Hopp, T.P., Prickett, K.S., Price, V.L., Libby, R.T., March, C.J., Cerretti, D.P., Urdal, D.L. and Conlon, P.J. (1988) A short polypeptide marker sequence useful for recombinant protein identification and purification. Bio/Technology 6, 1204-1210.
http://thomas-hopp.com/THscience1.html

A SHORT POLYPEPTIDE MARKER SEQUENCE USEFUL FOR RECOMBINANT PROTEIN IDENTIFICATION AND PURIFICATION

Thomas P. Hopp, Kathryn S. Prickett, Virginia L. Price, Randell T. Libby[1], Carl J. March, Douglas Pat Cerretti, David L. Urdal and Paul J. Conlon

Immunex Corporation, 51 University Street, Seattle, Washington 98101. [1]Present address: Departments of Genetics, SK-50, and Pathology, SM-30, University of Washington, Seattle, WA 98195.

Abstract: A small hydrophilic peptide of eight amino acids (AspTyrLysAspAspAspAspLys) was engineered onto the N-terminus of a variety of recombinant lymphokines for the purpose of aiding in their detection and purification from yeast supernatants or *E. coli* extracts. An antibody specific for the first four amino acids of this sequence was used as a detection reagent and for immunoaffinity purification of products under mild conditions. Because of the small size of the peptide moiety and its hydrophilic nature, the proteins were unaffected by its presence and retained a high level of biological activity. In addition, it was possible to remove the peptide via an enzymatic cleavage procedure using enterokinase.

INTRODUCTION

Among the many approaches taken to improve the yield and purity of recombinant proteins, one particularly useful procedure is to express the desired polypeptide as part of a larger fusion protein[1,2]. Fusion to a protein-export signal sequence has been used to cause secretion of products from yeast[3] and *E. coli*[4] cells. Furthermore, it has recently been recognized that an attached fusion polypeptide sequence might serve as an aid to identifying or purifying the product. For example, in several cases the added polypeptide segment contains a complete second protein that binds to affinity columns via its specific substrate or ligand. These include β-galactosidase fusion proteins that bind to aminophenylthiogalactosidyl Sepharose columns[5] and protein A fusion proteins that bind to immunoglobulin columns[6]. Such fusion proteins can be highly purified in good yields in a single step by passing cell extracts or supernatants over columns of an appropriate affinity matrix, then eluting the purified fusion protein by changing conditions so that binding is no longer possible. A related approach is to use an antibody directed against the added sequence as a detection or affinity purification reagent[7-9] although the high binding affinity of most antisera and monoclonal antibodies often requires the use of denaturing conditions for elution of the product.

The fusion protein approach as several drawbacks that have not been adequately addressed in the past. First, most fusion protein products fail to fold properly into a native, active state[10]. It is possible that the added polypeptide segment is responsible for this misfolding due to unfavorable interactions during folding of the fusion protein. This often necessitates treatment with strong denaturants such as 8M urea and 7M guanidine HCl followed by refolding procedures[11,12].

A further problem with fusion proteins is that it is often difficult or impossible to remove the additional amino- or carboxyl-terminal sequence from the desired protein product. One solution has been to use relatively drastic conditions and chemical cleavage agents such as CNBr in 70% formic acid[8,12] or low pH incubations[13] to effect cleavage. However, recent studies have attempted to use somewhat milder chemical cleavages such as hydroxylamine treatment at pH 9.0[14] or enzymatic cleavage procedures under physiological conditions. The blood clotting factor X_a has a proteolytic specificity for the tetrapeptide sequence

IleGluGlyArg, and has been used to liberate β-globin from a λcII protein fusion sequence[10]. A sequence allowing specific cleavage by collagenase has also been proposed[5]. Sassenfeld and Brewer[15] developed a facilitated ion-exchange purification technique by fusing their proteins to a C-terminal series of arginine residues that are subsequently removed by carboxypeptidase B treatment. These enzymatic processes have been successful in several instances, but often have been limited by poor cleavage yields or by unwanted cleavages that occur within the desired protein sequence[10].

We decided to create a recombinant protein detection and purification system that incorporated several aspects of the above mentioned procedures in order to produce a fusion sequence with a combination of the most desirable properties. Here we report the development of a short, N-terminal fusion sequence AspTyrLysAspAspAspAspLys, that we refer to as a marker sequence or "Flag™" segment for antibody mediated identification and purification of recombinant proteins. We also describe a monoclonal antibody that reacts with this sequence and can be used as an immunoaffinity purification reagent that purifies marker fusion proteins under very mild conditions. Finally, the marker sequence can be removed by treatment with the protease, enterokinase, which is specific for the five C terminal amino acids of the marker sequence[16]. No harsh treatments are required at any step in this process and proteins purified by this approach retained their biological activity throughout the purification, even while the marker sequence was attached. This paper describes the expression of several such fusion proteins in *Saccharomyces cerevisiae* and *Escherichia coli.*

RESULTS

In order to develop this system, we performed a series of interrelated steps. The eight amino acid marker peptide was engineered onto the N-terminus of the lymphokine interleukin 2 (IL-2)[17] by means of synthetic oligonucleotides. The fusion protein was expressed in yeast and the product purified by conventional means, then used as an immunogen to produce a monoclonal antibody (4E11) specific for the marker sequence. The antibody functioned as a reagent for a number of different immunochemical procedures including "Westerns," "dot blots," immunoprecipitations, and affinity purification when coupled to a solid support. Furthermore, the discovery that the 4E11 antibody would release its antigen when calcium was removed from the medium led to the development of a mild purification procedure for fusion protein elution from the affinity columns. Next, treatment of the fusion proteins with enterokinase demonstrated that the enzyme was capable of removing the marker segment efficiently, with little or no observable degradation of the desired protein product. Finally, measurements of specific activity demonstrated that, for all proteins tested, no appreciable loss of activity was caused by the presence of the marker segment on the N-termini of the recombinant products.

