IJBPAS, November, 2024, 13(11): 5420-5448

ISSN: 2277–4998

International Journal of Biology, Pharmacy and Allied Sciences (IJBPAS)

'A Bridge Between Laboratory and Reader'

www.ijbpas.com

EXPRESSION, PURIFICATION, CHARACTERIZATION AND FORMULATION OF RECOMBINANT PROTEINS

SINGH A[1*], KUMAR P[2], DIKSHA[2] AND VERMA N[2]

1: Research Scholar, School of Pharmaceuticals Science, IFTM University, Moradabad, Uttar Pradesh, India

2: Pharmacy Academy, IFTM University, Moradabad, Uttar Pradesh, India

***Corresponding Author: Mr. Alok Singh: E Mail: alokthevictor@gmail.com**

Received 24th Sept. 2023; Revised 25th Oct. 2023; Accepted 9th Jan. 2023; Available online 1st Nov. 2024

https://doi.org/10.31032/IJBPAS/2024/13.11.8389

ABSTRACT

Biopharmaceutical industry is booming and with the recent pandemic of COVID-19 and advancements in cell and gene therapy, recombinant protein production has come on the forefront of the industrial protein production. Apart from applications in various industries, recombinant proteins find applications as therapeutics, vaccines and diagnostics. Monoclonal antibodies and other biosimilars make up the largest products in the pipeline of most biotechnology industries. This review article consolidates the knowledge and advances in the process of expression, purification, characterization and formulation of recombinant proteins from various available literature resources.

Keywords: Expression, Purification, Characterization, Formulation, Protein

1. INTRODUCTION

The biopharmaceutical market is one of the fastest growing sector of pharmaceutical industry and is proposed to grow upto USD 856.1 billion by 2030 with a CAGR of

12.5% from 2022 to 2030. Biopharmaceuticals include a product produced from biological systems either *in vitro*, *ex vivo*, or *in vivo* and find applications as therapeutics (anti-cancer monoclonal antibodies, gene therapy products), vaccines, diagnostics, etc. Various products have been approved for therapeutic use by European Medicines Agency (EMA) in European Union (EU) and Food and Drugs Administration of USA (FDA) over the period of 2018-2022. These include monoclonal antibodies (97 products), hormones (19 products), nucleic acid or gene therapy-based products (16 products), and vaccines (16 products). Additionally other products including biosimilars like colony-stimulating factors (12 products), nine cell-based products, eight enzymes, seven fusion products and six different clotting factors were given approval for human use **[1]**. While a lot of proteins are being purified from biological fluids and wastes even today, production of proteins using biological systems is being extensively exploited. With the advent of recombinant DNA (rDNA) technology while leveraging the vast understanding of biological systems, heterologous expression of recombinant proteins has become a promising biotechnological tool. Recombinant protein is a protein expressed from a foreign gene introduced into a biological host.

Recombinant proteins are expressed using various different types of hosts like bacteria (*Escherichia coli* [*E. coli*]), fungi (*Saccharomyces cerevisiae* [*S. cerevisiae*]), insect cells (sf9) and mammalian cells (Chinese Hamster Ovarian [CHO] cells and Human Embryonic Kidney [HEK-293] cells) **[2]**. Downstream processing of the expressed proteins from the cellular milieu is carried out using various chromatographic methods and refolding of purified proteins **[2]**. Purified proteins are of use only when they are functional and have the correct conformational structure **[3]**. Stability of the expressed protein is due to optimum ionic strength of the buffers and the correct storage temperature. Thus, designing the optimal formulation and storage buffers for the expressed proteins renders them stable for longer periods of time and allows them to be used for various applications **[4]**. **Figure 1** describes the process of recombinant protein synthesis starting from vector designing to final formulation and storage stage.

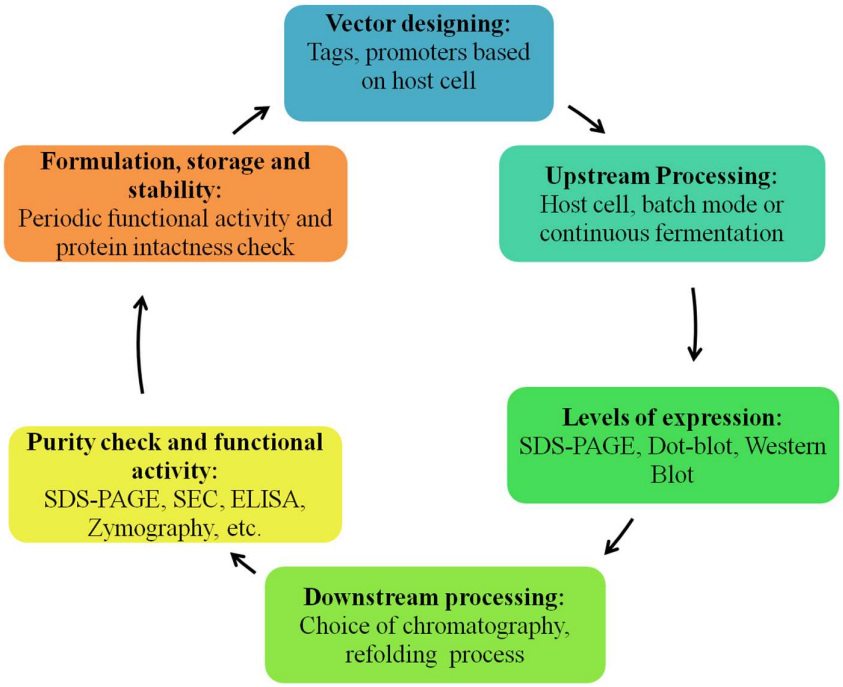

Figure 1: The overall process of recombinant protein expression

This review will be a consolidated study discussing in detail the hosts for expression of proteins, their advantages and disadvantages, focusing mainly on bacterial cells as the expression host. Further, various purification methods, underlying chemistry for optimal process development and protein refolding will be described. The review will also include details regarding the different methods used for characterization of purified proteins. Furthermore, the importance of formulation and applications of recombinant proteins will be discussed at length.

2. Hosts for heterologous protein expression

Various different biological systems are being exploited for heterologous expression of proteins. The choice of expression host depends upon the protein of interest as well as the economic aspects for culturing and expression of the protein. Various hosts are being currently used like bacterial cells, fungi, insect cells and mammalian cells **[5]**. All the expressions hosts have their own advantages and disadvantages for use in protein expression and production at industrial scales. This section will provide insights into the various expression's hosts used, their advantages and disadvantages. Furthermore, the section will focus on bacterial cells as expression host in details while also highlighting the various strategies to achieve optimal levels of expression and different techniques to monitor expression.

2.1 Bacterial Cells

Bacteria like *Escherichia coli* (*E. coli*), *Bacillus subtilis* (*B. subtilis*), *Pseudomonas*

Lactic acid bacteria (LAB)and others are one of the most favored hosts of expression and production of recombinant proteins **[5]**.

a. *E. coli*

E. coli is the most widely used bacteria for recombinant protein expression **[6]**. *E. coli* and other bacteria as hosts for protein expression provide a lot of advantages over other expression hosts. The major advantages include rapid cell growth, easy genetic manipulation, well understood cellular mechanisms, ease of handling and scaling up for industrial purposes, use of simple and economic growth medium, and high production of recombinant proteins. In *E. coli*, the recombinant proteins could be produced as either insoluble inclusion bodies (IBs), or soluble or secretory proteins **[4, 5]**. However, bacterial hosts also have various disadvantages like non-functional protein production due to improper refolding or absence of post-translational modifications in bacteria. Drawbacks specific to *E. coli* include use of unsolicited inducers for gene expression, which would require extensive purification for clearing the inducers, ineffective secretion, high probability of sequestration of the protein of interest into IBs, increased metabolic stress due to production of large amounts of recombinant proteins and presence of endotoxins in the protein product **[4, 5]**.

b. Other Bacteria

B. subtilis is a Gram-positive bacterium and can be used to overcome the issues of endotoxin, although the protein expression levels could be lower than in *E. coli*. This bacterium is mostly used for meta-genomic library construction **[7]**. *B. megaterium*, another Gram-positive bacterium, has been systematically optimized for production and secretion of high levels of recombinant proteins. The strain uses plasmids carrying the gene of interest with specific origin of replication and regulatory elements to get optimum recombinant protein production. It can be used for production of individual proteins and also for coproduction of upto 14 recombinant proteins **[8]**. Lactic acid bacteria like *Lactococcus lactis, Lactobacillus* have been used to safely deliver heterologous proteins like cytokines and antigens into the mucosa due to their food-grade properties **[9, 10]**. Other bacteria like *Pseudomonas fluorescens* and *Pseudomonas putida* have been shown to produce high levels of recombinant proteins in the soluble forms, indicating that they are a good alternative to *E. coli* **[5]**. *Ralstonia eutropha* has been shown to overcome the major drawback of *E. coli*, i.e. formation of IBs when high levels of recombinant proteins are produced. This bacterium can be used for high cell-density fermentation and can produce soluble recombinant proteins in amounts that are almost 100-

folds greater than that observed in *E. coli* **[11, 12]**.

