Pure & Appl. Chem., Vol. 64, No. 1, pp. 169–184, 1992.
Printed in Great Britain.
© 1992 IUPAC

INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY

APPLIED CHEMISTRY DIVISION
COMMISSION ON BIOTECHNOLOGY*

Interrelations of Chemistry and Biotechnology—III†

PURIFICATION TECHNIQUES FOR BIOLOGICAL PRODUCTS

(Technical Report)

Prepared for publication by
ERICH HOCHULI
Research and New Technologies Department,
Hoffman-La Roche AG, Basel, Switzerland

*Membership of the Commission during the period (1989–91) in which this document was prepared was as follows:

Chairman: M. Ringpfeil (FRG); *Vice-Chairman*: J. L. Fox (USA); *Secretary*: L. M. Gierasch (USA); *Titular Members*: A. M. Boronin (USSR); H. K. Kölbl (FRG); H. G. W. Leuenberger (Switzerland); Y. Yamada (Japan); *Associate Members*: J. W. Engels (FRG); J.-M. Masson (France); M. Moo-Young (Canada); B. Nagel (FRG); S. Ramachandran (India); *National Representatives*: M. A. F. El-Sheikh (Arab Republic of Egypt); R. P. Gregson (Australia); M. van Montagu (Belgium); V. N. Beshkov (Bulgaria); L. Kjaergaard (Denmark); H. Dellweg (FRG); L. Pénasse (France); J. Holló (Hungary); T. K. Ghose (India); I. Chet (Israel); E. Rizzarelli (Italy); B. G. Yeoh (Malaysia); G. B. Petersen (New Zealand); O. Yenigün (Turkey); P. N. Campbell (UK); V. K. Eroshin (USSR).

†The Commission has chosen this series with a view to intensify the interrelations of chemistry and biotechnology. By improving the knowledge of chemists in the field of biotechnology it is hoped to initiate more ideas for applying biological methods in chemistry, inspire more use of chemical knowledge in the biological sciences and help scientists in both fields to work closer together. In the articles in the series, to be published in this journal, outstanding experts in their respective fields will give (a) an overview on topics related to the practical use of biotechnological methods in organic chemistry, (b) an outlook on upcoming research topics, their impact on existing areas and their potential for future developments.

The Commission solicits comments as well as suggestions for future topics, and will aim to help in providing answers to any questions in this field.

Republication of this report is permitted without the need for formal IUPAC permission on condition that an acknowledgement, with full reference together with IUPAC copyright symbol (© 1992 IUPAC), is printed. Publication of a translation into another language is subject to the additional condition of prior approval from the relevant IUPAC National Adhering Organization.

Interrelations of Chemistry and Biotechnology—III. Purification techniques for biological products (Technical Report)

Abstract- The purification of a biological product, e.g. an antibiotic, a vitamin, or a recombinant protein, from fermentation broth or cell culture supernatant is an important part of biotechnological manufacturing. Downstream recovery represents a large part of the product cost and therefore efficient and robust separation techniques are essential for the production of biological products. Recovering a biological product from the contaminants and impurities requires a series of purification steps, each removing some of the impurities and bringing the product closer to the final specifications.

The demand for improved purification techniques has increased over the past decade, following the successful molecular cloning and expression of recombinant proteins for research, diagnostic or therapeutic applications. In response to this demand, research activity in downstream processing has increased. In this review some new and innovative methods for purification of recombinant proteins will be discussed.

INTRODUCTION

When the technology for biosynthesis of the product of interest has been developed, it is the task of the chemist or biochemist to isolate the product and convert it into a form suitable for its required purpose. First, various methods of isolation and purification are tested on a laboratory scale. It is important to become acquainted with the properties of the compound to be isolated, such as its stability under various conditions and its solubility. The isolation experiments must be based on suitable analytical methods, chemical or biological, which make it possible to ascertain the concentration of the product, its purity, toxicity, etc.

Isolation procedures differ considerably depending on whether the product in question is contained inside the producing cell or in the culture fluid and depending on its concentration and stability. Usually, in the biosynthesis of antibiotics, vitamins etc., the culture liquid contains only small amounts of the desired compound, (typically about 2 to 3 % of the soluble components). Atkinson and Mavituna (1) have divided biotechnological processes into three categories:

Large scale processes

Typical large scale processes are alcohol or citric acid production. The downstream processing of the products is well understood. Recovery of alcohol and similar products requires for example distillation (2). Disposal of by-products and wastes presents severe problems in large scale processes.

Medium scale processes

These processes are concerned with the production of commodity organic chemicals such as Cephalosporin, Penicillin G, Streptomycin or extracellular enzymes. These products are chemically robust and conventional process operations can be used (3,4,5). The downstream processing problems are essentially a matter of costs.

Small scale processes

A wide range of materials and downstream processes are involved in small scale processes. Typical products of small-scale processes are intracellular enzymes or vaccines (6). Over the past decade there has been a rapidly increasing demand for highly purified proteins, following the successful molecular cloning and expression of recombinant proteins. The production of such highly purified proteins for use in research, diagnostic and therapeutic applications, requires a complex series of operations. In response to this demand, research activities in downstream processing have increased. In this paper some new and innovative methods for purification of recombinant proteins will be reviewed and their application potential will be illustrated with selected practical examples.

