the unbound and the specifically eluted fractions was determined by RIA of serial twofold dilutions of the sample. Data are reported as percent of the applied activity.

[a]LcH binds N-linked carbohydrate (except highly sialylated); specific for alpha-mannose.

[b]WGA binds N- and O-linked carbohydrate; specific for N-acetyl-D-glucosamine or sialic acid.

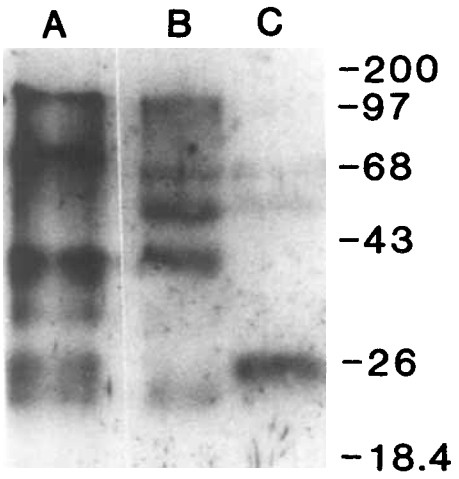

Fig. 2. Immunoblot of a reduced 10% SDS gel of MIA PaCa-2 M-CSF fractions from lectin affinity columns, developed with anti-rM-CSF rabbit antiserum and [^{125}I]-labeled goat antirabbit IgG. Aliquots of a calcium phosphate gel eluate of partially purified MIA PaCa-2 M-CSF were applied to lectin columns; the bound and eluted fraction from each column (see Table I) was concentrated by ultrafiltration and applied to SDS-PAGE. **Lane A:** Mixed-bed LcH/WGA column eluate. **Lane B:** LcH column eluate. **Lane C:** WGA column eluate.

glycosylation is the apparent cause of some of the MIA PaCa-2 M-CSF heterogeneity seen on immunoblots.

M-CSF Protein Heterogeneity Generated by Alternative mRNA Splicing

Alternative mRNA splicing has been shown to occur in the coding region of exon 6 of M-CSF, and cDNA clones coding for two divergent amino acid sequences have been sequenced [3,5,6]. The protein heterogeneity we observed on immunoblots might have resulted from alternative mRNA splicing as well as from the glycosylation heterogeneity indicated by the lectin binding data. Because of differential glycosylation of M-CSF, molecular weight on SDS-PAGE was not sufficient to determine polypeptide chain length. In order to determine whether M-CSF protein heterogeneity reflected alternative mRNA splicing, we developed antibodies specific for the protein products of long and short cDNA clones (see Fig. 3). The long cDNA clone codes for a 554-amino acid M-CSF precursor [5,6] while the short cDNA clone codes for a 256 amino acid precursor [3].

Small peptides have been shown to be useful immunogens in generating polyclonal antibodies to defined areas of a polypeptide sequence [35]. Potentially antigenic sequences were selected from M-CSF amino acid sequences in regions known to be present in post-translationally processed M-CSF from long or short clones. The regions chosen were distant from N-linked glycosylation sites which might mask peptide epitopes in glycosylated M-CSF molecules. The peptide antigens were conjugated to protein carriers and injected into the lymph nodes to maximize immunogenicity.

Two peptides were used to raise site-specific rabbit antibodies. Peptide I spanned predicted amino acids 146–159 of the mature protein specified by the short cDNA clone of M-CSF. Figure 3 shows a diagram of the M-CSF protein species predicted