Multiple strategies to overcome the drawbacks pertaining to use of *E. coli* or other bacteria for large scale protein production have been taken up and will be discussed further.

2.2 Yeasts

Since *E. coli* does not have post-translational mechanisms to produce functional mammalian proteins, eukaryotic hosts are used for production of mammalian proteins which require high amounts of glycosylation **[13]**. Yeasts have the advantages that they are unicellular organisms, and have lower nutrient requirements than insect and mammalian cells, making them more adaptable for industrial scale production of biologically active eukaryotic proteins. *S. cerevisiae* is being used for the commercial production of hormones like insulin and glucagon **[13]**. Other yeasts like *Pichia pastoris* (reclassified as *Komagataella phaffii*) are extensively used for the production of biologically active proteins due to their non-dependence of Crab-tree cycle and production of high amounts of recombinant therapeutic proteins such as human insulin, human serum albumin, hepatitis B vaccine, interferon-alpha 2b, trypsin, and collagen. *P. pastoris* can use methanol as a carbon source and this property is being used for the designing of expression vectors under the control of the inducible AOX1 promoter **[14–17]**. Although useful, *S. cerevisiae* does not generate correct type of N- and O- linked glycosylations as present inhuman proteins, thus rendering the proteins not completely useful. *P. pastoris* also generates glycosylation patterns which differ from the native human forms, but these variations are trivial and do not affect the functionality of the protein. Thus *P. pastoris* is more favorable choice as an expression host for recombinant proteins. Another drawback includes codon biasness in *S. cerevisiae*, which increases the work and cost of designing the expression cassette for the gene of interest **[5]**.

2.3 Filamentous Fungi

High mannose dependent glycosylation is observed in yeasts, and hence the use of filamentous fungi can overcome this shortcoming in the production of recombinant proteins. The versatile metabolic ability of fungi makes them an attractive and outstanding cell factory. Filamentous fungi have a strong capacity for secretion due to which they are also considered one of the most promising expression systems for recombinant protein production. Filamentous fungi like *Apsergillus niger, Trichoderma reesei* are being developed into platforms for production of industrially important

enzymes and recombinant therapeutic proteins like IFN-α2b **[18–22]**.

2.4 Insect cells

Insect cells are mainly used for the production of recombinant proteins to be used as vaccines or for diagnostic purposes. Insect cell lines are derived from the fall armyworm (*Spodoptera frugiperda*) or the cabbage looper, (*Trichoplusiani*). The commonly used cell lines include Sf9, Sf21 and High Five™. These cell lines can be used for large scale, efficient and rapid recombinant protein production by baculovirus infections **[23]**. Baculoviruses infect hosts limited to specific invertebrate species, thus making them safe to work with **[5]**. Baculovirus expression system generates recombinant baculoviruses carrying the gene under compatible promoters for high expression of the protein of interest. These recombinant viruses upon infection of the insect host cell lines can produce the soluble form of protein of interest in bulk. Insect cell lines produce glycosylation patterns varying from those of the mammalian cells, thus providing a different epitope to the antigen. Also, secretory proteins are not completely processed after translation and secretion rendering the protein non-fucntional. Baculovirus expression vectors can carry large amounts of foreign DNA and are being modified to include elements and other supporting proteins in order to mitigate these flaws. For example, it was shown that co-infection of a cell line with baculovirus expressing Activin A, a TGF-β family member, and Furin will allow high production of mature and functional Activin A. A modification to this was done by a group by incorporating both the genes into a single baculovirus expression vector, FlexiBAC, and obtaining similar levels of mature Activin A by a single infection **[5, 23–25]**. Another drawback includes discontinuous production of recombinant proteins due to the death of infected cells and the need to infect new cells for every production cycle, thus making the system inferior to prokaryotic and yeast systems in terms of continuous fermentation.

2.5 Mammalian cells

Mammalian cell lines like Chinese Hamster Ovarian (CHO) cell line, Mouse myeloma cell line (NS0) and Human Embryonic Kidney cell line (HEK) are the hosts of choice for the production of recombinant therapeutic proteins. Most of the therapeutic proteins contain complex glycan moieties attached for their efficient functionality. Thus, having correct N- and O-linked glycans is important for the therapeutic activities of these proteins and mammalian cells would be the right choice of hosts for the production of therapeutic proteins **[26]**. CHO cells are the most extensively used for the production of many biopharmaceuticals;

most of these are monoclonal antibodies being used as cancer therapeutics. CHO cells were derived from Chinese hamster ovarian fibroblast cells and multiple variants of the CHO cells have been developed to ease the production of biopharmaceuticals. The CHO cell lines present various advantages; they can be grown in suspension in serum-free medium and show high yields of production while being less susceptible to infection by human viruses. However, a major drawback of the CHO cell lines is the additions of potentially immunogenic glycan, which are non-human, although most of the glycans generated by CHO cells engineered for post-translational modificationsare "human-like". Another drawback is the difficulty to formulate a different medium for each variant of the CHO cell lines, requiring good development strategies for commercial purposes instead of using a specific variant as a platform cell line. HEK293 cells form a good alternative for the development of biopharmaceuticals with completely human type glycan addition, thus increasing their efficiency. HEK293 cell line has been used to produce clotting factor VIII (FVIII). With the introduction of cell and gene therapy products, HEK293 cells has gained importance for the production for recombinant Adeno-associated virus (AAV) and Lentivirses (LV) for product development. HEK293 cell line is an adherent cell line which requires fetal bovine serum for its growth, thus making it difficult to scale up cultures for large scale virus production. To circumvent this shortcoming, suspension variants growing in serum free medium have been developed. Another drawback of using human cell lines for protein production includes risk of infection by human viruses **[2, 5, 26–32]**.

2.6 Cell Free Protein Synthesis

Protein expression and synthesis in cell free systems have become one of the methods of choice. Cell free protein synthesis is an *in vitro* method for protein synthesis which removes then hindrances caused by cellular membranes. This method is scalable and results in high quantities of proteins being produced, with the advantages of being rapid and controllable. The cell free protein synthesis system could be of prokaryotic (*E. coli* or other bacteria like *B. subtilis*) or eukaryotic (yeast, wheat germ, rabbit reticulocytes and insect cell) origin. *E. coli* as the host species is the most preferred for most of the protein synthesis. In case of complex proteins requiring large amounts of post-translation modifications, eukaryotic cell free protein synthesis systems may be used. However, they could lead to lower protein yields. These systems use cell lysates of the host cells which provide the machinery for transcription and translation and plasmids which carry the gene of interest for protein. The whole process is

carried out in an appropriate vessel and the proteins are produced in the lysate mostly in the soluble form. The process or the system can be tweaked in order to avoid failures and hence is being extensively explored for industrial scale production of recombinant proteins. Cell free protein synthesis systems could also be explored to mix various components from different hosts and make an efficient one-stop system for complex protein production **[33–35]**.

3. Strategies to optimize expression of recombinant protein in *E. coli*

As discussed earlier, *E. coli* and other bacteria suffer from various shortcomings like absence of post-translational modifications, generation of IBs, incorrect disulfide bond formation, etc., despite having great advantages for use in industrial level production of recombinant therapeutic proteins. Over the years, a lot of research has been done to improvise the overall expression system in *E. coli* at different stages to obtain optimal production of the protein of interest.

