BioTechniques Vol. 7, No. 6 (1989) 580-589.

A Calcium-Dependent Antibody for Identification and Purification of Recombinant Proteins

Kathryn S. Prickett, David C. Amberg and Thomas P. Hopp

Immunex

ABSTRACT

We report a straightforward methodology for purification of recombinant proteins by incorporating a short hydrophilic peptide marker segment at their N-termini. A calcium-dependent antibody that reacts primarily with the first three amino acids of this peptide segment was used to affinity purify the fusion proteins in a single chromatographic step. The marker peptide could subsequently be removed by proteolysis with the enzyme enterokinase.

INTRODUCTION

Modern molecular biology has made it possible to clone and express proteins in a variety of host organisms. However, purification of recombinant proteins is an uncertain and often tedious task, usually requiring multiple steps, each of which is dependent upon the specific characteristics of the protein. In some cases harsh treatments are required during the purification, so that complicated renaturation procedures are necessary before biological activity can be restored. The yield of recombinant proteins can sometimes be increased by expressing them as part of a larger fusion protein. Such products have been made cytoplasmically (6,12) or secreted from yeast (4) or *E. coli* (19) cells. Although fusion *per se* seems to increase yield in many cases, it does nothing inherently to improve purification yield. A more sophisticated approach is to fuse the desired protein to a polypeptide sequence that is useful for identifying and purifying the product (1,5). In the past, the latter approach has had several drawbacks. First, the fusion protein products may fail to fold properly into a native, active state (15), and, second, it is often difficult or impossible to remove the additional amino or carboxyl-terminal sequence from the desired protein product.

To circumvent some of these problems, we have designed a fusion sequence that allows rapid purification of recombinant proteins under very mild conditions. The fusion sequence is a short hydrophilic peptide, AspTyrLysAspAspAspAspLys, that can be engineered onto the N-terminus of a protein using short oligonucleotides. This marker sequence, or Flag™ segment, can be identified by a monoclonal antibody that can also be used as an immuno-affinity gel reagent to purify the resulting fusion protein. The antibody reacts with specific amino acids of the marker sequence and is dependent on the presence of Ca^{++} for binding. This allows a mild, one-step purification of fusion proteins when the affinity gel is eluted by removing or chelating calcium from the washing buffer. The peptide can be subsequently detached from the protein by treatment with the protease enterokinase, which cleaves immediately after the -AspAspAspAspLys- sequence.

The antibody and affinity chromatography gel described here, as well as enterokinase, are under development by Immunex as the "Flag" system for recombinant protein identification and purification.

MATERIALS AND METHODS

Yeast Secretion of Interleukin 3 (IL-3) and Interleukin 4 (IL-4)

The vector used for protein expression in *Saccharomyces cerevisiae* has been described previously (10,17). Expression of recombinant proteins was controlled by the ADH2 promoter (3) and secretion was directed by the α-factor leader peptide (4). The α-factor pre-pro peptide segment, including the initiator Met and secretory signal sequence, was placed at the 5' end of the coding sequence, and the IL-3 coding cDNA was ligated at the 3' end of the coding sequence. The marker peptide segment was encoded by a pair of short oligonucleotides that allowed precise placement of the marker after the Lys-Arg proteolytic processing segment of α-factor, and immediately before the AlaProMet N-terminal protein coding sequence of IL-3 (20). The same process was used to generate the IL-4 marker fusion protein. Similar constructs have been used previously to secrete interleukin 2 (IL-2), granulocyte-macrophage colony stimulating factor (GM-CSF), colony stimulating factor-1 (CSF-1) and granulocyte colony stimulating factor (G-CSF) marker fusion proteins from yeast (10).

Protein production yeast cultures were grown in a rich medium consisting of 1% yeast extract, 2% peptone and 1% glucose supplemented with 80 μg/ml adenine and 80 μg/ml uracil. Cultures were grown at 30° C to stationary phase, then cells were removed by centrifugation and the medium was filtered through a 0.45 μ cellulose acetate filter. Sterile supernatants were stored at 4° C.

4E11 Hybridoma Antibody

Production and purification of the monoclonal antibody has been described previously (10). The 4E11 antibody is a monoclonal of isotype IgG_{2B}. It reacts with proteins bearing the marker peptide sequence in a variety of procedures, including ELISAs, dot blots, Western blots, immunoprecipitation and affinity chromatography. It reacts with all of the marker peptide fusion proteins that we have produced to date, and exhibits no reactivity with any component present in *E. coli* extracts, nor in yeast culture medium.

4E11 Column Preparation

Purified 4E11 immunoglobulin was concentrated by ultrafiltration. After dialysis against 0.1 M Hepes buffer, pH 7.5, at 4° C, the antibody was coupled to Affigel-10 (Bio-Rad, Richmond, CA) 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 of phosphate buffered saline (PBS), 15 ml 0.1 M glycine HCl, pH 3.0, and stored at 4° C in PBS containing 0.02% sodium azide.

4E11 Column Chromatography

Yeast culture filtrates were brought to physiological levels of salt and pH by adding 10-fold concentrated PBS, and made 0.5 or 1.0 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. 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 3-ml aliquots of PBS containing 1.0 mM $CaCl_2$. Elution was carried out with PBS lacking $CaCl_2$ and containing 2.0 mM Na_2 EDTA or with 0.1 M glycine HC1 pH 3.0 (these fractions were subsequently neutralized with 1 M Tris hydroxide). 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.