Secretion of fusion proteins from yeast. Figure 1 shows the plasmid pIXY8, used for production of the IL-2 fusion protein in yeast. Similar constructs with the ADH2 promoter replacing the α-factor promoter were used to express granulocyte colony stimulating factor (G-CSF)[18], interleukin 3 (IL-3)[19], interleukin 4 (IL-4)[20] and granulocyte-macrophage colony stimulating factor (GM-CSF)[21] fusion proteins. All of these proteins were secreted into culture media by yeast. Each had the expected molecular weight for the correctly processed form (i.e. with leader peptide removed) and yielded the expected sequence of the marker peptide on N-terminal amino acid sequence analysis. The IL-2 fusion protein was purified by HPLC for use as an immunogen, while the other fusion proteins were purified by the 4E11 antibody immunoaffinity chromatography procedure (below). For comparison, essentially identical vectors were prepared that expressed each protein without the marker peptide. These products were recovered from the yeast culture media and purified to homogeneity by conventional techniques including HPLC and ion exchange chromatography[22].

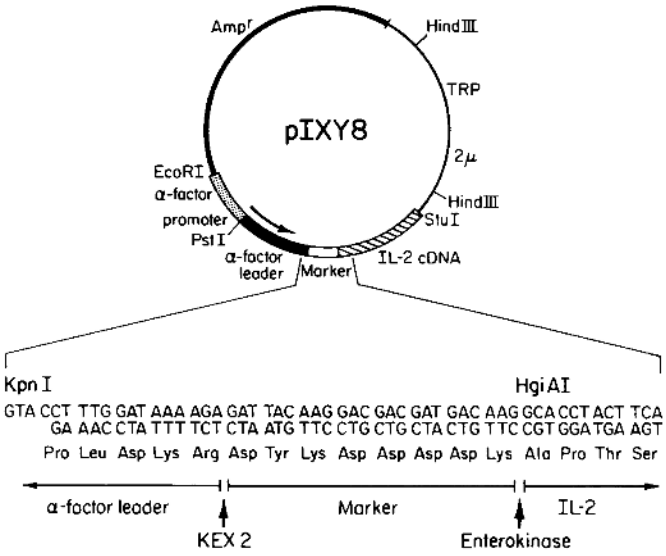

FIGURE 1. Plasmid pIXY8, for expression of the IL-2 fusion protein in yeast. The synthetic oligonucleotides used in constructing this plasmid extended from the KpnI site at the left of the sequence shown below the plasmid diagram, to a blunt end at the HgiAI site near the right side, ending with the codon for the C-terminal Lys residue of the marker peptide segment. The IL-2 coding sequence extended from a blunt end before the first codon of IL-2 (Ala) to a StuI site beyond the termination codon. The arrows below the amino acid sequence indicate the sites of cleavage by the KEX2 protease to remove the α-factor precursor sequence from the primary translation product, and by enterokinase, to remove the marker peptide from the product protein.

Expression of GM-CSF fusion protein in *E. coli.* The construction of the plasmid pRL6-6-87 for expression and secretion of the GM-CSF fusion protein in *E. coli* is outlined in Figure 2. This plasmid allows the secretion of the marker peptide GM-CSF fusion protein by means of the signal peptide from the outer membrane protein OmpA. The product obtained from *E. coli* cultures had the expected molecular weight for the marker peptide GM-CSF fusion and yielded an N-terminal amino acid sequence corresponding to the marker peptide sequence.

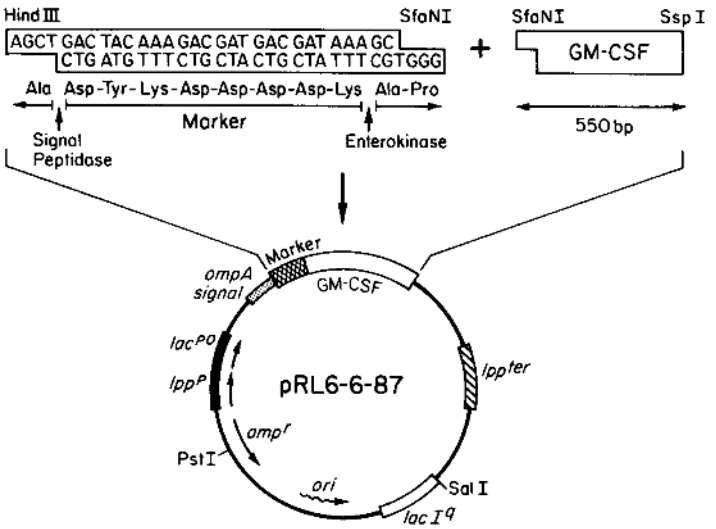

FIGURE 2. Plasmid pRL6-6-87, for expression of the GM-CSF fusion protein in *E. coli.* Abbreviations: *lpp*^{p}, lipoprotein promoter; *lac*^{po}, lactose promoter-operator region; *ompA* signal, outer membrane protein A signal peptide sequence; GM-CSF, human granulocyte-macrophage colony stimulating factor; *lpp*^{ter}, lipoprotein terminator; *lacI*^{q}, lactose repressor; *ori,* origin of replication.

The marker-specific antibody. The isotype of the 4E11 antibody is IgG 2B. It was found to be reactive with proteins bearing the marker peptide sequence in a variety of procedures, including ELISAs, dot blots, Western blots, immunoprecipitation and affinity chromatography, as described below. The antibody was found to react with all of the marker peptide fusion proteins that we have produced. The antibody exhibits no reactivity with the non-marker products, or any component present in *E. coli* extracts, or in yeast culture medium.

Purification. Figure 3 shows the results of typical affinity purification chromatograms on affinity columns made with the 4E11 antibody. In Figure 3A, a culture supernatant obtained by fermenting yeast bearing the GM-CSF expression vector was passed over the column to purify the fusion protein that had been secreted into the medium. Medium components were removed by washing with PBS containing 0.5 mM $CaCl_2$, while the GM-CSF remained

bound to the antibody. Subsequent elution with PBS containing EDTA dissociated the marker peptide-antibody complex and released the GM-CSF fusion protein as a purified product. The multiple molecular weight species eluting from the column are typical of glycoproteins secreted from yeast and result from heterogeneous glycosylation by yeast cells. All bands were identified as GM-CSF, based on Western analyses using 4E11 as well as anti GM-CSF monoclonal antibody as developing reagents.

Figure 3B shows the results of affinity chromatography of an extract of *E. coli* cells that had been transformed with pRL6-6-87 in order to produce the GM-CSF fusion protein. Chromatography was carried out as with the yeast GM-CSF fusion, except that 1 mM Ca^{+2} was used during washing and 0.1 M glycine HCl pH 3.0 buffer was used to elute the product. The GM-CSF fusion protein eluted as a single molecular weight species because *E. coli* does not glycosylate proteins. The product was essentially pure after this single chromatographic step.

