Separation Science and Technology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713708471

On the Recovery of Genetically Engineered Proteins from *Escherichia coli*

Satish K. Sharma[a]
[a] Biotechnology-Biopolymer Chemistry, The Upjohn Company, Kalamazoo, Michigan

To cite this Article Sharma, Satish K.(1986) 'On the Recovery of Genetically Engineered Proteins from *Escherichia coli*', Separation Science and Technology, 21: 8, 701 — 726
To link to this Article: DOI: 10.1080/01496398608056146
URL: http://dx.doi.org/10.1080/01496398608056146

SEPARATION SCIENCE AND TECHNOLOGY, 21(8), pp. 701-726, 1986

REVIEW

On the Recovery of Genetically Engineered Proteins from *Escherichia coli*

SATISH K. SHARMA

BIOTECHNOLOGY-BIOPOLYMER CHEMISTRY
THE UPJOHN COMPANY
KALAMAZOO, MICHIGAN 49001

Abstract

An overview of *Escherichia coli* (*E. coli*) as a host for the expression of useful eukaryotic proteins is presented. During the isolation of genetically engineered proteins from *E. coli*, one faces unique problems due to the precipitation of these proteins within cells. These problems include: 1) solubilization with strong denaturing agents, and 2) the removal of the denaturing agent under conditions optimal for protein folding. In addition there is the inherent inability of *E. coli* to perform various cotranslational and posttranslational events within its intracellular environment. Various approaches to solve some of the problems posed by the *E. coli* expression system for product recovery are critically evaluated and their usefulness and limitations are pinpointed. The impact of recombinant DNA technology on protein recovery from *E. coli* is discussed. Whether the intracellular expression in *E. coli* will continue to be the approach of choice for commercially useful proteins will depend upon our ability to find efficient and economical renaturation conditions as well as on the development of alternative expression systems.

I. INTRODUCTION

Escherichia coli (*E. coli*) is a well-characterized, gram-negative bacterium that has been used extensively in recombinant DNA technology (*1, 2*). This repetitive organism has been a favorite for at least three reasons. It has relatively simple genetics, a rapid growth rate, and is well characterized. It contains one large, circular thread of DNA. Smaller circular units called plasmids may also be present. Plasmids can be cut at specific sites by restriction endonucleases, thereby providing space for

 0149-6395/86/2108-0701$3.50/0

new DNA. Donor DNA may be derived from different sources: a chromosome of interest; known amino-acid sequence of a given protein; or complementary DNA (cDNA) may be synthesized using a template, purified messenger RNA (mRNA), and the enzyme reverse transcriptase. The donor DNA is annealed in the cleaved space of the plasmid by hydrogen bonding of complementary nucleotide sequences at the cleavage site. Once annealed, the donor DNA and the plasmid DNA are joined together by the enzyme DNA ligase to form a new plasmid called vector. The vector is introduced into the *E. coli* host through the cell envelope. Expression of a gene (a segment of DNA that codes for a protein) occurs in two steps: transcription of the DNA to mRNA, followed by translation of the message into the desired protein product. The focus here is on the recovery of the recombinant protein product from *E. coli.*

Most of the commercially important proteins which one would wish to produce by recombinant DNA technology are native to eukaryotic organisms, whereas the host microorganisms used in their production are prokaryotic organisms, for example, *E. coli.* Unfortunately, differences exist in the ways in which eukaryotes and prokaryotes express and process proteins. Thus, the application of *E. coli* as a host for a eukaryotic gene encoding a useful polypeptide or protein is not without problems. One problem with the use of *E. coli* as a host is that the eukaryotic proteins are found to be insoluble products in *E. coli.* This requires additional steps in the recovery process which allow for solubilization, renaturation, and in some cases, *in vitro* posttranslational modifications. In addition, some other modifications such as glycosylation cannot occur within *E. coli* cells. The second problem is that *E. coli* produces proteases which can destroy the foreign protein produced by *E. coli.* This destruction of protein product could lead to the wrong conclusion and imply that some cloned eukaryotic genes are expressed at low levels. The third problem which received early attention was the production of an endotoxin by *E. coli.* The fourth problem with the use of *E. coli* as a host is that the proteins produced are characterized by the addition of an extra methionine residue at their N-terminus. This occurs because translation is generally initiated at the AUG codon which codes for methionine. As the presence of an N-terminal methionine on eukaryotic proteins which normally do not possess this amino acid may cause an immune reaction when administered to mammals, it would be desirable to remove the N-terminal methionine, when applicable, thus producing the mature and authentic eukaryotic protein.

Proteins produced by recombinant DNA technology must be purified to homogeneity, particularly if they are to be employed for medicinal purposes. Not only must the desired protein be separated from any removed presequences, signal peptides, or conjugated proteins, but also it must be separated from other microbial proteins and endotoxins which are produced by the host microorganisms. Recovery of the desired protein product from recombinant *E. coli* in the purified and the active form has presented a number of problems. The main purpose of this review is to discuss general protein recovery problems posed by the *E. coli* expression system as well as to evaluate critically various options to overcome these problems. I have made sincere efforts to include every relevant publication in this article, but it is still possible that I have overlooked a few useful contributions. I wish to assure the readers that such omissions, if they have occurred, are unintentional.

II. INCLUSION BODIES IN *E. coli* CELLS

It was reported earlier that expression levels for human growth hormone have reached 1% of the total cell protein (*3*), and levels as high as 30 to 50% of the "soluble" protein may be attainable (*4*). Consequently, such high expression levels may allow for development of rather simple recovery schemes. At about the same time, Tarnowski (*5*) reviewed some general practical considerations involved in recovering heterologous "soluble" proteins produced by recombinant DNA technology. In contrast, it is well known now that under some conditions, and for most proteins, these heterologous proteins are frequently precipitated within the cells as inclusion bodies and constitute a significant portion of the total cell protein. These precipitated protein bodies appear as bright spots visible within the enclosure of the cell under a phase contrast microscope (*6*).

Inclusion bodies have been called many things ranging from refractile proteins to proteins encased in stainless steel balls. The formation of cytoplasmic inclusion bodies in *E. coli* producing biosynthetic human insulin was first reported by Williams et al. (*7*). Subsequently others have reported the presence of inclusion bodies in *E. coli* that have been genetically altered in order to produce proteins of commerical interest (*8–13*). For example, high-level expression of bovine growth hormone in *E. coli* results in the formation of distinct cytoplasmic granules that are visible under phase-contrast microscope (*14*). Intact granules have been

isolated from crude cell lysates by differential centrifugation and were further purified by a simple washing procedure that yields nearly homogeneous bovine growth hormone. It should be borne in mind that bovine growth hormone is not a unique protein in its ability to form intracellular granules. There are numerous reports (*15–17*) of native and fusion proteins that aggregate when expressed at a high level, and are found in the cell pellet fraction following centrifugation of cell lysates. It appears that all these proteins form similar granules, suggesting that aggregation and granule formation is a common property of even native *E. coli* proteins when they are overproduced. A second type of protein granule has been observed in *E. coli* comprised of abnormal proteins resulting from the incorporation of amino acid analogs or puromycin (*18, 19*). They resemble amorphous aggregates of proteins and are degraded by *E. coli* proteolytic enzymes (*19*), whereas the inclusion bodies from recombinant sources are stable.

It is not yet fully understood how and why these inclusion bodies are formed in *E. coli.* It appears that protein insolubility is the rule rather than the exception when high levels of intracellular expression are achieved. One theory is that it is because of overproduction of the recombinant protein in an environment which may not be conducive to proper protein folding. *In vitro* optimal concentrations for refolding are in the one micromolar range or below (*20*). If through genetic engineering one is producing 5% of the total cell protein as the product of interest, its intracellular concentration could be as high as 100 micromolar, and these higher concentrations may affect the folding process. This theory is supported by observations of inclusion body formation following expression of normally soluble *E. coli* proteins behind strong promoters (*21–23*).