SPECIFIC PROBLEMS RELATED TO THE RECOVERY OF RECOMBINANT PROTEINS

About 12 years ago, the only way to obtain a protein for study was to isolate it from its natural source. The mammalian proteins whose structures were known were purified from easily obtainable tissue or blood extract. Today, it is possible to isolate a gene, put it into an appropriate expression system and produce the protein in a bacterial or animal cell host. Most proteins can be produced in amounts that are orders of magnitude higher than in their native environment. Unfortunately, they are often produced in a denatured, inactive form that accumulates intracellularly. The formation of these inactive proteins in bacterial systems, for example, appears to be independent of the type of protein (7). However, the biological activity is dependent on whether or not the protein has been correctly folded. The transition from a denatured and inactive form of a protein to the native three-dimensional structure is called protein folding. It is well known that a protein molecule can develop its three-dimensional structure spontaneously under physiological conditions (8). However, certain proteins may assemble poorly in the cytoplasm of the host organism. Many proteins contain disulfide bonds and the E.coli cytoplasm, for example, does not seem to be favorable for disulfide bond formation(9). Despite intense research, a generalized mechanistic understanding of the folding transition remains unclear (10). Protein folding during the manufacture of a recombinant protein is still largely empirical. The problem is to find the physiological conditions for a particular protein which allow spontaneous formation of its correct three-dimensional structure.

The problems of protein purification are in principle the same whether the protein is extracted from a natural source or from a recombinant organism. However, the same genetic techniques used in cloning can also be used to modify and control the location of a protein. It is possible for a protein to be produced intracellularly or to be secreted into the medium (11). An intracellular protein can occur in one of three forms; it may be soluble in the cytoplasm (12), it may occur in an aggregated form as inclusion bodies (13), or it may be secreted into the periplasm (14). In terms of downstream processing, these different locations can have dramatic effects, as does the potential for increasing the expression of the desired protein by using suitable promoter sequences (15). A discussion of purification techniques for recombinant proteins makes only sense in connection with the different expression systems. Gene expression and downstream processing are closely correlated; they are integral parts of the manufacturing process.

SYSTEMS FOR HETEROLOGOUS GENE EXPRESSION

Today five major expression systems are used routinely:
Escherichia coli, Bacillus subtilis, yeast, mammalian cells and the baculovirus system in insect cells. Each of these systems has its advantages and disadvantages.

Bacterial expression systems

Escherichia coli

Over the past years a lot of knowledge has accumulated concerning E.coli expression.

Genes expressed in the cytoplasm of E.coli characteristically accumulate to levels ranging up to 25 % of total cell proteins.

For translation initiation in E. coli an ATG sequence must proceed the gene coding sequence. Thus the primary products of translation contain an N-terminal methionine residue. E.coli possesses enzymes which catalyse the efficient removal of the methionine residues from natural proteins when required, but these enzymes do not work with the same efficiency on recombinant proteins and, therefore, expressed proteins may have an non-natural N-terminus (16).

This N-terminal methionine may affect the immunogenicity of the protein.

However, almost any protein can be produced in E.coli as long as it is not too long (optimal size 100-300 amino acids) and does not contain too many cysteins. The reducing environment present in E.coli cells does not permit cysteine-rich proteins to form the disulfide bonds required for proper three-dimensional structure (9). This problem can sometimes be overcome by secretion into a more oxidizing environment. Secretion from E.coli is still very difficult and is most likely to work with a naturally secreted protein. High-level expression and secretion of mammalian proteins, such as human growth hormone, into the E.coli periplasm have been reported (14). Most secretory proteins in prokaryotic as well as in eukaryotic cells are initially synthesized as precursors that have an amino-terminal extension. This amino-terminal extension, called the signal peptide, is capable of directing the secretory protein across the cytoplasmic membrane. During the translocation process, the signal peptide is cleaved and the mature protein is transported across the membrane. For the secretion of human growth hormone into the periplasm, a vector containing the genes coding for E.coli OmpA signal peptide and human growth hormone was constructed. Purified product produced with this system was properly folded with correct amino terminus and correct disulfide bonds.

Attempts to construct expression systems in E.coli to secrete heterologous proteins into the culture medium have also been successful (11).

A two-component vector system was used. One plasmid expresses bacteriocin release protein which renders the outer membrane permeable. The other encodes the synthesis of human growth hormone fused to the OmpA signal peptide, allowing the hormone to enter the periplasm. There are advantages when a heterologoulsy produced protein is secreted; problems with unwanted N-terminal methionines are avoided, proteins tend to fold better and protein purification can be much easier. However, secretion of recombinant proteins in E.coli quite often does not work and the product is accumulated in the denatured form in the cytoplasm.

Another problem with dramatic effects on downstream processing is the proteolytic degradation of the recombinant product. Interferon α-2a synthesized in E.coli was purified by immunoaffinity chromatography. Beside the properly folded product, misfolded and a truncated Interferon α-2a variants were copurified on the affinity adsorbent. It could be shown that the truncated variant, an Interferon α-2a molecule lacking the 22 N-terminal amino acids (17) is a product of proteolytic cleavage (18). In E.coli some twenty proteases have been characterized (19). The normal function of these enzymes are very diverse and may include degradation of aberrant proteins, degradation or modification of proteins as a means of metabolic control and removal of the signal peptide. The traditional method of reducing the effects of proteases during protein processing are to apply low temperature and protease inhibitors. However, precautions/measures already taken during the biosynthesis stage, such as the use of protease negative mutant hosts, may be important (20).

Expression of recombinant proteins in E.coli quite often results in the formation of inactive protein that accumulates intracellularly as inclusion bodies. In this form the proteins are protected from proteolytic degradation (21). However, these insoluble products can only be solubilized using denaturants such as urea or guanidine hydrochloride which unfold the protein. The unfolded protein must then be refolded by removal of the denaturant. Denaturant removal is typically performed through the use of dialysis or rapid dilution with a buffer. The primary refolding problem is the reformation of inactive protein aggregates. Therefore, it is essential to choose conditions which minimize the formation of aggregates during refolding. However, the maximum protein concentration for efficient refolding often is very low; for Interleukin-2 the maximum was only 1µg/ml (22). The use of cosolvents in aqueous systems has been shown to enhance protein refolding and decrease aggregation (23). Every protein contaminant extracted from the host organism which is present during refolding increases the total dilution necessary to avoid aggregation. In addition, partially denatured proteins are excellent substrates for proteases. Therefore purification steps that can be used in the presence of the denaturants can be very effective. These include gel filtration and metal chelate chromatography (24). In vitro folding reactions are carried out on completed polypeptide chains. In the cell, the N-terminus of a nascent protein may start folding before the remaining C-terminus has been synthesized in the ribosome. Proteins have been identified which assist in the mechanism of folding of other proteins into their biologically active forms(25).