For the experiment to test the effect of other divalent cations, the following were prepared as 1 M or 0.5 M solutions in water and then substituted for $CaCl_2$ in the above procedure: $MgCl_2$, $ZnCl_2$, $CuCl_2$, $FeCl_2$, and $MnCl_2$. Equivalent aliquots of IL-4 yeast supernatants were used in each experiment.

Dot Blot

A semi-quantitative assay for the presence of the IL-4 fusion protein was performed using a dot blot procedure. One µl of each fraction was applied directly to a sheet of nitrocellulose (Schleicher and Schuell, Keene, NH) and allowed to dry. The sheet was then blocked for 30 min at room temperature with 3% bovine serum albumin (BSA) in 50 mM Tris HC1, 150 mM NaCl, pH 7.5 (TBS) containing 0.3 mM $CaCl_2$. After washing with TBS, the nitrocellulose sheet was incubated with 4E11 antibody in binding media containing 0.3 mM $CaCl_2$. After several washes, goat anti-mouse horseradish peroxidase (Bio-Rad) diluted 1:2000 into 3% BSA/TBS with 0.3 mM $CaCl_2$ was added and mixed for 1 h. The nitrocellulose sheet was again washed, the substrate (4-chloro-l-naphthol) added, and color development observed.

Western Blot

Immunoblot analysis of IL-3 marker fusion protein before and after cleavage with enterokinase was carried out in the following manner. Each sample was heated in Laemmli sample-reducing buffer for 5 min and analyzed by SDS-PAGE in a 10-20% gradient gel (13). The electrophoresed proteins on the gel were then electrophoretically transferred to a sheet of nitrocellulose (18). After transfer, the nitrocellulose was treated overnight with 1% BSA in PBS (PBSA) to block nonspecific binding sites. The sheet was then cut into two panels, one of which was incubated with a 1:200 dilution in PBSA of rabbit IL-3 antiserum, and the other with a 10 µg/ml solution of the 4E11 monoclonal antibody in PBSA, each for 1 h. The nitrocellulose panels were then washed with PBS and incubated for 1 h with their respective secondary antibody enzyme conjugate, either horseradish peroxidase goat anti-rabbit IgG (Bio-Rad) or horseradish peroxidase goat anti-mouse IgG (Bio-Rad), diluted 1:500 in PBSA. The panels were again washed with PBS and then incubated with substrate (4-chloro-l-naphthol) for 15 min to detect the presence of antibody.

Enterokinase Treatment

Enterokinase was provided by A. Light of Purdue University. For enterokinase treatment, the IL-3 fusion protein eluted from the antibody column was made 10 mM in Tris HC1 (pH 8). Bovine enterokinase, prepared as a stock solution at 100 µg/ml in 10 mM Tris HC1, pH 8, was added (10% by weight) and the reaction mixture incubated for 15 h at 37° C.

Peptide Synthesis

All peptides were synthesized either manually or on an Applied Biosystems 430A Peptide Synthesizer (Foster City, CA) using the Merrifield solid phase method (2) with N-α-tert-butyloxycarbonyl protected amino acids, standard side chain protection and either p-methylbenzhydrylamine or phenyl-acetamidomethyl (PAM) resins. After cleavage from the resin

with hydrofluoric acid, the free peptide was purified by HPLC on a C_{18} Vydac column (1 x 25 cm) equilibrated in 0.1% aqueous trifluoroacetic acid (v/v) and eluted with a gradient of acetonitrile containing 0.1% trifluoroacetic acid. The composition of the eluted peptide was confirmed by amino acid analysis.

Peptide Conjugates

Palmitic acid conjugated peptides were prepared as described previously (7) and had the sequence AspTyrLysAspAspAspAspLysGlyProLysLysGly to which palmitoyl moieties had been attached on the epsilon amino groups of the two C-terminal lysines (referred to as CDP-conjugate). A second synthetic peptide, AspTyrLysAspAspAspAspLysGlyCysGly, was coupled through its C-terminal sulfhydryl group to ovalbumin which had been dialyzed in 0.1 M Na/K phosphate buffer, pH 7.0, and activated with m-maleimido-benzoyl-N-hydroxysuccinimide ester (14). Conjugation proceeded for 3 h, then excess reactants were removed from the peptide-protein conjugate by extensive dialysis. The coupling efficiency, measured by amino acid analysis, was approximately five peptides per molecule of ovalbumin. This product is referred to as OVA-conjugate.

