EMERGING TRENDS IN DESIGNING SHORT AND EFFICIENT PROTEIN PURIFICATION PROTOCOLS

Saurabh Gautam, Joyeeta Mukherjee, Ipsita Roy and Munishwar Nath Gupta

Department of Chemistry, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India

Department of Biotechnology, National Institute of Pharmaceutical Education and Research (NIPER), Sector 67, S.A.S. Nagar, Punjab 160 062, India

ABSTRACT

Protein purification is carried out in both academic and industrial sectors. Both non-chromatographic approaches (precipitation, crystallization, partitioning) and chromatographic methods are used. When recombinant proteins are over expressed in foreign hosts, inclusion bodies are obtained. Many purification methods can be used to refold proteins. Many strategies allow one to avoid formation of inclusion bodies. The concept of affinity based separation extends to non-chromatographic methods as well. Affinity precipitation and Macro-(affinity ligand) Facilitated Three Phase Partitioning (MLFTPP) are two examples of this. Produced as fusion proteins, most of the proteins are purified via interaction of their affinity tag coupled to a solid matrix. Nanotechnology has made magnetic based methods even more powerful. Expanded bed chromatography, aqueous two phase systems and MLFTPP allow one to deal directly with feed containing particulate matter. Membrane based separations are another powerful option. Finally crystallization at the industrial level has further evolved to reoccupy an important place in protein purification. Overall, the main trends are: integrate upstream and downstream phases and reduce the number of steps required for purifying a protein. This has become possible by adopting variety of non-chromatographic methods. More important, bringing affinity based steps earlier in the purification have helped achieving this.

Keywords: Three Phase Partitioning, Macro-(Affinity Ligand) Facilitated Three Phase Partitioning, Expanded Bed Chromatography, Affinity Precipitation, Protein Refolding

1. INTRODUCTION

1.1. Early Developments

In the beginning, scientists did not believe that proteins were macromolecular in nature. This school of thought believed proteins to be “heterogenous aggregates of various small molecules, presumably peptides of moderate size” (Edsall, 1992). Hence, the early efforts of protein purification involved crystallization, just as chemists use crystallization to obtain pure compounds. Hermann Staudinger was probably the first one to proclaim that large molecular weight polypeptides can exist. Svedberg was a strong supporter of colloidal nature of proteins. Ironically, ultimately his ultracentrifugal study of proteins convinced him (and everybody else!) that proteins were indeed macromolecular in nature. This realization about the fundamental nature of proteins was necessary for subsequent development of various other techniques for protein purification (Edsall, 1992).

Nevertheless, the early purification efforts were motivated by scientists wanting to look at the structure of proteins by X-ray crystallography. So, crystallization continued to be the major technique for protein purification for quite some time. The crystallization of urease by Sumner and subsequent extensive results on
protein crystallization by Northrop and Kunitz are well documented (Dixon and Webb, 1964; Scopes, 1994). The paradigm shift in the area happened because of the world war. The serum proteins became a focus of study. Cohn had earlier accepted Edsall in his laboratory to work on muscle proteins. During the war years, along with Edsall, he carried out pioneering work on fractionation of plasma proteins (Edsall, 1992). The solubility properties and crystallization still dominated the design of early studies. Much of what we know about precipitation of proteins by metal ions and organic solvents was discovered by this group. The idea behind narration of these early studies is not merely to provide historical perspective. We will describe later how these early results form the basis of current trends in protein bioseparation.

It is obvious that approaches used by chemists for purification of organic compounds formed the basis of the early methods which were developed for protein purification. Precipitation and crystallization are routinely used by the chemists as purification strategies. So, it is not surprising that chronologically the next major approach used for protein purification was chromatography. As chemists started using silica and ion exchangers, protein purification was also attempted on similar chromatographic materials. Ion exchangers began to be used and have become indispensable in protein purification. With time, Whatman introduced cellulose based ion exchangers followed by introduction of dextran based ion exchangers by Pharmacia. Each event was some kind of a mini-revolution.

The only purification tool which is unique in nature and did not have its origin in chemical methods of purification is affinity chromatography. The idea of biological affinity is a central theme in biochemistry. Biocatalysis, signal transduction, DNA replication, translation of mRNA, active transport across membranes; all are based upon selective molecular recognition. The high selectivity inherent to affinity chromatography resulted in high resolution of protein mixtures. This was an important milestone.

For a couple of decades, a typical protein purification protocol invariably consisted of precipitation by ammonium sulfate, one or two ion exchange steps, gel filtration and finally an affinity chromatography step. The cost of affinity media used to be quite high (and still is, as compared to other chromatographic media). So, affinity chromatography was usually considered a “polishing step” with a small column to remove impurities from a small amount of purified protein.

The advent of recombinant DNA impacted purification in two ways:

- Firstly, cloning and protein engineering meant that sources of proteins were not limited to naturally occurring animals, plants and microbes. The protein production activities multiplied several fold. Scientists were looking for shorter protocols and faster methods
  - Use of “affinity tags” and production of proteins in the form of fusion proteins became possible

Today, in academic sector, protein purification is almost exclusively carried via commercially available kits which are based upon fusion protein technology. Like all technologies, this approach is great if it works. To be fair, more often than not it works. When it does not, biochemists have to rely upon vendors for troubleshooting. Enzymology as a discipline in academic sector is fast disappearing. In industrial sector as well fusion protein approach is used. However, industry has additional constraints like worrying about economics and scale of operation. So, much of what is discussed in the present review relates to developments which were necessitated because of requirements of the industry.

From that perspective, there is a need to appreciate that different industrial applications of proteins have different requirements of protein purity. The trade off paradigm between cost, purity and yield is shown in Fig. 1.

### 1.2. Selective Cell Permeabilization

Many enzymes are intracellular in nature. For example, Alcohol Dehydrogenase (ADH) catalyzes redox reactions and is being increasingly used in obtaining chiral building blocks (Kula and Kragl, 2000). This enzyme generally occurs intracellularly in many microorganisms. Another example is that of β-galactosidase, which finds applications in whey hydrolysis and in obtaining low lactose milk (Khare and Gupta, 1990; Adlercreutz and Straathof, 2000). Cell lysis to obtain such enzymes can be carried out in many ways: Mechanical, chemical and enzymatic methods have all been described (Hancock, 1984; Chisti and Mooyoung, 1986; Raghava and Gupta, 2009; Jamur and Oliver, 2010). As a part of integration between upstream and downstream processes, efforts have been made towards permeabilization of cells for selective release of proteins. Release of some cytoplasmic proteins in a selective way after osmotic shock has been fairly well documented (Hantash et al., 1997; Berrier et al., 2000; Ewis and Lu, 2005). Kuboi et al. (1995) have attempted disruption of E. coli to optimize selective release of β-galactosidase. Vazquez-Laslop et al. (2001) described a molecular sieve mechanism in which osmotic shock given to Ethylenediaminetetraacetic Acid (EDTA) treated E. coli cells allowed proteins <100 kDa to be released.
With an increase in the complexity of the purification protocol, there is an increase in the purity of the protein and the cost of the purification protocol. This also results in the decrease in the yield of the desired protein.

"Reprinted with permission from (Mondal et al., 2006). Copyright 2006 by American Chemical Society.