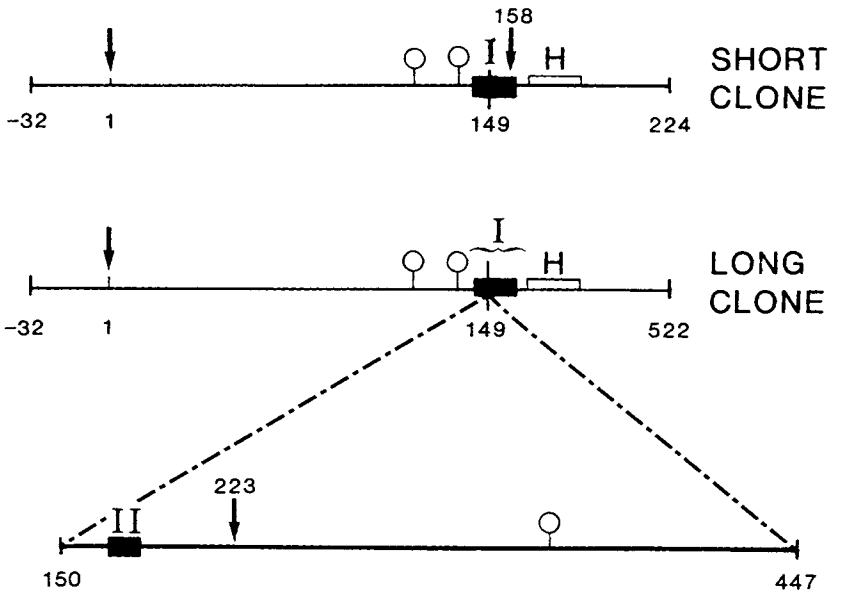

Fig. 3. Diagram of short and long prepro-M-CSF species, as predicted from cDNA clone sequencing, with synthetic peptides I and II highlighted. Diagram of proposed structure of M-CSF subunit as predicted from cDNA sequencing and N-terminal sequence analysis of protein. Short clone [3] and long clone [5,6] share identical amino acid sequences for the leader, amino acids 1–149, and the 74 amino acid C-terminal region. Long clone contains an additional 298 amino acids after amino acid 149 of the mature protein sequence that are not in short clone, due to alternative mRNA splicing in exon 6 [5,6]. ⚲ indicates predicted sites of N-linked glycosylation; H indicates the region of hydrophobic amino acids, a putative transmembrane domain [3]. Arrows indicate specific proteolytic processing sites reported for short clone [22] and long clone [5].

from sequencing of human M-CSF cDNA clones, with arrows indicating apparent proteolytic cleavage sites [3,5,6]. When the short cDNA clone was expressed in CV-1 monkey cells, the secreted protein ended at approximately amino acid 158 and contained the peptide I sequence at its C-terminus [22]. In the long cDNA clone sequence [5,6], the peptide I sequence is interrupted after the first four amino acids by a 298 amino acid insertion (See Fig. 3). Because the C-terminal proteolytic processing site(s) for the long clone occurs within this 298 amino acid region and is therefore N-terminal to the major portion of peptide I sequence [5], the processed, secreted long form protein is not expected to contain the major antigenic epitope of peptide I.

Peptide II spanned predicted amino acids 171–183 of the coding region of the long cDNA clone, which are contained in the 298 amino acid inserted region, and had a Cys residue added for coupling purposes. Because its sequence is present in the long cDNA clone only, peptide II antibody is expected to bind to long-clone-derived, but not short-clone-derived, M-CSF protein species. Together with the rM-CSF antibody, these peptide-specific antibodies permitted identification of the translation products of long and short mRNA precursors.

Figure 4 shows a dot-blot of rM-CSF standards developed with these three antibodies. Antibody to short clone rM-CSF recognized rM-CSF from both short and long clones (lanes A and B, respectively). As expected, affinity-purified antibody to

