

Review

Challenges in Expression and Purification of Functional Fab Fragments in *E. coli*: Current Strategies and Perspectives

Rucha S. Patil ¹, Anupa Anupa ² , Jaya A. Gupta ¹ and Anurag S. Rathore ^{1,*} 

¹ Department of Chemical Engineering, IIT Delhi, Hauz Khas, New Delhi 110016, India; patilrucha14@gmail.com (R.S.P.); jaya.2388@gmail.com (J.A.G.)

² School of Interdisciplinary Research, IIT Delhi, Hauz Khas, New Delhi 110016, India; anupa.pahuja@gmail.com

* Correspondence: asrathore@biotechcmz.com; Tel.: +91-9650770650

Abstract: Microbial host systems remain the most efficient and cost-effective chassis for biotherapeutics production. *Escherichia coli* is often the preferred host due to ease of cloning, scale-up, high product yields, and most importantly, cost-effective cultivation. *E. coli* often experience difficulties in producing biologically active therapeutics such as Fab fragments, which require protein folding and subsequent three-dimensional structure development. This paper outlines the recent improvements in upstream and downstream unit operations for producing Fab fragments in *E. coli*. Monoclonal antibody fragments (Fab) are a rising class of biotherapeutics and their production has been optimised using coexpression of molecular chaperones such as DsbC or DnaK–DnaJ–GrpE, as well as strain engineering for post-translational modifications such as disulphide bridging. Different media systems such as EnBase and combining nitrogen source supplementation with low-temperature cultivation have resulted in improvement in cell integrity, protein expression, and protein refolding. The recovery of native proteins from insoluble inclusion bodies can be improved by adjusting refolding conditions, as well as by incorporating multimodal and affinity chromatography for achieving high product yields in purification. Recent developments summarised in this review may tune the *E. coli* expression system to produce more complex and glycosylated proteins for therapeutic use in the near future.

Keywords: biotherapeutics; Fab fragment; protein folding; strain engineering; refolding; upstream and downstream unit operations



Citation: Patil, R.S.; Anupa, A.; Gupta, J.A.; Rathore, A.S. Challenges in Expression and Purification of Functional Fab Fragments in *E. coli*: Current Strategies and Perspectives. *Fermentation* **2022**, *8*, 175. <https://doi.org/10.3390/fermentation8040175>

Academic Editors: Oliver Spadiut, Julian Quehenberger, David Wurm, Julian Kopp and Stefan Pflügl

Received: 28 February 2022

Accepted: 1 April 2022

Published: 9 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Antibody fragments such as Fab and ScFv are small fragments that retain the parent antibody's antigen binding ability [1]. The Fab molecule has several advantages over a full antibody, the most important of which is its ease of penetration and clearance. As these molecules are not typically glycosylated, the bacterial expression system provides a compelling cost-effective production system [2].

The widespread use of the bacterial host system can be attributed to its several advantages over other options (such as mammalian cell culture). Fast growth, a well-known genome, and simple gene manipulation are some of these advantages [3]. While the mammalian expression system has a higher risk of contamination and higher costs for the media and subsequent processes, it has lower cell line stability [4]. Disulphide bond formation is, however, limited in prokaryotic expression systems. As such, protein misfolding, aggregation, and accumulation result in the formation of insoluble inclusion bodies, which, together with high chances of mutations, codon bias, inadequate translation, and protein degradation, are some of the major limitations of prokaryotic expression systems [5]. Significant advancements in the last few decades have allowed us to overcome these challenges and achieve sufficient expression and proper folding.

Many recent advances in strain engineering and cloning have increased the suitability of microbial hosts for biotherapeutic production. Bacterial cytoplasm has a reducing

environment, which prevents the formation of disulphide bonds under ordinary cytoplasmic conditions. Thus, researchers have tried to export complex polypeptides such as Fab fragments to the periplasmic space by several translocation pathways and periplasmic signal peptides [6]. The use of molecular chaperones to assist protein folding has been demonstrated to avoid the yield-limiting refolding process [7]. Various strains have been engineered to enhance heterologous protein production, including the development of a protease-deficient strain, modification of the outer cell membrane to secrete properly folded protein in the growth medium to avoid impurities from the host cell, and encoding genes for rare codons and molecular chaperones in the host cells [6]. Researchers have demonstrated that when anti-TNF Fab was coexpressed with the plasmid coding for the DnaK/DnaJ/GrpE chaperones, the soluble fraction increased [8].

Cells with a modified cytoplasm (reducing conditions) have been developed to enhance the formation of proper disulphide bonds and the subsequent folding of proteins [9]. Numerous process development studies have been conducted to improve the overall product yield. These studies included media screening, additives, and amino acid supplementation, along with optimisation of physical parameters [3]. Researchers have implemented a stress minimisation strategy, lowering the inducer concentration and growth temperature, to avoid certolizumab misfolding and mislocalisation [10]. Additionally, an increase in the active soluble fraction of Lucentis has been reported when expressed using the SUMO fusion technology in conjunction with the EnBase media cultivation method [1].

Along with the strategies outlined above, researchers have attempted to improve purification yield by optimising downstream process steps. Multistep chromatographic approaches have been attempted to eliminate product-related contaminants and improve purification, primarily for soluble Fab molecules. A three-dimensional chromatography method that comprises two affinity steps with protein G and protein L, followed by a cation exchange chromatography step has been proposed to identify, resolve, and characterise product-related impurities, hence facilitating improved purification [11].

Refolding is often the rate-limiting step and is mostly responsible for the low purification yield in the case of insoluble Fabs and aggregated inclusion bodies. Numerous refolding experiments have been undertaken to optimise physical factors such as temperature, pH, denaturant concentration, redox reagent concentration, and inclusion body concentration. A two-stage design of experiments strategy has been adopted to optimise oxidative refolding conditions maintained by an adequate dilution factor and redox reagent ratio based on their understanding of Fab unfolding, resulting in a refolding yield of 56 percent [12].

In the last several years, various strategies for Fab production have been explored. In view of this, we attempt to review major advancements in Fab fragment manufacturing in *E. coli*-based expression systems in this article.