In contrast, it has been shown that overproduction of a foreign protein hormone in *E. coli* may lead to a mixture of soluble as well as insoluble forms of the desired product (*24*). In this study, epidermal growth factor (urogastrone) was expressed as greater than 10% of the total cell protein, and at least 40% of the protein was soluble. It is not yet known if this is due to the fact that urogastrone is a small molecule and/or is not as complex as other proteins in terms of disulfide bonds. Therefore, in the case of larger proteins the possibility of entrapment of soluble protein molecules within inclusion bodies should also be considered. Recently, the author has shown that recombinant renin with a molecular weight of 37 kd and three disulfide bonds can be obtained in a soluble form, without denaturing agents, by high pressure in a French pressure cell (*25*). The data suggest that the recombinant renin may indeed be partially soluble when produced within *E. coli* cells. However, the soluble renin

appears insoluble due to its association with the insoluble material produced on lysis of *E. coli* cells and/or due to its entrapment within the insoluble renin molecules present in inclusion bodies.

III. ISOLATION OF THE DESIRED PROTEIN PRODUCT FROM *E. coli* CELLS

Basically, protein recovery from recombinant *E. coli* can be divided into two sections: isolation and purification. In isolation, the objective is to remove the product from the cells and other particulates. This involves breaking the cells by either mechanical or nonmechanical methods. A French pressure cell may be used for mechanical disruption of *E. coli* at 16,000 psi. Generally, three to four passes through the French pressure cell are required to obtain adequate cell breakage (*5*). One common large-scale technique is disruption of the microorganism in a Manton-Gaulin device (Gaulin Corp., Everett, Massachusetts). Table 1 shows the relative rates at which various microorganisms are broken by the Manton-Gaulin homogenizer. In addition to the differences between microbial species, the rate of release of protein from microorganisms is also influenced by the fermentation conditions. For example, the rate of release of β-galactosidase from *E. coli* by passage through a high-pressure homogenizer was faster when the culture was grown on glycerol in mineral salts than from a complex medium (*26*). The rate constant was highest for bacteria recovered early in the exponential phase of growth and decreased as further growth occurred until it was several times less for bacteria harvested in the stationary phase (*30*).

The most commonly used method for breaking *E. coli* cells on a bench-scale process development is by the enzyme lysozyme which digests the cell wall. Enzymatic procedures, however, may be expensive on a large scale. Moreover, the added enzyme is one more contaminant that must subsequently be removed during purification. Other chemical release methods offer several advantages for scaling up, including minimum requirement for operator attention, no mechanical energy input (heat), and elimination of possible mechanical failure. Some common chemical release methods include strong denaturing agents such as guanidine hydrochloride, urea, or detergents such as sodium lauryl sulfate.

Following disruption by mechanical or chemical procedures, residual cell debris must be removed from the total cell lysate. A word of caution is in order. The presence of nucleic acids in the cell extract can also effect subsequent isolation steps, and it may be necessary to eliminate them by precipitating or hydrolyzing them with added nucleases (*31*). Notably,

TABLE 1
Protein Release from Various Microorganisms in a High-Pressure Homogenizer

Microorganism	Rate constant (h^{-1})	Ref.
Escherichia coli	0.39	*26*
Saccharomyces cerevisiae	0.23	*27*
Baccillus brevis	0.28	*28*
Nocardia rhodochrous	0.0085	*29*

cell extracts of *E. coli* are usually much more viscous compared with other microorganisms such as baker's yeast.

Extraction of a recombinant protein, in its native form, from *E. coli* is a unique problem due to the fact that it is expressed in an insoluble form within cells. The precipitated protein can be liberated from the cells by employing means which disrupt the outer cell wall/membrane under conditions comprising sufficient ionic strength and proper pH. This is done so that host cell proteins (provided the cells are sufficiently disrupted) will be solubilized. Consequently, upon low-speed centrifugation the desired inclusion bodies will be accumulated in the pellet, and most of the contaminating proteins will remain in the supernatant. The pellet, however, may contain other proteins for at least two reasons. First, the original inclusion bodies may not have been totally comprised of the desired protein. Second, fragments of cell walls may be incompletely broken so that they remain with the pellet. However, the pellet which results could be predominantly the desired protein, and the problem becomes one of removing contaminants from a basically pure product, rather than isolating a small component of a complex mixture. It is worth mentioning here that in some cases the desired protein product can also undergo polymerization due to the formation of intermolecular disulfide bonds (*32*), requiring the use of reducing agents to form monomers before any recovery steps can be attempted.

Once an inclusion body preparation which is predominantly the desired protein is obtained, the next problem is that the protein must further be purified and then recovered in a form that is biologically active. Since the protein has been precipitated *in vivo* under cytoplasmic conditions, one concludes that the insoluble proteins can be dissolved only in strong denaturing agents such as ionic detergents, urea, guanidine hydrochloride, or a strong base such as sodium hydroxide. The situation is similar to developing an isolation process for egg albumin starting with hard-boiled eggs. In general, the techniques described (*11–13, 33–35*)

involve the denaturation of a protein followed by removal of the denaturant under conditions which are optimal for protein folding. Renaturation can be performed by dialyzing the solution containing the recombinant protein. The amount of time a protein can be held under denaturing conditions without significantly affecting renaturation sometime depends on the concentration of the denaturing solvent and the temperature. For example, it has been found that denaturation by NaOH (concentrations equal to or greater than 0.01 *M*), 8 *M* urea, or 6 *M* guanidine hydrochloride for periods of time in excess of a few minutes inhibits subsequent renaturation of prochymosin to a form capable of undergoing conversion to active chymosin (*35*).

These methods solubilize a significant percentage of the insoluble proteins found in inclusion bodies. However, there is great concern as to the quantitative aspect in terms of recovery of native protein. In some cases guanidine hydrochloride might solubilize all the material present in inclusion bodies, but only a portion may be converted into native protein after removal of the denaturant. On the other hand, alkali treatment alone may not result in complete solubilization and, in addition, may not allow complete renaturation to the native form of the protein. Therefore, it may be necessary at times to combine the two solubilization techniques in order to enhance yields of native protein. High recovery of active chymosin was demonstrated (*36*) when inclusion bodies were initially dissolved in 7 *M* urea or 6 *M* guanidine hydrochloride followed by dilution with an alkaline buffer.

Flow charts for the isolation and purification of human insulin (*37, 38*), chymosin (*12*), and urokinase (*32*) from recombinant *E. coli* are shown in Fig. 1. Obviously, isolation of recombinant proteins from inclusion bodies requires the additional steps which allow for complete solubilization with high efficiency. While on the surface these additional steps seem to be a major disadvantage for genetically engineered proteins in *E. coli,* the added benefit of inclusion formation is that the protein of interest is generally between 20–80% pure within these inclusion bodies. Therefore, following cell lysis, a centrifugation and washing of the pellet fraction sometimes offers a single and effective purification. However, if the lack of purity of the denatured protein affects the renaturation efficiency, then normal protein purification techniques can be applied to the denatured protein before the refolding step. Various general approaches for the purification and activity assurance of precipitated heterologous proteins have been described (*39–41*).

Many group specific adsorbents are available which can provide highly purified protein products. Adsorbents include monoclonal antibodies (*42*), organic dyes (*43*), thiol groups (*44*), and metal chelates (*45*).