The directed re-folding of a recombinant protein is still very empirical. However, by systematic variation of parameters influencing the folding and association of the constituent polypeptide chains, high yields of active eucaryotic proteins expressed in E.coli can be obtained. Recently, for example, the renaturation, purification and characterisation of recombinant murine antibody chains produced in E.coli were reported (26).

Bacillus subtilis

Bacillus subtilis is a well studied prokaryote with an extensive genetic system and well understood physiology. A further advantage of B. subtilis is its ability to secrete large amounts of proteins such as amylase and protease into the culture medium. Secreted proteins can be isolated without lysing the bacteria and therefore the purification should be easier. Expression systems in B. subtilis have been constructed. The E.coli β-lactamase was efficiently secreted from this host organism with the aid of the α-amylase signal sequence (27). However, some proteins seem to be efficiently secreted while others are poorly secreted. When human interferon α-2b was expressed in B. subtilis and secreted with the help of the α-amylase signal sequence, large amounts of the uncleaved precursor protein were accumulated in the cells (28). B. subtilis secretes several proteases into the culture medium which can potentially degrade the recombinant proteins that are also secreted into the medium. This problem of proteolytic degradation of the product can be solved by using a low protease mutant as host organism (27).

Yeast expression systems

The development of yeast genetic engineering has made possible the expression of heterologous genes and the secretion of their products from yeast (29).

Yeast, being eukaryotic, has much of the complex cell biology of multicellular organisms, including a highly compartmentalized intracellular organisation and a secretory pathway for proteins. Yeast grow rapidly, achieve high cell densities and carries out post-translational modifications, such as glycosylation, that bacterial systems are unable to do. Achieving high-level expression of a heterologous gene in yeast is possible at present. However, recovering large amounts of the desired gene product is sometimes very difficult. In yeast, the vacuole contains several endo- and exoproteinases.

These enzymes gain access to heterologous gene products and the proteins may be degraded after synthesis. Many of the genes that encodes these protease have been cloned and null mutants have been constructed. Use of these mutants can improve the yields of recombinant proteins that initially were degraded during purification.

The advantage of the yeast expression system is exemplified by the production of recombinant human serum albumin (30). This blood protein with a molecular weight of 65'000 has 17 disulfide linkages. It is misfolded and insoluble when produced intracellularly. Human serum albumin produced in yeast and secreted into the media is soluble and appears to have the same disulfide linkages as the human blood derived material.
Yeast is an excellent host system for heterologous gene expression but E.coli is still the first choice. However, in cases where the gene product of interest must be secreted and posttranslationally modified, yeast may be the next best expression system.

Mammalian cell expression systems

Microbial expression systems can be very effective and economical but when the proteins of interest are large and complicated in structure, the production of correctly folded and glycosylated biologically active molecules is very difficult or impossible. Mammalian cells secrete proteins very efficiently and they have the machinery to add and modify sugar residues such as those found in mammalian glycoproteins. These cell expression systems are therefore very attractive for the production of large glycoproteins and cell surface receptor proteins.
Chinese hamster ovary cells are often used for overexpression of recombinant proteins. High level of gene expression can be obtained with the method where gene amplification can be used to increase the vector copy number within the host cells. Titers of 180 mg/l product in shake flask cultures with the chinese hamster ovary cells system have been reported (31). However, stable mammalian cell expression is very time-consuming and the product titers usually are less then 1 mg/l.

Baculovirus system in insect cells

The baculovirus insect cell expression system is the most recent successful method for production of recombinant proteins (32). In this method baculovirus vectors are used for the expression of foreign genes in insect cell culture. Recombinant proteins produced in insect cells with baculovirus vectors are biologically active and seem to undergo post-translational processing to produce recombinant products very similar to the authentic proteins. In this system recombinant proteins can be secreted, N-glycosylated and O-glycosylated.
However, it is not clear whether the difference in the oligosaccharide processing pathways of insects and vertebrates is an important factor in the production of authentic mammalian glycoproteins in insect cells (33).

The purification of a protein is much easier if it can be isolated from cells that overproduce it. Therefore, the selection of the appropriate expression system to achieve high-level expression is very important for the purification of the recombinant protein. In general, the production of each recombinant protein presents its own peculiar set of problems that must be overcome to achieve high-level expression.

SEPARATION TECHNIQUES

The development of techniques for the purification of proteins has been an important prerequisite for many of the advance made in biotechnology.

Improvements in materials and equipment as well as in separation methods have made protein purification more predictable although it is still more an art than a science. Several useful reviews covering protein separation and purification have been written. In "Methods of Enzymology", volume 182 (34),for example, detailed application reports can be found.

With the development of a wide variety of chromatographic column packing materials such as gel-filtration media, ion exchanger, hydrophobic interaction adsorbents and affinity chromatography adsorbents a set of tools for efficient and selective protein fractionation are commercially available today. However, not all the problems in protein purification are solved by using these sophisticated tools. Difficulties still remain, for example, in finding the good conditions for protein extraction and for stabilizing the protein in solution as well as in choosing suitable methods for monitoring protein and biological activity. In addition, the physico-chemical characterisation of purified proteins is still not very sophisticated. Many minor variations are possible in protein molecules, such as deamidated asparagine or glutamine residues (35) and these can be very difficult to detect analytically. Such modifications can affect the biological or immunological properties of a protein. However, one of the most exciting recent protein characterisation methods is electrospray ionization mass spectrometry (36). This technique allows the determination of protein molecular weights down to an accuracy of 0,01 per cent mass units.