More recently, Three-Phase Partitioning (TPP) of cells was developed as a general method by which sieves in the cells become progressively bigger (Raghava and Gupta, 2009). This made it possible to separate proteins on the basis of their sizes. TPP is a versatile technique which has been used both at the cellular level as well as at the molecular level. The technique has found applications in edible oil extraction (Sharma et al., 2002), oil extraction for biodiesel formation (Shah et al., 2004), separation of anti-nutritional factor from plant proteins (Saxena et al., 2007), protein separation and protein refolding (Dennison and Loverien, 1997; Jain et al., 2004; Roy et al., 2004a; Raghava et al., 2008). The last two aspects would be discussed in more detail later in this review. TPP essentially consists of mixing the right amount of salt and an organic solvent to a protein mixture or cell mass. In the context of cell permeabilization, it was found that TPP treatment of bacterial or yeast cells led to the cells being obtained in the permeabilized form in the interfacial layer between the upper organic solvent rich phase and lower water rich phase. Pretreatment of the cells with organic solvent before TPP controlled the extent of permeabilization. While higher pre-incubation with organic solvent (Fig. 2A) released most of the intracellular proteins, the high molecular weight ADH was retained (Fig. 2B) and could be obtained in almost pure form by cell lysis of these highly permeabilized cells.

1.3. Aqueous Two Phase System (ATPS)

Albertsson (1972) described a protein purification approach based upon partitioning of macromolecules between two water rich (generally as high as 70-95%) phases. These two systems are obtained either by using two water soluble polymers (frequently used polymers are Polyethylene Glycol (PEG), dextran and modified starch) or a water soluble polymer like PEG with a salt like sodium phosphate. While the technique is quite popular with industry, it is seldom used for protein purification in academic sector. The fold purification achieved is not high. The main reason why it is preferred in the industry is that it is one of the very few techniques which can work well with feed containing suspensions. Based upon sound physical chemistry principles, it is also scalable in a more predictable way (Sutherland et al., 2011). The parameters which influence the selectivity in an ATPS include molecular weight of the polymers, pH, temperature and concentration of the electrolytes (Walter and Johansson, 1994; Cabezas, 1996).

Very soon, the selectivity of this process was enhanced by linking affinity ligands to one of the polymers. PEGylation is used in the pharmaceutical industry for drug delivery purposes (Veronese and Pasut, 2005). So, the chemistry of linking other molecules to PEG developed very fast (Veronese, 2001). It is not surprising that in most of the examples of aqueous two phase affinity systems, the affinity ligand was linked to PEG which is the upper phase in both PEG-dextran and PEG-salt phases.

Two main concerns in the use of ATPS in industrial enzymology have been (a) cost of polymers, (b) separating the polymer and the protein(s). The latter also impacts the first factor as it affects the recyclability of the polymers. Incorporating a smart polymer linked affinity ligand (instead of using a PEG or dextran linked affinity ligand) seems to be a good solution. This approach was first described by Kamihira et al. (1992) and involves integration of ATPS with affinity precipitation. The smart methylmethacrylate Eudragit S-100 was coupled with human IgG and incorporated into PEG-Reppal PES (a modified starch two phase system). Recombinant protein A was recovered from E. coli cell homogenates.
The affinity complex of Eudragit S-100-IgG with protein A was precipitated by lowering the pH to 4.5 and protein A eluted off the affinity complex precipitate by washing with the pH 2.5 buffer. The authors report <5% loss of Eudragit-IgG during each cycle but PEG could also be recovered and reused. In our laboratory at IIT Delhi, the approach was successfully used to purify α-amylase (Teotia and Gupta, 2001a), β-amylase (Teotia and Gupta, 2001a), pullulanase (Teotia and Gupta, 2001b), xylanase (Teotia and Gupta, 2001b), phospholipase D (Teotia and Gupta, 2004), chitinases (Teotia et al., 2004) and a few plant lectins (Teotia et al., 2006). As has been explained later in the review, smart polymers themselves were used for the “affinity” capture of the target protein. We believe that this approach has been overlooked and may be found more useful in industrial enzymology.

1.4. Use of Affinity Tags

Recombinant technology made it possible to attach an “affinity tag” while producing the protein. Such tags are also called “fusion tags” since such recombinant proteins are referred to as “fusion proteins”. The fusion protein then binds to an appropriate affinity media because of the presence of the “affinity tag”. The technique relies upon “affinity pairs” just like affinity chromatography. However, as only peptides and polypeptides can be linked by recombinant approach this
technique has given rise to search for new “affinity pairs”. An illustrative list of such “affinity pairs” is given in Table 1. Expressing proteins as fusion proteins has become the most powerful and most frequently used way of purifying recombinant proteins. This is one instance where scientists working in the academic sector have unwittingly exploited the virtue of using affinity interactions right early in the purification protocol. This strategy as an option has been fairly well reviewed from time to time (Lichty et al., 2005; Waugh, 2005; Malhotra, 2009).

Apart from acting as a convenient handle for purifying proteins, many affinity tags also result in enhancement of solubility of expressed proteins. Such “solubility tags” include Maltese Binding Protein (MBP), thioredoxin, NusA (N-utilizing substance a transcription anti-termination factor), Small Ubiquitin Like Modifier (SUMO), HaloTag, GB1 tag (Malhotra, 2009). In many cases, it is suggested that fusion proteins are expressed under other conditions which are known to improve protein folding (Wiseman, 1995; Sahdev et al., 2008). It should be added that the usefulness of any tag varies from protein to protein. So, one needs to screen different tags. So, it is not necessary that a fusion protein would not go into inclusion bodies. In fact, several MBP fusions are expressed as inclusion bodies and solubilized by refolding (Malhotra, 2009; Gautam et al., 2012a. MBP is considered as one of the best choices for enhancing the solubility of proteins. New England Biolabs Inc. (NEB) sells MBP tagging vectors for cytoplasmic and periplasmic expression of the fusion proteins (www.neb.com). Recently, Miladi et al. (2011) have described purification of Tobacco Etch virus protease as a Streptag II fusion protein and found it as a better tag than His6 in respect of enhancing solubility as well as overall purification. In many cases, such as for protein crystallization experiments, it is desirable to remove the tag. The tag removal can be carried out enzymatically or even chemically (www.qiagen.com; Fairlie et al., 2002; Arnau et al., 2006; Hu et al., 2008) or via an autocatalytic route, as in the case of INTEIN tag (Singh and Gupta, 2008; Wang et al., 2010).

Lichty et al. (2005) have compared eight tags from the perspective of protein purification and concluded that His tag generally provided good yields although purity varies from system to system. This is quite understandable and it may always be a prudent course to screen various tags for a specific purpose. It should be remembered that it is not just the tag, or even fusion protein. Results may vary depending upon the affinity media used. For example, different affinity media are available from GE Healthcare and NEB for purifying MBP fusions. Waugh (2005) mentions that combinatorial tagging is a good approach. Many crystallography and structural genomics groups use a dual His6-MBP tag, MBP for increasing the solubility, His6 for purification via Immobilized Metal Affinity Chromatography (IMAC). Cass et al. (2005) described purification of several proteins from mammalian tissue culture by tagging both Streptag II and His8 to the C-terminal ends of the recombinant proteins. Very high purity (>99%) with varying yields (29-81%) were reported. Yeliseev et al. (2007) described the use of dual affinity tags for expression and purification of a Cannabinoid receptor. Kenig et al. (2006) have cautioned that the fusion protein approach for protein purification works less efficiently with oligomeric proteins.