2. Major Developments in Upstream Processing of Fab Fragments in *E. coli*

Production of a Fab fragment in a bacterial host is an attractive option due to the above-mentioned advantages. Recent advances in cloning technologies and host cell engineering along with process development have resulted in a significant improvement in the expression of Fab molecules in *E. coli*.

2.1. Expression and Localisation of Fab Fragment

Production of Fab fragments in *E. coli* typically offers multiple challenges, including low expression, poor yield of properly folded Fab fragments, and unequal expression of heavy and light chain in case of bicistronic expression. Oxidising conditions that allow disulphide bond formation are critical for the soluble expression of Fab fragments. This is achieved by directing the expressed molecule to the periplasmic space, which provides the oxidising environment required for disulphide bond formation. Inefficient translocation is one of the major hurdles in periplasmic expression, and different Sec pathways and codon modulation in signal sequences have been tried to increase the expression of properly

folded soluble fractions efficiently. The pertinent cytoplasmic and periplasmic pathways have been illustrated in Figure 1. The PelB signal sequence library was created by codon modulation, and its effect on Fab expression has been reported [13]. Further, the fifth leucine position of light chain PelB has been proposed to be significant for Fab fragment expression. Fab leakage has been observed in the case of OmpA-signalled Fab, and researchers have used different signal sequences to replace heavy and light chain OmpA [14]. Further, DsbA has been used to prevent leakage and for re-routing of secretory pathways to rescue the poor performance, with this strategy demonstrating significant improvement in Fab expression. In a recent development, it has been observed that when OmpA of both chains is replaced with DsbA, uncleaved soluble light chains are observed in both cases and the amount of uncleaved soluble fragments is more in the plasmid-based system, compared with genome-integrated clones. Based on these observations, we can conclude that the use of different signal sequences for heavy and light chains can increase the periplasmic translocation, as well as cleavage efficiency, and can prevent overburden on the signal translocation pathway [15].

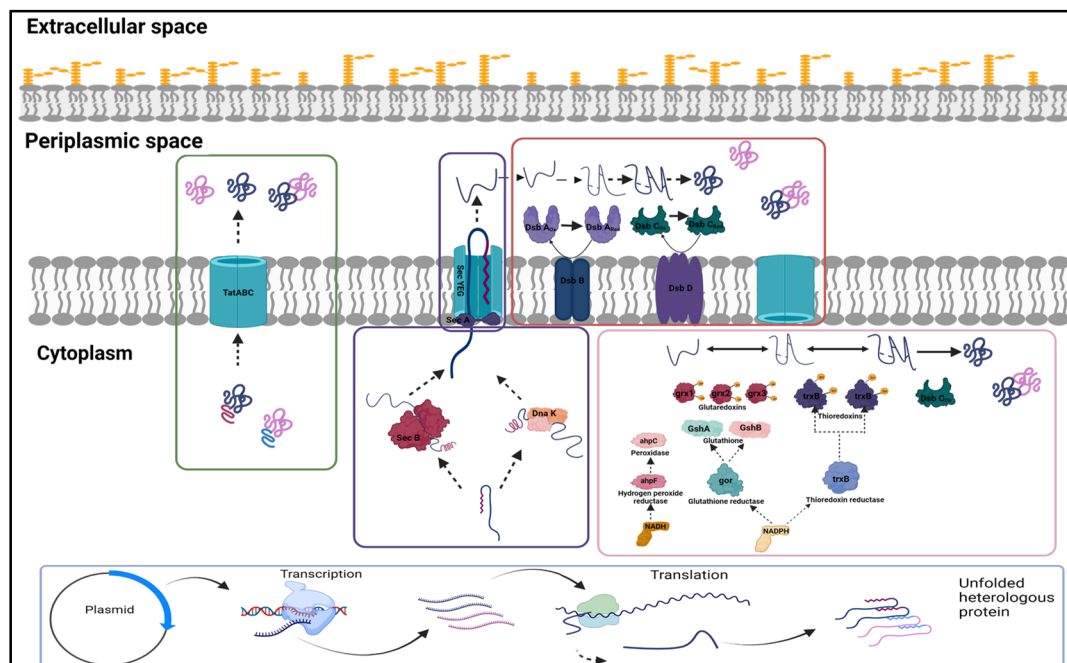


Figure 1. Schematic of (A) translocation pathways: Sec- and Tat-dependent protein transport pathways. The Sec pathway transports loosely folded or unfolded proteins across the plasma membrane. Cytoplasmic targeting factors such as SecB help control protein targeting and folding. On the other hand, the SecYEG channel and the translocation ATPase SecA use ATP as a driving force to push the protein through the membrane. The transmembrane proton gradient may also power translocation. Periplasmic protein folding is mediated by oxidases DsbA and DsbC and chaperones, which give final active and protease-resistant conformation of translocated protein. Unlike the Sec pathway, the Tat pathway transports fully folded cofactor proteins. TatA, TatB, and TatC may make up the Tat translocase. Transmembrane proton-motive force drives protein transport via Tat; (B) cytoplasmic disulphide bond formation: Thioredoxin (TrxB) and glutaredoxin (Gor) pathways are required for reducing cytoplasmic protein disulphide bonds. While one relies on glutathione reductase and tripeptide glutathione, the other relies on thioredoxin reductase and one of the two thioredoxin family members. They both use NADPH. Glutaredoxins (Grx1, Grx2, Grx3) and oxidised thioredoxins (TrxB) catalyse the reversible oxidation–reduction of protein disulphide groups. Cytoplasmic DsbC isomerises mis-oxidised proteins to their native correctly folded state. (Image created in BioRender.com).

While the cytoplasm is more reducing and not desirable for protein folding, it has many advantages over the periplasmic space, and to capitalise on these modified host cell mutants

that lack thioredoxin, reductase and glutathione reductase genes have been developed to provide an oxidative environment. Small ubiquitin-related modifier (SUMO) tagged technology has been used such that SUMO fused heavy and light chains are expressed to enhance protein expression, stability, and solubility in the cytoplasmic space [1]. Higher expression of heavy and light chains was observed in the soluble fraction in the case of SUMO-tagged molecules, compared with untagged molecules. However, the increased expression was limited to the SHuffle strain and remained unaffected in the case of the BL21 strain.