ELISA Inhibition Assay

Linbro/Titertek 96-well polystyrene plates (Flow Laboratories, McLean, VA) were coated with antigen, 100 µl of either CDP-conjugate or OVA-conjugate at 5 µg/ml in PBS, and allowed to incubate overnight at 37° C (antigen plates). A second set of plates (preincubation plates) were blocked with 200 µl/well of 1% bovine serum albumin in PBS (PBSA) containing 0.5 mM $CaCl_2$ for 1 h at room temperature and then were incubated with serial dilutions of peptide in PBSA/0.5 mM $CaCl_2$ in the presence of 4E11 at 300 ng/ml on a shaker at 4° C overnight. On the following day all plates were allowed to come to room temperature, the antigen plates were washed 6 times with PBS/0.5 mM $CaCl_2$, blocked for 30 min with PBSA/0.5 mM $CaCl_2$ and then incubated for 1 h at room temperature with the previously preincubated peptide/antibody solutions from the preincubation plates. The antigen plates were then washed with PBS, and an alkaline phosphatase-labeled goat anti-mouse antibody (Zymed, San Francisco, CA) was added. Following a 1-h incubation, the plates were washed several times with PBS/0.5 mM $CaCl_2$ and a colorimetric-indicating reagent (p-nitrophenyl phosphate) added. The plates were allowed to develop for 30 min at room temperature and the absorbance at 405 nm read on a multiscan ELISA plate reader (Titertek, Flow Laboratories, McLean, VA).

RESULTS

Secretion of Fusion Proteins from Yeast

Saccharomyces cerevisiae cells transformed with either the IL-3 or IL-4 plasmid were used to express the marker fusion proteins. Similar results were obtained in the expression of IL-3 and IL-4 (described here) as had been seen with the other lymphokines described previously (10). The IL-3 and IL-4 fusion proteins were secreted into culture media at levels of 15-20 mg/1 and 7 mg/1, respectively. The leader peptide had been removed during secretion so that the amino terminus of the marker segment was reactive with the 4E11 antibody on dot blots, which were used to monitor expression levels. Both proteins were apparently glycosylated during secretion (see below).

Purification

Figure 1 shows the result of purification of yeast-secreted IL-3 on an affinity column made with the 4E11 antibody. In Figure 1A, a culture supernatant obtained by fermenting yeast bearing the IL-3 expression vector was passed over the column. Medium components were removed by washing with PBS containing 1.0 mM $CaCl_2$, while the IL-3 fusion protein remained bound to the antibody. Subsequent elution with PBS containing EDTA dissociated the marker peptide-antibody complex and released the IL-3 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. The two bands running just below the 67 kD marker are artifacts of the electrophoresis, and are not derived from the samples (note that they occur in the standard lanes). All other bands were identified as the IL-3 fusion protein, based on Western analyses using 4E11 as well as anti IL-3 antibody as developing reagents. Furthermore, amino acid analysis confirmed that the product had the expected composition for the IL-3 marker fusion protein.

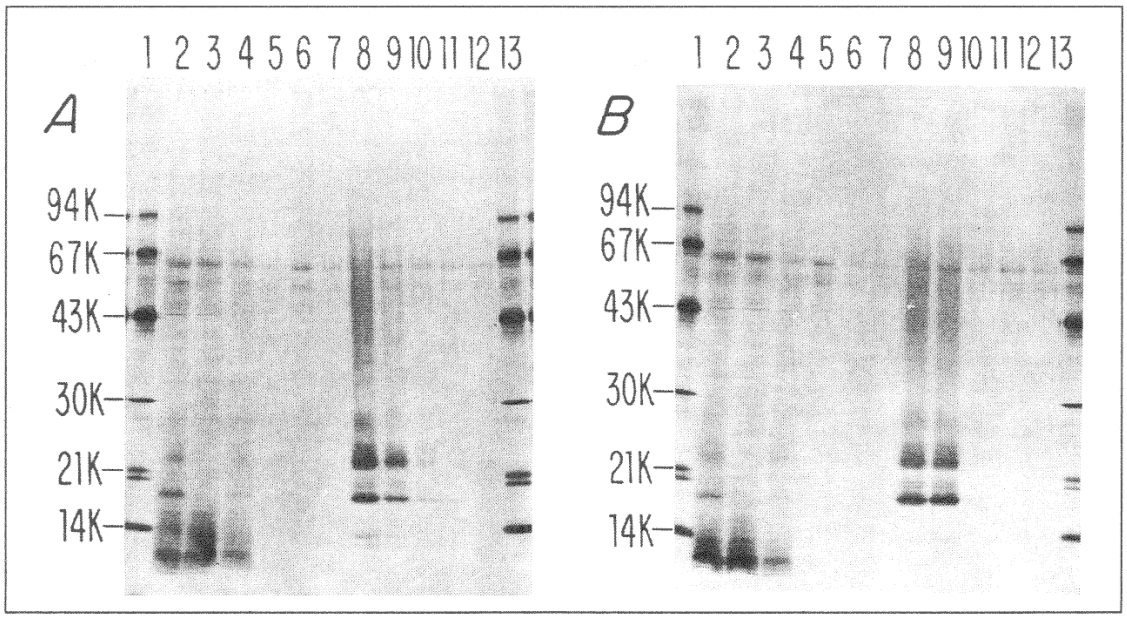

Figure 1. Purification of recombinant IL-3 by affinity chromatography. Panel A: Silver-stained polyacrylamide gel utilizing EDTA elution. Lanes are: 1: molecular weight standards; 2: yeast supernatant; 3: flow through material; 4-6: sequential PBS/1.0 mM $CaCl_2$ washes; 7-12: sequential PBS/2.0 mM EDTA elutions; 13: molecular weight standards. **Panel B:** Silver-stained polyacrylamide gel utilizing Gly-HCl elution. Lanes are: 1: molecular weight standards; 2: yeast supernatant; 3: flow through material; 4-6: sequential PBS/1.0 mM $CaCl_2$ washes; 7-12: 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 two bands that occur throughout the gels just below the 67 K marker are artifacts of the electrophoresis, and are not present in the chromatography samples.