Finally, it should be mentioned (see the following section on precipitation for more details) that it is not necessary that a chromatography step be used for the purification of a fusion protein. Very early, purification of His tagged protein was described by the Lund group (Lillius et al., 1991). We have described the affinity precipitation of an INTEIN tagged lipase by chitosan, a pH sensitive polymer (Singh and Gupta, 2008).

1.5. Affinity Precipitation for Protein Purification and Refolding

As has been mentioned in the introduction, precipitation has been used since early days as a tool for protein purification. Precipitation by polymers like dextran, PEG and polyacrylic acid (Ingham, 1990; Glatz, 2000; Patrickios et al., 1994) has been demonstrated. Precipitation by smart polymers offers additional advantages. These water soluble polymers are reversibly soluble-insoluble, i.e., a change in pH, temperature, light or the presence of a chemical species like Ca$^{2+}$ or K$^+$ can change their solubility in water drastically. In view of this behavior, such polymers are also called stimuli-sensitive polymers (Roy and Gupta, 2003a). It has been found that many of these smart polymers have inherent affinity for many enzymes/proteins (Senstad and Mattiasson, 1989; Gupta et al., 1993; Roy and Gupta, 2003a; Mondal et al., 2006). Hence many enzymes/proteins could be purified by precipitation in the form of their affinity complexes (polymer-protein) by altering the solubility of the polymer with an appropriate stimulus. The protein/enzyme could be dissociated from such complexes fairly easily by altering pH, ionic strength. Over the years, it was found that this process (called affinity precipitation for obvious reasons) could also refold thermally or urea denatured proteins (Roy and Gupta, 2003b; Yoshimoto et al., 2003; Freitag et al., 2007; Mondal et al., 2007).
Table 1. Different fusion tags for purification and enhancing solubility of the proteins. Number of hits in Google Scholar (as on 20th September, 2012) for each fusion tag as a rough indication of the use of each of them by the scientific community is given in each case.

<table>
<thead>
<tr>
<th>Affinity tags</th>
<th>Tag</th>
<th>Affinity resin</th>
<th>Number of hits in Google scholar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>His</td>
<td>Immobilized metal ions (Nickel, Copper, Zinc, Cobalt)</td>
<td>998,000</td>
</tr>
<tr>
<td></td>
<td>FLAG</td>
<td>Immobilized anti-FLAG monoclonal antibodies</td>
<td>184,000</td>
</tr>
<tr>
<td></td>
<td>GST (Glutathione S-transferase)</td>
<td>Immobilized glutathione</td>
<td>64,500</td>
</tr>
<tr>
<td></td>
<td>MBP (Maltose binding protein)</td>
<td>Amylose, immobilized dextrin</td>
<td>22,500</td>
</tr>
<tr>
<td></td>
<td>Strep II</td>
<td>Immobilized Strept-Tactin (engineered streptavidin)</td>
<td>18,200</td>
</tr>
<tr>
<td></td>
<td>CBD</td>
<td>Chitin</td>
<td>9,540</td>
</tr>
<tr>
<td></td>
<td>Halo</td>
<td>Immobilized chloroalkane</td>
<td>849</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Solubility tags</th>
<th>Tag</th>
<th>Affinity resin</th>
<th>Number of hits in Google scholar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MBP</td>
<td>Maltose binding protein</td>
<td>22,500</td>
</tr>
<tr>
<td></td>
<td>Trx</td>
<td>Thioredoxin</td>
<td>17,800</td>
</tr>
<tr>
<td></td>
<td>SUMO</td>
<td>Small ubiquitin related modifier</td>
<td>11,700</td>
</tr>
<tr>
<td></td>
<td>GB1</td>
<td>B1 domain of protein G</td>
<td>3,970</td>
</tr>
<tr>
<td></td>
<td>NusA</td>
<td>N utilizing substance A</td>
<td>2,720</td>
</tr>
</tbody>
</table>

Table 2. High-throughput screening of the affinity ligand (smart polymer) in 96-well plate for appropriate refolding of the proteins. The promising conditions are shown in bold. Columns 1 to 9 indicate inclusion bodies in 8 M urea or dilution control of solubilized inclusion bodies with appropriate buffer instead of the smart polymer or solubilized inclusion bodies with different smart polymers and rows A to E indicate different inclusion bodies of different proteins. The numbers in the parentheses below the fluorescence emission maxima (λ\text{max}) values indicate the ratio of the fluorescence intensity at that λ\text{max} (emission) to the fluorescence intensity of the protein solution in 8 M urea at λ\text{max} (emission). “Reprinted with permission from (Gautam et al., 2012a). Copyright 2012 by Elsevier”

<table>
<thead>
<tr>
<th>Fluorescence emission maxima (λ\text{max})</th>
<th>Proteins</th>
<th>Inclusion bodies in 8 M urea</th>
<th>Dilution control</th>
<th>Eudragit E (L-100)</th>
<th>Eudragit S (S-100)</th>
<th>Protanal LP</th>
<th>Algic Acid</th>
<th>Cationic Starch</th>
<th>κ-Carrageenan</th>
<th>Chitosan</th>
</tr>
</thead>
<tbody>
<tr>
<td>A CcdB-358</td>
<td>358 nm</td>
<td>355 nm</td>
<td>(1.093)</td>
<td>(1.081)</td>
<td>340 nm</td>
<td>(1.082)</td>
<td>351 nm</td>
<td>352 nm</td>
<td>354 nm</td>
<td>355 nm</td>
</tr>
<tr>
<td>F17P 356 nm</td>
<td>352 nm</td>
<td>342 nm</td>
<td>(1.080)</td>
<td>(1.082)</td>
<td>350 nm</td>
<td>(1.081)</td>
<td>352 nm</td>
<td>352 nm</td>
<td>354 nm</td>
<td>355 nm</td>
</tr>
<tr>
<td>malETx</td>
<td>357 nm</td>
<td>354 nm</td>
<td>340 nm</td>
<td>342 nm</td>
<td>352 nm</td>
<td>358 nm</td>
<td>353 nm</td>
<td>355 nm</td>
<td>354 nm</td>
<td>354 nm</td>
</tr>
<tr>
<td>CD4D12</td>
<td>360 nm</td>
<td>353 nm</td>
<td>354 nm</td>
<td>355 nm</td>
<td>340 nm</td>
<td>346 nm</td>
<td>352 nm</td>
<td>354 nm</td>
<td>352 nm</td>
<td>352 nm</td>
</tr>
<tr>
<td>ScFv b12</td>
<td>360 nm</td>
<td>352 nm</td>
<td>354 nm</td>
<td>355 nm</td>
<td>340 nm</td>
<td>345 nm</td>
<td>352 nm</td>
<td>354 nm</td>
<td>355 nm</td>
<td>355 nm</td>
</tr>
<tr>
<td>E ScFab b12</td>
<td>360 nm</td>
<td>352 nm</td>
<td>354 nm</td>
<td>355 nm</td>
<td>340 nm</td>
<td>345 nm</td>
<td>352 nm</td>
<td>354 nm</td>
<td>355 nm</td>
<td>355 nm</td>
</tr>
</tbody>
</table>

In more recent years, the process has been adapted for simultaneous purification and refolding of recombinant proteins when the latter were obtained as “inclusion bodies” (Singh and Gupta, 2008; Gautam et al., 2012a; 2012b). Table 2 shows a high throughput screen by which an appropriate smart polymer could be identified for a given protein. The correct “hit” was identified by the correct λ\text{max} emission of the refolded protein. Subsequent work showed that all these proteins would be refolded by affinity precipitation with these identified smart polymers (Gautam et al., 2012a). This was confirmed by extensive characterization of the refolded proteins by Circular Dichroism (CD), fluorescence, melting Temperature (T\text{m}). The process also resulted in
simultaneous purification as all the refolded proteins showed a single band on SDS-PAGE.