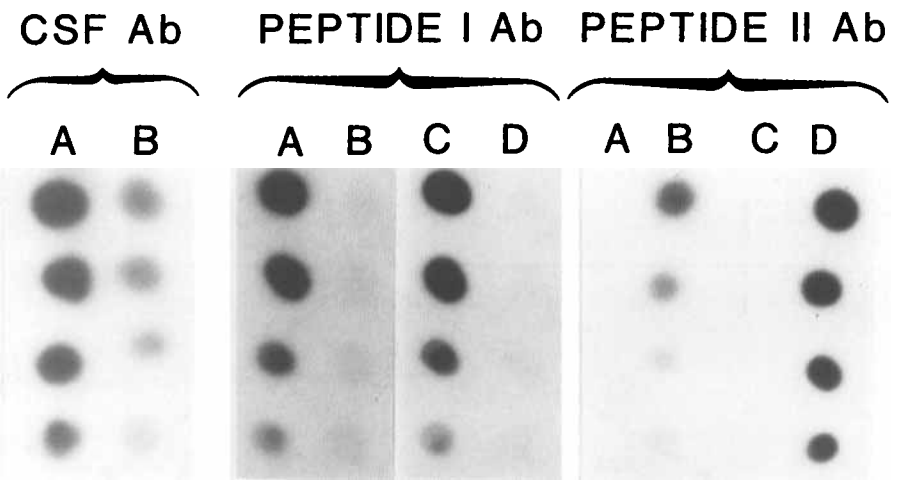

Fig. 4. Dot-blot of rM-CSF standards showing specificities of peptide antibodies. Samples (1 μl) were spotted on nitrocellulose and developed with [^{125}I]-labeled rM-CSF IgG, peptide I IgG, and peptide II IgG, as described for immunoblots in Experimental Procedures. Serial fivefold dilutions are shown of the following samples. **A:** Short clone rM-CSF (1.5×10^7 U/ml) secreted from CV-1 cells and purified [22]. **B:** Long clone rM-CSF (1×10^7 U/ml), a truncated 221 amino acid form expressed and purified from *E. coli* (Halenbeck et al., submitted). **C:** Purified peptide I/BSA conjugate (0.04 mg/ml). **D:** Purified peptide II/BSA conjugate (0.04 mg/ml).

peptide I reacted with rM-CSF purified from supernatants of CV-1 cells expressing the short cDNA clone (lane A), and did not react with rM-CSF expressed from the long cDNA clone (lane B). In contrast, affinity-purified antibody to peptide II did not react with short-clone-derived rM-CSF (lane A), but did react with long-clone-derived rM-CSF (lane B). Reactivity with peptide antibodies was specific for the chosen sequences, since neither peptide antibody reacted with the other peptide (lanes C, D) or with three other M-CSF synthetic peptides having unrelated sequences. In addition, neither antibody reacted with a 150-amino acid C-terminally truncated rM-CSF molecule expressed in *E. coli,* which did not contain peptide II sequence and contained only the first four amino acids of peptide I (data not shown). Since the major antigenic epitope of peptide I appears to be after amino acid 150, peptide I antibody is specific for secreted short-clone-derived M-CSF that has a C-terminus after amino acid 150 of the mature protein.

Some native MIA PaCa-2 M-CSF species that reacted with rM-CSF antibody (Fig. 5, lane A) also reacted strongly with peptide II antibody (Fig. 5, lane B). In contrast, no signal was seen when the same MIA PaCa-2 M-CSF sample was developed with peptide I antibody under the same conditions (lane C). After tenfold concentration of the sample and development as before, a 25 kd band was seen (lane D). Thus, secreted MIA PaCa-2 M-CSF (purified as described) contained predominantly long clone (peptide II) epitope and therefore could not have been derived from the short mRNA splice variant. The remainder reacted with short-form-specific antibody.