Extracellular expression offers multiple advantages, including reduced proteolysis, easy purification, and thereby reduced time and cost. The lack of secretory pathways in *E. coli* makes it less efficient. Very few pathways, such as haemolysin and pullulanase, are used for extracellular protein production. Researchers have developed a novel system in which they used the *phoA* promoter and STII sequence for extracellular Fab production [16]. A study carried out on five Fab molecules indicated that secretion efficiency increased in the case of the *phoA* promoter, compared with the T7 promoter, and the STII signal sequence showed higher secretion ability than PelB. The proteins produced were of high purity, integrity, and bioactivity.

2.2. Co-Expression of Soluble Expression Partners

Insoluble aggregate formation and coexpression of molecular chaperones with Fab fragments to enhance the solubility and proper folding of antibody fragments are being widely studied as an effective approach to overcome protein loss and reduce process time at the refolding step in the case of inclusion bodies.

The expression of periplasmic chaperones such as DsbC exhibits a synergistic effect on protein folding and disulphide bond formation. In the cytoplasm, coexpression of DnaK–DnaJ–GrpE chaperones facilitates protein folding and transport across the membrane (Figure 1). DnaK–DnaJ–GrpE chaperones coexpressed with anti-TNF- α Fab in the BL21 host system result in a marked increase in the periplasmic translocation and active soluble expression [17]. DsbC has a positive effect on the solubility of the proteins expressed in the cytoplasm but not in the proteins targeted to the periplasmic space. In another study, it has been observed that coexpression of the DsbC chaperon in periplasmic protease deficient strains such as wild-type W3110 and MXE001 strain resulted in high periplasmic Fab yield and increased levels of cell viability for four different types of Fab. A twofold yield increase in three Fab molecules (average 1.1 g/L to 2.25 g/L) and a fivefold (average 0.48 g/L to 2.6 g/L) increase in one Fab molecule was observed, as compared with wild type. Additionally, an increase in cell viability up to 40 h post-induction was observed in these modified strains (from around 80 OD to 105 OD) [18]. Since only a few pathways transfer folded or partially folded proteins to periplasmic space, researchers have examined the effect of coexpression of different sets of cytoplasmic chaperones on the yield of the Fab fragments and observed that the DnaKJE (DnaK–DnaJ–GrpE) had a positive effect on solubilising recombinant proteins expressed in the cytoplasm and had no effect on increasing the functionality of protein [6,17]. In the same study, it was reported that GroES–GroEL did not improve solubility and functionality, and it was concluded that GroES–GroEL chaperones are likely more effective in the case of cytoplasmic expression than the periplasmic expression.

In summary, the coexpression of chaperones and their effect on protein folding, solubility, and translocation may vary depending on Fab molecule, strain type, and protein localisation.

2.3. Strain Development and Protein Expression

Proteolytic degradation is the primary concern when expressing heterologous proteins in prokaryotic host systems. Recent advances in strain engineering have enabled us to improve the tolerance and applicability of host cells for the expression of heterologous proteins.

In the case of periplasmic expression of Fab fragments, researchers have developed Tsp, protease III, and DegP periplasmic protease deficient strains, and an increase in protein

expression was observed in the Tsp deficient strain. Coexpression of protease deficient strains with DsbC gave increased yield and also restored cell viability; the final yield of 2.4 g/L was achieved, compared with 1 g/L in wild-type strain [18]. This indicated the likely involvement of Tsp in the degradation of Fab in periplasmic space, which was tested in wild-type W3110 and MXE001 strains and showed consistent results in both cases.

The CyDisCo system, which is based on the coexpression of Erv1p, DsbB, or VKOR, and either DsbC or PDI, has been developed for the production of disulphide bond containing proteins in *E. coli*. Researchers have studied the efficiency of the CyDisCo system to express Fab and scFv in *E. coli*, and it has been reported that the CyDisCo system is able to generate high yields of folded, biologically active, antibody fragments in the cytoplasm of *E. coli* with more than 90% success rate. In one study, four Fabs were used, and a nearly 20-fold increased yield was observed. In addition, 42 mg/L folded purified Fab (Maa48) was obtained from a shake flask study, and overall general average Fab levels were 23 mg/L [19].

Genomic integration of the gene of interest has been attempted by performing site-directed integration of the gene of interest into the host cell genome [15]. When compared with the plasmid-based and genomic integration-based systems, the proposed approach resulted in yields ranging from 80% to 300%, whereas plasmid-based counterparts were considered at 100%. Soluble specific Fab yields ranged from 0.01 mg to 7.4 mg Fab per gram of cell dry mass. This study also proved that the genome-based system is a better option for expressing proteins translocated to the periplasm. The genomic integration lowers the concentration of mRNA, which prevents the overburdening of the translation, translocation, and folding machinery. This can also be used as an alternative to a plasmid-free, antibiotic selection pressure-free system.

As mentioned above, plasmid-based expression of recombinant protein imposes an excessive burden on the host cell machinery, and to overcome this, researchers have studied promoter engineering. In one study, a strong Ptac promoter was converted to a weak Ptic promoter by insertion of two amino acids between the −10 and −35 region to increase the gap between the two sites [20]. About a 35% decrease in expression was observed in the case of Ptic promoter, together with reduced intracellular accumulation, reduced leakage of Fab in the external environment, and increased cell viability. In this case, Fab expression was decreased, compared with wild type, from 0.32 to 0.23 mg Fab/g cell weight/h; additionally, the death rate was reduced from 18.1% to 11.6% at the end of fermentation. This proved that the engineered plasmid is significantly effective for periplasmic Fab production.

Leucine zipper-fused Fab, Zipbody, has been constructed to enhance the rate of formation of heterodimers [21]. In this study, an artificially designed pair of leucine zipper peptides, LZA and LZB, and the most widely studied leucine zipper pair, c-Jun and c-Fos, were assembled onto the C-termini of Fab, and it was observed that the rate of disulphide bond formation and association of heavy and light chain to form the heterodimer increased in case of the leucine Zipper. Different strain engineering and promoter engineering strategies have been developed to increase the soluble expression and yield of Fab fragments.