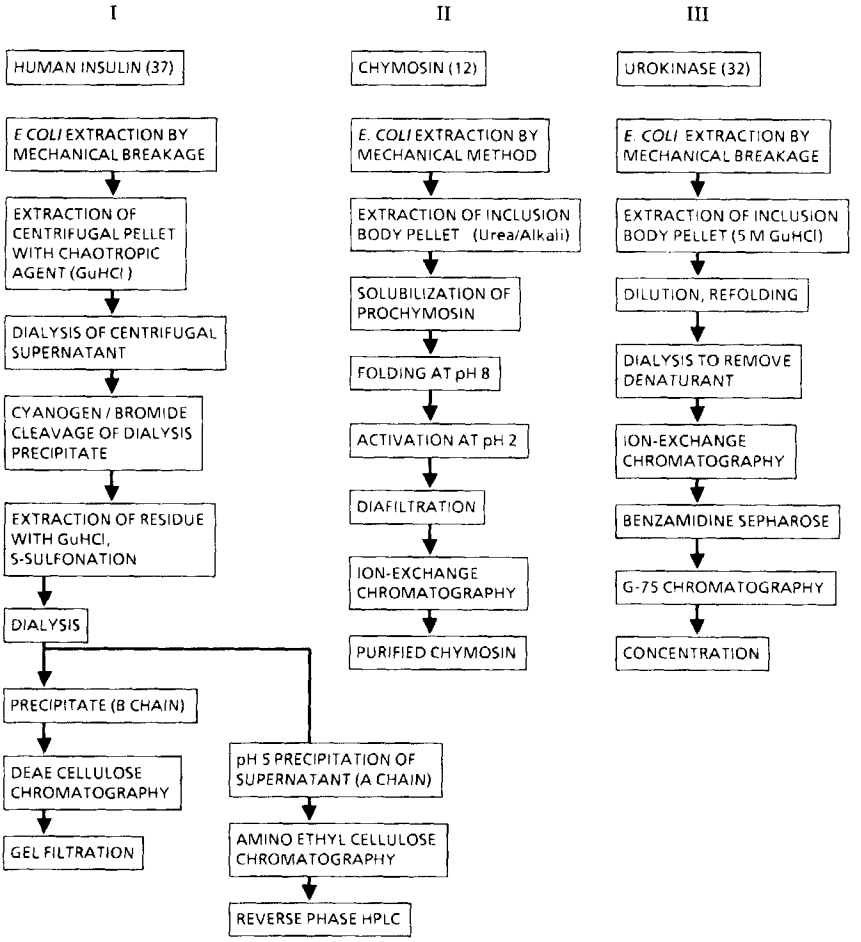

FIG. 1. Flow charts for the recovery of recombinant proteins from *E. coli*.

Recent advances in analytical methodology, such as reverse phase HPLC (*46*), have raised the expectations for purity and thus increased the processing challenge. The most common method employed to determine the purity of a protein, whether it be natural or recombinant, is SDS polyacrylamide gel electrophoresis followed by Coomassie Blue staining. Although it is used quantitatively, this method detects only protein impurities. Furthermore, there are protein-to-protein variations in Coomassie Blue staining intensity. A new staining technique using silver is claimed to be 100 times more sensitive than Coomassie Blue (*47, 48*).

Other criteria for purity and identity of proteins (*5, 49*) include tryptic mapping, sequence analysis, isoelectric focusing, HPLC, and immunological data.

In short, the recovery of a biologically active protein from recombinant *E. coli* is not a trivial task. There are some guidelines but usually a new purification and renaturation protocol must be tailor-made for each recombinant protein. In addition, there are other factors that can significantly affect the final recovery of the "desired" biologically active protein product; for example, endotoxin removal, disulfide bond formation, limited proteolysis, processing of N-terminal methionine, and *E. coli* proteases. These factors are discussed in detail in the following sections.

IV. INABILITY OF *E. coli* TO PERFORM CO- AND POSTTRANSLATIONAL MODIFICATIONS

Figure 2 shows some of the more common eukaryotic biochemical processing steps which do not take place in *E. coli*. In most cases more than one of these co- and posttranslational modification events occurs for any given protein. Signal processing and glycosylation are cotranslational events because they occur at a time when the protein is being secreted across a membrane while still attached to the ribosome (Fig. 2). Table 2 depicts structural features of some recombinant proteins produced in *E. coli*. Clearly, most of these proteins have pre- or prosignal amino acid sequences associated with the protein when synthesized by their natural host cells. These signal sequences are associated with all secretory proteins and aid in the transport of proteins across membrane barriers. Specific membrane associated signal peptidases exist which cleave these presequences and generate the mature protein. Bacteria also contain signal or presequence processing enzymes and substantial similarity exists between the eukaryotic and prokaryotic processing systems (*50*). The removal of the presequence in an expression system which utilizes intracellular expression of foreign proteins (e.g., *E. coli*) is accomplished by deleting the DNA that codes for the signal sequence. A difficulty with this approach is that an additional methionine is retained at the start of the DNA coding for the mature protein. Sometimes this additional N-terminal methionine is ignored because the presence of this extra amino acid does not appear to inhibit the biological activity of the desired product (*51*). Various approaches are discussed in the next section to overcome this problem.

A common type of cotranslational modification is glycosylation although investigators have also described hydroxylation, carboxylation,

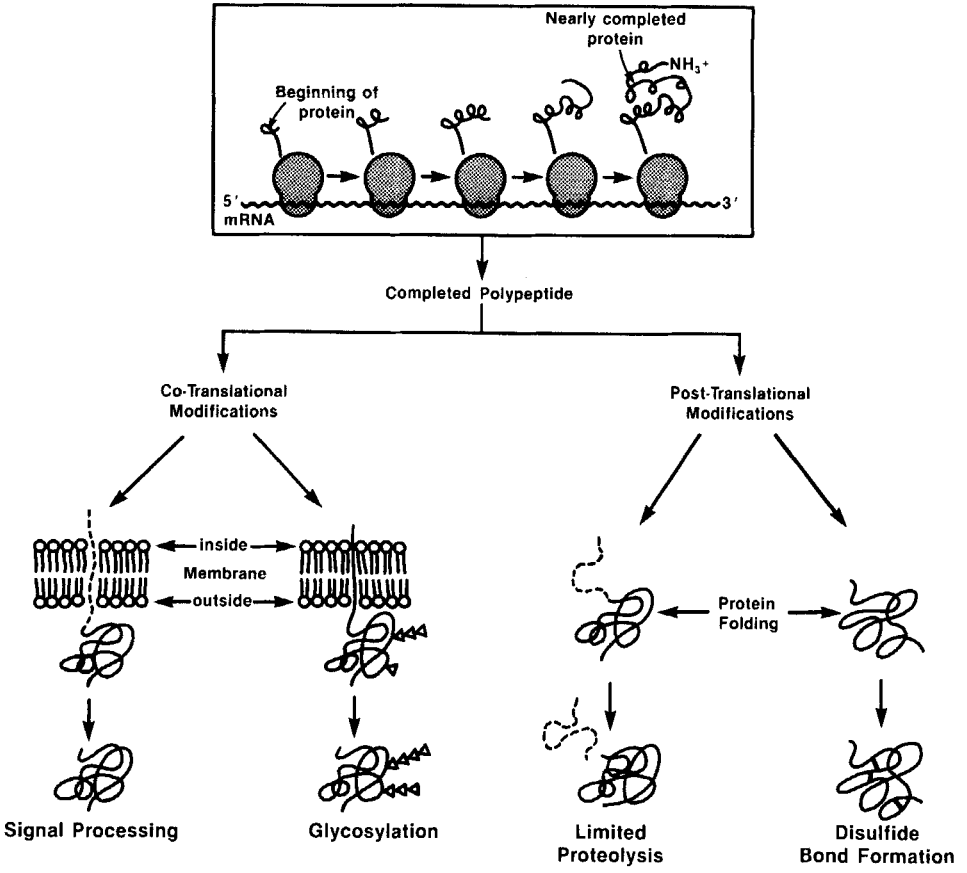

FIG. 2. Cotranslational and posttranslational events in protein biosynthesis.

phosphorylation, lipidation, methylation, and ADP-ribosylation modifications (*52*). A number of proteins which contain carbohydrates in their native structure are shown in Table 2. Notably, *E. coli* is incapable of carrying out the glycosylation process.