3.1 Plasmid design, promoters, and strain engineering

With technical advancements and a better understanding of various bacterial processes, newer promoters and regulatory elements are being included in the plasmid design for efficient protein expression. Majority of expression systems in *E. coli* use the inducible *lac* promoter/operator system to regulate gene expression. The native *lac* promoter is induced in presence of lactose and cyclic AMP (cAMP) generated due to glucose metabolism, but remains partially active in the usual growth media used due to the presence of glucose. Thus, this system is modified with the *lac*UV5 mutation to reduce catabolite dependence and being completely active when lactose or lactose analog like isopropyl β-D-1-thiogalactopyranoside (IPTG) is added to the medium. The *lac* promoter and the mutant version, *lac*UV5, are not strong promoters to produce large amounts of recombinant protein and also can show leaky expression, which is an issue when expressing toxic or complex proteins. To overcome this problem, dual expression controls and use of strains with lysozyme expression for controlled recombinant protein expression are being used. One such example is the use of the combination of pET expression vectors and *E. coli* BL21 (DE3). The expression of gene of interest is regulated by a promoter recognized by T7 phage RNA polymerase (high copy expression). The gene for T7 RNA polymerase is inserted in the *E. coli* gemone as a prophage (λDE3) and its expressionis driven by the *lac*UV5 promoter. Thus, induction by IPTG first leads to the expression of T7 RNA polymerase, which then drives the expression of the recombinant protein. The expression of

lacI^Q or T7 Lysozyme inhibits extraneous expression of the gene of interest by T7 RNA polymerase. For T7 lysozyme expression, special *E. coli* BL21 (DE3) strains have been developed, i.e.*E. coli* BL21 (DE3) pLysS and *E. coli* BL21 (DE3) pLysE. In these, the Lysozyme binds to T7 RNA polymerase and inhibits it. When the levels of T7 RNA polymerase surpass those of lysozyme, the expression of the gene of interest is carried forward by T7 RNA polymerase **[36, 37]**.

To overcome the issue of basal and leaky expression, another promoter with positive regulation has been exploited. The $araP_{BAD}$ promoter is present in the pBAD vector system, which is activated in the presence of arabinose through the fine regulation by the AraC protein. Another systems use the phage promoters like phage lambda leftwards promoter (pL), which is usually repressed by λcI protein during lysogenic cycle. Thus, a dual control system like *trp*/pL promoter system (λcI under control of *trp* promoter and gene of interest under pL) or a mutant of λcI gene which makes it sensitive to temperature or pH changes can be used. The *trp*/pL dual promoter system works on the principle that *trp* promoter is active in the absence of tryptophan, thus producing λcI, which represses expression of the gene of interest from pL. When tryptophan is added, *trp*promoter is repressed and λcI is ot produced, thus activating the expression of the gene of interest from Pl **[36]**.

The mutant λcI repressor protein (λcI^{857}) is temperature-sensitive and loses its activity at temperature above 37ºC. Thus, when the cultures are grown at 40-42ºC, the repression is lost and the expression of the gene of interest from pL is driven. Another temperature-sensitive system is the pCold vectors, which use *cspA* promoter (cold shock protein A promoter) to regulate gene expression. This promoter is active at temperature below 15ºC and produce large quantities of correctly folded recombinant protein. Using temperature sensitive expression systems can be advantageous industrially as it allows driving of expression by changing the temperature during fermentation and no inducing agent addition is required, decreasing the probability of contamination **[36, 37]**. For efficient expression of heterologous proteins in *E. coli*, codon optimization to avoid ribosome slippage due to rare human codons is done **[2]**.

In terms of strain engineering, *E. coli* has been engineered to carry out post-translational modifications like glycosylation, disulfide bond formation, phosphorylation, proteolytic cleavage and others. An *E. coli* host has been engineered specifically to produce glycosylated antibodies, while another strain has been developed for correct disulfide bond

formation in the periplasmic space. Shuffle is an *E. coli* strain developed for cytoplasmic disulfide bond formation and has been successfully used for IgG production. To overcome the issue of IBs, *E. coli* 'TatExpress' strain allows periplasmic delivery of the recombinant proteins and has been shown to produce large amounts of correctly folded, and active human growth hormone in the periplasmic space in fed-batch fermentation cycle **[36–45]**.

3.2 Use of fusion tags and proteins

Fusion tags and proteins are of great importance as they allow easier downstream purification of the protein of interest and also increase the solubility of the protein expressed. Some of the commonly used fusion tags include His-tag, FLAG-tag, and Streptavidin-tag. These tags allow detection of the expressed proteins using antibodies against them, most of which are commercially available and also allow purification using affinity chromatography.

Over time, a lot of new fusion proteins like Glutathione-S-transferase (GST), maltose binding protein (MBP), and members of the small ubiquitin-like modifier (SUMO) family of proteins have been explored as they have been shown to increase the solubility of the proteins and decrease the probability of IBs formation. The pET vector series uses the concept of 6x Histidine tag, wherein the gene for protein of interest is cloned with the His-tag either at the N- or the C-terminus and it allows expression of protein in the soluble form and its purification by metal-affinity column. Recently, multiple research groups have discovered new fusion partners and affinity tags like serine-lysine-isoleucine-lysine (SKIK tag) after the initial N-terminal Methionine, heparin-binding affinity tag (HB tag), truncated maltotriose-binding protein from *Pyrococcus furiosus* (MBP-Pyr), and small (<10 kDa) metal-binding proteins as fusion tags (SmbP and CusF). These tags were shown to enhance the solubility of the proteins and also allow ease of downstream processing of the expressed protein **[37, 46–53]**.

4. Purification of Recombinant Proteins

Recombinant proteins are expressed in the cellular milieu and hence require purification from other cellular proteins to be used as therapeutics, vaccines or for diagnostics. The first choice for purification of proteins are chromatographic methods, which could be either ion-exchange, affinity-based or hydrophobic interactions-based. For further purity levels, advanced methods of ultra-filtration, ultra-centrifugation and size exclusion chromatography are used. This section of the review will cover details of various methods used for the purification of recombinant proteins including their principles and different ways to optimize the processes to achieve maximum purity of the

protein of interest. This section will also include the different methods employed for refolding of proteins from IBs.

4.1 Cell Lysis

Recombinant proteins are produced in the cytoplasm of the host cells and thus cell lysis is required for releasing the protein of interest into solution for downstream processes. Usual industrial cell lysis processes include high-pressure homogenizers and bead mills. For purification of secreted proteins, concentration of the spent medium is required and is usually done using tangential flow-filtration (TFF). TFF uses a membrane system which allows concentration of the protein of interest while removing proteins and other cellular material of sizes below the pore size of the membrane used. For proteins which are sequestered into IBs, cell lysis is the most important step, followed by washing of IBs and refolding of the protein of interest **[2, 6, 41]**.

4.2 Chromatographic techniques for purification

4.2.1 Affinity Chromatography

Affinity chromatography is the preferred first step of protein purification wherein the property of the protein having affinity towards another protein, ligand or metal ion is exploited. It is selected for purification of antibodies, tagged proteins, DNA-based biologics, viral vectors, and viruses **[2]**. For example, C-reactive protein has a strong affinity to phospho-choline and hence agarose beads with immobilized phospho-choline are used to purify functional pentameric C-reactive protein **[54, 55]**. Similarly, metal-affinity of proteins is used when they carry His-tag, which has high affinity for Ni^{2+} ions. The imidazolium ring of the histidine residues forms coordinate bonds with the Ni^{2+} ions through which the protein binds to the resin. Increasing concentrations of imidazole replaces the protein, leading to its elution **[56, 57]**. Monoclonal antibodies (mAbs), which are the major biopharmaceuticals produced, are also purified employing their affinity to Protein A (mAbs with Fab fragment) and Protein L (mAbs with scFv and kappa light chain) resins. However, Protein A resins suffer from drawbacks like leaching and non-specific binding of host cell DNA and proteins, which could get eluted along with the mAbs of interest **[2, 41, 57]**. Other examples of affinity based purification methods use GST-tags, MBP-tags and heparin binding proteins **[41, 53]**. Recent advances in cell and gene therapy field have led to the development of affinity resins for AAV and LV for affinity based purification of viral vectors to be used for further therapeutic purposes **[57]**.

4.2.2 Ion-Exchange Chromatography

Ion exchange chromatographic techniques exploit the fact that proteins are charged molecules, especially in salt solutions of

different ionic strengths and pH away from their isoelectric point. Thus, a positively charged protein (cationic) can strongly or weakly adsorb onto negatively charged resin like carboxy-methyl cellulose (CMC) and can be eluted using increasing concentrations of salt solution; usually increasing concentrations of sodium ions (Na^+) is used, wherein Na+ ions replace the positively charged protein adsorbed onto the resin. Similarly, a negatively charged protein (anion) adsorbs onto a positively charged resin carrying diethyl-amino-ethyl (DEAE) group. The bound protein is eluted by a gradient of pH, which causes the change in surface charge of the protein **[2, 6, 58]**. These can also be used for polishing steps after affinity chromatography to reduce the load of host cell DNA in the protein of interest. Various recombinant proteins such as human growth hormone, human pro-renin, Japanese encephalitis (JE) virus domain III protein, malaria vaccine, and others have been purified using ion exchange chromatography **[6]**.