Extraction methods

Proteins which are secreted into the medium, e.g. proteins produced in mammalian or insect cells, can usually be obtained as a clear or nearly clear solution after removal of the cells by centrifugation or filtration. These solutions may be suitable for direct application to chromatography columns. Proteins which are accumulated intracellularly, e. g. proteins produced in bacteria, have to be extracted prior to purification by chromatography. Various methods are available to release proteins from the cells (37). Cell disruption techniques currently used are mechanical desintegrators, such as glass bead mills and high-pressure homoganizers. The extraction conditions are optimized by systematic variation of e. g. the composition of the extraction buffer, time, temperature and operating conditions of the desintegrator. The mechanical protein release methods have several drawbacks. One problem is that extensive fragmentation of the cells can make the subsequent clarification difficult (37). In addition, the high viscosity of the solution due to the released nucleic acids usually makes a nucleic acid removal step necessary (38). An alternative to mechanical disruption is to utilize chemicals which interact with the cell structure in a manner which causes it to become permeable. Guanidine hydrochloride is a chaotropic agent which is capable of solubilizing protein from membranes. Triton X-100 is a non-ionic detergent that has a high binding affinity for hydrophobic molecules. A combination of guanidine hydrochloride and Triton X-100 in the extraction buffer is a very effective permeabilization method (39). In addition to the chemical permeabilization, enzymatic lysis of microbial cells has found a number of important applications (40). The use of enzymes shows promise as a method for controlled lysis and selective product release.

The clarification of the cell homogenate or the crude extract obtained after cell permeabilization is usually done with a high speed centrifuge (15'000 x g). As a complement to centrifugation, tangential microfiltration has received increased attention (41) in recent years.

As a next step in protein purification, ultrafiltration has become an effective technique. Ultrafiltration membranes are availbale with different cut-off limits for separation of molecules from 1'000 Daltons up to 300'000 Daltons (42). The method is excellent for the separation of salts and other small molecules from a protein fraction with higher molecular weight. At the same time, the proteins can be concentrated gently and fast.

Chromatographic steps

Chromatography has proved to be a very efficient technique for the separation of proteins in biological extracts. Therefore the crude, clarified and concentrated extract of a recombinant protein is usually further purified by chromatographic steps. Separation by chromatography depends on the differential partition of proteins between a stationary phase (the adsorbent) and a mobile phase (the buffer solution). Normally, the stationary phase is packed into a vertical column of plastic, glass or stainless steel, whereas the buffer is pumped through this column. The sample to be fractionated is pumped on the top of the column and the various sample components travel with different velocities through the column and are subsequently detected and collected at the bottom of the column. A wide variety of adsorbents has been developed which exploit various properties of the protein for fractionation. The important techniques are summarized in Table 1.

Gel filtration

The use of column chromatography for the separation of proteins according to size became possible with the introduction of a cross-linked dextran material by Porath and Flodin (43). The separation in gel filtration depends on the different abilities of the various sample molecules to enter pores which contain the stationary phase. Very large molecules, which never enter the stationary phase, move through the chromatographic bed fastest. Smaller molecules, which can enter the gel pores, move more slowly through the column, since they spend a portion of their time in the stationary phase. Molecules are eluted according of decreasing molecular size. Gel filtration for the separation of biological molecules is used in the research laboratory as well as on a large scale with columns containing up to 500 l gel (44). Innumerable applications of gel chromatography for protein purification have been reported and reviewed in the literature (45).

Ion exchange chromatography

The basis for ion exchange chromatography is the competitive binding of proteins with differences in charge, to an oppositely charged chromatographic medium, the ion exchanger. The interaction between the proteins and the ion exchanger depends on several factors: net charge and surface charge distribution of the proteins, the ionic strength and the nature of the particular ions in the buffer, pH and other additives such as detergents. It is obvious that the more highly charged a protein is, the more strongly it will bind to an oppositely charged ion exchanger. Similarly, more highly charged ion exchangers bind proteins more effectively than weakly charged ones. Conditions, e.g. pH, which alter the effective charge on either the protein or the ion exchanger will affect their interaction and may be used to influence the ion exchange chromatography. Ion exchange chromatography for proteins was introduced in the middle 1950's with the cellulose ion exchangers. Since then a number of chromatographic media, such as cross-linked dextrans and agarose, synthetic hydrophilic polymers, have been derivatized with ion exchanging groups. A lot of empirical knowledge for this type of chromatography has been accumulated and actually ion exchange chromatography is the technique most often used in purification protocols (46).

Hydrophobic interaction chromatography

The knowledge of the three-dimensional structure of many proteins has revealed that the surfaces of globular proteins can have extensive hydrophobic patches in addition to the expected hydrophilic groups. In a solution favouring hydrophobic interaction, e.g. a solution containing a high salt concentration, these hydrophobic regions can bind to hydrophobic ligands such as alkyl or aryl side chains on the gel matrix. Fractionation of protein mixtures according to differences in the strength of interaction between the proteins and the hydrophobic ligands on the gel matrix is usually done by decreasing the salt concentration of the buffer system used. Hydrophobic interaction

TABLE 1: Chromatographic techniques

Molecular property exploited	*Type of chromatography*
- Size	Gel filtration
- Charge	Ion exchange chromatography
- Hydrophobicity	Hydrophobic interaction chromatography
- Biological affinity	Affinity chromatography (Immunoaffinity chromatography, Lectin affinity chromatography)
- Metal chelating groups	Immobilized metal ion affinity chromatography

TABLE 2: Affinity purification systems

Ligand	*Target protein*
Inhibitor, cofactor	Enzyme
Lectin	Glycoprotein Cell surface receptor
Hormone	Receptor
Antibody	Antigen

chromatography was first described by Porath (47) and Hjerten (48). Due to the stabilizing influence of salts, hydrophobic interaction chromatography is a mild method and recoveries are often high. However, the method is mainly used on a laboratory scale.