Figure 1B shows the results of an affinity purification carried out on another aliquot of the same yeast culture supernatant in which chromatography was carried out as above, except that the stronger eluting buffer, 0.1 M glycine HC1 pH 3.0, was used to elute the product. This product was identical to the protein eluted with EDTA but came off the column in fewer fractions.

Enterokinase Treatment

Figure 2 shows the results of enterokinase treatment of the IL-3 fusion protein. The sample was incubated for 15 h at 37° C. At this time point, the conversion of fusion protein to authentic protein was complete, and approximately identical amounts of authentic product had been formed.

There are no detectable cleavage products below the molecular weight of IL-3. This is consistent with enterokinase cleaving only at its specificity site between the marker and IL-3, and not at any point within the IL-3 sequence.

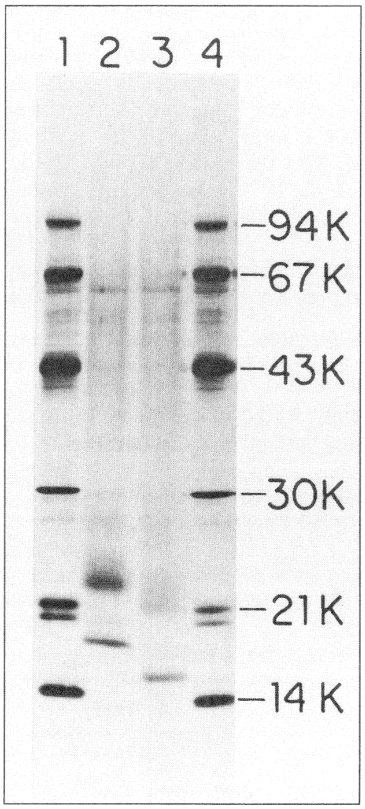

Figure 2. Removal of marker segment by enterokinase treatment. Silver-stained polyacrylamide gel of enterokinase digestion of the IL-3 fusion protein. Lanes are: 1: molecular weight standards; 2: IL-3 fusion protein, no enzyme; 3: fusion protein plus enterokinase; 4: molecular weight standards. Numbers at right indicate M_r values for the standard proteins (in kilodaltons).

Western blots were used to analyze the products of enterokinase cleavage, and Figure 3 shows the results obtained for IL-3. Before enzyme treatment, the multiple IL-3 products all reacted with both 4E11 and anti-IL-3 polyclonal sera. After treatment, only anti-IL-3 sera could react with the blots. This experiment shows that the marker segment can be removed completely from each of the different molecular weight species, and also is consistent with the concept that the size heterogeneity is entirely due to addition of different sized carbohydrate moieties to the same IL-3 marker fusion protein.

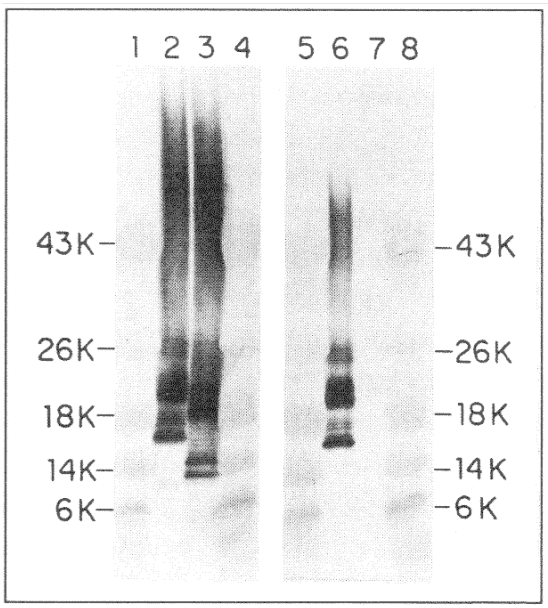

Figure 3. Western blot analysis of IL-3 marker fusion protein before and after enterokinase treatment. Lanes are: 1, 4, 5, 8: molecular weight standards; 2, 6: IL-3 fusion protein; 3, 7: IL-3 fusion protein plus enterokinase. Lanes 1-4 were reacted with IL-3 antiserum and lanes 5-8 with 4E11 antibody. Numbers on the right and left indicate M_r values for the standard proteins (in kilodaltons).

Inhibition of 4E11 Antibody by Peptide Analogs

Figure 4 illustrates the peptide analogs of the marker sequence that were synthesized. Each analog was preincubated with 4E11 antibody and then the solution was reacted in the ELISA with either CDP-conjugate or OVA-conjugate that had been coated onto the polystyrene plate (0.5 mM $CaCl_2$ was maintained throughout the assay). A graph of the inhibition obtained using several of these peptides is shown in Figure 5, and Table 1 lists the concentration of each peptide necessary to give 50% inhibition of 4E11 binding with each of the conjugates. Although different absolute levels of inhibitor were required depending on which carrier was used, the relative effects of the amino acid substitutions were closely similar for both carriers. This rules out the possibility that undesired interactions with the carrier might be influencing the results. Several peptides, Ac, S1, Al, A2, and A3, required concentrations greater than 200 µg/ml to give 50% inhibition, whereas the remainder required concentrations of 10 µg/ml or less. Substitution of Ala for either the Asp, Tyr, or Lys at the amino end of the marker sequence resulted in greatly reduced competition for 4E11 antibody, indicating that these amino acids are most important for antibody-peptide binding.