2.4. Fermentation Process and Media Development

After gene level modification and strain engineering, process development primarily focuses on media components and cultivation parameters. The former entails examining suitable combinations of carbon source, nitrogen source, and complex rich media with different combinations of yeast extract, peptone, and tryptone. Researchers have also optimised different process parameters such as pH, agitation, temperature, cell density, the concentration of inducers, induction time, and OD.

The use of a suitable carbon source and its concentration during different stages of expression is the main factor affecting cell growth and protein expression. Excessive glucose supplementation leads to higher acetate production, which affects cell growth

and product yield. The EnBase system, in which glucose is gradually released into the medium by enzymatic degradation of glucose polymer, was applied for the expression of Fab molecules, and it was observed that better protein yield and ratio of soluble to total protein were achieved in *E. coli* BL21 [22]. Higher biomass production and lower acetate accumulation resulted in improved recombinant protein expression. Similar observations have been reported by other researchers, in studies in which the use of the EnBase system resulted in a 3-times increase in biomass, compared with LB in both BL21 and TSHuffle host systems [1]. While a 19% increase in protein expression was reported in the TSHuffle strain, the difference was minor in the case of the BL21 strain. The highest achieved in this cultivation mode was 12 mg/L.

While low temperature is favoured for the expression of heterologous proteins, it is associated with a reduced growth rate and lower cell density, which often results in an increased burden of time and cost. Researchers have studied the effect of temperature and compared the synergistic effect of low temperature with nitrogen supplementation to overcome a lower growth rate [23]. In one study, it was confirmed that a lower temperature is important for soluble Fab production, whereas a higher aggregation rate was observed at a higher temperature. Using nitrogen supplementation, productivity increased by 60% (from 332 mg/L to 529 mg/L active form expression) at 25 °C, compared with 37 °C.

The stress minimisation approach has been explored by researchers, through which metabolic stress on host cells was reduced by decreasing the culture temperature or the inducer concentration [10]. This resulted in improved cell integrity, protein expression, and protein refolding.

3. Major Developments in Downstream Processing of Fab Fragments in *E. coli*

3.1. Refolding

Recombinant antibody fragments such as Fab molecules, when produced in *E. coli* in the form of misfolded aggregates called inclusion bodies (IBs), need to be refolded into a functionally active protein (Figure 2). Refolding is often a rate-limiting step and a major challenge in the production of such Fab molecules owing to the shuffling of disulphide bonds. These multidomain proteins carry intra- and interchain disulphide bonds, which makes them highly complex molecules, such that the refolding becomes an overall determining step for cost-effective and time-effective production of these biotherapeutics.

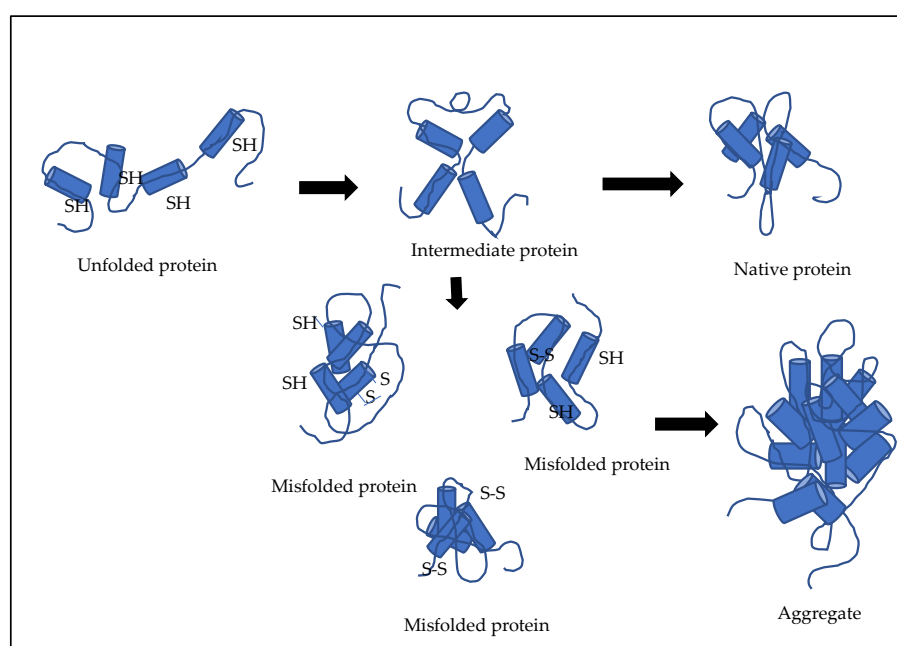


Figure 2. Illustration of refolding of multidomain proteins.

There are various events that occur during the refolding of these multidomain proteins. These include domain association, structural changes based on hydrophobicity, bond formation involving hydrogen or electrostatic interaction, and disulphide bond formation [12]. The association of light and heavy chains via disulphide bridges is highly inefficient in *E. coli* due to the large number of combinations that can occur in making and breaking the disulphide bonds [24,25]. The refolding efficiency is highly constrained due to the higher propensity of aggregation formation owing to higher-order reaction kinetics than refolding, which is a first-order reaction [26]. Thus, a deeper understanding of molecular behaviour is required to enhance the refolding of these multidomain proteins. Researchers have studied the in vitro unfolding of Fab molecules by applying two- and three-state thermodynamic models, followed by DoE-based optimisation [12]. The refolding kinetics was examined, and it was surmised that refolding follows a three-state mechanism, wherein the formation of intermediates by light and heavy chains was the rate-limiting step. For renaturing of proteins, redox agents are often employed, wherein the reduced and oxidised forms are added in an optimum ratio to favour the refolding process. These forms include the glutathione system (GSH/GSSG), cysteine/cysteine system, DTT/GSSG system in the molar ratio range of 1:1 to 10:1 [27]. To enhance refolding yield, various additives are added to the refolding buffer. These additives can act during the different conformations that the protein acquires during refolding, such as stabilising the native state, increasing the reaction kinetics towards correct folding, or suppressing aggregation during the formation of intermediates. Some of the commonly used additives are sugars, polyethylene glycol (PEG), urea, acetone, dimethyl sulfoxide (DMSO), and amino acids such as arginine [28,29].