4.2.3 Hydrophobic Interaction Chromatography

Hydrophobic interaction (HI) chromatography purifies proteins on the basis of their relative hydrophobicity. The recombinant proteins are usually captured onto the matrix with immobilized hydrophobic groups at high ionic strengths of approximately 1 to 1.5 M. Such high ionic strengths cause the hydrophobic patches of the proteins to open, facilitating the adsorption of the protein onto matrix. Elution is carried out by decreasing the ionic strength. HIC is used at analytical, preparatory and production scales to purify a variety of recombinant proteins, hormones, and enzymes **[6, 59–61]**.

4.2.4 Size-exclusion Chromatography

Size-exclusion chromatography (SEC) works on the basis of differences in sizes of the proteins to be separated. Mostly used for the final steps of purification, this chromatographic technique uses beads with pore sizes above or below the size of protein of interest and the impurities. If the protein of interest is a small sized molecule with large molecular weight impurities, gel matrix with pores sizes just above the protein size is used. During the flow of solution, the protein of interest is captured in the stationary phase, while the impurities are first ones to be eluted followed by the protein of interest. SEC has been successfully used to purify IgE-hyporeactive molecule and single chain variable fragment against type 1 insulinlike growth factor receptor, which were expressed in *E. coli* **[6]**.

4.3 Refolding of proteins from Inclusion Bodies

A number of recombinant proteins form insoluble inclusion bodies when expressed in *E. coli*. Inclusion bodies are protein

aggregates in non-native conformation. While inclusion bodies provide the advantage of presence of fewer amounts of contaminating proteins, the major shortcoming is the tedious process of refolding of proteins to achieve the native conformation and biologically active proteins. Various methods for protein refolding are used when faced with the issues of formation of IBs. The traditional method involves denaturation of proteins using high concentrations (6-8 M) of chaotropic agents like ammonium sulphate, urea, guanidinium hydrochloride or ionic detergents like sodium dodecyl sulfate (SDS), etc. Reducing agents like dithiothreitol (DTT) or 2-mercaptoethanol (β-ME) are added to reduce unwanted disulfide bonds formed in the protein aggregates. Once denatured, the proteins are refolded by dialysis into a compatible buffer to obtain correctly re-folded and active protein and removal of the high concentrations of denaturing agents. However, this method leads to a loss of large amount of protein as aggregates are formed during the dialysis process and hence the yield of biologically active protein is very low. Addition of different chemicals, such as sugars, polyethylene glygol (PEG), DMSO, Arginine, etc. in the refolding buffer has been shown to have a beneficial effect in the refolding of proteins. Diluting the solubilized protein to low concentrations in the refolding buffer also facilitates gradual refolding and results in high yields of active protein. Nevertheless, dilution method can be used for only small amounts of proteins and not large bulk amounts due to the requirements of large vessels and further concentration steps, which can increase the cost of protein production. For proteins with a lot of disulfide bonds, addition of glutathione (GSH)-oxidized glutathione (GSSG), DTT/GSSG or cysteine-cystine combinations prevent atmospheric oxidation of the protein cysteine residues and facilitate correct disulfide bond formation **[62–64]**.

Improved methods of protein refolding from IBs have been developed like pulse refolding method, size exclusion chromatography, adsorption chromatography and chaperone mediated refolding. In Pulse refolding method, small amounts of solubilized protein is added to the renaturation buffer at intervals to prevent aggregation of the protein and allow refolding of the small amount of protein added. This method helps in using the same tank for refolding a large amount of protein and thus brings down the requirement of buffer for refolding and improves the overall process. SEC separates large aggregates of proteins and the contaminating partners from the native form of the active protein, when used before addition of denaturing agents. This process can also be used to remove denaturing agents and refolding the

protein on the column simultaneously. The correct pore size of the matrix should be selected to separate intermediate folding variants of the proteins. SEC allows the separation of the refolded protein, partially refolded and denatured species and also prevents the interaction between various forms, decreasing the chances of aggregation. Thus, it is one of the methods for efficient separation and purification of refolded proteins. Lysozyme and carbonic anhydrase have been successfully refolded in large quantities using SEC via the sephacryl S-100 column. Adsorption chromatography in terms of ion exchange can be used for simultaneous buffer exchange, refolding and purification of proteins in IBs. The ion exchange columns allow removal of denaturing agents, while maintaining the protein solubility and preventing aggregation. Addition of refolding buffer allows the formation of native conformation of the protein and optimization of the elution buffer will provide a pure and active form of the recombinant protein. Chaperone mediated refolding method involves the use of chaperonins, GroES and GroEL, for the refolding of denatured proteins **[62–64]**.

4.4 Strategies to improve and optimize purification

The major parameter to keep in mind while purifying the recombinant protein is the chemistry of the protein. Depending on the various properties of the protein like isoelectric pH, hydrophobicity, presence of tags and co-factors and size, the method of purification may be selected. In all types of chromatographic separation of the protein, the pH of the buffers and their ionic strength are of utmost importance. Other major considerations are the type of resin used, the pore size of the matrix in case of SEC, binding capacities of the affinity resins, the load of protein and the flow rate of the column for optimal purification of the recombinant protein **[6]**.

5. Strategies for Characterization of Recombinant Proteins

The use of recombinant proteins as therapeutics, vaccines or in diagnostics purposes requirescharacterization of the protein at each step of the process of expression and purification. Furthermore, the identity, conformation and biological activity of final purified protein has to be determined to confirm that the process of expression and purification is optimal and the batch of protein produced conform to the required guidelines for biopharmaceuticals. Various techniques are used for monitoring the expression of recombinant proteins in the host cells like gel electrophoresis (SDS-PAGE), and dot-blot. Once expression levels are determined, purity of the recombinant proteins is determined. This section of the review will give insights into the different methods employed to monitor

the purity of the proteins at various levels of purification. Additionally this section will also provide an overview about the various techniques used for physicochemical and functional characterization of the recombinant proteins.

5.1 Strategies to monitor expression levels of recombinant proteins

Various pharmacopeia guidelines lay down the importance of gel electrophoresis for the preliminary purity check of the proteins. SDS-PAGE can be used to monitor the levels of expression of the recombinant protein in host cells. Though a qualitative method, the samples of whole cell lysates from different time points and conditions can be run on a SDS-PAGE followed by staining the gel with coomassie brilliant blue. This allows one to determine which condition is the best for optimal protein expression. Dot-blot is a quick method whichuses antibodies against the protein of interest or the tag. This method can be used the same way as SDS-PAGE and is a qualitative way to determine the optimal condition for recombinant protein expression **[65]**.

5.2 Strategies to determine the purity of recombinant proteins

Purity of biopharmaceutical products is of utmost importance for further use. Various methods are used for determining the level of purity of the recombinant proteins. SDS-PAGE and Western blotting are the qualitative methods for purity check of a protein. Denatured proteins samples are run on SDS-PAGE gels and hence, they are useful to estimate molecular weight of the purified protein. Silver staining of SDS-PAGE gels allows the detection of protein contaminants up to ~10 ng, thus being a handy and quick tool to check the purity. Isoelectric focusing helps in determining the presence of any charge variants in the purified proteins. Western blotting uses antibodies against the protein of interest or a tag to determine if the correct epitopes are present on the purified protein and if the protein preparation has any higher molecular weight forms of the protein of interest **[65, 66]**.

UV-visible spectroscopy is normally used to check for nucleic acid contamination in the protein. A ratio of absorbance at 260 nm and 280 nm should be ~0.57 or lower, a higher ratio indicates presence of nucleic acid contamination **[66]**. Endotoxin levels have to be controlled for most of the therapeutic proteins. Various conventional methods and newer methods have been put in place to determine the levels of endotoxins, which have been reviewed elsewhere **[67]**. Furthermore, protein homogeneity, i.e. the absence of aggregates and the presence of the same type of protein molecule in a batch, is usually measured using UV-visible and fluorescence spectroscopy, differential light scattering (DLS) and SEC. DLS is a reliable,

sensitive and rapid technique to determine the presence of higher-order structures and aggregates and also study protein-protein and protein-nucleic acid interactions. SEC is normally used to remove contaminating proteins and protein aggregates in a batch of recombinant proteins **[66]**.