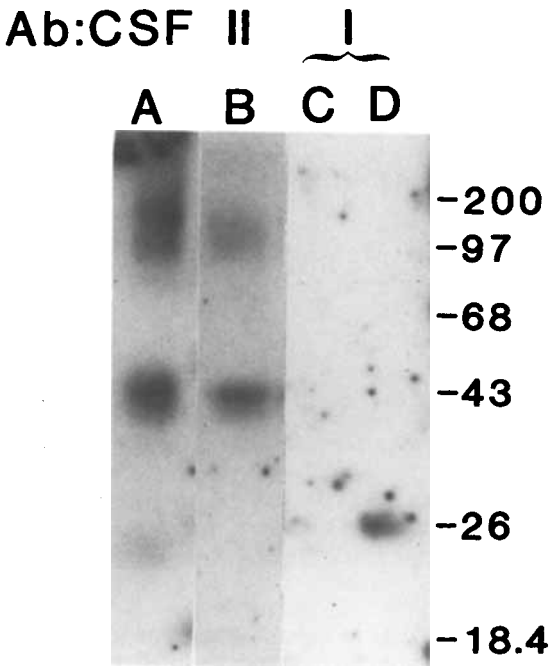

Fig. 5. Most MIA PaCa-2 M-CSF protein reacts with long-clone-specific peptide II antibody. MIA PaCa-2 M-CSF, purified by mixed-bed lectin chromatography and immunoaffinity, was immunoblotted after 10% reducing SDS-PAGE. **Lane A:** Developed with [^{125}I]-labeled rM-CSF IgG. **Lane B:** Developed with labeled peptide II antibody (specific for long form M-CSF). **Lanes C,D:** Developed with labeled peptide I antibody (specific for short-form M-CSF). In lanes A–C, 2,000 units of M-CSF were applied; in lane D, the same sample was concentrated using a Centricon-10 filter and 15,000 units were applied to the gel. No M-CSF activity was detectable in the fraction which passed through the filter.

Cell-Associated M-CSF

As shown in Figure 3, both cDNA clones contain a putative transmembrane region. To examine the possibility that proM-CSF forms might be detectable in the cell fraction, we analyzed cell lysates by RIA and immunoblotting. Several researchers had detected M-CSF antigen in cell lysates by using RIA [23,24, reviewed in 2]. As Table II shows, detergent lysates of monocytes, HL-60, and MIA PaCa-2 cells contain a small amount of RIA-detectable M-CSF. The amount detected represents about 0.1–1% of total antigen in the culture, assuming a similar reactivity with antibody. The quantity of cell-associated M-CSF in monocytes increased with time in culture, and addition of PMA to the culture increased the levels of both cell-associated and secreted M-CSF by 4–5-fold (Table II). Lymphocytes isolated from peripheral blood by adherence removal of monocytes did not contain RIA-detectable M-CSF.

Cell-associated M-CSF differed from secreted M-CSF in size and exhibited little heterogeneity on immunoblots (compare Fig. 6 to Fig. 1, lanes D–F). In monocytes and HL-60 cells, a single 60 kd species of M-CSF was seen on immunoblots of reduced SDS gels (Fig. 6, lanes A,B). For comparison, CV-1 cells expressing the short

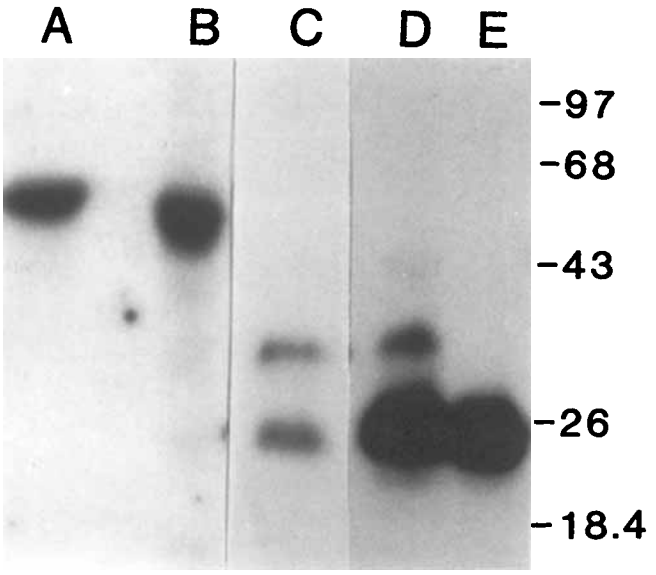

Fig. 6. Cell-associated M-CSF analyzed on immunoblots. Immunoblot of a 10% reduced SDS gel, developed using [^{125}I]-anti-CSF IgG. Samples prepared as described in Materials and Methods. **Lane A:** Monocyte lysate. **Lane B:** HL-60 lysate. **Lane C:** Lysate from recombinant CV-1 cells expressing short-clone form of M-CSF. **Lane D:** Sample of lane C, to which secreted CV-1 rM-CSF has been added. **Lane E:** Secreted CV-1 rM-CSF.

form of rM-CSF having a predicted 224 amino acid sequence were extracted and analyzed and species of 24 kd and 30 kd were seen (Fig. 6, lane C). Addition of purified secreted CV-1 rM-CSF to the cell-associated fraction (Fig. 6, lane D) showed that the smaller species comigrated with secreted CV-1 rM-CSF (lane E) on SDS-PAGE. The data suggest that rM-CSF is synthesized in these cells as a larger polypeptide which is proteolytically processed to the secreted 24 kd form, as would be expected if it represented an intermediate in post-translational processing.