In hosts other than *E. coli*, refolding has also been a limitation in yeast, *Pichia pastoris* [30], and insect cell-based expression systems [31,32]. Even though mammalian-based expression systems can produce a refolded biotherapeutic, the cost of production increases considerably. Therefore, there is a need to address the challenges posed by the refolding unit operation for the efficient production of these recombinant biotherapeutic proteins.

A variety of approaches have been suggested by researchers for the refolding of Fab biotherapeutics. These approaches involve techniques such as drip dilution, dialysis, and on-column refolding (Figure 3).

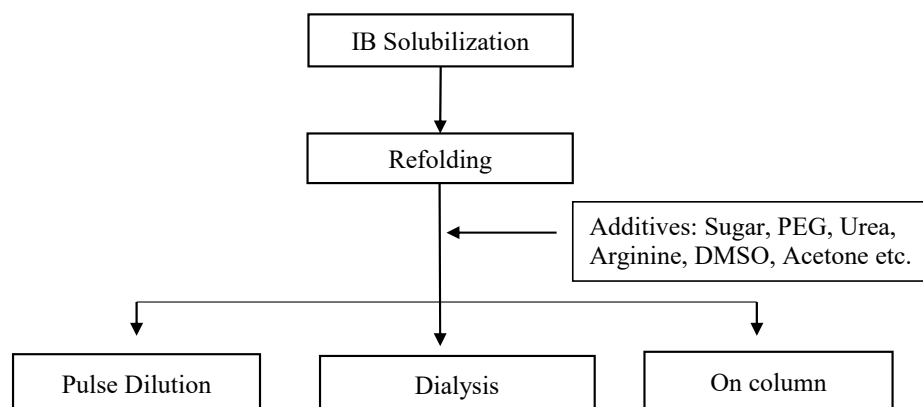


Figure 3. Different methods of refolding and various additives utilised in refolding.

To enhance the refolding efficiency, the solubilisation and unfolding behaviour of Fab molecules need to be studied. A group of scientists determined the unfolding events using nano-differential scattering fluorimetry (nano-DSF) [12]. The unfolding of proteins usually follows two-state models, which have native and unfolded forms. In three-state models, there is also an intermediate form along with native and unfolded forms. Understanding these forms can help improve refolding efficiency. In addition, other parameters also determine refolding such as the concentration of chaotrophs, pH of refolding buffer, the temperature during refolding, and the ratio of redox reagents, while the presence of refold-

ing enhancers such as arginine hydrochloride (ArgHCl) determines the overall refolding efficiency. Dilution based in vitro refolding of Fab molecule has been optimised by using the DoE approach, and a refolding yield of 56% has been achieved in 120 h by appropriate dilution in the presence of redox reagents. A recently filed patent covers the process for the production of Fab molecule ranibizumab, in which the refolding of protein was performed by diluting solubilised IBs into refolding buffer, and pH and temperature shift was performed at suitable time intervals to attain the refolded product, with a refolding yield of 13.6% in 22 h [33].

A matrix-associated method of refolding has also been explored; in this approach, the solubilised IBs are subjected to a refolding-based affinity column, also called the matrix-associated refolding method. Here, the elimination of chaotropic agents occurs by giving a linear gradient of refolding buffers. The refolded antibody fragments are then eluted and analysed by SDS–PAGE. The researchers compared the matrix-associated refolding strategy with the drip dilution method and found that the former method generated structurally folded and active fragments in terms of amount, as compared with the drip dilution method when performed in absence of redox pair [34]. Further, matrix-associated refolding has been explored with different molecules. For example, a group of researchers performed refolding of a recombinant human growth hormone fused with glutathione S transferase. This solid-phase refolding utilises the expanded-bed adsorption chromatography, wherein the unwanted particulates and denaturants are removed, and the target unfolded protein binds to the column. The results demonstrated an enhanced refolding yield of around 80%, which is significantly higher than that obtained with the conventional dilution method, and reduced aggregation was also observed, which usually has a high propensity in solution-phase refolding [35].

In another study, refolding of Fab molecules by the dialysis method was performed along with the screening of various additives to enhance the refolding yield, and taurine was found to be the most effective additive to attain 60 mg of refolded human Fab from a 1 L culture harvest [36]. Other researchers isolated the IBs of light and heavy chains, and the solubilised IBs were refolded by following a dialysis approach, reducing the concentration of chaotropic agent sequentially in three stages [37].

Many research groups have explored the refolding of Fab molecules to enhance the refolding yield, but understanding the molecular behaviour during refolding needs to be explored further to enhance the refolding efficiency. The chemistry behind the intermediates that form during refolding needs to be studied to move the reaction in the forward direction towards the folded protein.

3.2. Purification of Fab Molecules

The refolded Fab solution contains a large amount of product-related impurities (residuals of light and heavy chains, intermediates, and aggregates) and process-related impurities (HCPs and HcDNA) that must be removed in order to make a purified drug substance. An extensive study is required for the purification of Fab fragments, as the conventional protein A chromatography performed for mAbs is not applicable due to the lack of an Fc region in Fab molecules (Figure 4).

Affinity resins such as protein L and protein G are available for the purification of Fab molecules [38–42]. Traditional purification schemes involve the use of ion exchange and multimodal chromatography after affinity chromatography [40,41,43,44]. Various chromatography techniques have been employed by researchers to purify the Fab fragments. The purification of Fab molecules has been demonstrated using two multimodal chromatography resins, and the resulting product had a product purity of 99.50% and an overall process yield of 32.55% [45]. In one study, researchers performed IgG digestion and isolated Fab fragments using protein A and protein L affinity resins along with ultrafiltration/diafiltration to remove small molecular weight impurities. The overall yield for Fab isolation was approximately 50%. This was followed by resin screening of cation exchange and multimodal resins in the microfluidic device. Their findings suggest that

multimodal resins are most suitable for the purification of Fab molecules due to their high salt tolerance and efficient binding of Fab molecules. Furthermore, the microfluidic strategy allows for rapid initial screening of various parameters for designing a purification process. It helps in determining the salt concentration, pH, and adequate resin in a rapid and cheaper way [40]. Another approach that has been attempted involves the use of cation exchange chromatography (SP Toyopearl resin) to purify the refolded Fab molecule [36]. An affinity-based apo-B-100-coupled antigen column was successfully used to purify a recombinant Fab molecule specific for apolipoprotein B-100, and a refolded molecule with the production of ~3 mg of purified protein from 1 L *E. coli* harvest was generated specific to apolipoprotein B-100, which offered high selectivity for Fab molecule [37].