5.3 Strategies for structural characterization of recombinant proteins

Heterologously expressed proteins have to be characterized for their molecular weight, primary structure, presence and identity of glycans and structural conformity for further use. Molecular weight analysis can be done using SDS-PAGE and mass spectrometry (MS). Using advanced MS techniques like matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) and/or electrospray ionization quadrupole time-of-flight (ESI-QTOF) MS, the amino acid sequence and hence identity of the protein can be determined. Protein samples are digested to smaller peptides using proteases and these resulting peptides are separated by HPLC and analyzed by MS. MS allows determination of the amino acid sequence, estimation of disulfide bonds and the type of post-translational modifications that have occurred on the protein. For glycoprotein based biopharmaceuticals, the identity of glycans have to be determined. The usual technique for monosaccharide determination is gas chromatography mass spectrometry (GC-MS), while oligosaccharide patterns can be studied using LC/MS, or MADI-TOF MS. Structural conformation of the protein is studied using techniques like nuclear magnetic resonance (NMR) and X-ray diffraction. While NMR can determine the tertiary and quaternary structures of proteins in solution, X-ray diffraction requires crystallization of proteins before analysis. All the advanced methods require long time for analysis and final read-outs **[65, 66]**.

5.4 Strategies for functional characterization of recombinant proteins

The recombinant proteins can be used as therapeutics, vaccines or diagnostic material only if they are functionally active. Various methods are used to determine if the protein is biologically active. The effector functions of the mAbs are investigated to confirm their biological activity through tests like antibody-dependent cellular cytotoxicity. Biological activity of the protein could be studied using protein-protein interactions *in vitro* and protein interactions with the cellular receptors using techniques like ligand binding assay, enzymatic reaction rates, surface plasmon resonance (SPR), etc. Zymography is another method to study the activity of enzymes while running native PAGE gels, which contain the substrate for the enzyme. Another method for functional characterization of mAbs and proteins to be

used as vaccines is Enzyme linked immunosorbent assay (ELISA). ELISA is a quantitative method which uses the antibody-antigen interactions. mAbs could be used as coating antibodies while the candidate vaccine proteins can be used as the antigen **[65, 66]**.

ELISA is of three different types **[68]**

a. Direct ELISA: antigen is coated and antibody binds directly to the antigen. Detection could be through primary antibody, which is enzyme tagged for colorimetric estimation.
b. Indirect ELISA: antigen is coated and antibody binds directly to the antigen. A primary antibody binds to the antigen and then the detection is done using an enzyme tagged secondary antibody againstthe primary antibody.
c. Sandwich ELISA: Primary antigen is coated on the plate, and it captures the antigen. The detection is through tagged antibody against the same or different epitope of the antigen.
d. Competitive ELISA: test for the presence of antibodies in a test sample. The method utilizes mixing of antigen, an enzyme tagged antibody and the test sample to check for the antibodies present in the test sample. Presence of antibodies in test sample decreases the colorimetric read-out, while its absence shows high read-out.

5.5 Strategies for studying the stability and degradation of recombinant proteins

Long term storage is the key to making the production of recombinant proteins economical. Thus, high stability of protein at its storage temperature is a must. One of the methods to monitor the stability of proteins is to measure the functional activity of the protein. Furthermore circular dichroism (CD) spectroscopy is a method to determine the stability of the structural conformation of the protein. Advanced methods using fluorescence dyes have been developed, wherein the fluorescent dyes intercalate with the hydrophobic patches of the protein and fluoresce. In case of degradation, the fluorescence of the dye increase, while stable proteins do not show any change in fluorescence. Real time analysis using these dyes can be done as it is usually done in a qPCR assay **[66]**.

6. Formulation of Recombinant Proteins

Safety and efficacy of the recombinant proteins depend on its structure and hence to maintain the protein structure, the optimal formulation of the protein is necessary to be developed. The biopharmaceutical industry suffers from various challenges for formulating the recombinant proteins for their long term stability. This section of the

article reviews the challenges in the formulation of biopharmaceuticals, the methods employed to formulate therapeutics and various strategies to optimize formulation buffers for long term stability of the products.

6.1 Challenges in formulation of recombinant proteins

Biopharmaceuticals are macromolecules with complex structures which render them difficult to formulate for delivery and long term storage. The structural changes in these macromolecules and hence loss of activity could occur due to changes in physical conditions (moisture or temperature or viscosity) or their chemical environment (pH or ionic strength). Long term stability of the biopharmaceutical products is important for the safety and efficacy of these products.

The major parameters to be considered for the formulation of a recombinant protein are the structure, physicochemical properties, the overall buffer composition and the type of excipients used **[69]**.

6.2 Strategies for formulation of recombinant proteins

As discussed earlier, to maintain the structural conformation and activity of the recombinant protein, optimal composition of the formulation buffer is a must. Various different components are present in the formulation buffer to facilitate stability of the recombinant protein. The following table provides an insight into the various components used for formulation of recombinant proteins, and their roles in the formulation **[4, 69]**.

Table 1: List of different components of formulation buffers

Sr. No.	Component	Chemicals used	Role
1	Buffers	Salts (Sodium hydroxide, Hydrochloric acid), amino acids (arginine, Histidine, glycine), acids like phosphoric acid, citric acid	Maintenance of pH slightly above or below the isoelectric pH of the protein to prevent aggregation. Usually maintained around physiological pH (7.35-7.45) to avoid any issues of pain or acidity while administration.
2	Osmotic agents	Salts like Sodium chloride and potassium chloride	Maintain isotonicity to avoid any tissue damage. Agents not added to proteins eign lyophilized to avoid salt concentration, instead added to diluents or reconstitution buffer.
3	Bulking agents	Sugars like sucrose, mannose, mannitol, and amino acids like glycine, arginine	Usually added to proteins being lyophilized to stabilize protein structural conformity.
4	Surfactants	Ionic and non-ionic detergents like SDS, Tween 20 and Tween 80	Prevent protein aggregation by intermolecular interactions and enhance solubility of proteins.
5	Preservatives	Benzyl alcohol, phenol	reduce the probability of bacterial contamination.
6	Anti-oxidants	GSH, acetylcysteine, ascorbic acid	Avoid degradation due to oxidative stress.
7	Adjuvants	Aluminum salts, squalene-based oil-in-water emulsions, monophosphoryl lipid A	Modification of immune response to the therapeutic proteins.

6.3 Techniques for formulation of recombinant proteins

Recombinant proteins are formulated either as liquid solutions or as lyophilized product. Liquid solutions have been used for recombinant insulin, Peg-filgrastim and mAbs like adalimumab and aflibercept. However, liquid formulation suffers from the drawback of lower stability over longer periods of storage. To overcome this shortcoming, lyophilization of the proteins has been explored. Lyophilization is a three step process of freezing, primary drying, and secondary drying. During the freezing step, the water gets crystallized at a temperature below the freezing point of each component of the buffer (eutectic temperature). In the primary drying, crystallized water and all the water molecules not bound to the protein are lost due to sublimation under reduced pressure. During secondary drying stage, the water molecules bound to the protein molecules and Excipients are removed. The resulting dry powder has protein and all the salts and other components, which could be easily reconstituted. The drawback of using lyophilized powders is the possibility of a change in the pH which could lead to change in protein structure when reconstituted for administration **[4]**. Spray drying is used to prepare dry powder from liquid solution, and then fabricating the powder into tablets, capsule or other forms for oral usage. Usually bulking agents like mannitol are used for spray drying process. One of the examples where spray drying was successfully used is the thyrotropin releasing hormone **[69]**.

7. Application of Recombinant Proteins

Recombinant proteins are being widely used as therapeutics, vaccines, in developing diagnostic kits and even in industries. Industries employ recombinant enzymes and biopolymers for various purposes.

7.1 Recombinant proteins as therapeutics and vaccines

One of the oldest examples of a recombinant protein therapeutic is insulin. Currently, a lot of different recombinant proteins like erythropoietin, clotting factors, hormones, etc. are being used as therapeutics. The biopharmaceutical industries are thriving on the production and sales of various monoclonal antibodies, which are being used as a therapy for a lot of life-threatening diseases like cancer. Furthermore, with the advancement in the field of cell and gene therapy, recombinant virus-like protein (AAV and LV) generation from large scale cultures has seen a major boost in the technology. These will be used for curing genetic disorders and cancers **[1]**. Recombinant growth factors and cytokines like GCSF, IL-2, and IL-15 etc. are being used for *in vitro* cell cultures, which can be then used for cell based therapies.

Recombinant proteins are being used as vaccines. Recombinant protein vaccines

against Hepatitis B virus (HBV), human papilloma virus (HPV) are already available. The recent pandemic of COVID-19 called for prompt and speedy research and development of a large number of recombinant protein based vaccine candidates across the globe **[70, 71]**. A lot of recombinant proteins like HBsAg, HCG, CRP, etc. are used for the development of *in vitro* diagnostic kits and point-of-care diagnostics.