### 1.6. Three Phase Partitioning (TPP)

This technique was first described by Lovrein et al. (1987) group at University of Minnesota. It consists of mixing in appropriate amounts of ammonium sulfate and t-butanol to a solution of protein(s). Within few minutes, three phases appear. The upper phase consists of t-butanol rich phase (which can remove any colored impurity present in the protein solution), lower water rich phase and interfacial protein precipitate. The process is illustrated in Fig. 3 with precipitation of Green Fluorescent Protein (GFP).

Three phase partitioning was originally conceived as an alternate to conventional ammonium sulfate and was developed mostly as a concentration step (Lovrein et al., 1987). It uses a lot less salt than required in the conventional salting-out process by ammonium sulfate. As with ammonium sulfate fractionation, limited extent of purification was also reported with several proteins/enzymes. The article on TPP by Dennison and Lovrein (1997) also contains one of the most rigorous discussions on protein-ammonium sulfate interactions and needs wider appreciation. Somehow, it has failed to capture adequate attention. Around early 2000, our laboratory successfully employed this simple technique to purify several proteins/enzymes to near homogeneity (Sharma et al., 2000; Sharma and Gupta, 2001a; 2001b; 2001c; Jain et al., 2004). In recent years, it is now slowly drawing wider attention (Coimbra et al., 2010; Kumar and Raphael, 2011; Harde and Singhal, 2012).

In 2004, it was shown that TPP could be used for refolding urea denatured enzymes (Roy et al., 2004a). Fig. 4 shows the activity recovery and simultaneous purification which could be achieved in the case of cellulase present in a commercial pectinase preparation (Roy et al., 2005a). Figure 4 shows how the results varied with different starting protein concentrations. As aggregation competes with refolding, refolding with higher concentrations of proteins is more challenging. Gel filtration, which estimates the presence of aggregates, is a good technique for assessing the refolding efficiency in a given case (Fig. 5).

Recently, we described refolding of 12 different recombinant proteins which are obtained as inclusion bodies (Raghuva et al., 2008). These proteins varied widely in terms of molecular weight, isoelectric point and number of disulfide bridges. In each case, TPP refolded the proteins successfully with higher yields than obtained with conventional methods. Simultaneous purification of the protein was also achieved.

Hence, TPP is a simple technique which has proven to be useful both for purification and refolding of proteins. It somehow has not attracted attention as much as, say, expanded bed chromatography (see the section on expanded bed chromatography). Heralded as a new unit process for downstream processing, expanded bed chromatography relies upon fairly costly chromatographic media. Backed by Pharmacia (now GE Healthcare), the technique has been extensively used. TPP, on the other hand, is slowly beginning to be used by a few groups (Chaiwut et al., 2010; Kumar et al., 2011; Kurmudle et al., 2011). Expanded bed chromatography has the virtue of being capable of directly dealing with crude feed containing suspended material. TPP has another version called Macro-(Affinity Ligand) Facilitated Three Phase Partitioning (MLFTPP), which can also do the same.

### 1.7. Macro-(Affinity Ligand) Facilitated Three Phase Partitioning (MLFTPP)

It was found that water soluble polymers could also be separated out as an interfacial precipitate during TPP (Mondal et al., 2004). Hence, these could be used as macro-(affinity ligands) to precipitate the desired protein as polymer-protein complex (Sharma and Gupta, 2002). The process can be viewed as analogous to affinity precipitation except that in this case, the precipitate floats as an interface and hence is even easier to separate. Ammonium sulfate and t-butanol together are used to precipitate the affinity complex.

Rest of the steps involved is similar to affinity precipitation. Figure 6 illustrates the earliest application of MLFTPP which exploited the earlier known affinity of Eudragit S-100 for xylanase (Gupta et al., 1994). Subsequently, MLFTPP has been successfully used for protein purification in several cases. Just like affinity precipitation and TPP, MLFTPP was also shown to refold urea denatured proteins/enzymes (Sharma et al., 2004). Recently, it was shown that MLFTPP could also refold a variety of recombinant proteins obtained as inclusion bodies. MLFTPP has been known to be faster than TPP and is expected to be more selective (being an affinity based process) (Gautam et al., 2012c).

MLFTPP certainly deserves wider attention. It is only when it is tried by several groups, a true assessment of this simple, easily scalable and non-chromatographic technique can emerge. As mentioned earlier, MLFTPP is one of the only three known techniques (ATPS and expanded bed chromatography being the other two) which are capable of dealing directly with crude feed containing particulate material. This eliminates the need for a centrifugation/membrane separation step which could be costly and tedious at the industrial scale. Also, these techniques are a part of the strategy to integrate upstream and downstream processes in protein purification.
Fig. 3. Three phase partitioning of Green Fluorescent Protein (GFP). (a) GFP solution (1 mL) was placed in a quartz cuvette under UV light. (b) 60% (w/v) ammonium sulfate and t-butanol (1:1 v/v) were added. The solution was once more placed under UV-light. (c) After 3 min phase separation has begun. After 30 min, three phases viz. lower aqueous phase, the interfacial precipitate and the upper layer of t-butanol are clearly seen. The aqueous phase once fluorescent is now a clear solution, while the interfacial precipitate containing GFP shows green fluorescence. “Reprinted with permission from (Roy et al., 2007). Copyright 2008 by Elsevier”
Fig. 4. Effect of various conditions of three phase partitioning on the renaturation efficiency of cellulase. The different conditions used were:

- Condition 1: Ammonium sulfate (A.S.) saturation (w/v) 20, t-butanol (v/v) 1:1, temperature (°C) 25;
- Condition 2: A.S. saturation (w/v) 20, t-butanol (v/v) 1:2, temperature (°C) 25;
- Condition 3: A.S. saturation (w/v) 30, t-butanol (v/v) 1:1, temperature (°C) 25;
- Condition 4: A.S. saturation (w/v) 30, t-butanol (v/v) 1:2, temperature (°C) 25;
- Condition 5: A.S. saturation (w/v) 40, t-butanol (v/v) 1:1, temperature (°C) 25;
- Condition 6: A.S. saturation (w/v) 40, t-butanol (v/v) 1:2, temperature (°C) 25;
- Condition 7: A.S. saturation (w/v) 20, t-butanol (v/v) 1:1, temperature (°C) 37;
- Condition 8: A.S. saturation (w/v) 20, t-butanol (v/v) 1:2, temperature (°C) 37;
- Condition 9: A.S. saturation (w/v) 30, t-butanol (v/v) 1:1, temperature (°C) 37;
- Condition 10: A.S. saturation (w/v) 30, t-butanol (v/v) 1:2, temperature (°C) 37;
- Condition 11: A.S. saturation (w/v) 40, t-butanol (v/v) 1:1, temperature (°C) 37;
- Condition 12: A.S. saturation (w/v) 40, t-butanol (v/v) 1:2, temperature (°C) 37.