DISCUSSION

Alternative mRNA splicing within the coding region of M-CSF is predicted to generate protein molecules that share a 32 amino acid leader sequence, a 149 amino acid N-terminal region, and a 74 amino acid C-terminal region that contains a hydrophobic sequence [3,5,6]. Some M-CSF mRNA species apparently specify up to 298 amino acids of additional sequence that follows amino acid 149 of the mature protein. C-Terminal proteolytic processing, which is thought to be necessary for secretion, occurs some distance after amino acid 149 and apparently occurs at different sites among the expressed products of the M-CSF splice variants [3,5,6].

We have shown that heterogeneous forms of human M-CSF protein are secreted by several cell types in tissue culture. These forms react specifically with an antibody to a truncated, 150 amino acid long, rM-CSF purified from *E. coli* (Kawasaki et al., submitted). This portion of the M-CSF sequence is present in the protein expressed

TABLE II. Cell-Associated M-CSF Detectable by RIA*

Cell type	M-CSF Concentration (RIA)	
	Cell lysate (units/10^7 cells)	Medium (units/ml)
Blood monocyte		
Freshly isolated	490	—
24 h	890	<50
24h + PMA	3,700	2,070
HL-60 cells		
24 h + PMA	2,400	1,360
MIA PaCa-2 cells		
PMA + RA	90	1,400
Blood lymphocyte		
24 h + PMA	<50	Not done

*Cells (5×10^7) were extracted on ice with buffer containing Triton X-100, as described in Materials and Methods. Extracts were clarified and diluted immediately into buffer containing 1 mg/ml BSA for RIA. (The RIA had a limit of detection of 50 units/ml.)

from all known cDNA clones of M-CSF and probably contains the domain required for M-CSF receptor interaction (see Fig. 3) [3–6,20]. Expression of a truncated cDNA clone coding for the N-terminal 150 or 158 amino acids of human M-CSF in mammalian cells results in secretion of active rM-CSF protein (Kawasaki and Ladner, personal communication) [36]. Recently, purification of a 158 amino acid long rM-CSF from CV-1 cells expressing the short cDNA clone has confirmed that such a truncated molecule has full biological activity [22]. The MIA PaCa-2 cell line, from which M-CSF was originally cloned, secretes several M-CSF protein species which are bioactive and partially separable using lectin affinity chromatography. Size heterogeneity is also seen in native M-CSF from HL-60, PANC-1, and peripheral blood monocytes cultured in vitro, suggesting that heterogeneity is not an artifact of transformed cells.

M-CSF protein has been purified to apparent homogeneity from both MIA PaCa-2 cells and human urine. In each case, the purified M-CSF protein was reported to consist of a single species of 23 to 28 kd on reduced SDS-PAGE [9–14,16,18]. However, about half of the starting M-CSF activity was lost during lectin affinity chromatography. As we have shown here, use of a mixed-bed column of two different lectins doubles the yield of M-CSF from MIA PaCa-2 cell supernatants and apparently retains all species that were present in the culture supernatant (as determined by immunoblotting, data not shown). Our data suggest that use of a single lectin affinity column to purify M-CSF has greatly reduced M-CSF heterogeneity and yield in previous studies by fractionating differently glycosylated species. The M-CSF species that bound to LcH presumably contain N-linked complex carbohydrate, the structure that has been previously reported for M-CSF [9,16]. The WGA-binding M-CSF may include both highly sialylated N-linked species and O-linked species. Indeed, Csejtey and Boosman saw minor protein contaminant species of 40, 33, and 28 kd in their final purified sample [14]. Our analysis suggests that bands of these sizes react with antibodies to rM-CSF and reflect variable glycosylation. In agreement with our data, the

long splice variant of M-CSF has been recently reported to have O-linked carbohydrate [42]. The mixed-bed lectin column utilized in our purification method provides an efficient means of purifying heterogeneous M-CSF species that differ greatly in charge and size.