7.2 Application of recombinant proteins in industries

Recombinant proteins like enzymes are used in various industries like food industry (amylases, pectinase), biofuel industry (cellulose, xylanase, pectinase), leather industry (proteases and lipases) and biopharmaceutical industry (recombinant endonucleases, PCR enzymes, etc.) **[72, 73]**. Another avenue is the biopolymers, which may be naturally or synthetically obtained. Biopolymers could be proteins, nucleic acids, carbohydrates and other derivatives which are naturally present in living organisms like plants, animals, insects and micro-organisms **[74]**. Biopolymers are finding applications in various fields like food packaging, medicines, drug delivery, biomedical device coatings, etc. **[75]**. Algae are the most commonly used biomass for the production of biopolymers. Various bacteria can be metabolically engineered to produce biopolymers of choice to prevent the burden on natural resources, *viz. E. coli*, *Akaligenes eurrophus* and *Bacillus megaterium* **[76, 77]**. One of the studies have developed a fermentation based production system for Poly(3-hydroxybutyrate-co-3-hydroxyvalerate) (PHBV), a biodegradable plastic polymer, using genetically engineered *Bacillus megaterium* **[76]**.

7.3 Application of recombinant proteins in agriculture

Transgenic plants have been developed over years for better yields of crops and fruits or vegetables and making them resistant to different stress like heat, drought, pests, etc. The classic example of transgenic plants for agriculture is the Bt cotton, where in the cotton seeds carry the Cry toxin gene of *Bacillus thuringiensis* (Bt). When grown, the cotton plants are resistant to cotton pests, like *Helicoverpa armigera* **[78]**. Apart from agriculture, plants are also used for recombinant production of biopharmaceuticals on large scale as they are easy to handle and can be grown in large amounts in a field, a practice called molecular farming **[79]**. Genetically engineered tobacco plants were used for the expression of three monoclonal antibodies (c13C6, c2G4, and c4G7) which were a part of the experimental Ebola drug cocktail, ZMapp **[80]**.

8. CONCLUSION

The review provides and in depth insight into the various aspects of recombinant

protein production process. The major considerations that impact the choice of host for expression with a major focus on *E. coli* as the expression host were discussed. The downstream processes employed for purification of the recombinant protein and their characterization at each step of purification was reviewed in detail. Further the review also shed light on the challenges of formulation and the various strategies employed for formulation and storage of recombinant proteins. Lastly, the review enlisted the different applicatiosn of recombinant proteins in medicine and industries. While a lot of advances have been made in the field of recombinant protein production, there are countless avenues where further research and development could facilitate the whole process making it more economical. Also, the use of genetically modified organisms involves regulatory hurdles, especially plants systems. The challenges need proper investigation and action from researchers, industries and regulatory bodies for a sustainable, eco-friendly, economic and quick system for recombinant protein production.

REFERENCES

[1] G. Walsh, E. Walsh, Biopharmaceutical benchmarks 2022, Nat. Biotechnol. 40 (2022) 1722–1760. https://doi.org/10.1038/s41587-022-01582-x.

[2] N.K. Tripathi, A. Shrivastava, Recent Developments in Bioprocessing of Recombinant Proteins: Expression Hosts and Process Development, Front. Bioeng. Biotechnol. 7 (2019) 420. https://doi.org/10.3389/fbioe.2019.00420.

[3] G. Ben-Nissan, S. Vimer, S. Warszawski, A. Katz, M. Yona, T. Unger, Y. Peleg, D. Morgenstern, H. Cohen-Dvashi, R. Diskin, S.J. Fleishman, M. Sharon, Rapid characterization of secreted recombinant proteins by native mass spectrometry, Commun. Biol. 1 (2018) 213. https://doi.org/10.1038/s42003-018-0231-3.

[4] R. Ribeiro, T.R. Abreu, A.C. Silva, J. Gonçalves, J.N. Moreira, Insights on the Formulation of Recombinant Proteins, in: A.C. Silva, J.N. Moreira, J.M.S. Lobo, H. Almeida (Eds.), Curr. Appl. Pharm. Biotechnol., Springer International Publishing, Cham, 2019: pp. 23–54. https://doi.org/10.1007/10_2019_119.

[5] A.R. Gomes, S.M. Byregowda, B.M. Veeregowda, V. Balamurugan, An Overview of Heterologous

Expression Host Systems for the Production of Recombinant Proteins, Adv. Anim. Vet. Sci. 4 (2016) 346–356. https://doi.org/10.14737/journal.aavs/2016/4.7.346.356.

[6] N.K. Tripathi, Production and Purification of Recombinant Proteins from *Escherichia coli*, ChemBioEng Rev. 3 (2016) 116–133. https://doi.org/10.1002/cben.201600002.

[7] C. Luan, H.W. Zhang, D.G. Song, Y.G. Xie, J. Feng, Y.Z. Wang, Expressing antimicrobial peptide cathelicidin-BF in Bacillus subtilis using SUMO technology, Appl. Microbiol. Biotechnol. 98 (2014) 3651–3658. https://doi.org/10.1007/s00253-013-5246-6.

[8] R. Biedendieck, A Bacillus megaterium System for the Production of Recombinant Proteins and Protein Complexes, in: M.C. Vega (Ed.), Adv. Technol. Protein Complex Prod. Charact., Springer International Publishing, Cham, 2016: pp. 97–113. https://doi.org/10.1007/978-3-319-27216-0_7.

[9] Y. Le Loir, V. Azevedo, S.C. Oliveira, D.A. Freitas, A. Miyoshi, L.G. Bermúdez-Humarán, S. Nouaille, L.A. Ribeiro, S. Leclercq, J.E. Gabriel, V.D. Guimaraes, M.N. Oliveira, C. Charlier, M. Gautier, P. Langella, Protein secretion in Lactococcus lactis : an efficient way to increase the overall heterologous protein production, Microb. Cell Factories. 4 (2005) 2. https://doi.org/10.1186/1475-2859-4-2.

[10] T. Allain, N.M. Mansour, M.M.A. Bahr, R. Martin, I. Florent, P. Langella, L.G. Bermúdez-Humarán, A new lactobacilli *in vivo* expression system for the production and delivery of heterologous proteins at mucosal surfaces, FEMS Microbiol. Lett. 363 (2016) fnw117. https://doi.org/10.1093/femsle/fnw117.

[11] S. Srinivasan, G.C. Barnard, T.U. Gerngross, A Novel High-Cell-Density Protein Expression System Based on *Ralstonia eutropha*, Appl. Environ. Microbiol. 68 (2002) 5925–5932. https://doi.org/10.1128/AEM.68.12.5925-5932.2002.

[12] G.C. Barnard, G.E. Henderson, S. Srinivasan, T.U. Gerngross, High level recombinant protein expression in Ralstonia eutropha

using T7 RNA polymerase based amplification, Protein Expr. Purif. 38 (2004) 264–271. https://doi.org/10.1016/j.pep.2004.09.001.

[13] A. Vieira Gomes, T. Souza Carmo, L. Silva Carvalho, F. Mendonça Bahia, N. Parachin, Comparison of Yeasts as Hosts for Recombinant Protein Production, Microorganisms. 6 (2018) 38. https://doi.org/10.3390/microorganisms6020038.

[14] F.W. Krainer, C. Dietzsch, T. Hajek, C. Herwig, O. Spadiut, A. Glieder, Recombinant protein expression in Pichia pastoris strains with an engineered methanol utilization pathway, Microb. Cell Factories. 11 (2012) 22. https://doi.org/10.1186/1475-2859-11-22.

[15] M. Karbalaei, S.A. Rezaee, H. Farsiani, *Pichia pastoris* : A highly successful expression system for optimal synthesis of heterologous proteins, J. Cell. Physiol. 235 (2020) 5867–5881. https://doi.org/10.1002/jcp.29583.

[16] N. Kaushik, U. Lamminmäki, N. Khanna, G. Batra, Enhanced cell density cultivation and rapid expression-screening of recombinant Pichia pastoris clones in microscale, Sci. Rep. 10 (2020) 7458. https://doi.org/10.1038/s41598-020-63995-5.

[17] Y. Pan, J. Yang, J. Wu, L. Yang, H. Fang, Current advances of Pichia pastoris as cell factories for production of recombinant proteins, Front. Microbiol. 13 (2022) 1059777. https://doi.org/10.3389/fmicb.2022.1059777.

[18] C.P. Landowski, E. Mustalahti, R. Wahl, L. Croute, D. Sivasiddarthan, A. Westerholm-Parvinen, B. Sommer, C. Ostermeier, B. Helk, J. Saarinen, M. Saloheimo, Enabling low cost biopharmaceuticals: high level interferon alpha-2b production in Trichoderma reesei, Microb. Cell Factories. 15 (2016) 104. https://doi.org/10.1186/s12934-016-0508-5.