Different concentrations of starting protein were:

- (a) 0.56 mg mL$^{-1}$
- (b) 1.12 mg mL$^{-1}$
- (c) 5.6 mg mL$^{-1}$
- (d) 11.2 mg mL$^{-1}$

The shaded bars represent the activity regained after renaturation whereas the unshaded bars represent the fold purification of refolded cellulase. "Reprinted with permission from (Roy et al., 2005a). Copyright 2005 by Elsevier"

1.8. Magnetic Particles for Protein Purification

Few may remember but it was Whitesides et al. (1983) who very early discussed the promise of magnetic particles in the context of purification of biological molecules. The use of magnetism is especially valuable when the protein of interest is present in a medium of high viscosity. High speed centrifugation or membranes along with high pressure would be less attractive options in such cases. Micron sized particles, notably of Fe$_3$O$_4$ have been described many years back. Safarik’s group has done considerable work in this area and extensively reviewed the literature related to that phase (Safarikova and Safarik, 1995; Safarik and Safarikova, 1999; 2009; Safarikova et al., 2011). In our lab, we entrapped micron sized Fe$_3$O$_4$ particles in polymeric affinity ligands to separate some enzymes (Teotia and Gupta, 2001c; 2002; Safarikova et al., 2003).
Fig. 5. Gel filtration pattern of purified and renatured precipitate of cellulase, obtained under condition 8 (as described in Fig. 4). The enzyme precipitates were dissolved in 1 mL of 0.05 M acetate buffer, pH 4.8 and loaded on a pre-equilibrated Superdex 75 gel filtration column (120 cm×1.25 cm; pre-washed with sodium acetate buffer, pH 4.8). The column was eluted with the same buffer. The fractions obtained with (a) pure cellulase; (b) interfacial precipitate obtained with 0.56 mg mL$^{-1}$ starting protein concentration under condition 8; (c) interfacial precipitate obtained with 1.12 mg mL$^{-1}$ starting protein concentration under condition 8; (d) interfacial precipitate obtained with 5.6 mg mL$^{-1}$ starting protein concentration under condition 8 and (e) interfacial precipitate obtained with 11.2 mg mL$^{-1}$ starting protein concentration under condition 8, were checked for cellulase activity and protein. *Reprinted with permission from (Roy et al., 2005a). Copyright 2005 by Elsevier.

In recent years, attention has switched over to nanosized magnetic particles. Mostly Fe$_3$O$_4$ and composite material containing Fe$_3$O$_4$ have been used. Wang (2006) has provided excellent reviews on the behavior of Fe$_3$O$_4$ nanoparticles which are <30 nm diameter in the context of support materials for designing nanobiocatalysts. Fe$_3$O$_4$ particles of <30 nm size show superparamagnetism and Brownian motion. The advantage of the nanodimension in the context of bioseparation also lies in high surface area to volume ratio.

Unfortunately, as one of us has pointed out recently (Gupta, 2012), reporting multidisciplinary efforts require caution. In a very recent work reported on thiol functionalized magnetic nanoparticles (Lee et al., 2012). The chemistry which was believed to be involved in the binding of BSA and lysozyme is unfortunately missing. Even if one assumes either disulfide bond formation or thiol disulfide exchange reactions, it is not clear why NaCl was expected to elute.
Fig. 6. MLFTPP of xylanase using Eudragit S-100 as the affinity macroligand. Xylanase enzyme was purified from the commercial preparation of A. niger using eudragit S-100 as the smart affinity macroligand. “Reprinted with permission from (Roy et al., 2007). Copyright 2007 by Elsevier.

Hatton’s group (Bucak et al., 2003) described the use of phospholipid coated magnetic nanoparticles as ion exchangers for separation of binary mixtures of cytochrome c and Soybean Trypsin Inhibitor (STI) and ternary mixtures of α-chymotrypsinogen, myoglobin and albumin. Very high capacities (upto 1200 mg protein per mL exchanger) were reported. More important, unlike porous ion exchangers (which are in current use), the particles show no diffusional resistance.

Liao and Chen (2002) described a polyacrylic acid bound iron oxide as a high capacity cation exchanger for basic proteins. Later on, binding and elution of bromolain with these particles were described (Chen and Huang, 2004).

Gu et al. (2006) synthesized monodispersed nanoparticles of Au, FePt and Co compound core shells, all with terminal nitrilotriacetate groups. His6-tagged GFP was used as a model protein and results showed that the magnetic nanoparticles could be reused efficiently. Pham et al. (2008) showed that gold coated Fe3O4 composite particles could be stabilized by citrate anion. The IgG could bind to such particles via ion exchange mechanism. However, no elution was tried as the objective was to use IgG for bioaffinity immobilization purposes. Kim et al. (2010) have described nanocomposite spheres with NiO nanoparticles decorating silica shell with a magnetic core. The NiO nanoparticles bound His-tagged proteins and were a substitute for Ni chelates used in conventional mode of IMAC. Again, His-tagged GFP was used as a model protein and the affinity material could be recycled four times effectively. The system worked well even for separating His-tagged alpha synuclein from E. coli cell lysates. While it is an excellent piece of chemistry, its superiority over easily available commercial magnetic nanoadsorbents for the purification of His-tagged proteins remains to be proven.

Few other similar magnetic core shell particles have been described (Fang et al., 2010; Kim et al., 2010). Chang et al. (2010) showed the binding of BSA to silica-coated magnetic nanoparticles modified with different alkyl groups. This, obviously, is an effort to develop magnetic media equivalent to the media used for hydrophobic interaction chromatography (Agarwal and Gupta, 1995).

Okada et al. (2011) have described the use of Au/Fe3O4 nanoparticles (2-4 nm Au particles bound to Fe3O4 nanoparticles) for binding to proteins with tags containing cysteine and methionine via Au-S linkages. Again, fusion with GFP was used as a model system. The purification efficiency was claimed to be comparable to “His-tagged protein purification systems”. (Also see discussion on affinity tagged proteins).

Recently, concanavalin A (Con A), a lectin was immobilized on “functionally activated” magnetic chitosan nanoparticles by “a relatively well known and
widely used procedure” (Kavaz et al., 2012). The immobilization procedure or its relative success is not unfortunately shared. These particles were used to purify IgG from fresh human plasma.

Fischer and Franzreb (2012) have described an interesting approach in which commercially available magnetic nanoparticles were employed as cation exchangers. The single phase system consisting of these particles and a nonionic surfactant Eumulgin ES separates into micelle rich and micelle depleted phases on raising the temperature from 20-26°C. The ion exchanger nanoparticles partition into the micellar phase. Thus, lysozyme (basic protein) was separated from ovalbumin (acidic protein) with a purity of 100% as only the former protein would bind to the cation exchanger.

In another recent interesting approach (Diao et al., 2012), “stacked-cup carbon nanotubes” with magnetic Fe nanoparticles trapped in their tips were functionalized with-COOH groups and subsequently derivatized with 3-aminophenylboronic acid. Oriented immobilization of goat anti-human IgG was achieved. These were used to separate the corresponding antibodies. This is perhaps the first report of oriented immobilization of a protein affinity ligand on a nanoparticle and hence deserves a mention.

It is obvious that all the rich chemistry (and some new ones!) developed with the micron sized separation media is being revisited in the context of nano-sized materials. Ion exchangers, HIC media, IMAC media are all being used. However, as pointed out here, a caution is needed in this multidisciplinary area.