Not only has it been difficult to ascribe a unique function to the carbohydrate moieties, but it seems that each glycoprotein must be evaluated individually to determine the importance of glycosylation to its stability and function. Although the functions of glycosylation are not entirely predictable, general differences between deglycosylated molecules and their natural counterparts have been defined. These include antigenicity, stability, solubility, and tertiary structure. The recent work

TABLE 2
Structural Features of Some Proteins Produced in *E. coli* by Recombinant DNA Technology

Protein	Sugars	Presequence	Prosequence	Number of cysteine residues	Number of disulfide bonds	Theoretical random pairing	Refs.
Interleukin 2	Yes	Yes	No	3	1	3	*10*
Insulin	No	Yes	Yes	4	2	3	*37, 38*
Bovine growth hormone	No	Yes	No	4	2	3	*14*
Human renin	Yes	Yes	Yes	6	3	15	*54, 55*
Chymosin	No	Yes	Yes	6	3	15	*11-13*
Urokinase	Yes	Yes	Yes	24	12	3.1×10^{11}	*32*
Immunoglobulin	Yes	Yes	No	34	17	6.3×10^{18}	*56, 57*
Tissue plasminogen activator	Yes	Yes	Yes	35	17	2.2×10^{20}	*58*

by Sairam and Bhargavi (*53*) confirms the link between glycosylation and functional activity of gonadotropic hormones. From a biotechnological point of view, this finding is important because it is the first time that a specific functional role for the carbohydrate moiety of a protein has been assigned. In contrast, some of the proteins listed in Table 2 can be shown to be functional *in vitro* in their deglycosylated forms (*10, 32, 58*). However, it has not been shown whether the presence of carbohydrates would increase yields of refolded material, in view of the fact that carbohydrate increases the amount of folded material in the case of α-subunit of the bovine glycoprotein hormones (*59*).

It is now known that the protein as synthesized on the ribosome (Fig. 2) is often not the form that is isolated. It undergoes posttranslational modifications as shown in Fig. 2. Disulfide bond formation and limited proteolysis are posttranslational modifications because they occur primarily after the protein is released from the ribosome and complete or partial folding has occurred. As shown in Table 2, in most cases both of these posttranslational modification events occur for any given protein. Therefore, protein isolation from recombinant *E. coli* must include additional steps to compensate for the inability of *E. coli* to carry out these posttranslational modification events.

Disulfide bonds generally do not occur naturally within intracellular proteins. It is believed that in general no disulfide bonds are formed in the cytoplasm of *E. coli.* These bonds function to stabilize the structure of extracellular proteins which encounter more varied physiological environments. Obviously, if intracellular expression systems such as *E. coli* are used, steps in the isolation process must be included for the correct formation of these bonds. As shown in Table 2, the number of possibly incorrect disulfide bonds increases as the number of disulfides increases. The real problem in practice is the refolding process after complete unfolding of the protein. Certainly the conditions are very different; local pH, ionic strength, and ion concentrations on the ribosome are certainly determining parameters for directing the folding process. The *in vitro* refolding of a protein is a process very sensitive to these experimental conditions (*60*). The secretory proteins usually undergo a number of steps during formation in their natural cell type. For example, these steps include synthesis, segregation, transport, concentration, storage, and discharge (*61*). It is unclear to what extent the various folding steps occur in each of these compartments. Can *E. coli* exactly duplicate the natural folding environment for a foreign protein? The answer appears to be no, and consequently the development of highly efficient folding conditions for genetically engineered proteins in *E. coli* is indeed a general problem.

In consideration of *in vitro* refolding, one must cope with protein–protein interactions and the insolubility of the polypeptide chain. Protein–protein interactions and insolubility problems can be overcome to some extent by carrying out folding experiments at low protein concentration (10^{-7} *M*). From a commercial viewpoint this may not be desirable because of the difficulty in handling large volumes containing μg/mL of the product. Thus, the final recovery of the biologically active protein product may largely depend upon one's ability to concentrate the product in active form and in high yield.

Although a number of recombinant enzymes have been successfully refolded (*11–13, 32, 39–41*), at least to the point of regaining activity, this is an empirical observation, and there is at present no way of predicting whether an isolated enzyme is capable of refolding or, more precisely, since refolding always takes place, whether proteins can refold to form the active structure rather than a "wrong" one.

Some proteins also have prosequences attached to the mature protein (Table 2). Although the physiological function of these prosequences is not always clear, they appear to inhibit the biological activity of the molecule until its activation by limited proteolysis at a site distal to the site of synthesis. In the case of chymosin it seems clear that these prosequences are essential for efficient folding and disulfide bond formation (*11–13, 35–36*). As a result, the DNA coding for these sequences cannot be removed without affecting the yield of the biologically active molecule.

This means that *in vitro* processing steps must be included to remove the prosequence and generate the required active protein. In the case of chymosin it is accomplished by acid activation after the refolding step (*11–13, 35, 36*). The proteolytic processing steps are more difficult to achieve in the case of proinsulin (*62, 63*). In general, specific cleavage of the "pro" sequence after the folding step appears to be a difficult task, if not impossible. This step could in turn have significant impact on the recovery of the final biologically active protein product.

V. PROCESSING OF N-TERMINAL METHIONINE

a. Specificity of Aminopeptidase

The biosynthesis of all proteins from all living cells begins with methionine. The additional methionine raises some concerns for the production of pharmaceutical proteins in *E. coli* because it may represent

an antigenic determinant not found in the native protein. However, the amino-terminal methionine, in some cases, can be cleaved from *E. coli* proteins by a naturally occurring methionine aminopeptidase. The aminopeptidase removes amino-terminal residues of methionine when they precede certain amino acids, with a specificity that appears to be determined largely by the residue adjacent to the methionine residue at the amino terminus. The results from published sequences of proteins (*64*) from a wide range of prokaryotes and eukaryotes suggest that the aminopeptidase usually cleaves amino-terminal methionine when it precedes residues of alanine, cysteine, glycine, proline, serine, threonine, and valine but not when it precedes other amino acids. Recent studies with the mutationally altered iso-1-cytochrome c suggest that the specificity is almost always determined simply by the size of the side chain of the penultimate residue; methionine is usually cleaved from residues with a side chain having a radius of gyration of 1.22 Å or less, but is not cleaved from residues with larger side chains. In contrast, recent N-terminal data on highly purified recombinant interleukin-2, expressed in *E. coli,* showed that methionine at the 0 position was found in 90% of the molecules and was not completely removed in posttranslational processing (*65*). This is despite the fact that alanine is the amino acid next to this N-terminal methionine. Sherman et al. (*66*) have also noted exceptions to the pattern of methionine cleavage. Therefore, it would be desirable to process recombinant proteins *in vitro* to remove the N-terminal methionine. Indeed, the N-terminal methionine processing enzyme has been isolated and shown to catalyze the removal of methionine from recombinant bovine and human growth hormones and their derivatives (*67*).

b. Site-Directed Mutagenesis

If the naturally occurring N-terminus is not recognized by aminopeptidase(s) found in *E. coli,* then a site-directed mutagenized form of the protein can be obtained where the natural N-terminal residue is changed to meet the specificity requirement for this enzyme. This would result in a recombinant protein with an altered amino terminal. This approach might be appropriate in cases where the N-terminal amino acid is not essential for determining the structure and/or function of the recombinant protein.