[19] K.M.H. Nevalainen, V.S.J. Te'o, P.L. Bergquist, Heterologous protein expression in filamentous fungi, Trends Biotechnol. 23 (2005) 468–474. https://doi.org/10.1016/j.tibtech.2005.06.002.

[20] F. Ntana, U.H. Mortensen, C. Sarazin, R. Figge, Aspergillus: A Powerful Protein Production Platform, Catalysts. 10 (2020) 1064.

https://doi.org/10.3390/catal10091064.

[21] Q. Wang, C. Zhong, H. Xiao, Genetic Engineering of Filamentous Fungi for Efficient Protein Expression and Secretion, Front. Bioeng. Biotechnol. 8 (2020) 293. https://doi.org/10.3389/fbioe.2020.00293.

[22] A.A. Sakekar, S.R. Gaikwad, N.S. Punekar, Protein expression and secretion by filamentous fungi, J. Biosci. 46 (2021) 5. https://doi.org/10.1007/s12038-020-00120-8.

[23] W. Weber, M. Fussenegger, Insect Cell-Based Recombinant Protein Production, in: Cell Tissue React. Eng., Springer Berlin Heidelberg, Berlin, Heidelberg, 2009: pp. 263–277. https://doi.org/10.1007/978-3-540-68182-3_6.

[24] R.P. Lemaitre, A. Bogdanova, B. Borgonovo, J.B. Woodruff, D.N. Drechsel, FlexiBAC: a versatile, open-source baculovirus vector system for protein expression, secretion, and proteolytic processing, BMC Biotechnol. 19 (2019) 20. https://doi.org/10.1186/s12896-019-0512-z.

[25] E.A. McKenzie, W.M. Abbott, Expression of recombinant proteins in insect and mammalian cells, Methods. 147 (2018) 40–49. https://doi.org/10.1016/j.ymeth.2018.05.013.

[26] K.M. Heffner, Q. Wang, D.B. Hizal, Ö. Can, M.J. Betenbaugh, Glycoengineering of Mammalian Expression Systems on a Cellular Level, in: E. Rapp, U. Reichl (Eds.), Adv. Glycobiotechnology, Springer International Publishing, Cham, 2018: pp. 37–69. https://doi.org/10.1007/10_2017_57.

[27] D.L. Hacker, ed., Recombinant Protein Expression in Mammalian Cells: Methods and Protocols, Springer New York, New York, NY, 2018. https://doi.org/10.1007/978-1-4939-8730-6.

[28] M.-E. Lalonde, Y. Durocher, Therapeutic glycoprotein production in mammalian cells, J. Biotechnol. 251 (2017) 128–140. https://doi.org/10.1016/j.jbiotec.2017.04.028.

[29] T. Nishida, S. Kubota, M. Takigawa, Production of Recombinant CCN2 Protein by Mammalian Cells, in: M. Takigawa (Ed.), CCN Proteins, Springer New York, New York, NY, 2017: pp. 95–105. https://doi.org/10.1007/978-1-4939-6430-7_10.

[30] B. Owczarek, A. Gerszberg, K. Hnatuszko-Konka, A Brief Reminder of Systems of Production and Chromatography-Based Recovery of Recombinant Protein Biopharmaceuticals, BioMed Res. Int. 2019 (2019) 1–13.

https://doi.org/10.1155/2019/4216060.

[31] J.J. Priola, N. Calzadilla, M. Baumann, N. Borth, C.G. Tate, M.J. Betenbaugh, High-throughput screening and selection of mammalian cells for enhanced protein production, Biotechnol. J. 11 (2016) 853–865. https://doi.org/10.1002/biot.201500579.

[32] Q. Wang, B. Yin, C.-Y. Chung, M.J. Betenbaugh, Glycoengineering of CHO Cells to Improve Product Quality, in: P. Meleady (Ed.), Heterologous Protein Prod. CHO Cells, Springer New York, New York, NY, 2017: pp. 25–44. https://doi.org/10.1007/978-1-4939-6972-2_2.

[33] D. Garenne, M.C. Haines, E.F. Romantseva, P. Freemont, E.A. Strychalski, V. Noireaux, Cell-free gene expression, Nat. Rev. Methods Primer. 1 (2021) 49. https://doi.org/10.1038/s43586-021-00046-x.

[34] K. Khambhati, G. Bhattacharjee, N. Gohil, D. Braddick, V. Kulkarni, V. Singh, Exploring the Potential of Cell-Free Protein Synthesis for Extending the Abilities of Biological Systems, Front. Bioeng. Biotechnol. 7 (2019) 248. https://doi.org/10.3389/fbioe.2019.00248.

[35] L. Zhang, X. Lin, T. Wang, W. Guo, Y. Lu, Development and comparison of cell-free protein synthesis systems derived from typical bacterial chassis, Bioresour. Bioprocess. 8 (2021) 58. https://doi.org/10.1186/s40643-021-00413-2.

[36] G.L. Rosano, E.A. Ceccarelli, Recombinant protein expression in Escherichia coli: advances and challenges, Front. Microbiol. 5 (2014). https://doi.org/10.3389/fmicb.2014.00172.

[37] G.L. Rosano, E.S. Morales, E.A. Ceccarelli, New tools for recombinant protein production in *Escherichia coli* : A 5-year update, Protein Sci. 28 (2019) 1412–1422. https://doi.org/10.1002/pro.3668.

[38] T. Baumgarten, A.J. Ytterberg, R.A. Zubarev, J.-W. De Gier, Optimizing Recombinant Protein Production in the Escherichia coli Periplasm Alleviates Stress, Appl. Environ. Microbiol. 84 (2018) e00270-18. https://doi.org/10.1128/AEM.00270-18.

[39] S.K. Gupta, A.K. Dangi, M. Smita, S. Dwivedi, P. Shukla, Effectual Bioprocess Development for Protein Production, in: Appl. Microbiol. Bioeng., Elsevier, 2019: pp. 203–227. https://doi.org/10.1016/B978-0-12-815407-6.00011-3.

[40] S.K. Gupta, P. Shukla, Advanced technologies for improved expression of recombinant proteins in bacteria: perspectives and applications, Crit. Rev. Biotechnol. 36 (2016) 1089–

1098. https://doi.org/10.3109/07388551.2015.1084264.

[41] S.K. Gupta, P. Shukla, Sophisticated Cloning, Fermentation, and Purification Technologies for an Enhanced Therapeutic Protein Production: A Review, Front. Pharmacol. 8 (2017) 419. https://doi.org/10.3389/fphar.2017.00419.

[42] S.K. Gupta, P. Shukla, Glycosylation control technologies for recombinant therapeutic proteins, Appl. Microbiol. Biotechnol. 102 (2018) 10457–10468. https://doi.org/10.1007/s00253-018-9430-6.

[43] S.K. Gupta, P. Shukla, Microbial platform technology for recombinant antibody fragment production: A review, Crit. Rev. Microbiol. 43 (2017) 31–42. https://doi.org/10.3109/1040841X.2016.1150959.

[44] M. Liu, B. Wang, F. Wang, Z. Yang, D. Gao, C. Zhang, L. Ma, X. Yu, Soluble expression of single-chain variable fragment (scFv) in Escherichia coli using superfolder green fluorescent protein as fusion partner, Appl. Microbiol. Biotechnol. 103 (2019) 6071–6079. https://doi.org/10.1007/s00253-019-09925-6.

[45] J. Lobstein, C.A. Emrich, C. Jeans, M. Faulkner, P. Riggs, M. Berkmen, SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm, Microb. Cell Factories. 11 (2012) 753. https://doi.org/10.1186/1475-2859-11-56.

[46] Y. Han, W. Guo, B. Su, Y. Guo, J. Wang, B. Chu, G. Yang, High-level expression of soluble recombinant proteins in Escherichia coli using an HE-maltotriose-binding protein fusion tag, Protein Expr. Purif. 142 (2018) 25–31. https://doi.org/10.1016/j.pep.2017.09.013.

[47] T. Ojima-Kato, S. Nagai, H. Nakano, N-terminal SKIK peptide tag markedly improves expression of difficult-to-express proteins in Escherichia coli and Saccharomyces cerevisiae, J. Biosci. Bioeng. 123 (2017) 540–546. https://doi.org/10.1016/j.jbiosc.2016.12.004.

[48] V. Paraskevopoulou, F. Falcone, Polyionic Tags as Enhancers of Protein Solubility in Recombinant Protein Expression, Microorganisms. 6 (2018) 47. https://doi.org/10.3390/microorganisms6020047.