Affinity chromatography

Biological processes are related to specific interactions between molecules. These interactions may occur between a protein and a low molecular weight compound, e.g. an enzyme and an inhibitor or between two proteins, such as an antibody and an antigen. In affinity chromatography such biological affinities are exploited for the purification of proteins. One partner of the interaction pair, the ligand, is covalently bound to a solid matrix. The resulting affinity resin, packed into a chromatographic column, is then used to adsorb the protein of interest (the counterligand) from the crude extract that is passing the chromatographic column. With an appropriate elution buffer the purified protein can then be recovered from the affinity resin. Examples of affinity purification systems are listed in table 2.

The idea of purifying proteins by using a specific interaction between biological molecules was first published in 1910. Starkenstein attempted to purify α-amylase with insoluble starch (49). However, the development of affinity chromatography was delayed until satisfactory chemical methods for attaching ligands covalently to the matrix became available (50). Since then, very many proteins have been purified by affinity chromatography (51). An impressive example is the purification of 10 mg of the vitamin B_{12} transport protein from 40 kg of plasma using a column with immobilized cobalamin as ligand (52).

Lectins are proteins isolated from plant and animal sources which have a wide range of carbohydrate binding properties. The ability of immobilized lectins to interact specifically with sugars makes them excellent tools for purifying glycoproteins (53). Immobilized lectins are often used for the purification of recombinant glycoproteins from cell cultures.

The high specificity of monoclonal antibodies for their antigens makes them very useful ligands for affinity chromatography. Immobilized monoclonal antibodies are known as immunoadsorbents. The high costs of immunoadsorbents and the high risk of irreversible denaturation and proteolytic degradation, which is a problem with all adsorbents based on immobilized proteins, are some disadvantages of this affintiy technique. However, immunoaffinitiy chromatography has been widely used for the recovery of recombinant proteins.

The manufacture of recombinant Interferon α-2a synthesized in E.coli has been facilitated by affinity chromatography using immobilized monoclonal antibodies (54). Hybridoma techniques have been used for the production of monoclonal antibodies against partially purified human alpha interferons from induced leukocytes. Subsequently these antibodies were immobilized on crosslinked agarose and Interferon α-2a synthesized in bacteria was purified on this immunoadsorbent.

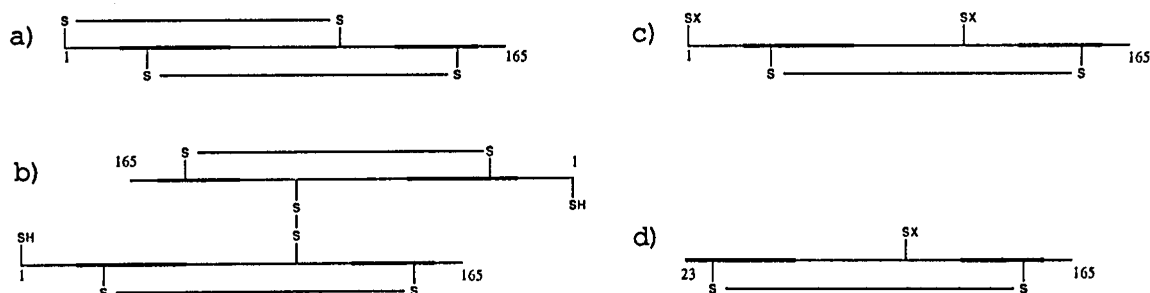

Fig 1: Schematic structure of the interferon α-2a variants eluted from the immunoadsorbent. The thick lines represents the sequence of the antigenic determinant.

a: correctly structured molecule
b: dimer
c: molecule with only one disulfide bridge
d: fragment

During scale up of the production process it was found that the Interferon α-2a eluted from the immunoaffinity column was not homogeneous. Analysis of the preparations by non-reducing sodium dodecyl sulfate-polyacrylamide gel electrophoresis revealed the presence of different molecular forms. Oligomers (dimers, trimers, etc.) and two monomeric forms were observed. Furthermore an interferon fragment was detected (17). The amino acid sequence of the antigenic determinant of the interferon molecule was determined in binding experiments using Interferon α-2a fragments prepared by cyanogen bromide cleavage. This antigenic sequence was in all the different molecular forms of Interferon α-2a eluted from the immunoadsorbent (figure 1). It is therefore understandable that the interferon fragment, the oligomers formed via intermolecular disulfide cross-linkage and the monomer with incorrect disulfide arrangement are recovered together with the correct interferon α-2a molecule. This example demonstrates a general limitation of immunoadsorption chromatography. An immobilized antibody will always recognize and bind a fragment of the target protein. It will not discriminate between molecules which have been modified or partially degraded. Therefore a one-step purification of a recombinant protein is very unlikely and other separation techniques have to be applied after immunoadsorption chromatography.

Immobilized metal ion affinity chromatography

Immobilized metal ion affinity chromatography (IMAC) for the purification of peptides and proteins was introduced in 1975 by Porath and co-workers (55).
The principle of this technique is the coordination between the electron donor groups on a protein (peptide) surface and immobilized transition metal ions. The tridentate chelator, iminodiacetic acid or the quadridentate chelator nitrilotriacetic acid (56), are coupled via a spacer arm to a solid support and used for the immobilization of metal ions such as Ni, Cu or Zn (figure 2).