1.9. Expanded Bed Chromatography

Chromatographic methods are the protocols of choice while purifying proteins. One major drawback of such methods is that the feed needs to be clarified; columns loaded with particulate matter-containing crude material tend to clog, rendering the matrix useless for purification. Hence, chromatographic procedures are invariably preceded by solid-liquid clarification steps. The extent of purification required is decided by the final application of the protein (Mondal et al., 2006; Roy et al., 2007). This increases the number of unit steps in the purification protocol, which has a direct effect on the cost of the final product and an adverse outcome on its yield (Fig. 1). The technique of fluidized bed chromatography, or its modified version, that of expanded bed chromatography, surmounts the problem of multiple steps. It is capable of utilizing unclarified crude extract as the feed material, eliminating the need for multiple unit steps and potentially reducing the cost of the final product (Anspach et al., 1999; Roy et al., 2007). Unlike the matrices used in conventional chromatography where size variation of the beads has to be within an acceptable limit, the matrices used in expanded bed chromatography are designed to have a gradation in size. This polydispersity in size overcomes the problems of channeling and turbulence in the column by giving rise to a stable fluidized bed (Chase and Draeger, 1992). It also reduces back mixing and enhances resolution. The breakthrough capacities are similar to those in packed bed columns, giving rise to very high absorption capacities (Chase, 1994; Willoughby et al., 2000). A second difference is that the inlet has a different position than in packed bed chromatography. The feed is pumped in from below the column, which fluidizes or expands the column. The interparticle distance increases and particulate matter can easily pass through the column while the target protein is bound to the matrix. The column is washed in the expanded bed mode and then allowed to settle down. Elution of the bound protein is usually carried out in the packed bed configuration. All the different formats which are commonly used in packed bed chromatography are also employed in this format. For example, recombinant hepatitis B core antigen was purified by expanded bed affinity chromatography using M13 phage displaying an optimized heptapeptide sequence as the affinity ligand (Ng et al., 2012). Unclarified E. coli lysate was used as the feed. The yield could be improved by recycling the feed; however, the purity was compromised. No host cell protein could be detected on the matrix. The rate of leakage of the phage ligand also remained very low throughout the run. Other capsid proteins like the recombinant nucleocapsid protein of the Nipah virus (NCP-NiV) have been purified in the hydrophobic interaction mode (Chong et al., 2010). DNase I-pretreated unclarified E. coli cell lysate was loaded on Phenyl-Streamline® and binding was carried out at a degree of expansion of 2. The yield of the target protein, carried out with a decreasing gradient of ammonium sulphate, was~59% which could be increased to >80% when glycerol was incorporated in the wash buffer, without compromising on the purity of the preparation (Chong et al., 2010). Apart from these, formats such as immobilized metal ion affinity chromatography have been used for the purification of His6-hepatitis B core antigen with high yield (Yap et al., 2010). Tailored matrices have been designed in the laboratory, which have yielded excellent results. Crosslinked alginate-guar gum beads showed excellent fluidization properties and have been used for the purification of jacalin from partially clarified jackfruit seed extract (Roy et al., 2005b). Haemagglutination activity with glutaraldehyde-
treated trypsinized RBCs showed ~88% recovery of the protein, with high purity. Other carbohydrate matrices like alginate beads have been used for the purification of phospholipase D from partially clarified peanut extract (Sharma et al., 2001). Fluidized beds of carbohydrate matrices have also been exploited for one step purification of α-amylases (Roy et al., 2000a; 2000b), cellulase (Roy et al., 2000a), pectinase (Roy et al., 2004b) and pullulanase (Roy and Gupta, 2002) from crude extracts. Another advantage of the technique is illustrated by the purification of polyphenol oxidase. Since such crude extracts are rich in polyphenols, enzymatic conversion to quinines can occur which undergo Schiff base formation with all proteins present in the extract, thus inactivating them. Hence, a “quick-and-dirty” approach is required to separate the enzyme from inactivating agents. Aqueous seed extracts of Duranta plumieri were loaded on a Streamline® DEAE column and partial purification of polyphenol oxidase and more importantly, separation from contaminating polyphenols, was achieved in a single step (Roy et al., 2002). Group-specific affinity chromatography has been carried out by immobilizing Cibacron blue 3GA on CELBEADSTM (Roy and Gupta, 2000). Alkaline phosphatase could be purified by loading crude chicken intestine extract. Both yield and purity of the eluted enzyme were higher when the protein was purified by expanded bed affinity chromatography as compared to the batch mode using the same matrix (Roy and Gupta, 2000).

Cleaning-In-Place (CIP) is one area which needs special attention in expanded bed chromatography as unclarified broth is fed directly into the column. If cell debris is retained by the column, this may affect its performance. The silica-based matrices which had been design initially had very good fluidization properties but could not withstand the high pH required for removing the non-specifically bound protein (Dasari et al., 1993). Thus, new matrices were designed, both in the industry and in academic laboratories.

1.10. Membrane-Based Separation

Membranes have always formed a part of the traditional bioseparation protocol for proteins. The major use has been that of solid-liquid clarification and sterile filtration, more commonly referred to as microfiltration. Ultrafiltration refers to the exchange of buffers in protein solutions and their concentration. Microfiltration is usually the first step after protein extraction since it separates the solid material from the culture broth and is the first step of concentration before chromatographic steps can be initiated. Although traditionally considered to be a low-resolution process, the pore size of the sieves can be adjusted to achieve some degree of primary separation. The technique is also employed for the separation of adventitious viruses during production of therapeutic proteins. There are some excellent reviews available regarding the twin techniques of microfiltration and ultrafiltration (Reis and Zydney, 2001; Roush and Lu, 2008).

Fouling of the membrane is a major problem with the technique. This results in loss of membrane material and has an adverse effect on protein yield. The emphasis lately has thus moved to the development of materials with high throughput and ability to withstand high flow rates so that removal of particulate matter is faster. Another factor that is often ignored is the possible interaction between the target protein and any excipient present in solution which may alter the clearance rate of the former. For example, the presence of citrate during filtration of a Fab product led to an aberrant clearance plot for the small molecule on a semi-logarithmic scale whereas the plots for the other two excipients, Tris and Hepes, remained unaffected (Harinarayan et al., 2009). A possible interaction between the trivalent citrate and the protein molecule (pI 8.3) was suggested as the reason. A less known problem is that of protein degradation during microfiltration. In an in vitro study with different membrane materials used in a haemodialyzer for filtration of plasma proteins, it was shown that more proteins were retained and carboxylated on cellulose diacetate and polymethylmethacrylate than on ethylene vinyl alcohol membrane (Pavone et al., 2010).

1.11. Avoiding Formation of Inclusion Bodies

Because of the favourable cost and yield calculations, as well as the vast knowledge base available regarding its growth, metabolism and genome-level characterization, Escherichia coli has emerged as the system of choice for heterologous protein expression. The high level of the foreign protein, mostly of human origin, leads to the formation of particulate structures commonly referred to as inclusion bodies (Burgess, 2009; Wang, 2009; Wang et al., 2008). These share nomenclature as well as structural features with neuronal inclusions generally found in various neurodegenerative diseases (Garcia-Fruitos et al., 2011). The formation of inclusion bodies has been viewed differently across generations. In the early days of recombinant era, formation of inclusion bodies was considered to be useful as the protein of interest was localized and concentrated in a single mass, making its isolation easier (Hockney, 1994). Since the target protein forms the major component of the inclusions, the purification
process was considered to be fairly straightforward. The use of strong promoters results in high local concentration of the target protein. Combined with an adverse redox potential of the prokaryotic cytosol, these conditions cause the proteins in inclusion bodies to frequently misfold and aggregate, leading to functionally inactive protein. The first step is to release the target protein from inclusion bodies. Solubilization of the aggregates is generally achieved by the use of high concentrations of chaotropes like urea or guanidine, which unfold the protein completely (Carrio and Villaverde, 2002). This is followed by dilution of the denaturant to facilitate refolding of the protein. This step has been shown to increase the total production cost of proteins. In case of refolding of rhGM-CSF (granulocyte macrophage colony-stimulating factor), for example, this step contributed to 60-75% of the total operational costs (Lee et al., 2006). Thus, new strategies are continuously evolving, which combine refolding with purification steps to reduce costs. An older strategy, which is finding fresh converts, is that of developing approaches to avoid formation of inclusion bodies altogether. This generally takes three forms, which are discussed below.