TABLE 3
Linkers Cleavable by Enzymes and Chemical Techniques

Linker sequence[a]	Cleavage by	Refs.
-Met-x-	Cyanogen bromide	*37*
-Asp-Pro-x	Acid	*68*
-Asn-Gly-x-	Hydroxylamine	*69*
-Ile-Glu-Gly-Arg-x-	Factor X_a	*70*
-Arg(Lys)-x-	Trypsin	*71*
-Pro α-2 Collagen-x-	Collagenase	*72*
$-(Asp)_4-Lys-x-$	Enterokinase	*73*
His-Pro-His-Pro-His-Pro-x	Aminodipeptidase IV	*74*
x-Arg-Arg-Arg-Arg	Carboxypeptidase B	*75, 76*

[a]x stands for any given protein sequence.

c. Introduction of Cleavable Linker Sequences

Another approach to remove N-terminal methionine is to insert specific amino-acid sequences which can be removed by proteolytic enzymes. Some well-known examples of such linkers are shown in Table 3. In some instances linkers cleavable by chemical methods, such as cyanogen bromide, have been found to be useful for the processing of N-terminal methionine. Introduction of cleavable linker sequences, however, may result in some cases in the formation of hybrid proteins whose consequences are discussed below.

VI. CONSEQUENCES OF EXPRESSING FUSION PROTEINS IN *E. coli*

Once a fusion polypeptide is produced, methods must be found for converting it to the natural product. Rutter (*77*) suggested the use of endopeptidases with extended sites as a general method of cleaving peptide fusions. Examples of linkers cleavable by enzymes are shown in Table 3. The rationale is that the greater the number of amino-acid residues in the specific cleavage sequence, the more unlikely the possibility of such a sequence present within the desired recombinant protein product. As for example, enterokinase cleaves after $(Asp)_4$-Lys sequence (*73*). However, the use of enzyme-based linkers should be viewed with some caution, because the efficiency of cleavage with these

enzymes may be considerably reduced by the three-dimensional structure of the fusion protein, or additional amino acids may be left on the desired protein. Moreover, high expression of recombinant proteins leads to the formation of inclusion bodies which can be dissolved only in strong denaturing agents (*39–41*). Therefore the cleavage step might prove to be extremely difficult under denaturing conditions, if not impossible.

Other methods for fusion sequence removal have also been reported (*37, 68, 69*). For example, β-galactosidase (β-gal) fusions with insulin A and B chains were cleaved chemically at a methionine residue linking the two polypeptides (*37*). This was feasible because human insulin does not contain any methionine residues. In another example, β-endorphin was removed from β-galactosidase fusion using trypsin (*71*). This was successful because β-endorphin has no arginine residues and its lysines were chemically protected from enzymatic cleavage. For practical purposes, the specificity of the desired cleavage site need not be always exclusive with respect to other potential cleavage sites in the desired product. If the desired cleavage site is highly favored kinetically, that is, the desired site is cleaved preferentially with respect to other sites, a reasonable yield of the desired protein can be obtained. For example, folded fused prorenin from recombinant *E. coli* can be activated with trypsin (*78*). Presumably the trypsin cleavage site is on the surface of folded protein and therefore becomes the preferred site.

VII. ROLE OF *E. coli* PROTEASES IN ISOLATING A RECOMBINANT PROTEIN

E. coli has been shown to contain at least eight soluble proteolytic activities (*79*) and there may be more. It is also known that *E. coli* selectively degrades polypeptides with abnormal conformations (*80*). However, not all recombinant proteins are necessarily unstable in *E. coli,* although recombinant proteins would no doubt be recognized by the cell as being abnormal. At this time it is not possible to predict whether a protein will be rapidly turned over or not. The variation in half lives for normal *E. coli* proteins is very large. Only 7% of all proteins have half lives of less than 15 min, another 20 to 30% are not broken down except under starvation conditions, while the remainder are not turned over at all (*81*). The structural features of proteins which affect their half life are not known.

In most of the isolation and purification schemes, attempts are made to overcome the problem of proteolysis by using phenyl methyl sulfonyl fluoride, an effective inhibitor of serine proteases. Earlier work showed

that of the eight *E. coli* protease activities reported (*79*), six were of the serine type and two were metalloenzymes. This strongly suggests that the serine protease inhibitor can only partially overcome the problem of proteolysis by *E. coli* proteases.

The *E. coli* strain chosen as the host for a recombinant plasmid can greatly influence the levels of expression of a cloned gene. In most instances there is, as yet, no logical explanation for this effect, but in some cases it is a reflection of the level of protease activity within the cell. It may be desirable to find mutants that lack these proteolytic activities. For example, mutations in the lon gene that encode a defective ATP-dependent protease La (*82*) reduce rates of degradation of abnormal polypeptides. On the other hand, data suggest that inhibition of protein breakdown by protease inhibitors can result in induction of protease La (*83*). Thus, use of protease inhibitors may cause, in some cases, the accumulation of and/or the generation of abnormal proteins.

An alternative approach is to include in the cloning vector the antiprotease gene of phage T_4 (*84*). This phage gene product reduces proteolysis, and its use can result in increased levels of expression. The process described in this publication involves the use of the apparently hybrid plasmids transformed into the *E. coli* host. One plasmid contains the T_4 pin gene whereas the other contains the eukaryotic gene encoding a desired protein product. However, it is not a desirable commercial process because it is difficult to maintain two different plasmids in the *E. coli* host. Recently, a new approach which avoids the issue of plasmid incompatibility has been described elsewhere (*85*).

One can also choose to increase expression to such an extent that the cellular proteases are saturated by the substrate, therefore allowing a high proportion of recombinant protein to remain intact. With this approach a relatively rapid burst of expression extending over no more than one or two generations is preferable to continuous low expression over an extended period.

The synthesis of recombinant somatostatin (*86*) in *E. coli* was undetectable, presumably due to the product degradation by *E. coli* proteases. In this particualr case the classical solution has been to protect the recombinant protein by fusing it to β-galactosidase. This enzyme affords considerable protection for the low molecular weight somatostatin. The major advantage of this approach is that the host fusion leader has its own Shine-Dalgarno site which ensures successful initiation of translation. The disadvantage is that a fusion polypeptide is produced which may not be acceptable for commercial use.

Thus, from the point of view of those trying to isolate recombinant proteins from *E. coli* "lysates," they should remember that proteases will

be present at all times and that even highly purified proteins may still contain them. Whether the proteins are active or not will depend on the conformation of the desired protein and probably on some other factors yet unknown, which may still be significant, in *E. coli* lysates.

VIII. *E. coli* ENDOTOXINS

Lipopolysaccharide (LPS) or endotoxin is a macromolecular substance tightly bound to the surface of gram-negative bacteria, forming an integral part of the membrane structure (*87*). Endotoxins are pyrogenic, that is, they produce fever in man and other mammals. Even if the bacteria are killed, their removal is essential for the safe parenteral administration of products produced by natural sources or by recombinant DNA technology. The sensitivity of mammals to endotoxins is extraordinary, and contamination levels of less than 1 ng/mL elicit a strong fever response and can even result in death (*88*).

Administering human proteins derived by recombinant DNA technology has the potential of unknown risks. The questions that are relevant are: What difficulties will one face with contamination by extraneous bacterial proteins and lipopolysaccharide such as endotoxin? What are the clinical implications? How serious are such problems? Bacterial contaminants, even at a relatively low concentration, could give rise to undesirable side effects. Such contamination has been experienced in the recovery of L-asparaginase from *E. coli* (*89*). Therefore, in the production of biopharmaceuticals by recombinant DNA technology the challenge is not only to produce large quantities of protein, but also to approach a biologically active product of 100% purity. Various techniques for endotoxin detection and elimination have been recently reviewed (*90*). Efficient elimination of endotoxins from a recombinant protein produced in *E. coli* could significantly affect the final yield of the desired product and, therefore, the economics of the recovery process.