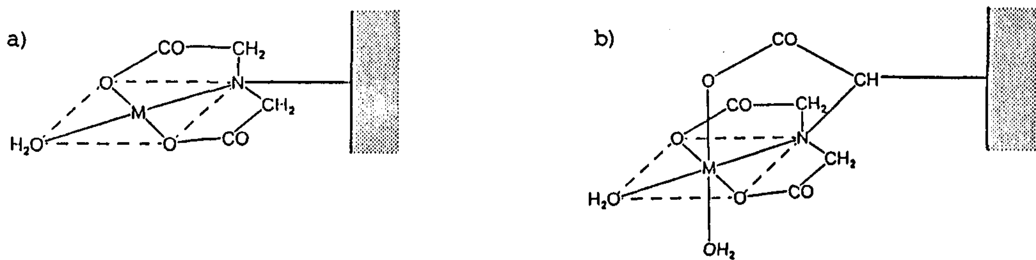

Fig. 2: Schematic structure of metal chelate adsorbents.

a) The tridentate iminodiacetic acid adsorbent charged with a metal ion having coordination number 4 (e.g. Cu^{2+}). One valency is occupied by H_2O and remains for biopolymer interaction.

b) The quadridentate nitrilotriacetic acid adsorbent charged with a metal ion of coordination number 6 (e.g. Ni^{2+}). Two valencies are occupied by H_2O and remain for biopolymer interaction.

Porath postulated that the histidine, cysteine and tryptophan residues in proteins (peptides) are most likely to form stable coordination bonds with metal chelates at neutral pH. Present experience (57) indicates that histidine residues on protein surfaces are the predominant electron donor groups.

We have used this affinity technique in our laboratory for the purification of many proteins and peptides. A recent example is the purification of "Anantin", a peptide antagonist of the atrial natriuretic factor (ANF), on a copper chelate affinity column (58). Anantin is a peptide consisting of 17 amino acids including one histidine.

IMAC is also used in industrial manufacture of recombinant protein drugs. A copper chelate affinity step is used for the separation of misfolded and truncated Interferon α-2a variants from the properly folded product (59).

Histidine is a rarely occurring amino acid, accounting for about 2 % of the amino acids in globular proteins (60). Only about half are exposed on the protein surface and may be available for immobilized metal ion affinity chromatography. Although this fact is limiting the utility of IMAC in the purification of natural proteins, the use of metal-affinity separation for purification of recombinant fusion proteins containing an artificial histidine affinity "tag" is becoming very attractive (see below).

GENETIC APPROACH TO FACILITATE PURIFICATION OF RECOMBINANT PROTEINS

In recent years very attractive genetic approaches to facilitate purification of recombinant proteins have been developed. The use of genetic engineering techniques has allowed the synthesis of hybrid proteins. By fusing the coding sequence of a protein of interest with the coding sequence of a polypeptide with high affinity to a ligand, a hybrid protein with an affinity tag can be produced directly by a microorganism. The affinity tag can then be used to recover the product from a culture medium or a cell lysate by affinity chromatography. It is also possible to introduce a specific cleavage site at the junction between the protein parts, which enables cleavage of the hybrid molecule to yield the protein of interest free of affinity tag. The concept is shown schematically in figure 3.

Purification tags have been developed to be applied in ion exchange, hydrophobic interaction, affinity, immunoaffinity and immobilized metal ion affinity chromato-graphy (IMAC). Table 3 lists some of the affinity tags and chromatographic techniques used.

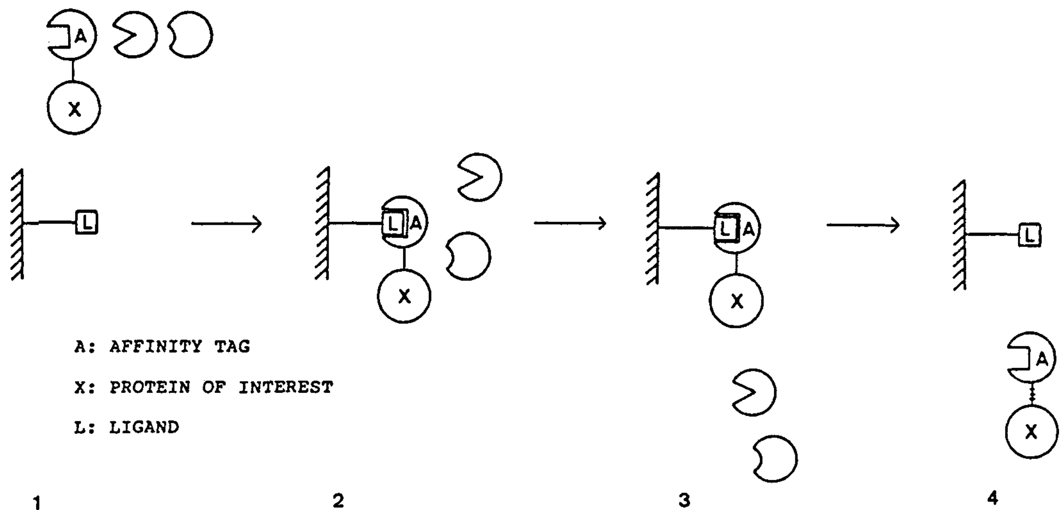

Fig. 3. Schematic illustration of the purification system.
A crude extract containing a hybrid protein consisting of the desired protein X fused with an affinity tag A is passed through an affinity column containing a specific ligand L (1). Only the protein molecules which have an appreciable affinity for the immobilized ligand will be retained (2) and all the contaminants are washed out (3). After elution, the purified fusion protein can be cleaved at the designed cleavage site (4) to give the desired protein X and the affinity tag A, which can then be removed by a second passage on the affinity column.

One of the first applications of this approach was the use of a polyarginine tag at the C-terminus of recombinant urogastrone which caused the protein to bind very strongly to a cation exchanger (61). In a similar manner galactosidase tagged with polyphenylalanine was purified on phenylsepharose (62). β-Galactosidase fusion proteins have been purified on a p-aminophenyl-β-D-thiogalactosidase affinity resin (63). Protein A fusion proteins rely on the specific interaction between staphylococcal protein A and the constant region of immunoglobulins. The fusion proteins are purified using an immobilized IgG (64). By using a short antigenic peptide at the N-terminus, proteins can be purified by using an immobilized monoclonal antibody specific to the peptide (65). The nitrilotriacetic acid metal chelate adsorbent charged with Ni^{2+} ions (figure 1) binds proteins with surface-adjacent histidnes. The addition of either an N-terminal or a C-terminal polyhistidine tag permits purification of the fusion protein on the metal chelate adsorbent (66). This chromatography permits purification in the presence of high concentration of guanidinium chloride and is, therefore, very useful for proteins extracted from inclusion bodies.