[49] G. Vandemoortele, S. Eyckerman, K. Gevaert, Pick a Tag and Explore the Functions of Your Pet Protein, Trends Biotechnol. 37 (2019) 1078–1090. https://doi.org/10.1016/j.tibtech.2019.03.016.

[50] T. Vargas-Cortez, J.R. Morones-Ramirez, I. Balderas-Renteria, X. Zarate, Expression and purification of recombinant proteins in Escherichia coli tagged with a small metal-binding protein from Nitrosomonas europaea, Protein Expr. Purif. 118 (2016) 49–54. https://doi.org/10.1016/j.pep.2015.10.009.

[51] T. Vargas-Cortez, J.R. Morones-Ramirez, I. Balderas-Renteria, X. Zarate, Production of recombinant proteins in Escherichia coli tagged with the fusion protein CusF3H+, Protein Expr. Purif. 132 (2017) 44–49. https://doi.org/10.1016/j.pep.2017.01.006.

[52] M. Vedadi, C.H. Arrowsmith, A. Allali-Hassani, G. Senisterra, G.A. Wasney, Biophysical characterization of recombinant proteins: A key to higher structural genomics success, J. Struct. Biol. 172 (2010) 107–119. https://doi.org/10.1016/j.jsb.2010.05.005.

[53] C.L. Young, Z.T. Britton, A.S. Robinson, Recombinant protein expression and purification: A comprehensive review of affinity tags and microbial applications, Biotechnol. J. 7 (2012) 620–634. https://doi.org/10.1002/biot.201100155.

[54] T.B. Gang, D.J. Hammond, S.K. Singh, D.A. Ferguson, V.K. Mishra, A. Agrawal, The Phosphocholine-binding Pocket on C-reactive Protein Is Necessary for Initial Protection of Mice against Pneumococcal Infection, J. Biol. Chem. 287 (2012) 43116–43125. https://doi.org/10.1074/jbc.M112.427310.

[55] D. Thompson, M.B. Pepys, S.P. Wood, The physiological structure of human C-reactive protein and its complex with phosphocholine, Structure. 7 (1999) 169–177. https://doi.org/10.1016/S0969-2126(99)80023-9.

[56] H. Diibeli, R. Gent, E. Hochuli, GxHis-Ni-NTA Chromatography as a Superior Technique in Recombinant Protein ExpressiodEkrification, (n.d.).

[57] K.M. Łącki, F.J. Riske, Affinity Chromatography: An Enabling Technology for Large-Scale Bioprocessing, Biotechnol. J. 15 (2020) 1800397. https://doi.org/10.1002/biot.201800397.

[58] Y.M. Moustafa, R.E. Morsi, Ion Exchange Chromatography - An Overview, in: D. Martin (Ed.), Column Chromatogr., InTech, 2013. https://doi.org/10.5772/55652.

[59] J.T. McCue, Chapter 25 Theory and Use of Hydrophobic Interaction Chromatography in Protein Purification Applications, in: Methods Enzymol., Elsevier, 2009: pp. 405–414. https://doi.org/10.1016/S0076-6879(09)63025-1.

[60] M.R. Mirani, F. Rahimpour, Thermodynamic modelling of hydrophobic interaction chromatography of biomolecules in the presence of salt, J. Chromatogr. A. 1422 (2015) 170–177. https://doi.org/10.1016/j.chroma.2015.10.019.

[61] H.P. Jennissen, Hydrophobic Interaction Chromatography, Encycl. LIFE Sci. (2001).

[62] S.M. Singh, A.K. Panda, Solubilization and refolding of bacterial inclusion body proteins, J. Biosci. Bioeng. 99 (2005) 303–310. https://doi.org/10.1263/jbb.99.303.

[63] H. Yamaguchi, M. Miyazaki, Refolding Techniques for Recovering Biologically Active Recombinant Proteins from Inclusion Bodies, Biomolecules. 4 (2014) 235–251. https://doi.org/10.3390/biom4010235.

[64] A.S. Rathore, P. Bade, V. Joshi, M. Pathak, S.K. Pattanayek, Refolding of biotech therapeutic proteins expressed in bacteria: review: Refolding of biotech therapeutic proteins expressed in bacteria, J. Chem. Technol. Biotechnol. 88 (2013) 1794–1806. https://doi.org/10.1002/jctb.4152.

[65] M.K. Parr, O. Montacir, H. Montacir, Physicochemical characterization of biopharmaceuticals, J. Pharm. Biomed. Anal. 130 (2016) 366–389. https://doi.org/10.1016/j.jpba.2016.05.028.

[66] C. Oliveira, L. Domingues, Guidelines to reach high-quality purified recombinant proteins, Appl. Microbiol. Biotechnol. 102 (2018) 81–92. https://doi.org/10.1007/s00253-017-8623-8.

[67] E.C. Dullah, C.M. Ongkudon, Current trends in endotoxin detection and analysis of endotoxin–protein interactions, Crit. Rev. Biotechnol. 37 (2017) 251–261. https://doi.org/10.3109/07388551.2016.1141393.

[68] S. Catherine, Types of ELISA-Enzyme Linked Immunosorbent Assays, (n.d.).

[69] V. Truong-Le, P.M. Lovalenti, A.M. Abdul-Fattah, Stabilization Challenges and Formulation Strategies Associated with Oral Biologic Drug Delivery Systems, Adv. Drug Deliv. Rev. 93 (2015) 95–108. https://doi.org/10.1016/j.addr.2015.08.001.

[70] M.T. De Pinho Favaro, J. Atienza-Garriga, C. Martínez-Torró, E. Parladé, E. Vázquez, J.L. Corchero, N. Ferrer-Miralles, A. Villaverde, Recombinant vaccines in 2022: a perspective from the cell factory, Microb. Cell Factories. 21 (2022) 203. https://doi.org/10.1186/s12934-022-01929-8.

[71] J. Pollet, W.-H. Chen, U. Strych, Recombinant protein vaccines, a proven approach against coronavirus pandemics, Adv. Drug Deliv. Rev. 170

(2021) 71–82. https://doi.org/10.1016/j.addr.2021.01.001.

[72] J. Puetz, F.M. Wurm, Recombinant Proteins for Industrial versus Pharmaceutical Purposes: A Review of Process and Pricing, Processes. 7 (2019) 476. https://doi.org/10.3390/pr7080476.

[73] M.F. Adegboye, O.B. Ojuederie, P.M. Talia, O.O. Babalola, Bioprospecting of microbial strains for biofuel production: metabolic engineering, applications, and challenges, Biotechnol. Biofuels. 14 (2021) 5. https://doi.org/10.1186/s13068-020-01853-2.

[74] J. Baranwal, B. Barse, A. Fais, G.L. Delogu, A. Kumar, Biopolymer: A Sustainable Material for Food and Medical Applications, Polymers. 14 (2022) 983. https://doi.org/10.3390/polym14050983.

[75] R. Gheorghita, L. Anchidin-Norocel, R. Filip, M. Dimian, M. Covasa, Applications of Biopolymers for Drugs and Probiotics Delivery, Polymers. 13 (2021) 2729. https://doi.org/10.3390/polym13162729.

[76] M. Akdoğan, E. Çelik, Enhanced production of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) biopolymer by recombinant Bacillus megaterium in fed-batch bioreactors, Bioprocess Biosyst. Eng. 44 (2021) 403–416. https://doi.org/10.1007/s00449-020-02452-z.

[77] S. Mohan, O.S. Oluwafemi, N. Kalarikkal, S. Thomas, S.P. Songca, Biopolymers – Application in Nanoscience and Nanotechnology, in: F.K. Parveen (Ed.), Recent Adv. Biopolym., InTech, 2016. https://doi.org/10.5772/62225.

[78] L. Liu, R. Guo, Q. Qin, J. Fu, B. Liu, Expression of Bt Protein in Transgenic Bt Cotton Plants and Ecological Fitness of These Plants in Different Habitats, Front. Plant Sci. 11 (2020) 1209. https://doi.org/10.3389/fpls.2020.01209.

[79] B. Shanmugaraj, C.J. I. Bulaon, W. Phoolcharoen, Plant Molecular Farming: A Viable Platform for Recombinant Biopharmaceutical Production, Plants. 9 (2020) 842. https://doi.org/10.3390/plants9070842.

[80] K.Y. Park, S.J. Wi, Potential of plants to produce recombinant protein products, J. Plant Biol. 59 (2016) 559–568. https://doi.org/10.1007/s12374-016-0482-9.