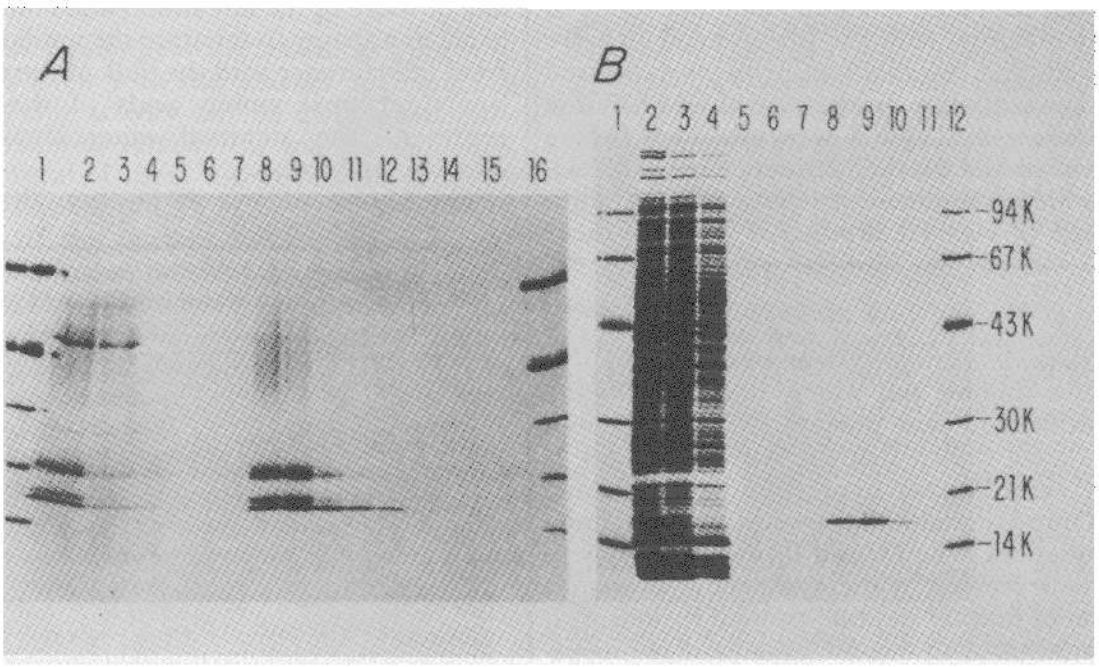

FIGURE 3. Affinity purification of recombinant fusion proteins. Panel A: Silver-stained polyacrylamide gel of yeast GM-CSF fusion protein purification. Lanes are: 1: molecular weight standards; 2: yeast supernatant; 3: flow through material; 4–6: sequential PBS/0.5 mM $CaCl_2$ washes; 7–15: sequential PBS/2.0 mM EDTA elutions; 16: molecular weight standards. Panel B: Silver-stained gel of *E. coli* GM-CSF fusion protein purification. Lanes are: 1: molecular weight standards; 2: *E. coli* supernatant; 3: flow through material; 4–6: sequential PBS/0.5 mM $CaCl_2$ washes; 7–11: sequential 0.1 M Gly-HCl pH 3 elutions; 12: molecular weight standards. Numbers at right indicate M_r values for the standard proteins (in kiloDaltons). The 94K standard was omitted from the gel shown in panel A.

The binding of the 4E11 antibody to the marker peptide is dependent on the presence of calcium. This property has been reported for a few other antibodies in the past[23]. We observed that if insufficient levels of calcium were present in washing buffer, then the fusion proteins would leak from the affinity column, even though they had bound quantitatively when yeast supernatant or E. coli extract was passing over the column. By testing the effect of various concentrations of $CaCl_2$ in washing buffer, we determined that concentrations of calcium above 0.3 mM were necessary to retain the fusion proteins on the column. Given this calcium dependence, it was found that rapid elution of fusion proteins could be achieved by using EDTA in the elution buffer. It was also possible to elute proteins simply by using a calcium free elution buffer after the columns had been washed with PBS containing calcium. However, under these conditions the eluted fusion protein tended to spread through more fractions than when the elution buffer contained EDTA.

Enterokinase treatment. Figure 4 shows the results of enterokinase treatment of the IL-2 fusion protein. Increasing amounts of enterokinase were added to identical aliquots of the IL-2 fusion protein, then the samples were incubated for 16 hr at 37°. As the concentration of enzyme was increased, a component appeared at the molecular weight of the authentic IL-2. At the highest concentrations of enzyme, the conversion of fusion protein to authentic protein was complete, and an approximately identical amount of authentic product had been formed. Western blot analyses using 4E11 and anti-IL-2 monoclonal antibodies confirmed the identities of the products seen on the silver gel. Only the higher molecular weight species was reactive with 4E11, but as expected both the higher and lower molecular weight species were reactive with the anti-IL-2 antibody. Amino terminal sequence analysis on the resulting cleavage products indicated that the IL-2 fusion protein was cleaved after the second lysine of the marker peptide yielding the expected sequence for the authentic IL-2 N-terminus. As can be seen in Figure 4, the cleavage is complete, and in this case results in only small amounts of detectable lower molecular weight byproducts formed by enterokinase cleavage within the IL-2 sequence. Essentially the same cleavage pattern was obtained with the other fusions as well. In no case were degradation products present in quantities greater than 10% of total protein (determined by sequence analysis).

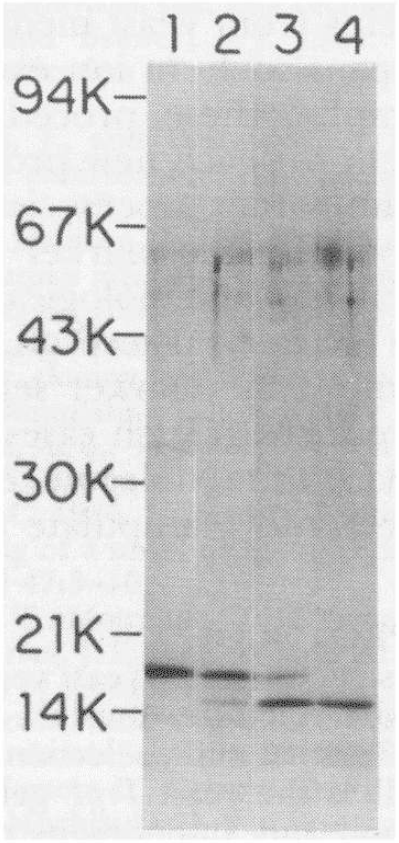

FIGURE 4. Enterokinase digestions. The IL-2 fusion protein was incubated with increasing amounts of bovine enterokinase and the digestion was analyzed by silver-staining PAGE. Lanes are: 1: purified IL-2 fusion protein (200 ng) no enzyme; 2: fusion protein + 2 ng enterokinase; 3: fusion protein + 10 ng enterokinase; 4: fusion protein + 20 ng enterokinase.