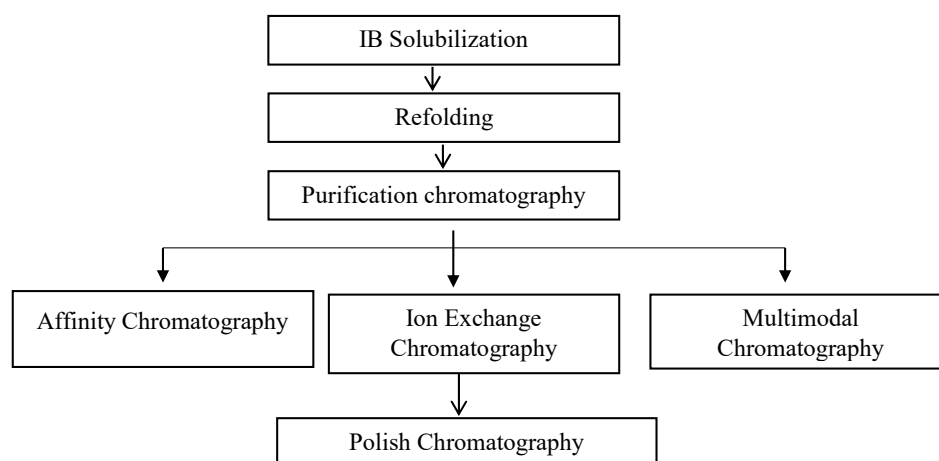


Figure 4. Conventional process flow chart for the purification of Fab molecules.

Recently, a study was performed to extract and characterise the Fab molecule and the associated impurities. A three-dimensional chromatography method was designed, involving two steps of affinity chromatography (protein G and protein L), followed by cation exchange chromatography and subsequent analysis by mass spectrometry [11]. This 3D approach offers rapid optimisation of the preparative purification process. In this study, Fab molecules were purified from crude *E. coli* samples that contained Fab along with other fragments. Automation was carried out such that the elute from affinity chromatography was loaded directly onto high-resolution cation exchange resins via loops, to form a three-dimensional chromatography system generating a purified sample. Thus, an automated process was designed, wherein the separation and characterisation of target Fab molecule, as well as the impurities associated with them, were determined, which was mainly found to be a fragment of light chain.

Another group identified the purification challenge for Fab molecules associated with affinity purification. Protein L resins can only bind some of the subclasses of variable regions of the light-kappa chain ($V_{L\kappa}$ 1, 3, 4) but no binding to $V_{L\kappa}$ 2. The authors presented a camelid-based affinity resin that can bind the constant region of the lambda light chain. This ensured 100% selectivity for the lambda light chain, thus bridging the gap between the purification of various Fab molecules. This affinity chromatography medium, lambda FabSelect, has been developed with its effectiveness equivalent to that of protein A for mAbs [46].

Thus, various groups have explored different types of strategies for refolding these multidomain antibody fragments and purifying them. These antibody fragments produced in *E. coli* still need to be explored for their refolding pathways, which can help in generating high amounts of refolded protein and thus increases yield and productivity.

4. Remaining Challenges in Production of Fab Fragments in *E. coli*

E. coli expression system is preferred for the production of Fab fragments due to the many advantages, as discussed in previous sections. Industrial production of biopharmaceuticals offers several challenges that need to be solved. Production of Fab fragments or high molecular weight proteins imposes a physiological burden and toxicity on the host cell, which can adversely affect cellular physiology. Reduced metabolic burden, stable host, and process predictability are critical for achieving higher yield and stable expression. However, in order to achieve these improvements, changes must be made at both the genetic and process levels.

The generation of genetically stable clones is critical for constant product formation. For the production of larger protein molecules such as Fab, which trigger metabolic burden on the cell, the selection of an appropriate vector or promoter with moderate strength that remains stable is key to success. Promoter engineering can also be used to tailor the promoter to the required strength. Continuous culturing of *E. coli* requires a stable and consistent clone, either plasmid-based or genome-integrated. Long-term stability, no plasmid loss, and the ability to create an antibiotic-free system are all advantages of genome-integrated systems. However, there are significant drawbacks, such as lower product titre, which must be compensated for by site-specific integration, modifying/optimising fermentation conditions, or continuous production. The design-of-experiments approach is increasingly being used to explore multiparameter experimental space and understand the interactions between different parameters and their effect on yield and other responses.

A comprehensive approach that takes into account both the upstream and downstream processes in an integrated manner is required. The physical and biological state of cells, process parameters, feeding strategy, process robustness, protein localisation, inefficient translocation and incomplete signal sequence processing, and many other parameters in the upstream process can all affect the downstream steps. Small changes in the upstream process can have a significant impact on the downstream operations and performance. As a result, in order to develop a productive and economic process, both upstream and downstream processes need to be investigated as an integrated system.

5. Summary

In this review, we presented a comprehensive list of possible solutions for the production of Fab molecules in microbial systems. It was noted that different growth media and feeding strategies, as well as the synergistic effect of different parameter combinations and screening strategies, can significantly improve Fab molecule expression. Coexpression of multiple chaperone-encoding genes and careful selection of modifying secretion pathways can result in increased yields of correctly folded proteins. A deeper understanding of the stress-related degradation pathway may be effective for preventing protein misfolding and mislocalisation. Various methods are being developed to identify high-yielding scalable refolding conditions for multidomain proteins with interdomain disulphide bonds. Varied purification approaches have also been developed, including an automated purification method for a rapid and knowledge-based purification process, as well as its improvement to identify and separate product-related impurities. While some strategies for troubleshooting recombinant protein expression are protein-specific, there remains space for interpretation. The development of a suitable host with a stable expression platform, as well as the development of a general systematic, media, and fermentation process, are critical for robust process development and stable production. Recent advances in strain engineering, proteomics, and metabolomics can provide a better understanding of cellular pathways. Knowledge of gene expression, secretion pathways, and protein folding can be applied to improve the expression of functional Fab molecules. A newly modified strain that can perform oxidative protein folding in the cytoplasm, as well as interesting findings from cell-free experiments, offer new avenues for future developments. A holistic approach that takes into account both the upstream and downstream processes in an integrated manner is required.