H_2N-ASP-TYR-LYS-ASP-ASP-ASP-ASP-LYS-$CONH_2$	Flag™ Peptide
$CH_3\overset{O}{\overset{\|}{C}}NH$- - - - - - - - -	Ac Peptide
-SER- - - - - - - -	S1 Peptide
-ASN- - - - - - - -	N1 Peptide
-ALA- - - - - - - -	A1 Peptide
- -ALA- - - - - - -	A2 Peptide
- - -ALA- - - - - -	A3 Peptide
- - - -ALA- - - - -	A4 Peptide
- - - - -ALA- - - -	A5 Peptide
- - - - - -ALA- - -	A6 Peptide
- - - - - - -ALA- -	A7 Peptide
- - - - - - - -ALA-	A8 Peptide
- - - - - - - - -COOH	OH Peptide

Figure 4. Chemically synthesized analogs of the marker peptide. Each of the analogs were prepared with the single indicated change to the Flag peptide sequence. The control peptide was Gly-Pro-Ile-Arg-Asn-Gly-Ser-Glu-Val-Arg-Asp-Pro-Gly. Peptides possessed a free N-terminal amino group and an amidated C-terminal carboxyl unless otherwise indicated.

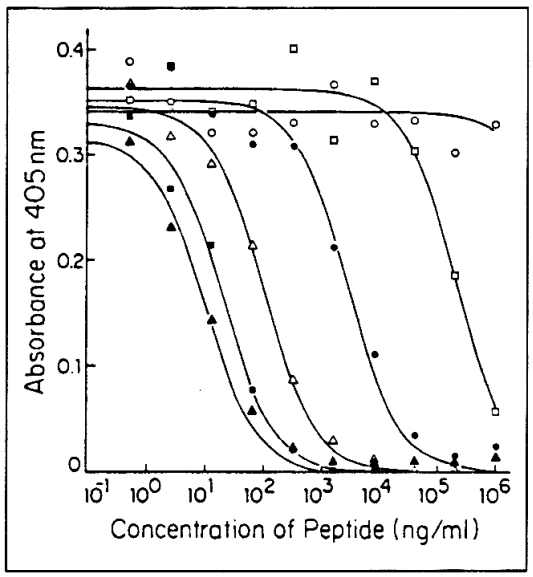

Figure 5. Plot of selected peptide inhibition experiments. The curves represent the ability of several peptides preincubated with 4E11 antibody to inhibit the binding of the antibody to OVA-conjugate coated on polystyrene plates. The best-fit curves were calculated from an equilibrium binding equation. Symbols represent: (■), Flag peptide; (Δ), OH peptide; (❑), Ac peptide; (●), A4 peptide; (▲), A5 peptide; (❍), control peptide.

Table 1. Ability of Peptide Analogs to Inhibit Binding of the 4E11 Antibody to the Flag Sequence

	50% Inhibition (μg/ml) ELISA Coating Antigen	
Peptide	CDP-Conjugate	OVA Conjugate
Flag	0.070	0.021
Ac	566.419	207.593
S1	2004.597	481.860
N1	9.382	4.179
A1	4050.252	543.538
A2	1874.982	729.742
A3	778.780	260.520
A4	10.059	2.900
A5	0.019	0.010
A6	0.102	0.019
A7	7.793	0.858
A8	0.031	0.009
OH	0.606	0.104
Control	$>1 \times 10^6$	$>1 \times 10^6$

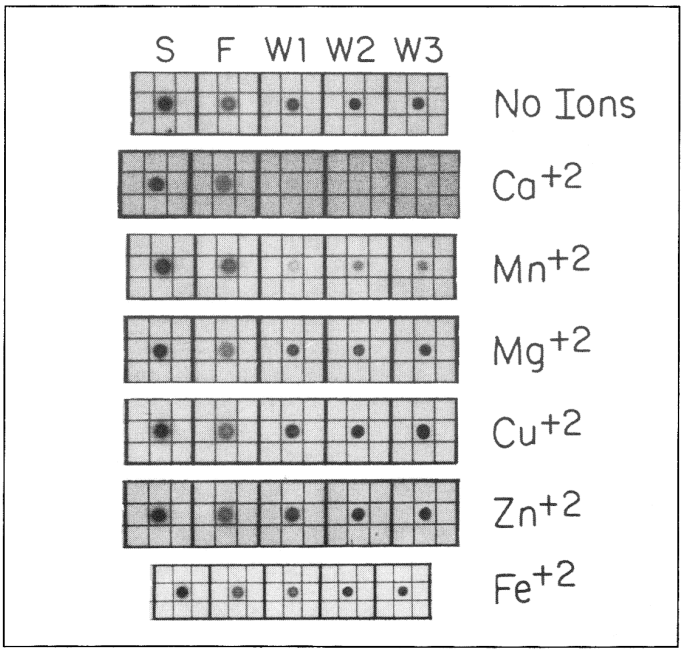

Figure 6. Metal specificity chromatograms. Dot blots of IL-4 affinity column wash fractions were developed with 4E11 antibody. Fractions are indicated as S: yeast supernatant; F: flow through; W1-W3: sequential washes of PBS containing the indicated cations at 0.5 mM or PBS only. These correspond to fractions identical to those seen in lanes 2 through 6 of the IL-3 affinity chromatogram in Figure 1.