Retention of biological activity. The marker peptide fusion proteins expressed in yeast and *E. coli* were all biologically active despite the presence of the marker segment, as is seen in Table 1. In all cases the levels of biological activity obtained with the fusion proteins were comparable to wild-type recombinant proteins expressed without the marker sequence. In the case of GM-CSF the specific activity values in Table 1 were obtained before and after removal of the marker sequence by enterokinase. The nearly identical specific activities obtained indicate that the yield of cleaved product is probably near 100%.

TABLE 1. Expression level and specific activity of marker fusion proteins.

		Specific Activity (U/mg × 10^{-6})	
Protein	mg/l[a]	With Marker	Without Marker[b]
Expressed in Yeast			
IL-2	1-3	240 ± 20	280 ± 40
G-CSF	10-20	63 ± 10	12 ± 1.0
IL-3	15-20	49 ± 19	42 ± 12
IL-4	7	5.4 ± 1.6	N.D.[c]
Expressed in *E. coli*			
GM-CSF[d]	15-20	150 ± 60	140 ± 60

[a]Determined by dot blot assay using 4E11 antibody to detect marker peptide containing material.

[b]Proteins in this column (except *E. coli* GM-CSF) were produced without the marker segment and purified by conventional means.

[c]Not determined.

[d]In this case, specific activities were determined on the same sample before and after enterokinase treatment to remove the marker segment.

DISCUSSION

The marker peptide fusion system described in this report comprises a unique and widely useful technique for protein identification and purification. In addition, two observations have shown that the marker peptide is compatible with heterologous expression systems: one, both the *E. coli* OmpA signal and the yeast pre-pro alpha factor leader sequences, when fused to the marker peptide, are correctly processed by their respective proteases allowing secretion of marker peptide-protein fusions with the correct N-terminus. And second, because many investigators have reported problems in the N-terminal processing of microbially expressed mammalian proteins[24–26], the ability of the marker peptide to protect the N-terminus of the desired product may be another important feature of this system.

The capability to produce authentic N-termini upon enterokinase treatment is an improvement over a number of other fusion protein approaches. Those that require chemical cleavages using Asp-Pro[12] or Asn-Gly[13] directed reagents, for example, must necessarily leave a proline or glycine at the N-terminus of the product. Although we have not yet tested the ability of enterokinase to cleave the marker sequence from N-termini containing all of the 20 possible amino acids, we have found that it is capable of cleaving products with N-terminal Glu, Ala, Thr, Leu and Ile residues. This suggests that this procedure will be useful for a wide range of N-termini including charged and uncharged, hydrophobic and hydrophilic residues.

There are several requirements that should be met by an efficient detection and purification system based upon fusion polypeptide expression: First, the added marker segment should not interfere with the native folding of proteins to which it is attached. Second, the marker peptide sequence should be intrinsically water soluble and should retain a high degree of exposure in the aqueous environment of the protein, so that it can readily interact with the affinity purification substrate. Third, it should be useful in an affinity purification step that requires only very mild media, and be elutable with a non-denaturing and inexpensive eluant.

Finally, the marker peptide should be easy to remove and the product protein should not have any amino acids added or deleted once the marker peptide has been removed. The Flag™ peptide fusion system was designed to possess all of these properties, and our data with several recombinant proteins suggest that it may prove to be a universal purification system for proteins expressed in heterologous organisms.

Several factors were considered in choosing the specific sequence of the marker peptide moiety. We chose to limit the marker peptide sequence to only eight amino acids because it can easily be encoded in a single synthetic oligonucleotide, and because the longest trypsinogen pro-sequences are of this length. We therefore could be reasonably sure that the trypsin-activating enzyme, enterokinase, would work efficiently to release the peptide. Additionally, because antibodies require up to six or seven amino acids for avid binding interactions, we reasoned that eight amino acids should be the minimum sequence capable of strong binding to an antibody while allowing one or more of the last amino acids on the C-terminal end to act as a spacer to separate the antibody binding portion from interference with the bulk of the protein. Finally, the five C-terminal amino acids of the marker sequence represent the minimal enterokinase specificity site, AspAspAspAspLys.

The choice of Lys at position three of the marker sequence causes the marker peptide to contain the hexapeptide sequence, LysAspAspAspAspLys, that has a maximum value on the hydrophilicity scale of Hopp and Woods[27]. Such maximally hydrophilic sequences have been proven to express strong antigenicity and are correspondingly likely to adopt a highly exposed conformation in the three dimensional folding of a protein[28]. As can be seen in Figure 5, it is impossible for any other region of a protein to have a higher hydrophilicity value than this maximally hydrophilic sequence, so the marker segment is virtually guaranteed to be exposed at the surface of any fusion protein. Therefore it can always be expected to be available for binding to antibody. Perhaps most importantly, the strong predilection for externalization should guarantee that the marker segment will not interfere in the adoption of a native conformation by the remainder of the protein.

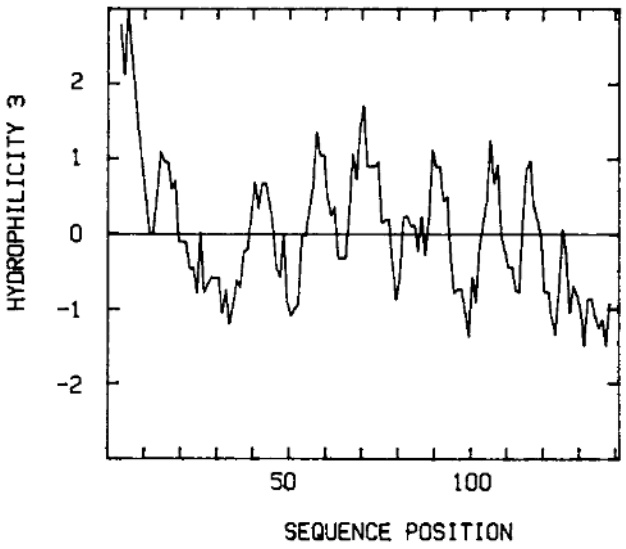

FIGURE 5. Hydrophilicity plot of the IL-2 fusion protein. The profile was generated using the updated HYDRO3 program of Hopp[27,28]. The scale is oriented so that hydrophilic is at top; hydrophobic is at bottom. Valley regions are expected to be buried portions of the polypeptide, whereas peaks are expected to be exposed at the surface of the protein. The prominent peak near the N-terminus results from the extremely hydrophilic hexapeptide, LysAsp-AspAspAspLys, contained within the marker peptide sequence.