Author Contributions: Data collection and literature analysis, writing—original draft preparation, R.S.P. and A.A.; writing—review and editing, J.A.G.; supervision, writing—review and editing, A.S.R. All authors have read and agreed to the published version of the manuscript.

Funding: Authors would like to acknowledge funding support from the Department of Biotechnology, Ministry of Science and Technology (File number BT/COE/34/SP15097/2015) and Biotechnology Industry Research Assistance Council (Reference number BT/NBM0160/04/19).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Author Rucha S. Patil would like to thank the Department of Biotechnology, Ministry of Science and Technology, New Delhi, for providing the DBT-JRF Fellowship grant during her PhD (Fellow Number DBT/2017/IIT-D/964).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rezaie, F.; Davami, F.; Mansouri, K.; Amiri, S.A.; Fazel, R.; Mahdian, R.; Davoudi, N.; Enayati, S.; Azizi, M.; Khalaj, V. Cytosolic expression of functional Fab fragments in *Escherichia coli* using a novel combination of dual SUMO expression cassette and EnBase® cultivation mode. *J. Appl. Microbiol.* **2017**, *123*, 134–144. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Bates, A.; Power, C.A. David vs. Goliath: The Structure, Function, and Clinical Prospects of Antibody Fragments. *Antibodies* **2019**, *8*, 28. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Sandomenico, A.; Sivaccumar, J.P.; Ruvo, M. Evolution of *Escherichia coli* Expression System in Producing Antibody Recombinant Fragments. *Int. J. Mol. Sci.* **2020**, *21*, 6324. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Lee, Y.J.; Jeong, K.J. Challenges to production of antibodies in bacteria and yeast. *J. Biosci. Bioeng.* **2015**, *120*, 483–490. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Fakruddin, M.; Mazumdar, R.M.; Mannan, K.S.; Chowdhury, A.; Hossain, M.N. Critical Factors Affecting the Success of Cloning, Expression, and Mass Production of Enzymes by Recombinant *E. coli*. *Int. Sch. Res. Not.* **2013**, *2013*, 590587. [\[CrossRef\]](#)
6. Levy, R.; Weiss, R.; Chen, G.; Iverson, B.L.; Georgiou, G. Production of correctly folded Fab antibody fragment in the cytoplasm of *Escherichia coli* trxB gor mutants via the coexpression of molecular chaperones. *Protein Expr. Purif.* **2001**, *23*, 338–347. [\[CrossRef\]](#)
7. Baeshen, M.N.; Al-Hejin, A.M.; Bora, R.S.; Ahmed, M.M.M.; Ramadan, H.A.I.; Saini, K.S.; Baeshen, N.A.; Redwan, E.M. Production of Biopharmaceuticals in *E. coli*: Current Scenario and Future Perspectives. *J. Microbiol. Biotechnol.* **2015**, *25*, 953–962. [\[CrossRef\]](#)
8. Farajnia, S.; Ghorbanzadeh, V.; Dariushnejad, H. Effect of Molecular Chaperone on the Soluble Expression of Recombinant Fab Fragment in *E. coli*. *Int. J. Pept. Res. Ther.* **2020**, *26*, 251–258. [\[CrossRef\]](#)
9. Rosano, G.L.; Ceccarelli, E.A. Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Front. Microbiol.* **2014**, *5*, 172. [\[CrossRef\]](#)
10. Hsu, C.C.; Thomas, O.R.; Overton, T.W. Periplasmic expression in and release of Fab fragments from *Escherichia coli* using stress minimization. *J. Chem. Technol. Biotechnol.* **2016**, *91*, 815–822. [\[CrossRef\]](#)
11. Schimek, C.; Kubek, M.; Scheich, D.; Fink, M.; Brocard, C.; Striedner, G.; Cserjan-Puschmann, M.; Hahn, R. Three-dimensional chromatography for purification and characterization of antibody fragments and related impurities from *Escherichia coli* crude extracts. *J. Chromatogr. A* **2021**, *1638*, 461702. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Gani, K.; Bhambure, R.; Deulgaonkar, P.; Mehta, D.; Kamble, M. Understanding unfolding and refolding of the antibody fragment (Fab). I. In-vitro study. *Biochem. Eng. J.* **2020**, *164*, 107764. [\[CrossRef\]](#)
13. Kulmala, A.; Huovinen, T.; Lamminmäki, U. Improvement of Fab expression by screening combinatorial synonymous signal sequence libraries. *Microb. Cell Factories* **2019**, *18*, 157. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Schofield, D.M.; Sirka, E.; Keshavarz-Moore, E.; Ward, J.; Nesbeth, D.N. Improving Fab' fragment retention in an autonucleolytic *Escherichia coli* strain by swapping periplasmic nuclease translocation signal from OmpA to DsbA. *Biotechnol. Lett.* **2017**, *39*, 1865–1873. [\[CrossRef\]](#)
15. Fink, M.; Vazulka, S.; Egger, E.; Jarmer, J.; Grabherr, R.; Cserjan-Puschmann, M.; Striedner, G. Microbioreactor Cultivations of Fab-Producing *Escherichia coli* Reveal Genome-Integrated Systems as Suitable for Prospective Studies on Direct Fab Expression Effects. *Biotechnol. J.* **2019**, *14*, e1800637. [\[CrossRef\]](#)
16. Luo, M.; Zhao, M.; Cagliero, C.; Jiang, H.; Xie, Y.; Zhu, J.; Yang, H.; Zhang, M.; Zheng, Y.; Yuan, Y.; et al. A general platform for efficient extracellular expression and purification of Fab from *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 3341–3353. [\[CrossRef\]](#)
17. Dariushnejad, H.; Farajnia, S.; Zarghami, N.; Aria, M.; Tanomand, A. Effect of DnaK/DnaJ/GrpE and DsbC Chaperons on Periplasmic Expression of Fab Antibody by *E. coli* SEC Pathway. *Int. J. Pept. Res. Ther.* **2019**, *25*, 67–74. [\[CrossRef\]](#)