IX. IMPACT OF RECOMBINANT DNA TECHNOLOGY ON PROTEIN RECOVERY

Generally, the impact of recombinant DNA technology on protein recovery can include the following:

(a) Higher expression of desired protein (*3*)
(b) Selection of suitable microorganisms (*49*)
(c) Genetic manipulation to aid recovery (*75, 76, 91, 92*)

All of the above factors must be balanced to achieve the objectives of downstream processing. These objectives are high recovery, high purity, reproducibility at small scale, and, when scaled up, low cost and practicality. A number of excellent reviews have been published on general downstream processing of enzymes and proteins to which reference should be made for more detailed background material (*93–95*).

With the application of recombinant DNA methods it is now possible to increase substantially the level of the desired protein product in *E. coli,* and this should ease the problems of purification and recovery. In some cases, however, unfolding and refolding of protein chains must take place during reprocessing. This is a new and potentially difficult unit operation that may cause biochemical engineers some problems.

The ability to transfer structural and regulatory genes between microorganisms increases the potential freedom of choice of microorganisms (see Section X). In view of the large differences between microorganisms in their ease of handling, this choice could lead to marked improvement in protein recovery and purification protocols. There are several ways in which recombinant DNA techniques can be directly used to improve recovery and purification of protein products. For example, how can the properties of a recombinant protein be reversibly altered to enable a simple and effective purification? Both natural and synthetic genes have been expressed in *E. coli,* but isolation and purification of the resultant gene product from bacterial lysates can be difficult. Moreover, instability of the desired protein product may lead to its degradation by *E. coli* proteases (see Section VII). One of the earlier methods described to overcome this problem was to produce fusion of proteins or peptides to β-galactosidase which stabilizes the fused protein in *E. coli* (*86*). As illustrated below, ion-exchange chromatography or affinity chromatography combined with genetic approaches might provide powerful systems for protein recovery.

Sassenfeld and Brewer (*75*) reported how a C-terminal fusion polyarginine can facilitate the purification of recombinant proteins produced in *E. coli.* A synthetic DNA sequence that codes for five additional arginines at the carboxy terminus was added to the β-urogastrone gene. A substantial purification was achieved by ion-exchange chromatography due to the unusual basicity of the polyarginine-fused protein. The polyarginine tail was removed by carboxypeptidase B and the desired product was isolated by rechromatography on the same column. In addition, this approach allows one to assay the fused recombinant protein simply and accurately. For example, polyarginine can be assayed using chemical reagents (*96*). The only drawback of this approach is that

N-terminal methionine is still retained on the final product, which may not be acceptable for pharmaceutical proteins.

Genetic approaches which allow two or more genes to be spliced together to yield fusion proteins have also been considered (*72*). While in some instances the fusion proteins retain the biological activities of the target protein (*97, 98*), it is reasonable to think that in other situations protein fusion may interfere with the biological activity of the target protein. It is, therefore, very important to develop a general method for separating the target protein from the fusion protein. One approach that has been used is a combination of genetic fusion and site-specific proteolysis (*72*). The technique consists of fusing the gene for the target protein to the DNA of a marker protein via a piece of DNA that codes for a linker peptide. The tripartite protein, β-galactosidase-pro α2 collagen-Rbk replication initiator, was rapidly purified by selective binding to and elution from a β-galactosidase specific affinity column. The target protein is then released from the marker protein by controlled digestion with collagenase. In another study (*99*), the usefulness of fusing the protein A gene and the *E. coli* genes encoding the enzyme β-galactosidase or alkaline phosphatase has also been described. Protein A is well suited for affinity purification due to its specific binding to the F_c part of immunoglobulins of many species including man. The problem with these types of approaches is that they cannot be of general use because different hosts will utilize different fusion sequences.

While designing a genetic construction for purification purposes, the following points should be kept in mind:

(a) The fusion should allow a simple, rapid, and cheap purification by ion-exchange or preferably affinity chromatography
(b) If a linker peptide is used, efficiency of the cleavage should be evaluated
(c) The fusion peptide must have a negligible effect on the protein folding and no permanent effect on the biological activity
(d) It must be readily and specifically removed after purification

A more general protein recovery system will be essential for rapid future developments in the area of protein engineering, which is characterized by the production of modified protein catalysts using recombinant DNA technology. Both the production and characterization of families of mutant proteins should come rapidly if a general purification system is available.

X. ALTERNATIVE EXPRESSION SYSTEMS

Recombinant DNA technology used for the production of pharmaceutically useful polypeptides such as insulin (*37, 38*), human growth hormone (*49, 66, 67*), and the interferons (*9, 34*) has thus far mainly been focused on the *E. coli* expression system. However, within the last few years alternative expression systems such as yeast (*100*), tissue culture (*101*), *Bacillus subtilis* (*102*), *Pseudomonas* (*103*), and streptomyces (*104*) have also attracted the interest of applied scientists. Since these systems are relatively new and have not been characterized extensively, a great deal of basic research examining gene expression and regulation is needed before these organisms can be harnessed for the recombinant DNA industry.

The focus here is on one of these alternative organisms, the yeast *Saccharomyces cerevisiae.* There are several possible advantages that yeast expression systems may have over *E. coli* systems. The major advantages are the finely developed fermentation science, the possibility of secretion, the absence of inducible or contaminating viruses which might result in cell lysis during production, and the lack of endotoxins. Although yeast can glycosylate proteins, so far the possibility for precise glycosylation does not look promising. Since the polysaccharide additions in yeast (*105*) are not exactly like that of mammalian cells, such additions could be more of a disadvantage than an advantage. A slight disadvantage of yeast, as compared to *E. coli* or other bacterial host systems, is the difficulty of breaking the cells to obtain the product. However, this difficulty may in fact be the greatest advantage of yeast if the desired product is secreted into the culture medium. The normal media proteins represent only 0.5% of the total cellular proteins and consist of 5 to 8% having molecular weights greater than 50,000 daltons, which make up about 90% of the protein content of the media. Therefore, if 5% of the cellular protein would be secreted as a desirable protein product, the product would be of 90% purity. The relatively protein-free medium, combined with the resistance of yeast to external stresses, may make it an ideal system for secretion. With regard to this, it has been shown that homologous protein signal sequences, such as those for yeast invertase (*106*) and yeast α-factor (*107*), attached to the heterologous proteins result, in some cases, in secretion of the properly processed heterologous proteins into the culture media. Such systems do not only allow easier purification but also produce natural products that do not begin with an amino-terminal methionine. It should be kept in mind, however, that the

future of yeast in the biotechnology industry will also require yield improvements.

Recently it has also been shown (*108*) that the secretory apparatus of *E. coli* functions well in the secretion of a human polypeptide derived from bacterially synthesized precursor with the *E. coli* signal-peptide sequence. In this approach the purification is easy and recovery of the product is high because the periplasmic proteins comprise only 4% of the total proteins of *E. coli* (*109*). Another attractive possibility is the use of gram-positive hosts, such as *Bacillus subtilis, streptomyces,* and *S. aureus.* Secretion in these hosts allows the formation of disulfide bridges which cannot be formed in the reducing environment of most bacteria. However, in many cases intracellular accumulation might be the only alternative since most intracellular proteins will not be transported across the membrane even when the protein contains a functional signal sequence. β-Galactosidase with such a sequence ends up in the membrane (*110, 111*) and is therefore susceptible to proteolysis. For some constructs *E. coli* may be a better host because the protein might be more susceptible to degradation in other hosts. Clearly, there is the need for alternative expression vectors as well as host organisms for expressing heterologous proteins of commercial interest.

In this section I have tried to indicate the influence that decisions about the choice of microorganisms by molecular biologists can have on the performance and therefore yield of protein recovery steps. Ongoing developments in molecular biology can be expected to improve yields for desired proteins. Other developments will include production of modified proteins and analogues via protein engineering techniques. These changes, as well as technological improvements in fermentation, will all have a direct practical effect on the recovery of proteins from recombinant microorganisms.