In addition to the hydrophilic effects of Lys at position three, several other considerations influenced the choice of amino acids at the N-terminus of the marker peptide. Aromatic amino acids have been recognized as major factors in antigen-antibody interactions[29] so a tyrosine was placed at position 2, flanked by charged amino acids. Recent evidence suggests that aromatic residues that are flanked by charged sequences are more likely to be involved in

antigenic sites than are other aromatic residues in less polar environments[28]. The decision to place an Asp residue at the N-terminus was made in part because the negative charge on the Asp should aid in exposing the Tyr to antibody, as mentioned above, and in part because, with the inclusion of the Asp at position 1, a total of eight charges are to be found on the marker peptide moiety, including the N-terminal amino group. This preponderance of charged residues was expected to make it likely that antibody binding would be heavily dependent on charge-charge interactions, and therefore might be highly susceptible to elution with commonly used salt solutions such as 2 M $MgCl_2$ or 1 M NaCl[9]. In the end, the serendipitous discovery that Ca^{+2} was involved in the charge-dependent binding of the marker sequence to the 4E11 antibody made even these mild salt treatments unnecessary.

We have seen that fusion proteins retain the appropriate specific activity even with the marker segment still attached, and that this activity can be maintained after enterokinase treatment to remove the marker sequence. Comparisons of several of these fusion protein products with their natural counterparts (Table 1) demonstrated that the presence of the marker did not decrease the specific activity of the fusion proteins relative to the same proteins with no extraneous amino acids added. We have recently begun using a larger version of the 4E11 column to prepare proteins in milligram quantities. This level of scale-up required no special procedures or equipment, and can still be done as a bench top experiment. Further scale up for production of gram or kilogram quantities is contemplated, and will be limited only by availability of antibody and enterokinase. One advantage of the 4E11 antibody is that it can be purified on an affinity column comprised of chemically synthesized marker peptide attached to a solid support and eluted with EDTA.

One area that remains potentially problematic is the provision of adequate enterokinase for this process, both in terms of quantity and quality. We sampled commercial sources of enterokinase, but found that our fusion proteins were digested into small fragments, presumably by contaminating chymotrypsin, trypsin and elastase that are likely to be present in these partially purified preparations. Our own crude bovine intestinal preparations also caused substantial unwanted hydrolysis, until we used the protocol of Liepnieks and Light[30] for removing these other proteases. The fact that a minor amount of unwanted cleavage is seen in some cases when the marker peptide is removed (Fig. 4) suggests that traces of contaminating proteolytic activity may still be present in our purified preparations of enterokinase. An ideal solution would be to clone enterokinase and express it in a recombinant organism. This would provide starting material with much lower levels of contaminating proteases, and much higher levels of enterokinase. Furthermore, with the enterokinase gene in hand, it might be possible to engineer a smaller form of the molecule, lacking the hydrophobic portion that binds it to the membranes of the intestinal villi. This would simplify purification and decrease the mass of enterokinase needed for cleavage of the marker segment.

Despite the need for further development of enterokinase, this system for fusion protein detection and purification already represents a useful technique. It offers the possibility of using a single procedure for the purification of multiple fusion proteins. Although it is also possible to purify fusion proteins from yeast medium or *E. coli* cells by conventional means such as ion exchange or reverse phase chromatography, these procedures require new method development for each new protein, whereas with marker fusions, the same process is applicable to all proteins. Finally, because the marker segment does not appear to have decreased the biological activity of any of the proteins that it has been placed on, it may not always be necessary to remove the marker segment in order to obtain an active product. In such cases, this useful “handle” can be retained on the molecule, enabling investigators to readily detect and manipulate their recombinant protein products.

EXPERIMENTAL PROTOCOL

Yeast plasmid construction. The yeast vector used for protein expression has been described previously[31]. This vector contains sequences from pBR322 that allow selection (Amp^r) and replication in *E. coli,* as well as the yeast *Trp1* gene and 2μ origin of replication for selection and autonomous replication in a trpl yeast strain. Expression of foreign genes is under control of the α-factor promoter[3] or the ADH2 promoter[32] and secretion is directed by the α-factor leader peptide. To generate the IL-2 expression vector pIXY8, the mature coding region of IL-2[17] was fused in-frame to the marker peptide and the α-factor leader by means of a synthetic oligonucleotide linker encoding the five C-terminal amino acids of the α-factor leader and the eight amino acids of the marker peptide (Fig. 1). The vectors that directed the expression of the other products were generated by two modifications of pIXY8. First, the α-factor promoter was removed by restricting the plasmid with EcoRI and PstI, then inserting the ADH2 promoter using a synthetic oligonucleotide linker. Second, the marker and appropriate protein coding sequences were added in place of the IL-2 sequence (Fig. 1) and linked with a synthetic oligonucleotide that extended to the HpaI site.

Growth of yeast strains. *S. cerevisiae* strain XV218/(a/α-trp-1) was grown in either selective medium [YNB $^-$trp, consisting of 0.67% Yeast Nitrogen Base (Difco), 0.5% Casamino acids, 2% glucose, 10 μg/ml adenine and 20 μg/ml uracil] or rich medium (YPD, consisting of 1% Yeast Extract, 2% peptone and 1% glucose supplemented with 80 μg/ml adenine and 80 μg/ml uracil). Yeast transformations were done by selecting for Trp^+ transformants[33]. Cultures were grown for biological assay by inoculating 20-50 ml of rich medium with the appropriate strain and growing the cultures at 30°C to stationary phase. Cells were then removed by centrifugation and the medium was filtered through a 0.45μ cellulose acetate filter. Sterile supernates were stored at 4°C. Larger scale fermentations were done in a 10 liter New Brunswick Microferm fermentor. Cells were removed from the medium using a Millipore Pellicon filtration system.