Washing the Column with Other Cations

When the yeast-expressed IL-4 marker fusion protein was applied to a 4E11 antibody column and subsequently eluted under the same conditions as for the IL-3 marker fusion protein, essentially identical behavior was seen. Based upon this behavior, it was possible to develop a test

for the binding of metals other than calcium, by comparing their ability to retain the IL-4 marker fusion protein on the affinity column during the PBS washing phase. Accordingly, the IL-4 marker fusion protein was loaded onto a series of affinity columns, then washed with 0.5 mM solutions of either $CaCl_2$, $MgCl_2$, $ZnCl_2$, $CuCl_2$, $FeCl_2$, or $MnCl_2$ in PBS, or PBS with no added divalent cation. Figure 6 shows a dot blot comparison of the applied supernatant, flow-through and washes of each experiment. With Ca^{++} present in the wash buffer, no leakage of the fusion protein from the column occurred during washing, whereas all of the other metals allowed loss of the protein from the column comparable to PBS alone. Of these, only Mn^{++} had a detectable effect, causing a slight reduction in the amount of leakage from the column.

DISCUSSION

The Flag System

The marker peptide fusion system described in this report comprises a unique and widely useful technique for protein identification and purification. The purification and proteolytic processing of IL-3 are typical of results obtained using this procedure. The affinity chromatograms shown in Figure 1 demonstrate several features of the process. First, the yeast culture medium is a complex mixture that includes peptides, proteins and byproducts of yeast metabolism, as can be seen in lane 2 of the gel. Essentially all of this unwanted material passes through the column without sticking (lane 3). After washing, the eluate is free of proteins (lanes 4-6). Finally, when conditions are changed to favor release of the product from the antibody (lanes 7-12), the IL-3 is eluted in pure form. In this case the product exhibits multiple molecular weight species as expected for a glycoprotein expressed by yeast. Western blots with 4E11 or anti-IL-3 polyclonal sera yielded an identical pattern, demonstrating that all products are indeed IL-3, and that no extraneous proteins are present.

Second, purification of the higher molecular weight forms of the IL-3 glycoprotein, along with the smaller forms, confirms that the affinity purification column is capable of binding large, as well as small proteins. As can be seen in Figure 1, yeast cells may attach a wide size range of carbohydrate moieties to the protein, leading to molecular weight forms of IL-3 that approach 100,000 daltons, in contrast to the normal molecular weight of 16,000 daltons. The observation that all forms up to 100,000 daltons can be bound and eluted from the column demonstrates that steric factors are favorable both in allowing the antibody to bind large fusion protein products, and in terms of the suitability of Affigel-10 for use as a solid support. More experimentation will be required, however, in order to determine the maximum size permissible in fusion proteins to be purified by this technology.

It appears that enterokinase also works satisfactorily on larger versions of glycosylated marker fusion proteins. This can be inferred because several of the multiple molecular weight forms of IL-3 are seen to simultaneously decrease by the expected 1,000 molecular weight during enterokinase treatment (Figure 2). It is not possible to see a change in the molecular weight of the highly glycosylated forms (40-100 kD) because they are a continuum of increasing molecular weight species, and a change of 1,000 would cause their mobility on SDS gels to increase by only an insignificant amount. Further proof that enterokinase can cleave the highly glycosylated forms is seen in Figure 3, where marker antigenicity disappears completely after digestion.

The marker peptide fusion forms of IL-3 and IL-4 were biologically active despite the presence of the marker segment, as was the case for all of the other proteins that we have expressed in yeast or *E. coli* (10,16). In both cases, the levels of biological activity obtained with the fusion proteins were comparable to wild-type recombinant proteins expressed without the marker sequence.

4E11 Antibody Specificity

Several of the experiments described here were carried out in order to improve our understanding of the interactions that give the 4E11 monoclonal antibody its unique binding properties. Although we had previously noted the dependence on calcium (10), it was not known whether other divalent cations could serve a similar function. The results of the metal specificity studies (Figure 6) indicate that the requirement for calcium is indeed quite specific. None of the other metals that we tested were capable of promoting strong binding of the IL-4 marker fusion protein to the column. However, in the case of manganese, a slight binding ability was demonstrated by reduced washout of IL-4. The partial ability of Mn^{++} to promote binding is interesting in light of the fact that the ionic radius of Mn^{++} (0.80 Å) is closer to that of Ca^{++} (0.99 Å) than any other metal tested. All of the other metal ions are smaller. Future experiments will test whether other divalent metal ions that more nearly approximate the radius of calcium will also promote strong binding.