### 1.12. Medium Engineering

Manipulation of the culture medium by changing physical parameters such as pH, temperature or osmolarity, is probably the most straightforward approach to alter the solubility of recombinant proteins. It has been known for a long time that the growth of E. coli at sub-optimal temperatures, e.g., 28-30°C, results in significantly higher production of the soluble target protein. Most of the activity of a recombinantly produced human interferon-2 in E. coli could be recovered in the soluble fraction when cells where grown at 30°C, whereas the protein was mostly concentrated in the pellet fraction when cells were grown at 37°C (Schein and Noteborn, 1988). In fact, these authors showed that incubating pure interferon-2 with non-transformed E. coli lysates at 37°C resulted in precipitation of the target protein whereas the protein was able to retain its solubility when incubation with the cell lysates was carried out at 30°C. Incubation with buffer containing BSA alone (so-called “mock lysate”) was not able to cause insolubilization, showing that some “factors” from the host cell are involved in this temperature-dependent process (Schein and Noteborn, 1988). One of the more obvious reasons could be the temperature-dependent collision between the protein molecules, which increases at higher temperature and supports the propensity of the protein to aggregate. However, since the growth rate of the cells itself is adversely affected as the temperature is lowered from the optimal one, this approach is mostly limited to the laboratory scenario.

The existence of the Hofmeister series and the vast literature available regarding compatible solutes/osmolytes make it clear that the solubility of a protein can be modulated by additives. Addition of pseudochaperones like sorbitol and arginine was able to solubilize green fluorescent protein expressed from pBAD.GFPwtS99, which otherwise formed inclusion bodies (Prasad et al., 2011). Solubilization occurred only in those cases where the solute could permeate the cell membrane and was a potential participant in the generation of ATP in the cells. The presence of trehalose was able to solubilize the model amyloid-forming protein, HypF-N, which is normally expressed in inclusion bodies in E. coli (Prasad et al., 2011).

One major drawback of prokaryotic expression systems is that proteins with disulphide bonds cannot be expressed in the correct conformation. In eukaryotic cells, disulphide bonds are formed in the oxidative environment of the endoplasmic reticulum. In the absence of any overt compartmentalization, the reducing cytosolic environment in prokaryotes does not support disulphide bond formation. In the absence of disulphide bonds, the nascent polypeptide chains misfold and then form aggregates. The oxidative environment of the periplasmic space, on the other hand, provides favourable conditions for the formation of disulphide bonds. One of the ways in which yields of proteins with disulphide bonds can be increased in the periplasm is by the addition of compatible solutes. Induction of osmotic stress by the addition of different osmolytes led to the expression of functional chimeric immunotoxins RFTS (scFv) and Ki-4(scFv) fused to the modified Pseudomonas aeruginosa exotoxin A gene in the periplasm (Barth et al., 2000). The best yields were obtained with glycine betaine as a stress inducer. The high yield of the active proteins is likely due to the overexpression of chaperones and heat shock proteins in response to osmotic stress. This was more clearly observed by the addition of the membrane fluidizer benzyl alcohol in the growth medium for bacterial cells expressing the folding refractory protein O35 (Marco et al., 2005). Benzyl alcohol invoked a strong heat shock response in the cells and induced the overexpression of the bacterial chaperone DnaK, resulting in a 48-fold increase in the amount of soluble protein. However, scaling up of the process resulted in inhibition of cell growth in the presence of benzyl alcohol (Marco et al., 2005), pointing to the need to optimize the role of benzyl alcohol in individual cases.
It is well-documented that taking up osmolytes from the environment is energetically economical for the bacterial cell than to synthesize them intracellularly (Marco et al., 2005), presumably via stress-activated mechanisms. Thus, keeping in mind the scale-up considerations, this is probably the most viable strategy to alter the solubility characteristics of a heterologously expressed protein in the prokaryotic host system.

1.13. Protein Engineering

The solubility characteristics of proteins can be altered by either mutating aggregation-prone amino acids to others or by creating genetic constructs with tags which increase the solubility of the target protein. This has been discussed in an earlier section. However, here the approach is illustrated with various examples. Fusion of scFvs (against human collagen type VI and human desmin) with Maltose-Binding Proteins (MBPs) led to increased production of the chimeric protein in the periplasmic space (Martin et al., 2006). However, binding assays with ELISA showed diminished functionality. It is likely that although MBP improved of the solubility of the passenger proteins, it may have interfered with the correct folding of the scFvs, hiding the epitopes and resulting in low binding efficiencies in ELISA. A common partner in fusion constructs, especially for the production of recombinant antibodies, has been alkaline phosphatase. In the above system, expression of alkaline phosphatase with MBP as a fusion tag resulted in a functional protein (Martin et al., 2006). Fusion of alkaline phosphatase with the scFvs also led to the production of high levels of scFvs in functionally active forms.

Fusion constructs have also been created with the outer membrane protein or bacteriocin signal sequences as tags to ensure extracellular production of recombinant proteins in the soluble form. Other tags have also been employed. Thioredoxin tag has been reported to result in soluble functional proteins (Yanga et al., 2009; Bao et al., 2012). Due to its inherent chaperonin activity, ubiquitin has been commonly used a fusion tag to improve yield of recombinantly expressed proteins like metallothionin and -subunit of the GTP-binding stimulatory protein of adenylate cyclase in E. coli (Butt et al., 1989). Fusion constructs of ecotin, mouse ubiquitin b and a target FLS peptide genes were used to ensure expression of the passenger protein in the periplasmic space (Paal et al., 2009). Hydrolysis of the N-terminal Ecotin-Ubiquitin-Tag (ECUT) was carried out in vitro using ubiquitin C-terminal hydrolase L-3. Concentrations of 18 mg of fusion protein per gram of dry cell weight were reported.

1.14. Host Cell Engineering

Bacterial cells exhibit a number of drawbacks because of which they may not appear to be ideal hosts for recombinant protein production. At the same time, their short doubling time, ease of manipulation of genome and the relatively low cost of production, make them attractive candidates for production of heterologous proteins. The unfavourable redox potential of the bacterial cytosol is maintained by the combined action of thioredoxins (trx) and glutaredoxins (glutathione oxidoreductase, gor) (Prinz et al., 1997; Schlapschy and Skerra, 2011). Double deletion of trxB and gor results in cells which are available commercially as Origami (Novagen). These allow the expression of disulphide bond-containing proteins in the bacterial cytosol (Subedi et al., 2012; Sun et al., 2012). Another expression strain, referred to as SHuffle (New England Biolabs), is a double mutant (trxB-gor-) and additionally overexpresses the cytoplasmic chaperone protein DsbC. The latter is involved in disulphide isomerization and facilitates the formation of correct pairs of disulphide bonds.