PANC-1 cells secreted an M-CSF antibody-reactive species that shifted from a diffuse 100–200 kd smear to a 100 kd band upon reduction (Fig. 1A,B). Whereas CV-1 cells expressing rM-CSF with the short coding region did not have a 100 kd band, M-CSF of this size was seen in PANC-1 culture supernatants as well as some immunoaffinity-purified MIA PaCa-2 samples. The unusual band might represent a new form of M-CSF, such as a hyper-glycosylated species, or another mRNA splice variant [46]. An alternative interpretation might be an unusually stable aggregate or complex containing M-CSF.

We have developed synthetic peptide antibodies as specific probes to identify the protein products of alternative mRNA splicing of M-CSF on immunoblots. These antibodies, which are specific for C-terminal regions of M-CSF, do not inhibit M-CSF activity, supporting the hypothesis that receptor binding involves the N-terminal region of M-CSF. MIA PaCa-2 and PANC-1 cells express several M-CSF mRNA species that range in size from 1.4 to 4 kb [3,27] and that comprise at least two predicted coding sequences [3,5,6]. Our immunoblot data suggest that relatively little of the short form of M-CSF protein is secreted by MIA PaCa-2 cells, the source from which the short cDNA species was originally cloned [3]. Since the corresponding mRNA species is detectable by Northern blotting, the low level of short form protein detected may reflect differences in rates of translation or secretion of the long and short forms [37,38, 46]. Several species of various molecular weights reacted with rM-CSF antibody and with peptide II (long coding region) antibody on immunoblots. Variability in M-CSF glycosylation might interfere with the ability of some peptide antibodies to bind to their epitopes. Because all of the bands recognized by a polyclonal M-CSF antibody were recognized by one of the antipeptide antibodies, it is unlikely that a major form was overlooked by this analysis. However, still other splicing variants may exist, that have yet to be characterized.

In agreement with previous reports [21,23,24], we were able to detect M-CSF in lysates of whole cells using RIA. We also analyzed this material on immunoblots and demonstrated that cell-associated M-CSF differs from secreted M-CSF in size. HL-60 cells and peripheral blood monocytes had a single 60 kd species, while expression of a short cDNA clone of human M-CSF in the CV-1 cell line [22,29] gave cell-associated M-CSF species of 24 kd and 30 kd (Fig. 6). The 24 kd species comigrated with M-CSF secreted by CV-1 cells and may represent trapping of the secreted 24 kd species in these virus-infected cells, proteolytic processing of the 30 kd precursor during sample handling, or an intracellular pool of releasable rM-CSF. The 30 kd species may represent the 224 amino acid pro-CSF species prior to C-terminal processing, while the 60 kd species detected in the HL-60 and monocyte lysates may be a product of a different, longer mRNA species. Indeed, monocytes have been shown to express a 4 kb M-CSF mRNA species that corresponds to the long cDNA clone [28,39,40].

Like epidermal growth factor [41], M-CSF is proteolytically processed both N- and C-terminally from a membrane-spanning preprohormone [3,20]. rM-CSF from the short coding region cDNA clone has been reported to be cleaved and released from the plasma membrane [3,5,21,24], while the long coding region protein may be

cleaved within the cell and secreted by a different mechanism [37,38,46]. Polypeptide sequence heterogeneity at or near the C-terminus of mature, secreted M-CSF, generated by alternative mRNA splicing and proteolytic processing, contributes directly to the size heterogeneity seen in human M-CSF derived from several cell lines. It may also affect size indirectly by altering glycosylation [37]. The presence of what appears to be pro M-CSF in the cell-associated fraction indicates that this molecule is available for local release from the cell surface or cell-to-cell interaction [21].

Sequence analysis of trypsin-digested human urinary M-CSF [5] has clearly demonstrated that the long form of this protein is present in vivo. It is not yet known whether the M-CSF product from the coding region of the short mRNA splice variant is also present, because it differs in only a short sequence at the C-terminus. Both long and short forms are active on mouse bone marrow cells in vitro, and comparison of rM-CSF species suggests that both forms compete similarly in a receptor-binding assay (M.-T. Lee, P. Shadle, R. Halenbeck, unpublished).