We studied the effects of amino acid substitutions in the Flag sequence in order to delineate the portions critical to antibody (and metal) binding. Peptides with single alanine substitutions (A1-A8) gave a spectrum of inhibition of the 4E11 antibody, depending on the position of the substitution (Table 1). Peptides Al, A2, and A3 were unable to effectively block the binding of the antibody except at extremely high concentrations. This result suggests that each of the first three amino acids are critical for binding. The requirement for aspartic acid at the N-terminus is further substantiated by peptides S1 and N1, although the substitution of Asn for Asp allows the antibody to bind to at least 100 fold better than substituting an Ala. This may be due to the similarity in the shapes and electrostatic nature of the side groups of Asp and Asn. At physiological pH, Asp will have a formal negative charge, whereas Asn will have a partial negative charge on the carbonyl oxygen. The importance of the N-terminus for binding was further emphasized by the finding that blocking the free α-amino group by acetylation (Ac peptide) also caused a major decrease in the inhibitory ability of the peptide. The importance of the free α-amino group is further underscored by our findings that 4E11 does not bind marker fusion proteins bearing a formyl-Met moiety attached to Asp1 (unpublished).

As can be seen in Table 1, the effects of amino acid substitutions are much less significant among positions 4-8 of the marker sequence. Asp 4 and Asp 7 may play minor roles in binding, because changes to Ala at these positions led to a decrease of approximately two orders of magnitude in inhibitory ability. This is reproducible with either CDP or OVA conjugates and probably reflects a stabilizing role for these amino acids in the interactions of the peptide with the antibody or metal. Changing Asp6 has almost no effect, and this residue may be assumed to play little or no role in the binding process. The substitutes of Ala for Asp5 or Lys8 are of more interest. These substitutions actually increase the inhibitory ability of the peptide by 2 to 3 fold. This may reflect a tendency for these charged residues to destabilize the interaction of the peptide with antibody or metal, or conversely, the Ala side chain may be able to improve binding by a more favorable interaction with the antibody or metal. More work will be required in order to further define the specificity of this binding reaction.

Marker Segment Design

The Flag marker system appears to be a versatile method for recombinant protein identification and purification, and this is due in part to the unique properties of the 4E11 antibody. However, much of the usefulness of this process resides in the specific sequence of the marker peptide as well. This sequence was designed based on several criteria. It was necessary to have an antigenic site and a proteolytic cleavage site in as close proximity as possible in order to minimize the length of the synthetic peptides and oligonucleotides used in this work. As it turned out, the antigenic site and enzyme specificity site actually overlap somewhat because, as

mentioned, Asp4 and Asp7, which are involved in the enterokinase site, also have some influence on antibody binding.

The tendency of fusion proteins to denature (15) was avoided by designing the marker segment to be extremely hydrophilic so that it would remain highly exposed during protein folding, simultaneously avoiding denaturation and promoting its availability for antibody binding. Figure 7 shows that the marker sequence actually achieves a maximum value when plotted by the procedure of Hopp and Woods (11). According to these authors, the segment of maximum hydrophilicity in any protein is always highly exposed. Because the marker has a maximum value, it is assured of always being the most hydrophilic segment of any fusion protein to which it is attached, and therefore of always being exposed when folding is complete. Such an exposed, hydrophilic segment is also expected to be a good target for limited proteolysis, as has been noted previously (8). This accounts for the efficient removal by enterokinase, and also for the observation that processing enzymes such as yeast Kex2, and *E. coli* signal peptidase are effective in liberating the marker sequence from the secretory precursor polypeptides of IL-3, IL-4 and a number of other proteins made using this system (10). Recently we have obtained preliminary information that mammalian signal peptidase will also cleave in front of the Flag sequence during secretion from cultured mammalian cells (Doug Cerretti, unpublished observations).

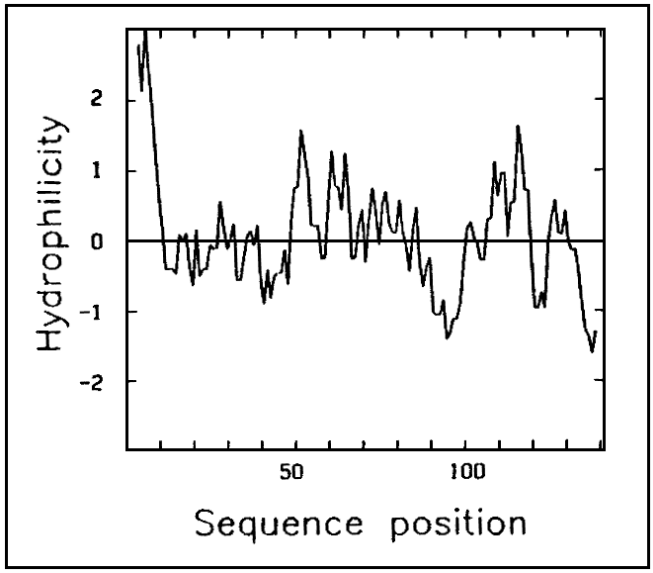

Figure 7. Hydrophilicity profile for the IL-3 marker fusion protein. The profile was generated using the HYDRO3 program of Hopp (9). The scale is oriented so that hydrophilic is at top; hydrophobic is at bottom. Valley regions are expected to be buried portions of the polypeptide, whereas peaks are expected to be exposed at the surface of the protein. The prominent peak near the N-terminus results from the extremely hydrophilic hexapeptide, LysAspAspAspAspLys, contained within the marker peptide sequence.