Construction of the *E. coli* vector. Plasmid pIN-III-$OmpA_3$ is an *E. coli* expression vector regulated by the tandem lpp^p (lipoprotein promoter)/lac^{po} (lactose-promoter-operator) that contains a synthetic *ompA* leader sequence for protein secretion[34]. Construction of the GM-CSF expression vector was accomplished by restriction digestion at the unique BamHI site of pIN-III-$OmpA_3$, followed by its conversion to blunt ends by treatment with reverse transcriptase (Boehringer-Mannheim). The vector was subsequently restricted with HindIII, and used in a three-way ligation with a synthetic oligonucleotide encoding the marker peptide sequence and cDNA encoding GM-CSF to produce pRL6-6-87 as outlined in Figure 2.

Growth of *E. coli.* Plasmid pRL6-6-87 was introduced by transformation into *E. coli* strain JM107, *(Δlac, pro, thi, strA, endA, sbcB15, hspR4/F', traD36, proAB*$^+$, *lac*I^q-ZΔM15) which was grown at 37°C in M9 minimal medium containing 1 % (w/v) methionine assay medium (DIFCO) and ampicillin (50 μg/ml) to an OD_{600} of 0.25. Cultures were induced following addition of isopropyl-β-D-thiogalactopyranoside (IPTG) and cyclic 3'-5' adenosine monophosphate (cAMP) to 2mM and 4mM, respectively, and allowed to accumulate the GM-CSF fusion protein for 2-4 hr. Cells were harvested by centrifugation, and pellets either stored at −70°C or directly processed for extraction and purification of marker fusion proteins. *E. coli* pellets were extracted by the following procedure: the pellet from 500 ml of culture was suspended in 50 ml of 150 mM NaCl, 50 mM NaH_2P0_4, pH8.4, to which 1 mM phenylmethylsulfonylfluoride had been added immediately prior to mixing. After freezing (−70°C) and thawing three times to lyse the *E. coli* cells the sample was incubated at 37°C for 30 min to complete lysis and to extract the protein product. The viscous extract was treated by dounce homogenization to achieve a uniform solution, then centrifuged at 25,000 g for 45 min at 4°C. The supernatant was adjusted to 0.5 mM $CaCl_2$, recentrifuged if necessary to

remove any resulting precipitate, then applied to the affinity column.

Preparation of immunogens. Palmitic acid conjugated peptides were produced by solid phase chemical synthesis as described previously[35]. The antigenic marker peptide had the sequence AspTyrLysAspAspAspAspLysGlyProLysLysGly to which palmitoyl moieties had been attached on the epsilon amino groups of the two C-terminal lysines. It is referred to as CDP-marker (C-terminal dipalmitoyl marker peptide). A second palmitoyl peptide, NDP-GM1 (N-terminal dipalmitoyl GM-CSF peptide 1) was used as a non-specific binding control. It had the structure LysGlyGlyGluSerPheLysGluAsnLeuLysAspPheLeuValGly, and also possessed two palmitoyl moieties, in this case attached to the two amino groups of the N-terminal lysine residue. For purification of the IL-2 marker peptide fusion protein, supernatants of yeast expressing the IL-2 fusion protein were applied to a reverse phase HPLC column. The IL-2 fusion product was eluted from the column using a gradient of acetonitrile, as described previously[22].

Immunization. BALB/c female mice were purchased from the Jackson Laboratories (Bar Harbor, ME) and maintained in our animal facility. Mice were immunized subcutaneously with 250 μg of IL-2 fusion protein emulsified in Freund's complete adjuvant, followed with 125 μg of the same protein emulsified in Freund's incomplete adjuvant four weeks later. Two weeks after the second inoculation, a serum antibody titer to IL-2 and the marker peptide was measured by "dot-blot". The animal was then challenged with 10 μg of protein intravenously four days prior to fusion.

Hybridoma derivation. Four days after the intravenous boost, the animals were sacrificed, their spleens removed, and a single cell suspension prepared. The splenocytes were fused to the HAT sensitive myeloma cell, NS-1. The resulting hybridomas were then assayed for the production of antibodies to the marker peptide seven to ten days later by ELISA (see below). One hybridoma antibody consistently produced a positive reaction specifically with the marker peptide moiety. This cell line, designated 4E11, was then cloned by limiting dilution, isotyped, and further characterized.

ELISA. Various peptide solutions (CDP-marker, or NDP-GM1) were applied to HA plates (Millipore, Bedford, MA) at a concentration of 40 ng per well and allowed to incubate for 30 min at room temperature. Nonspecific protein binding sites were blocked by an incubation with 3% bovine serum albumin in Tris buffered saline, pH 7.0 (TBSA) for 1 hr at room temperature. Hybridoma supernatants were added and the plates incubated for 1 hr. Following this incubation, the plates were washed with PBS and an alkaline phosphatase labeled goat anti-mouse antibody (Sigma Chemical, St. Louis, MO) was added. Following a 1 hr incubation, the plates were washed several times with PBS and a colorimetric indicating reagent was added (substrate tablets, Sigma Chemical). Contents of each HA plate were then transferred to a polystyrene 96 well plate (Linbro/Titertek, Flow Laboratories, McLean, VA) and the absorbance at 405 nm determined on a Titerscan (Flow Laboratories).

Production and purification of 4E11 antibody. Pristane-primed BALB/c mice were injected IP with 1×10^6 hybridoma cells. Ten to twenty days later, the ascitic fluid was recovered, centrifuged at $1000 \times g$ for 30 minutes at 4°C, passed through cotton gauze, and the supernatant stored at −20°C until needed. The monoclonal antibody 4E11 was purified from ascites fluid using MAPS II Protein A Affigel (Bio-Rad, Richmond, CA) affinity chromatography. The purified antibody was found to be homogeneous by SDS-PAGE analysis.

4E11 column preparation. Purified 4E11 immunoglobulin was concentrated by ultrafiltration. After dialysis against 0.1M Hepes buffer, pH 7.5 at 4°C the antibody was coupled to Affigel-10 (Bio-Rad) in accordance with the manufacturer's instructions. A

typical antibody-coupled gel contained from 1.5 to 4.5 mg antibody/ml of gel. Columns of 4E11 coupled gel of 1.5 ml bed volume were prepared in polypropylene columns (Bio-Rad) and washed with 15 ml PBS, 15 ml 0.1M glycine HCl, pH 3.0, and stored at 4°C in PBS/0.02% sodium azide.