We have shown that considerable M-CSF protein heterogeneity is generated after translation. Glycosylation differences, such as those demonstrated here, could affect turnover, tissue targeting, and stability of the protein. The existence of distinct mRNA splicing events that lead to predicted amino acid sequences that diverge [5,6] raises the possibility that M-CSF protein species having differing functions are expressed from the single M-CSF gene in vivo [21,42]. Both hematopoietic and fibroblastic cells have been shown to express M-CSF mRNA [28,39,43,44]. It is possible that the different coding region forms of M-CSF, first observed in tumor cell lines, are tissue-specific [44]. Indeed, the recent finding that M-CSF mRNA expressed in mouse uterus during pregnancy consists of a 2.3 kb species (distinct from the 4–4.5 kb seen in monocytes) supports this notion [45]. Alternatively, these different M-CSF forms may play different roles in the same tissues. The synthetic peptide antibodies we have developed will be useful probes to address these questions. In any event, M-CSF expressed by cells could functionally diverse, and both the secreted and the cell-associated forms mediate various interactions in vivo.

NOTE ADDED IN PROOF

After this paper was submitted, Cerretti et al. (Cerretti DP, Wignall J, Anderson D, Tushinski RJ, Gallis BM, Stya M, Gillis S, Urdal DJ, and Cosman D: *Mol. Immunol.* 25:761–770, 1988) reported isolation of an intermediate-size human M-CSF cDNA clone (M-CSFγ) that arises from alternative splicing of the M-CSF transcript. We believe that the processed protein form of M-CSFγ should be indistinguishable from long clone in our Western analyses because it is predicted to contain the same peptide sequence that we used to generate antibody specific for M-CSF long clone.

ACKNOWLEDGMENTS

Mei-Ting Lee and Adam Sampson-Johannes for monocyte and HL-60 cells and culture supernatants; Robert Halenbeck for antibodies to rM-CSF; Richard Stanley for monoclonal antibody to M-CSF; Judit Csejtey for peptide purifications; Angela Guadamuz for technical assistance in peptide synthesis; and the Cetus Assay Lab for RIA determinations and bioassay.