It is likely then, that this system works well because of the combination of a properly designed hydrophilic marker segment, and a monoclonal antibody of unique properties and unusual specificity. These facts, coupled with our findings using a variety of proteins and cell types, offer hope that the Flag system may be a universal approach for efficient identification and purification of secreted recombinant proteins.

ACKNOWLEDGMENTS

The authors wish to thank Jennifer Martin and Bill Clevenger for excellent technical assistance and the Hybridoma Facility at Immunex for providing antibodies. We would also like to thank Linda Troup for preparation of the manuscript.

REFERENCES

1. Abrahamsen, L., T. Moks, B. Nilsson and M. Uhlen. 1986. Secretion of heterologous gene products to the culture medium of *Escherichia coli.* Nucl. Acids Res. 14:7487-7500.
2. Barany, G. and R.B. Merrifield. 1979. *In* E. Gross and J. Meinhofer (Eds.), The Peptides. Academic Press, New York.
3. Beier, D.R. and E.T. Young. 1982. Characterization of a regulatory region upstream of the ADH2 locus of *S. cerevisiae.* Nature 300:724-728.
4. Brake, A.J., J.P. Merryweather, D.G. Coit, U.A. Heberlein, F.R. Masiarz, G.T. Mullenback, M.S. Urdea, P. Valenzuela and P.J. Barr. 1984. α-Factor directed synthesis and secretion of foreign proteins in *Saccharomyces cerevisiae.* Proc. Natl. Acad. Sci. USA 81:4642-4646.
5. Germino, J. and D. Bastia. 1984. Rapid purification of a cloned gene product by genetic fusion and site-specific proteolysis. Proc. Natl. Acad. Sci. USA 81:4692-4696.
6. Goeddel, D.V., D.G. Kleid, F. Bolivar, H.L. Heyneker, D.G. Yanzura, R. Crea, T. Hirose, A. Kraszewski, K. Itakura and A.D. Riggs. 1979. Expression in *Escherichia coli* of chemically synthesized genes for human insulin. Proc. Natl. Acad. Sci. USA 76:106-110.
7. Hopp, T.P. 1984. Immunogenicity of a synthetic HBSAg peptide. Enhancement by conjugation to a fatty acid carrier. Mol. Immunol. 21:13-16.
8. Hopp, T.P. 1986. Protein surface analysis: Methods for identifying antigenic determinants and other interaction sites. J. Immunol. Meth. 88:1-18.
9. Hopp, T.P. 1988. The use of hydrophilicity plotting procedures to identify protein antigenic segments and other interaction sites. Meth. Enzymol. (in press)
10. Hopp, T.P., K.S. Prickett, V.L. Price, R.T. Libby, C.J. March, D.P. Cerretti, D.L. Urdal and P.J. Conlon. 1988. A short polypeptide marker sequence useful for recombinant protein identification and purification. Bio/Technology 6:1204-1210.
11. Hopp, T.P. and K.R. Woods. 1981. Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA 78:3824.
12. Itakura, K., T. Hirose, R. Crea, A.D. Riggs, H.L. Heyneker, F. Bolivar and H.W. Boyer. 1977. Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. Science 198:1056-1063.
13. Laemmli, V.K. 1970. Determination of protein molecular weight in polyacrylamide gels. Nature 277:680.
14. Liu. F.-T., M. Zinnecker, T. Hamoaka and D.H. Katz. 1979. New procedures for preparation and isolation of conjugates of proteins and a synthetic copolymer of D-amino acids and immunochemical characterization of such conjugates. Biochemistry 18:690.
15. Nagai, K. and H.C. Thogerson. 1987. Synthesis and sequence specific proteolysis of hybrid proteins produced in *Escherichia coli.* Meth. Enzymol. 153:461-481.
16. Park, L.S., D. Friend, V. Price, D. Anderson, J. Singer, K.S. Prickett and D.L. Urdal. 1989. Heterogeneity in human interleukin-3 receptors: A subclass that binds human GM-CSF. J. Biol. Chem. 264:5420-5427.
17. Price, V., D. Mochizuki, C.J. March, D. Cosman, M.C. Deeley, R. Klinke, W. Clevenger, S. Gillis, P. Baker and D. Urdal. 1987. Expression, purification and characterization of recombinant murine granulocyte-macrophage colony-stimulating factor and bovine interleukin-2 from yeast. Gene 55:287-293.
18. Towbin, H., T. Staehelin and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc. Natl. Acad. Sci. USA 76:4350.
19. Villa-Komaroff, L., A. Efstratiadis, S. Broome, P. Lomedico, R. Tizard, S.P. Naber, W.L. Chick and W. Gilbert. 1978. A bacterial clone synthesizing proinsulin. Proc. Natl. Acad. Sci. USA 75:3727-3731.
20. Yang, Y.-C., A.B. Ciarletta, P.A. Temple, M.P. Chung, S. Kovacic, J.S. Witek-Giannotti, A.C. Leary, R. Kriz, R.E. Donahue, G.G. Wong and S.C. Clark. 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.

Address correspondence to:

Kathryn S. Prickett
Immunex Corporation
51 University Street
Seattle, WA 98101

DYKDDDDK
Flag Epitope Tag