Acknowledgments

I wish to thank my colleagues Dr E. R. Olson, Dr G. S. Gill, and Dr M. Deibel for their helpful comments. I am grateful to Debbie Piper and Paula Lupina for typing the manuscript.

REFERENCES

1. J. L. Gueriguian (ed.), *Insulins, Growth Hormone, and Recombinant DNA Technology,* Raven, New York, 1981, pp. 1-227.
2. A. P. Bollon (ed.), *Recombinant DNA Products: Insulin, Interferon and Growth Hormone,* CRC Press, Boca Raton, Florida, 1984, pp. 1-191.

3. M. J. Ross, in Ref. *1,* pp. 33–48.
4. G. F. Miozzari, in Ref. *1,* pp. 13–32.
5. S. J. Tarnowski, *Pharm. Technol.,* p. 70 (February 1983).
6. M. J. Carrier, M. F. Nugent, W. C. A. Tacon, and S. B. Primrose, *Trends Biotechnol., 1,* 109 (1983).
7. D. C. Williams, R. M. VanFrank, W. L. Muth, and J. P. Burnett, *Science, 215,* 687 (1982).
8. D. C. Paul, R. M. VanFrank, W. L. Muth, J. W. Ross, and D. C. Williams, *Eur. J. Cell Biol, 31,* 171 (1983).
9. G. Simons, E. Remaut, B. Allet, R. Devos, and W. Fiers, *Gene, 28,* 55 (1984).
10. R. Devos, G. Plaetinck, H. Chérontre, G. Simons, W. Degrave, J. Tavernier, E. Remaut, and W. Fiers, *Nucleic Acids Res., 11,* 4307 (1983).
11. M. T. McLaman, W. H. Andrews, and J. G. Files, *J. Biotechnol., 2,* 177 (1985).
12. F. A. O. Marston, P. A. Lowe, M. T. Doel, J. M. Schoemaker, S. White, and S. Angal, *Bio/Technol., 2,* 800 (1984).
13. Y. Kawaguchi, N. Shimizu, K. Nishimori, T. Uozumi, and T. Beppu, *J. Biotechnol., 1,* 307 (1984).
14. R. G. Schoner, L. F. Ellis, and B. F. Schoner, *Bio/Technol., 3,* 151 (1985).
15. E. Remaut, P. Stanssens, and W. Fiers, *Nucleic Acids Res., 11,* 4677 (1983).
16. M. Zabeau and K. K. Stanley, *EMBO J., 1,* 1217 (1982).
17. P. W. Gray, D. W. Leung, D. Pennica, E. Yelverton, R. Najarian, C. Simonsen, R. Derynck, P. J. Sherwood, D. M. Wallace, S. L. Berger, A. D. Levinson, and D. V. Guedell, *Nature, 294,* 503 (1982).
18. W. F. Prouty, M. J. Karnovsky, and A. L. Goldberg, *J. Biol. Chem., 250,* 1112 (1975).
19. W. F. Prouty and A. L. Goldberg, *Nature, New Biol., 240,* 147 (1972).
20. A. Light, *BioTechniques, 3,* 298 (1985).
21. M. Gribskov and R. R. Burgess, *Gene, 26,* 109 (1983).
22. Y.-S. Ho, M. Lewis, and M. Rosenberg, *J. Biol. Chem., 257,* 9128 (1982).
23. Y.-S. Cheng, D. Y. Kwoh, T. J. Kwoh, B. C. Soltvedt, and D. Zipse, *Gene, 14,* 121 (1981).
24. S. I. Sumi, A. Hasegawa, S. Yagi, K. I. Miyoshi, A. Kanezawa, S. Nakagawa, and M. Suzuki, *J. Biotechnol., 2,* 59 (1985).
25. S. K. Sharma, *J. Biotechnol., 4,* 119 (1986).
26. P. P. Gray, P. Dunnill, and M D. Lilly, *Fermentation Technol. Today,* p. 347 (1972).
27. P. J. Hetherington, M. Follows, P. Dunnill, and M. D. Lilly, *Trans. Inst. Chem. Engr., 49,* 142 (1971).
28. D. C. Augenstein, K. Thrasher, A. J. Sinskey, and D. I. C. Wang, *Biotechnol. Bioeng., 16,* 1433 (1974).
29. B. C. Brickland, W. Richmond, P. Dunnill, and M. D. Lilly, in *Industrial Aspects of Biochemistry* (B. Spencer, ed.), EFSB, Amsterdam, 1974, pp. 65–79.
30. M. D. Lilly, *Appl. Biochem. Bioeng., 2,* 1 (1979).
31. J. J. Higgins, D. J. Lewis, W. H. Daly, F. G. Mosqueria, P. Dunnill, and M. D. Lilly, *Biotechnol. Bioeng., 20,* 159 (1978).
32. M. E. Winkler, M. Blaber, G. L. Bennett, W. Holmes, and G. A. Vehar, *Bio/Technol., 3,* 993 (1985).
33. P. A. Lowe, F. A. O. Marston, S. Angal, and J. A. Schoemaker, UK Patent Application GB 2138004A (1983).
34. T. Arakawa, N. K. Alton, and Y. R. Hsu, *J. Biol. Chem., 260,* 14435 (1985).
35. J. Uren, D. E. Robinson, and C. J. Scandella, European Patent Application 84306001.3 (1985).

36. P. A. Lowe, F. A. O. Marston, S. Angal, and J. A. Schoemaker, International Patent Application WO 84/03711 (1984).
37. D. V. Goeddel, D. G. Kleid, F. Bolivar, H. L. Heynecker, D. G. Yansura, R. Crea, T. Hirose, A. Kraszewski, K. Itakura, and A. D. Riggs, *Proc. Natl. Acad. Sci. U. S. A., 76,* 106 (1979).
38. R. E. Chance, J. A. Hoffmann, E. P. Kroeff, M. G. Johnson, E. W. Schirmer, and W. W. Bromer, in *Peptides: Synthesis, Structure and Function* (Proceedings of the Seventh American Peptide Symposium) (D. G. Rich and E. Gross, eds.), Pierce Chemical Co., Rockford, Illinois, 1981, pp. 729–738.
39. S. E. Builder and J. R. Ogez, U.S. Patent 4,511,502 (1985).
40. K. C. Olson and Rong-Chang Pai, U.S. Patent 4,511,503 (1985).
41. A. J. S. Jones, K. C. Olson, and S. J. Shire, U.S. Patent 4,512,922 (1985).
42. T. Staehelin, D. S. Hubbs, H. G. Kung, C. Y. Lai, and S. Pestka *J. Biol. Chem., 256,* 9750 (1981).
43. R. L. Easterday and I. Easterday, in *Immobilized Biochemicals and Affinity Chromatography* (R. B. Dunlap, ed.), Plenum, New York, 1974, pp. 123–124.
44. R. Axen, H. Brevin, and J. Carlsson, *Acta Chem. Scand., B29,* 471 (1975).
45. J. Porath, J. Carlsson, I. Olsson, and G. Belfuge, *Nature, 258,* 598 (1975).
46. F. E. Regnier and K. M. Gooding, *Anal. Biochem., 103,* 1 (1980).
47. B. R. Oakley, D. R. Kirsch, and N. R. Morris, *Ibid., 105,* 361 (1980).
48. J. H. Morrisey, *Ibid., 117,* 307 (1981).
49. W. C. McGregor, *Ann. N. Y. Acad. Sci., 413,* 231 (1983).
50. D. Perlman and H. Halvorson, *J. Mol. Biol., 167,* 138 (1983).
51. N. G. Mayne, H. M. Hsiung, J. D. Baxter, and R. M. Belagaje, in Ref. 2, pp. 135–144.
52. C. Ghelis and J. Yons (eds.), in *Protein Folding,* Academic, New York, 1983, p. 3.
53. M. R. Sairam and G. N. Bhargavi, *Science, 229,* 65 (1985).
54. T. Ima, H. Miyazaki, S. Hirose, H. Hori, T. Hayashi, R. Kageyama, H. Ohkubo, S. Nakanishi, and K. Muradami, *Proc. Natl. Acad. Sci. U. S. A., 80,* 7405 (1983).
55. P. S. Kaytes, N. Y. Thériault, R. A. Poorman, K. Murakami, and C.-S. C. Tomich, Unpublished Data.
56. M. A. Boxx, J. H. Kenten, C. R. Wood, and J. S. Emtage, *Nucleic Acids Res., 12,* 3791 (1984).
57. S. Cabilly, A. D. Riggs, H. Pande, J. E. Shively, W. E. Holmes, M. Rey, L. J. Perry, R. Wetzel, and H. L. Heyneker, *Proc. Natl. Acad. Sci. U. S. A., 81,* 3273 (1984).
58. D. Pennica, W. E. Holmes, W. J. Kohr, R. N. Harkins, G. A. Vehar, C. A. Ward, W. F. Bennett, E. Yelverston, P. H. Seeburg, H. L. Heyneker, and D. V. Goeddel, *Nature, 301,* 214 (1983).
59. T. W. Stickland and J. G. Pierce, *J. Cell Biochem., Suppl. 9B, 124,* Abst. No. 0724 (1985).
60. C. Ghelis, in Ref. 52, pp. 223–296.
61. G. Palade, *Science, 188,* 347 (1975).
62. R. Chance and J. Hoffman, U.S. Patent 4,421,685 (1983).
63. B. Frank, U.S. Patent 4,430,266 (1984).
64. S. Tsunasawa, J. W. Stewart, and F. Sherman, *J. Biol. Chem., 260,* 5382 (1985).
65. H. W. Lahm and S. Stein, *J. Chromatogr., 326,* 357 (1985).
66. F. Sherman, J. W. Stewart, and S. Tsunasawa, *Bioassays, 3,* 27 (1985).
67. S. Blumberg, D. Ben-Meir, D. Hadary, T. Vogel, M. Gorecki, H. Aviv, A. Levanon, J. Hartman, E. Zeelon, D. Kanner, M. Zeevi, D. Bartfeld, A. Nimrod, Z. Yavin, U. Factor, and B. Amit, *Biotech. Rep.,* p. A285 (1984).