REFERENCES

1. Metcalf D: Blood 67:257–267, 1986.
2. Stanley ER: Methods Enzymol 116:564–587, 1985.
3. Kawasaki ES, Ladner MB, Wang AM, Van Arsdell J, Warren MK, Coyne MY, Schweickart VL, Lee MT, Wilson KJ, Boosman A: Science 230:291–293, 1985.
4. Ralph R, Warren MK, Ladner MB, Kawasaki ES, Boosman A, White TJ: Cold Spring Harbor Symp Quant Biol (US) 51:679–683, 1986.
5. Wong GG, Temple PA, Leary AC, Witek-Giannotti JS, Yang YC, Ciarletta AB, Chung M, Murtha P, Kriz R, Kaufman RJ, Ferenz CR, Sibley BS, Turner KJ, Hewick RM, Clark SC, Yanai N, Yokota H, Yamada M, Saito M, Motoyoshi K, Takaku F: Science 235:1504–1508, 1987.
6. Ladner MB, Martin GA, Noble JA, Nikoloff DM, Tal R, Kawasaki ES, White TJ: EMBO J 6:2693–2698, 1988.
7. Das SK, Stanley ER, Guilbert LJ, Forman LW: Blood 58:630–641, 1981.
8. Stanley ER, Metcalf D: Pse B 37:1029–1031, 1971.
9. Wang FF, Goldwasser E: J Cell Biochem 21:263–275, 1983.
10. Sakai N, Tsuneoka K, Shikita M: J Cell Physiol 118:1–5, 1984.
11. Hatake K, Motoyoshi K, Ishizaka Y, Sato M, Takaku F, Miura Y: J Chromatogr 344:339–344, 1985.
12. Wu MC, Cini JK, Yunis AA: J Biol Chem 254:6226–6229, 1979.
13. Wu MC, Yunis AA: J Clin Invest 65:772–775, 1980.
14. Csejtey J, Boosman A: Biochem Biophys Res Commun 138:238–245, 1986.
15. Waheed A, Shadduck RK: J Lab Clin Med 94:180–193, 1979.
16. Stanley ER, Guilbert LJ: J Immunol Methods 42:253–284, 1981.
17. Gaffney EV, Dell'Aquila ML, Lingenfelter SE, Huffnagle GB, Wiest DL: J Leukocyte Biol 39:409–421, 1986.
18. Shieh J-H, Cini JK, Wu M-C, Yunis AA: Arch Biochem Biophys 253:205–213, 1987.
19. Waheed A, Shadduck RK: Blood 60:238–244, 1982.
20. Ralph P, Ladner MB, Wang AW, Kawasaki ES, McConlogue L, Weaver JF, Weiss SA, Shadle PJ, Koths K, Warren MK: In Webb DR, Pierce CW, Cohen S (eds): "The Molecular Basis of Lymphokine Action." Clifton, NJ, Humana Press, 1987, pp 295–311.
21. Rettenmier CW, Roussel MF, Ashmun RA, Ralph P, Price K, Sherr CJ: Mol Cell Biol 7:2378–2387, 1987.
22. Halenbeck R, Shadle PJ, Lee P-J, Lee M-T, Koths K: J Biotechnol 8:44–58, 1988.
23. Cifone M, Defendi V: Nature 252:151–153, 1974.
24. Stanley ER: Proc Natl Acad Sci USA 76:2969–2973, 1979.
25. Merrifield RB: Science 232:341–347, 1985.
26. Aldwin L, Nitecki DE: Anal Biochem 164:494–501, 1987.
27. Ralph P, Warren MK, Lee MT, Csejtey J, Weaver JF, Broxmeyer HE, Williams DE, Stanley ER, Kawasaki ES: Blood 68:633–639, 1986.
28. Horiguchi J, Warren MK, Ralph P, Kufe D: Biochem Biophys Res Commun 141:924–930, 1986.
29. Weaver J, McCormick F, Manos MM: Bio/Technology 6:287–290, 1988.
30. Laemmli UK: Nature 227:680–685, 1970.
31. Burnette WN: Anal Biochem 112:195–203, 1981.
32. Hunter WM, Greenwood FC: Nature 194:495–496, 1962.
33. Burgess AW, Metcalf D, Kozka IJ, Simpson RJ, Vairo G, Hamilton JA, Nice EC: J Biol Chem 260:16004–16011, 1985.
34. Allen AK, Desai NN, Neuberger A: Biochem J 155:127–135, 1976.
35. Anderer FA, Schlumberger MD: Biochim Biophys Acta 97:503–509, 1965.
36. Heard JM, Roussel MF, Rettenmier CW, Sherr CJ: Oncogene Res 1:423–440, 1987.
37. Roussel MF, Rettenmier CW, Sherr CJ: Blood 71:1218–1225, 1988.
38. Rettenmier CW, Roussel MF: Mol Cell Biol 8:5026–5034, 1988.
39. Rambaldi A, Young DC, Griffin JD: Blood 69:1409–1413, 1987.
40. Haskill S, Johnson C, Eierman D, Becker S, Warren K: J Immunol 140:1690–1694, 1988.
41. Scott J, Urdea M, Quiroga M, Sanchez-Pescador R, Fong N, Selby M, Rutter WF, Bell GI: Science 221:236–240, 1983.

42. Clark SC, Kamen R: Science 236:1229–1237, 1988.
43. Bartocci A, Pollard JW, Stanley ER: J Exp Med 164:956–961, 1986.
44. Rajavashisth TB, Eng R, Shadduck RK, Waheed A, Ben-Avram CM, Shively JE, Lusis AJ: Proc Natl Acad Sci USA 84:1157–1161, 1987.
45. Pollard JW, Bartocci A, Arceci R, Orlofsky A, Ladner MB, Stanley ER: Nature 330:484–486, 1987.
46. Manos MM: Mol Cell Biol 88:5035–5039, 1988.
47. Shadle PJ, Allen JI, Geier MD, Koths K: Exp Hematol 17:154–159, 1989.