Removal of affinity tag

To split off the affinity tag from the protin of interest, a specific chemical or enzymatic cleavage site may be engineered into the fusion proteins. Table 4 lists some possibilities.

These methods are successful in several instances, but are often limited by poor cleavage yields or by unwanted cleavages that occur within the desired protein sequence. If the desired protein already contains the amino acid sequence of the cleavage site, the product will be degraded. When enzymatic methods are used, there is the problem of purifying the desired product from the contaminating protease. Chemical reagents in this respect can easily be separated from the cleavage mixture; a dialysis or desalting step can separate

TABLE 3: Affinity tags and separation techniques used to facilitate purification of recombinant proteins

Affinity tag	**Type of chromatography**	**Reference**
Polyarginine	Ion exchange	61
Polyphenylalanine	hydrophobic interaction	62
β-Galactosidase	affinity	63
Protein A	IgG-affinity	64
Antigenic peptide	Immunoaffinity	65
polyhistidine	IMAC	66

TABLE 4: Chemical and enzymatic methods to cleave fusion proteins

Cleavage reagent	**Cleavage site**	**Reference**
Cyanogen bromide	Met ↓	(71)
Formic acid	Asp ↓Pro	(67)
Hydroxylamine	Asn ↓Gly	(64)
Collagenase	Pro-Val ↓Gly-Pro	(68)
Factor Xa	Ile-Glu-Gly-Arg ↓	(69)
Enterokinase	Asp-Asp-Asp-Lys ↓	(65)
Renin	His-Pro-Phe-His-Leu-Leu ↓	(70)
Carboxypeptidase A	C-terminal aromatic amino acids	(66)
Carboxypeptidase B	C-terminal basic amino acids	(61)

the reagent from the protein mixture. The affinity tag can then be removed with a second run on the affinity column. From present experience it can be concluded that there is no generally applicable cleavage method. The optimal procedure for each protein of interest must be selected on a case-by-case basis, after analysis of the amino acid sequence. However, removal of the affinity tag is not always required.

CONCLUSIONS

The purification of a biological product from fermentation broth or cell culture supernatant is an important part of biotechnological manufacturing. Efficient separation processes are essential for the production of biochemical products. In recent years, the demand for improved techniques to purify proteins for use in research, diagnostic and therapeutic applications has increased following the successful expression of recombinant proteins. Following this need, research in downstream processing has been intensified,

focusing on the improvement of existing techniques as well as on the development of new and innovative methods.

Significant improvements in equipment using microprocessor based instruments and in materials such as new porous resin supports and new cross-linked beaded agarose have been made. However, the most exciting development in downstream processing was made possible by genetic engineering. Genetic engineering now provides means to tailor a protein in such a way that it can be purified more efficiently. Genetic techniques can be used, on one hand, to modify and control the location of a protein in the host cell (e.g. periplasm, culture medium) and, on the other hand, to produce fusion proteins which can be purified by specific chromatographic techniques. Many of these methods can be expected to be included in future manufacturing procedures.

REFERENCES

1. B. Atkinson and F. Mavituna (1983) Biochemical Engineering and Biotechnology Handbook (The Nature Press USA) 889.
2. A. Serra, M. Poch and C. Solà (1987) Process Biochemistry 22, 154.
3. R. Prasad, A.K. Gupta and R.K. Bajpai (1980) J. Chem. Techn. Biotechnol. 30, 324.
4. W. Voser (1982) J. Chem. Techn. Biotechnol. 32, 109.
5. R.C. Pandey, C.C. Kalita, R.J. White and M.W. Toussaint (1979) Process Biochemistry 14, 6.
6. A.L. van Wezel (1981) 1st Engineering Foundation Conference on Advances in Fermentation Recovery Process Technology, Banff.
7. F.A.O. Marston (1986) Biochem. J. 240, 1.
8. C.B. Anfinson (1973) Science 181, 223.
9. S. Pollit and H. Zalkin (1983) J. Bact 153, 27.
10. J.S. Fetrow, M.H. Zehfus and G.D. Rose (1988) Bio/Technology 6, 167.
11. H. M. Hsiung, A. Cantrell, J. Luirink, B. Oudega, A.J. Veros and G.W. Becker (1989) Bio/Technology 7, 267.
12. T. Staehelin, D.S. Hobbs, H.F. Kung and S. Pestka (1981) Methods Enzymol. 78, 505.
13. D.L. Hartley and J.F. Kane (1988) Biochem.Soc. Trans. 16, 101.
14. H.M. Hsiung, N.G. Mayne and G.W. Becker (1986) Bio/Technology 4, 991.
15. M. Lanzer and H. Bujard (1988) Proc. Natl. Acad. Sci. USA 85, 8973.
16. A. Ben-Bassat and K. Bauer (1987) Nature 326, 315.
17. E. Hochuli, D. Gillessen and H.P. Kocher (1987) J. Chromatogr. 411, 371.
18. S.J. Tarnowski, S.K. Roy, R.A. Liptak, D.K. Lee and R.Y. Ning (1986) Methods Enzymol. 119, 153.
19. K.H.S. Swamy and A.L. Goldberg (1982) J. Bacteriol. 149, 1027.
20. L.K. Strauch and J. Beckwith (1988) Proc. Natl. Acad. Sci USA 85, 1576.

21. S.-O. Enfors, H. Hellebust, K. Köhler, L. Strandberg and A. Veide (1990) Adavances in Biochemical Engineering/Biotechnology (A. Fiechter, ed.) Vol 43, 32. Springer-Verlag, Berlin.