Overexpression of the chaperonin GroELS resulted in high level expression of recombinant ribonuclease inhibitor in the active form (Siurkus and Neubauer, 2011). Depending on the growth temperature, ~3.7-fold jump in the production of the active protein was reported. The other chaperone system, that of DnaKJE, also facilitated folding and solubilisation of the target protein. However, this preparation was functionally inactive (Siurkus and Neubauer, 2011). Thus, a general purpose chaperone-overexpressing system for expression of soluble recombinant proteins cannot be proposed. Since the two chaperone systems function at different stages of protein folding, it is likely that depending on the stage at which misfolding and aggregation of the target protein are initiated, the roles and usefulness of the chaperone systems would differ.

Of late, our traditional view of structure and activity/function of inclusion bodies is undergoing a radical change. Inclusion bodies have been shown to have well-characterized cross β-sheet structures, reminiscent of amyloid fibrils (Wang et al., 2008; Groot et al., 2009; Jurgen et al., 2010). This has also been confirmed by Thioflavin T fluorescence spectroscopy as well as measurement of quenched hydrogen/deuterium exchange by solution NMR spectroscopy (Wang et al., 2008). In a study with overexpressed VP1 capsid protein-GFP fusion protein which formed inclusion bodies, the pellet fraction exhibited GFP fluorescence and was enzymatically active (Garcia-Fruitos et al., 2005). Thus, in a few cases, it has been speculated that mere
insolubility of a protein does not necessarily rule out an active protein (Gonzalez-Montalban et al., 2007; Garcia-Fruitos et al., 2012). Thus, a rethink on the solubilisation of inclusion bodies may be in order, mostly in those cases where there is no obvious cause of precipitation, such as lack of disulphide bonds or absence of an interaction partner. In cases where high local concentration of the recombinant target protein is the sole cause of precipitation, the activity of the insoluble fraction may be checked directly or by employing non-denaturing conditions, e.g., addition of arginine (Tsumoto et al., 2010), to solubilize inclusion bodies.

1.15. Extracellular Protein Production

Another approach, which has shown successful results, is that of extracellular production of recombinant proteins. This has the twin advantages of continuous production of target proteins since there is no disruption of the cell wall and improved yield due to low proteolytic digestion. Extracellular secretion of heterologous proteins can be facilitated in a number of ways. Since E. coli does not secrete proteins into the environment except for some defence proteins and lysisins, protein production in the extracellular space is directed using physical, chemical or genetic methods. In the first case, harvested cells are subjected to osmotic or mechanical stress using chemicals or biochemicals or enzymes, which disrupt the outer membrane of the bacterial cell. The concentration of the agents used needs to be carefully regulated so that cell lysis does not occur and purity and/or yield of the target protein are not compromised. Chemicals like polyethylene glycol, Triton, EDTA, Ca\(^{2+}\), Mg\(^{2+}\) have been used for the extracellular production of different proteins (Zhang et al., 2009; Wiese and Schmitz, 2011). Mechanical means like sonication are used quite commonly. Doubts have however been raised about the feasibility of scaling up such a process as it requires handling large volumes of the culture medium.

A second method, which finds widespread use in the laboratory, is the creation of gene fusion of the target protein with outer membrane proteins (OmpA, OmpF, OsmY) (Jeong and Lee, 2002; Zhang et al., 2006). After optimization of a variety of promoters and growth conditions, a high throughput system was employed to harvest a number of human proteins from the extracellular space using the OsmY fusion tag under the control of the araBAD promoter (Kotzsch et al., 2011). The Haemolysin (HlyA) transport system (HlyB-HlyD-TolA translocator) utilizes a protein channel running through the inner and outer membranes of E. coli (Gentschev et al., 2002). The toxin is secreted in the extracellular environment due to the presence of the secretory signal sequence. Expression of high level of cutinase from Thermobifida fusca has been reported using this system (Su et al., 2012). Periplasmic proteins can also be released into the bacterial culture medium by fusing the target gene sequences with the colicin releasing gene kil1 (Steidler et al., 1994). Release of most of the Bacteriocin Release Proteins (BRPs) compromises the integrity of the cell membrane and frequently leads to cell death. Using a hybrid-glucanase from Bacillus amylo liquefaciens and Bacillus macerans as a model protein, expression and secretion of kil1 gene was monitored under two stationary phase promoters, viz. p589, a weak natural promoter and p594, a strong synthetic promoter (Beshay et al., 2007). The yield of the soluble target protein in the culture medium increased by 221% when a strong promoter for kil1 gene was used and was further improved in the presence of a strong promoter for -glucanase gene. Extracellular production of soluble recombinant proteins thus appears to be a viable strategy provided cell death can be avoided and the culture volume can be concentrated prior to protein harvesting.

1.16. Crystallization Revisited

Some techniques are evergreen. We started this review with what can be dismissed off as a historical account. We mentioned that crystallization was the earliest technique which was used so extensively in the pre-chromatographic era. That does not mean that crystallization is no longer a modern technique for protein purification. Przybycien et al. (2004), in an excellent review, have summarized the evolution of protein crystallization into an industrially relevant technique. Both glucose isomerase and insulin were mentioned as illustrative examples of the proteins in this context. More recently, Etzel (2006) has discussed “bulk protein crystallization” along with case histories of Rubisco, Subtilisin, Aprotinin and insulin.

2. CONCLUSION

In 2004, one of the volumes of the journal “Biotechnology and Bioengineering” carried an editorial entitled “The importance of Bioseparations: Giving credit where it is due”. It said, “And finally the explosion in new high level expression systems for production of recombinant proteins has decreased up-stream processing costs to the point where product concentration and
puriﬁcation steps are now dominating the overall manufacturing cost for many protein therapeutics and most industrial enzymes” (Haynes, 2004). This comment has become more relevant with the passage of time.

It is not yet appreciated that protein puriﬁcation means different things to different people. Enzymologists working in the academic sector purify enzymes for understanding mechanism or metabolic pathway or signal transduction. In such cases, there is no great need to worry about economics of protein production. Purity means ≥95% purity on SDS-PAGE. This has little relevance to the applied biocatalysis or production of pharmaceutical proteins. In industrial enzymology (reborn in a bigger version as white biotechnology), one has to worry about cost, time and similar mundane factors. For pharmaceutical proteins, one has to follow stringent rules laid down by regulatory agencies and rule out the presence of endotoxin, host cell and media proteins, DNA, protein variants including aggregated forms, (Anicetti and Hancock, 1994). Non-chromatographic approaches are not adequate in such cases, invariably chromatographic method with one (or even more) afﬁnity chromatographic step is obligatory.

The present review has tried to discuss the strategies used for protein puriﬁcation in both industrial and academic sectors. The trend is to integrate upstream and downstream stages and thereby make puriﬁcation protocols shorter and economical.

3. ACKNOWLEDGEMENT

We acknowledge ﬁnancial support provided by Department of Biotechnology (DBT), Government of India. Financial support provided by Council of Scientiﬁc and Industrial Research to JM and SG in the form of Senior Research Fellowship and Junior Research Fellowship, respectively, is also gratefully acknowledged.

4. REFERENCES


Berrier, C., A. Garrigues, G. Richarme and A. Ghazi, 2000. Elongation factor Tu and DnaK are transferred from the cytoplasm to the periplasm of Escherichia coli during osmotic downshock presumably via the mechanosensitive channel mscL. J. Bacteriol., 182: 248-251. PMID: 10613892