68. B. Nilsson, E. Holmgren, S. Josephson, S. Gatenbeck, L. Philipson, and M. Uhlen, *Nucleic Acids Res., 13,* 1151 (1985).
69. J. B. Stern and R. C. Jackson, *Arch. Biochem. Biophys., 237,* 244 (1985).
70. K. Nagai and H. C. Thøgersen, *Nature, 309,* 810 (1984).
71. J. Shine, I. Fettes, N. C. Y. Roberts, and J. D. Baxter, *Ibid., 285,* 456 (1980).
72. J. Germino and D. Bastia, *Proc. Natl. Acad. Sci. U. S. A., 81,* 4692 (1984).
73. A. Light, H. S. Savithri, and J. J. Liepnieks, *Anal. Biochem., 106,* 199 (1980).
74. G. Keil, L. Haiml, and G. Suchanek, *Eur. J. Biochem., 111,* 49 (1980).
75. H. M. Sassenfeld and S. J. Brewer, *Bio/Technol., 2,* 76 (1984).
76. J. C. Smith, R. B. Derbyshire, E. Cook, L. Dunthrone, J Viney, S. J. Brewer, H. M. Sassenfeld, and L. D. Bell, *Gene, 32,* 321 (1984).
77. W. J. Rutter, European Patent Application 81300826.5 (1981).
78. D. B. Evans and S. K. Sharma, Unpublished Results.
79. K. H. S. Swamy and A. L. Goldberg, *Nature, 292,* 652 (1981).
80. A. L. Goldberg and A. C. St. John, *Annu. Rev. Biochem., 45,* 747 (1976).
81. K. Nath and A. L. Koch, *J. Biol. Chem., 246,* 6956 (1971).
82. C. H. Chung and A. L. Goldberg, *Proc. Natl. Acad. Sci. U. S. A., 78,* 4931 (1981).
83. S. A. Goff and A. L. Goldberg, *Cell, 41,* 587 (1985).
84. L. D. Simon, B. Randolph, N. Irwin, and G. Binkowski, *Proc. Natl. Acad. Sci. U. S. A., 80,* 2059 (1983).
85. D. A. Shub and N. J. Casna, European Patent Application 84305169.9 (1985).
86. K. Itakura, T. Hirose, R. Crea, A. D. Riggs, H. L. Heyneker, F. Bolivar, and H. W. Boyer, *Science, 198,* 1056 (1977).
87. D. Galanos and O. Lüderitz, in *Chemistry of Endotoxin* (E. T. Rietschel, ed.), Elsevier, New York, 1984, pp. 46–58.
88. L. L. Nelsen, *Pharm. Tech., 48,* 46 (1978).
89. E. L. Grinnan, in Ref. 1, pp. 99–108.
90. S. K. Sharma, *J. Appl. Biochem., 8,* 5 (1986).
91. A. Wang, S. D. Lu, and D. F. Mark, *Science, 224,* 1431 (1984).
92. S. J. Brewer and H. M. Sassenfeld, *Trends Biotechnol., 3,* 119 (1985).
93. C. R. Lowe, *Biotechnology, 1,* 3 (1984).
94. K. H. Kroner, H. Schutte, H. Hustedt, and M. R. Kula, *Process Biochem., 19,* 67 (1984).
95. G. Stree, *CRC Crit. Rev. Biotechnol., 1,* 59 (1983).
96. R. E. Smith and R. MacQuarrie, *Anal. Biochem., 90,* 246 (1978).
97. L. Guarente, G. Lauer, T. M. Roberts, and M. Ptashne, *Cell, 20,* 543 (1980).
98. B. Muller-Hill and J. Kania, *Nature, 249,* 561 (1974).
99. J. W. Guding, *J. Immunol. Methods, 20,* 241 (1978).
100. R. A. Hitzeman, F. E. Hagie, H. L. Levine, D. V. Goeddle, G. Ammerer, and B. D. Hall, *Nature, 293,* 717 (1981).
101. D. H. Hamer and P. Leder, *Ibid., 281,* 35 (1979).
102. I. Palva, M. Sarvas, M. L. Sibakov, and L. Kääriäinen, *Proc. Natl. Acad. Sci. U. S. A., 79,* 5582 (1982).
103. G. Gray, K. McKeown, A. Jones, P. Seeburg, and H. Heyneker, *Biotechnology, 2,* 161 (1984).
104. C. J. Thompson, J. M. Ward, and D. A. Hopwood, *J. Bacteriol., 151,* 668 (1982).
105. C. E. Ballou, in *The Molecular Biology of Yeast Saccharomyces* (J. N. Strathern, E. W. Jones, and J. R. Broach, eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1982, p. 335.
106. M. Carlson and D. Botstein, *Cell, 28,* 145 (1982).
107. J. Kurjan and I. Herskowitz, *Ibid., 30,* 933 (1982).

108. T. Oka, S. Sakamoto, K. I. Miyoshi, T. Fuwa, K. Yoda, M. Yamasaki, G. Tamura, and T. Miyake, *Proc. Natl. Acad. Sci. U. S. A., 82,* 7212 (1985).
109. N. G. Nossal and L. A. Heppel, *J. Biol. Chem., 241,* 3055 (1966).
110. T. J. Silhavy, H. A. Shuman, J. Beckwith, and M. Schwartz, *Proc. Natl. Acad. Sci. U. S. A., 74,* 5411 (1977).
111. M. Uhlén, B. Nilsson, B. Guss, M. Lindberg, S. Gatenbeck, and L. Philipson, *Gene, 23,* 369 (1983).

Received by editor January 13, 1986