22. M.P. Weir and J. Sparks (1987) Biochem. J. 245, 85.

23. J.L. Cleland and D.I.C. Wang (1990) Bio/Technology 8, 1274.

24. E. Hochuli (1990) Genetic Engineering, Principle and Methods (J.K. Setlow, ed.) Vol. 12, 87.

25. A.L. Horwich, W. Neupert and F.-U. Hartl (1990) TIBTECH 8, 126.

26. J. Buchner and R. Rudolph (1991) Bio/Technology 9, 157.

27. I. Palva, M. Sarvas, P. Lehtovaara, M. Sibakov and L. Kääriäinen (1982) Proc. Natl. Acad. Sci USA 79, 5582.

28. C.H. Schein, K. Kashiwagi, A. Fujisawa and Ch. Weissmann (1986) Bio/Technology 4, 719.

29. S.D. Emr (1990) Methods Enzymol. 185, 231.

30. D. Sleep, G.P. Belfield and A.R. Goodey (1990) Bio/Technology 8, 42.

31. M.I. Cockett, C.R. Bebbington and G.T. Yarranton (1990) Bio/Technology 8, 662.

32. V.A. Luckow and M.D. Summers (1988) Bio/Technology 6, 47.

33. T.D. Butters and R.C. Hughes (1981) Biochim.Biophys. Acta 640, 655.

34. Methods in Enzymology (1990) volume 182, M.P. Deutscher, ed.

35. V.R. Anicetti, B.A. Keyt and W.S. Hancock (1989) TIBTECH 7, 342.

36. R.D. Smith, J.A. Loo, Ch.G. Endmonds, Ch.J. Barinaga and H.R. Udseth (1990) Anal. Chem. 62, 882.

37. I. Agerkvist and S.-O. Enfors (1990) Biotechnol. Bioeng. 36, 1083.

38. J.J. Higgins (1978) Biotechnol.Bioeng. 20, 159.

39. D.H. Hettwer and H. Wang (1989) Biotechnol. Bioeng. 33, 886.

40. B.A. Andrews and J.A. Asenjo (1987) TIBTECH 5, 273.

41. D. Derise and V. Gekas (1988) Process Biochemistry 23, 105.

42. H. Strathman (1985) Trends in Biotechnology 3, 112.

43. J. Porath and P. Flodin (1959) Nature 183, 1657.

44. D. Ambrose, E. Bayer, O. Samuelson (1974) Pure and Appl. Chem. 37, 437.

45. W.W. Yan, J.J. Kirkland, D.D. Bly (1979) Modern Size-Exclusion Liquid Chromatography Practice of Gel Permeation and Gel Filtration Chromatography. Wiley-Interscience, New York.

46. J. Bonnerjea, S.Oh, M. Hoare and P. Dunhill (1986) Bio/Technology 4, 954.

47. J. Porath, L. Sundberg, N. Fornstedt, I. Olson (1973) Nature 245, 465.

48. S. Hjerten, J. Rosengren, S. Pahlman (1974) J. Chromatogr. 101, 281.

49. E. Starkenstein (1910) Biochem. Z. 24, 210.

50. J. Porath, R. Axen and S. Ernback (1967) Nature 215, 1491.

51. M. Wilchek, T. Miron, J. Kohn (1984) Methods in Enzymol. 104, 3.

52. R.H. Allen, P.W. Majerus (1972) J. Biol. Chem. 247, 7709.

53. D.M. Swallow, L. Evans and D.A. Hopkinson (1977) Nature 269, 261.

54. T. Staehelin, D.S. Hobbs, H.F. Kung, C.Y. Lai and S. Pestka (1981) J. Biol. Chem. 256, 9750.

55. J. Porath, J. Carlsson, J. Olsson and G. Belfrage (1975) Nature 258, 598

56. E. Hochuli, H. Döbeli and A. Schacher (1987) J. Chromatogr. 411,177.

57. S.E. Hemdan, Y. Zhao, E. Sulkowski and J. Porath (1989) Proc. Natl. Acad. Sci. USA 86, 1811.

58. W. Weber, W. Fischli, E. Hochuli, E. Kupfer and E.K. Weibel (1991) J. Antibiotics 44, 164.

59. E. Hochuli (1986) Chimia 40, 408.

60. M.H. Klapper (1977) Biochem. Biophys. Res. Comm. 78, 1018.

61. H.M. Sassenfeld and S.J. Brewer (1984) Bio/Technology 2, 76.

62. M. Persson, M.G. Bergstrand, L. Bulow and K. Mosbach (1988) Anal. Biochem. 172, 330.

63. A. Ullman (1984) Gene 29, 27.

64. T. Moks, L. Abrahmsen, B. Österlöf, S. Josephson, M. Östling, S.-O. Enfors, I. Persson, B. Nilsson and M. Uhlen (1987) Bio/Technology 5, 379.

65. T.P. Hopp, K.S. Prickett, V.L. Price, R.T. Libby, C.J. March, D.P. Cerretti, D.L. Urdal and P.J. Conlon (1988) Bio/Technology 6, 1204.

66. E. Hochuli, W. Bannwarth, H. Döbeli, R. Gentz and D. Stüber (1988) Bio/Technology 6, 1321.

67. P.R. Szoka, J. Murthy, S. Stoufer, A. Schreiber and H. Chan (1985) DNA 4, 82.

68. J. Germino and D. Bastia (1984) Proc. Natl. Acad. Sci. USA 81, 4692.

69. K. Nagi and H.C. Thogerson (1987) Methods Enzymol. 153, 461.

70. M.L. Haffey, D. Lehman and J. Borger (1987) DNA 6, 565.

71. K. Itakura, T. Hirose, R. Crea, A. Riggs, H.L. Heyneker, F. Bolivar and H.W. Boyer (1977) Science 198, 1056.