4E11 column chromatography. Yeast culture filtrates were brought to physiological levels of salt and pH by adding 10X PBS, and made 0.5 mM in $CaCl_2$ by adding 1 M $CaCl_2$, and then loaded onto the 1.5 ml column of 4E11 coupled Affigel 10 under gravity flow. *E. coli* extracts did not require any further additions because the extraction medium contained physiological levels of salt and pH as well as 0.5 mM $CaCl_2$. Up to 100 ml of filtrate were passed over the column, depending on the level of expression of the recombinant protein. After loading, the column was washed with three to five aliquots of 3 ml of PBS containing 0.5 mM $CaCl_2$-Elution was carried out with PBS lacking $CaCl_2$ and containing 2.0 mM Na_2 EDTA or with 0.1M glycine HCl pH 3.0. Each elution fraction was 1 ml. Yields of purified proteins were determined by amino acid analysis, and were typically 15-40% of the theoretical maximum assuming a 2:1 antigen to antibody binding ratio.

Enterokinase treatment. Enterokinase was purified from bovine intestine by the procedure of Liepnieks and Light[30]. Samples were also provided by A. Light of Purdue University. For enterokinase treatment, fusion proteins eluted from the antibody column were made 10 mM in Tris-HCl (pH 8) and adjusted to pH 8.0 by addition of 1N NaOH. For certain samples, the reaction mixture was made 40 mM in octyl-β-D-glucoside. Following the addition of an appropriate amount of bovine enterokinase (1-10% by weight; typically 0.2-2% by molarity), the reaction mixture was incubated for 16 hours at 37°C. Enterokinase dilutions were made from a 1 mg/ml stock solution of enzyme in 10 mM Tris-HCl, pH 8 kept at −70°C.

Bioassays. The activity of IL-2 was measured using the murine IL-2 dependent T-cell line CTLL-2[36]. The activity of GM-CSF was measured in a human bone marrow proliferation assay[37] and the activity of IL-3 was measured by FDC-P2 cell proliferation[38]. IL-4 and G-CSF were assayed as described[39,40]. Specific activities were derived by measuring the biological activities of purified samples of each protein, after quantifying by amino acid analysis.

ACKNOWLEDGMENTS

The authors wish to thank David Amberg, Janet Merriam, Aline Akelis, Norman Boiani, Bruce Hess, and David McMasters for excellent technical assistance. We would also like to thank Linda Troup for preparation of the manuscript.

REFERENCES

1. Itakura, K., Hirose, T., Crea, R., Riggs, A. D., Heyneker, H. L., Bolivar, F., and Boyer, H. W. 1977. Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. Science 198:1056-1063.
2. Goeddel, D. V., Kleid, D. G., Bolivar, F., Heyneker, H. L., Yansura, D. G., Crea, R., Hirose, T., Kraszewski, A., Itakura, K., and Riggs, A. D. 1979. Expression in *Escherichia coli* of chemically synthesized genes for human insulin. Proc. Natl. Acad. Sci. USA 76:106-110.
3. Brake, A. J., Merryweather, J. P., Coit, D. G., Heberlein, U. A., Masiarz, F. R., Mullenback, G. T., Urdea, M. S., Valenzuela, P., and Barr, P.J. 1984. α-Factor-directed synthesis and secretion of foreign proteins in *Saccharomyces cerevisiae.* Proc. Natl. Acad. Sci. USA 81:4642-4646.
4. Villa-Komaroff, L., Efstratiadis, A., Broome, S., Lomedico, P., Tizard, R., Naber, S. P., Chick, W. L., and Gilbert, W. 1978. A bacterial clone synthesizing proinsulin. Proc. Natl. Acad. Sci. USA 75:3727-3731.
5. Germino, J. and Bastia, D. 1984. Rapid purification of a cloned gene product by genetic fusion and site-specific proteolysis. Proc. Natl. Acad. Sci. USA 81:4692-4696.
6. Abrahmsen, L., Moks, T., Nilsson, B., and Uhlen, M. 1986. Secretion of heterologous gene products to the culture medium of *Escherichia coli.* Nucleic Acids Res. 14:7487-7500.
7. Munro, S. and Pelham, H. R. B. 1984. Use of peptide tagging to detect proteins expressed from cloned genes. EMBO J. 3:3087-3093.

8. Ikehara, M., Ohtsuka, E., Tokunaga, T., Nishikawa, S., Uesugi, S., Tanaka, T., Aoyama, Y., Kikyodani, S., Fujimoto, K., Yanase, K., Fuchimura, K., and Morioka, H. 1986. Inquiries into the structure-function relationship of ribonuclease T1 using chemically synthesized coding sequences. Biochemistry 83:4695-4699.
9. Wilchek, M., Miron, T., and Kohn, J. 1984. Affinity chromatography. Methods Enzymol. 104:3-55.
10. Nagai, K. and Thogerson, H. C. 1987. Synthesis and sequence specific proteolysis of hybrid proteins produced in *Escherichia coli.* Methods Enzymol. 153:461-481.
11. Schoner, R. G., Ellis, L. F., and Schoner, B. E. 1985. Isolation and purification of protein granules from *Escherichia coli* cells overproducing bovine growth hormone. Bio/Technology 3:151-154.
12. Sung, W. L., Yao, F. L., Zahab, D. M., and Narang, S. A. 1986. Short synthetic oligodeoxyribonucleotide leader sequences enhance accumulation of human proinsulin synthesized in *Escherichia coli.* Proc. Natl. Acad. Sci. USA 83:561-565.
13. Nilsson, B., Holmgren, E., Josephson, S., Gatenbeck, S., Philipson, L., and Uhlen, M. 1985. Efficient secretion and purification of human insulin-like growth factor I with a gene fusion vector in *Staphylococci.* Nucleic Acids Res. 13:1151-1162.
14. Moks, T., Abrahmsen, L., Osterlof, B., Josephson, S., Ostling, M., Enfors, S., Persson, I., Nilsson, B., and Uhlen, M. 1987. Large-scale affinity purification of human insulin-like growth factor I from culture medium of *Escherichia coli.* Bio/Technology 5:379-382.
15. Sassenfeld, H. M. and Brewer, S.J. 1984. A polypeptide fusion designed for the purification of recombinant proteins. Bio/Technology 2:76-81.
16. Maroux, S., Baratti, J., and Desnuelle, P. 1971. Purification and specificity of porcine enterokinase. J. Biol. Chem. 246:5031—5039.
17. Taniguchi, T., Matsui, H., Fujita, T., Takaoka, C., Kashima, N., Yoshimoto, R., and Hamuro, J. 1983. Structure and expression of a cloned cDNA for human interleukin-2. Nature 302:305-310.
18. Nagata, S., Tsuchiya, M., Asano, S., Kaziro, Y., Yamazaki, T., Yamamoto, O., Hirata, Y., Kubota, N., Oheda, M., Nomura, H., and Ono, M. 1986. Molecular cloning and expression of cDNA for human granulocyte colony stimulating factor. Nature 319:415-417.
19. Yang, Y-C., Ciarletta, A. B., Temple, P. A., Chung, M. P., Kovacic, S., Witek-Giannotti, J. S., Leary, A. C., Kriz, R., Donahue, R. E., Wong, G. G., and Clark, S. C. 1986. Human IL-3 (multi-CSF): Identification by expression cloning of a novel hematopoietic growth factor related to murine IL-3. Cell 47:3-10.
20. Yokota, T., Otsuka, T., Mosmann, T., Banchereau, J., DeFrance, T., Blanchard, D., DeVries, J. E., Lee, F., and Arai, K-I. 1986. Isolation and characterization of a human cDNA clone, homologous to mouse B-cell stimulatory factor 1, that expresses B-cell and T-cell stimulating activities. Proc. Natl. Acad. Sci. USA 83:5894-5898.
21. Cantrell, M. A., Anderson, D., Cerretti, D. P., Price, V., McKereghan, K., Tushinski, R. J., Mochizuki, D. Y., Larsen, A., Grabstein, K., Gillis, S., and Cosman, D. 1985. Cloning, sequence, and expression of a human granulocyte/macrophage colony stimulating factor. Proc. Natl. Acad. Sci. USA 82:6250-6254.
22. Urdal, D. L., Mochizuki, D., Conlon, P. J., March, C. J., Remerowski, M. L., Eisenman, J., Ramthun, C, and Gillis, S. 1984. Lymphokine purification by reversed-phase high performance liquid chromatography. J. Chromatography 296:171-179.
23. Maurer, P. H., Clark, L. G., and Liberti, P. A. 1970. Antigenicity of polypeptides (polyαamino acids): Calcium-dependent and independent antibodies. J. Immunol. 105:567-573.
24. Hitzeman, R. A., Leung, D. W., Perry, L. J., Kohr, W.J., Levine, H. L., and Goeddel, D. V. 1983. Secretion of human interferons by yeast. Science 219:620-625.
25. Aston, R., Cooper, L., Holder, A., Ivanyi, J., and Preece, M. 1985. Monoclonal antibodies to human growth hormone can distinguish between pituitary and genetically engineered forms. Molec. Immunol. 22:271-275.
26. Kronheim, S. R., Cantrell, M. A., Deeley, M. C., March, C. J., Glackin, P. J., Anderson, D. M., Hemenway, T., Merriam, J. E., Cosman, D., and Hopp, T. P. 1986. Purification and characterization of human interleukin-1 expressed in *Escherichia coli.* Bio/Technology 4:1078-1082.
27. Hopp, T. P. and Woods, K. R. 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA 78:3824-3828.
28. Hopp, T. P. 1986. Protein surface analysis: Methods for identifying antigenic determinants and other interaction sites. J. Immunol. Methods 88:1-18.
29. Sela, M. and Arnon, R. 1960. Studies on the chemical basis of the antigenicity of proteins. Biochem. J. 75:91-102.
30. Liepnieks, J. J. and Light, A. 1979. The preparation and properties of bovine enterokinase. J. Biol. Chem. 254:1677-1683.
31. Price, V., Mochizuki, D., March, C.J., Cosman, D., Deeley, M. C., Klinke, R., Clevenger, W., Gillis, S., Baker, P., and Urdal, D. 1987. Expression, purification and characterization of recombinant murine granulocyte-macrophage colony-stimulating factor and bovine interleukin-2 from yeast. Gene 55:287-293.
32. Beier, D. R. and Young, E. T. 1982. Characterization of a regulatory region upstream of the ADR2 locus of *S.cerevisiae.* Nature 300:724-728.
33. Hinnen, A., Hicks, J. B., and Fink, G. R. 1978. Transformation of yeast. Proc. Natl. Acad. Sci. USA 75:1929-1933.
34. Libby, R. T., Braedt, G., Kronheim, S. R., March, C. J., Urdal, D. L., Chiaverotti, T. A., Tushinski, R. J., Mochizuki, D. Y., Hopp, T. P., and Cosman, D. 1987. Expression and purification of native human granulocyte-macrophage colony stimulating factor from an *Escherichia coli* vector. DNA 6:221-229.
35. Hopp, T. P. 1984. Immunogenicity of a synthetic HBsAg peptide: Enhancement by conjugation to a fatty acid carrier. Molec. Immunol. 21:13-16.
36. Gillis, S., Ferm, M. M., Ou, W., and Smith, K. 1978. T cell growth factor: parameters of production and a quantitative microassay for activity. J. Immunol. 120:2027-2033.
37. Mochizuki, D. Y., Eisenman, J., Conlon, P. J., Park, L. S., and Urdal, D. L. 1986. Development and characterization of antiserum to murine granulocyte-macrophage colony stimulating factor. J. Immunol. 136:3760-3709.
38. Prestidge, R. L., Watson, J. D., Urdal, D. L., Mochizuki, D., Conlon, P., and Gillis, S. 1984. Biochemical comparison of murine colony stimulating factors secreted by a T cell lymphoma and a myelomonocytic leukemia. J. Immunol. 133:293-298.
39. Grabstein, K., Eisenman, J., Mochizuki, D., Shanebeck, K., Conlon, P., Hopp, T., March, C., and Gillis, S. 1986. Purification to homogeneity of B-cell stimulating factor: A molecule that stimulates proliferation of multiple lymphokine-dependent cell lines. J. Exp. Med. 163:1405-1414.
40. Metcalf, D. 1986. The molecular biology and functions of granulocyte-macrophage colony stimulating factors. Blood 67:257-267.