18. Ellis, M.; Patel, P.; Edon, M.; Ramage, W.; Dickinson, R.; Humphreys, D.P. Development of a high yielding *E. coli* periplasmic expression system for the production of humanized Fab' fragments. *Biotechnol. Prog.* **2017**, *33*, 212–220. [\[CrossRef\]](#)
19. Gaciarz, A.; Veijola, J.; Uchida, Y.; Saaranen, M.J.; Wang, C.; Hörkö, S.; Ruddock, L.W. Systematic screening of soluble expression of antibody fragments in the cytoplasm of *E. coli*. *Microb. Cell Factories* **2016**, *15*, 22. [\[CrossRef\]](#)
20. Schofield, D.M.; Templar, A.; Newton, J.; Nesbeth, D.N. Promoter engineering to optimize recombinant periplasmic Fab' fragment production in *Escherichia coli*. *Biotechnol. Prog.* **2016**, *32*, 840–847. [\[CrossRef\]](#)
21. Ojima-Kato, T.; Fukui, K.; Yamamoto, H.; Hashimura, D.; Miyake, S.; Hirakawa, Y.; Yamasaki, T.; Kojima, T.; Nakano, H. 'Zipbody' leucine zipper-fused Fab in *E. coli* in vitro and in vivo expression systems. *Protein Eng. Des. Sel.* **2016**, *29*, 149–157. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Azarian, B.; Azimi, A.; Sepehri, M.; Fam, V.S.; Rezaie, F.; Talebkhan, Y.; Khalaj, V.; Davami, F. Proteomics investigation of molecular mechanisms affected by EnBase culture system in anti-VEGF fab fragment producing *E. coli* BL21 (DE3). *Prep. Biochem. Biotechnol.* **2019**, *49*, 48–57. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Kim, S.J.; Ha, G.S.; Lee, G.; Lim, S.I.; Lee, C.M.; Yang, Y.H.; Lee, J.; Kim, J.E.; Lee, J.H.; Shin, Y.; et al. Enhanced expression of soluble antibody fragments by low-temperature and overdosing with a nitrogen source. *Enzym. Microb. Technol.* **2018**, *115*, 9–15. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Corisdeo, S.; Wang, B. Functional expression and display of an antibody Fab fragment in *Escherichia coli*: Study of vector designs and culture conditions. *Protein Expr. Purif.* **2004**, *34*, 270–279. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Venturi, M.; Seifert, C.; Hunte, C. High level production of functional antibody fab fragments in an oxidizing bacterial cytoplasm. *J. Mol. Biol.* **2002**, *315*, 1–8. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Kiefhaber, T.; Rudolph, R.; Kohler, H.-H.; Buchner, J. Protein Aggregation in vitro and in vivo: A Quantitative Model of the Kinetic Competition between Folding and Aggregation. *Nat. Biotechnol.* **1991**, *9*, 825–829. [\[CrossRef\]](#)
27. Chakravarthi, S.; Jessop, C.E.; Bulleid, N.J. The role of glutathione in disulphide bond formation and endoplasmic-reticulum-generated oxidative stress. *EMBO Rep.* **2006**, *7*, 271–275. [\[CrossRef\]](#)
28. Singh, S.M.; Panda, A.K. Solubilization and refolding of bacterial inclusion body proteins. *J. Biosci. Bioeng.* **2005**, *99*, 303–310. [\[CrossRef\]](#)
29. Yasuda, M.; Murakami, Y.; Sowa, A.; Ogino, H.; Ishikawa, H. Effect of Additives on Refolding of a Denatured Protein. *Biotechnol. Prog.* **1998**, *14*, 601–606. [\[CrossRef\]](#)
30. Lange, S.; Schmitt, J.; Schmid, R. High-yield expression of the recombinant, atrazine-specific Fab fragment K411B by the methylotrophic yeast *Pichia pastoris*. *J. Immunol. Methods* **2001**, *255*, 103–114. [\[CrossRef\]](#)
31. Furuta, T.; Ogawa, T.; Katsuda, T.; Fujii, I.; Yamaji, H. Efficient production of an antibody Fab fragment using the baculovirus–insect cell system. *J. Biosci. Bioeng.* **2010**, *110*, 577–581. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Mori, K.; Hamada, H.; Ogawa, T.; Ohmuro-Matsuyama, Y.; Katsuda, T.; Yamaji, H. Efficient production of antibody Fab fragment by transient gene expression in insect cells. *J. Biosci. Bioeng.* **2017**, *124*, 221–226. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Somani, S.; Pandey, A.; Mishra, A.; Mody, R.S. Refolding Process for Antibody's Fragments. US20180230206A1, 22 September 2020.
34. Ban, B.; Sharma, M.; Shetty, J. Optimization of Methods for the Production and Refolding of Biologically Active Disulfide Bond-Rich Antibody Fragments in Microbial Hosts. *Antibodies* **2020**, *9*, 39. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Cho, T.H.; Ahn, S.J.; Lee, E.K. Refolding of protein inclusion bodies directly from *E. coli* homogenate using expanded bed adsorption chromatography. *Bioseparation* **2001**, *10*, 189–196. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Fujii, T.; Ohkuri, T.; Onodera, R.; Ueda, T. Stable Supply of Large Amounts of Human Fab from the Inclusion Bodies in *E. coli*. *J. Biochem.* **2007**, *141*, 699–707. [\[CrossRef\]](#)
37. Lee, M.-H.; Kwak, J.-W. Expression and functional reconstitution of a recombinant antibody (Fab') specific for human apolipoprotein B-100. *J. Biotechnol.* **2003**, *101*, 189–198. [\[CrossRef\]](#)
38. Humphreys, D.P.; Carrington, B.; Bowering, L.C.; Ganesh, R.; Sehdev, M.; Smith, B.J.; King, L.M.; Reeks, D.G.; Lawson, A.; Popplewell, A.G. A plasmid system for optimization of Fab' production in *Escherichia coli*: Importance of balance of heavy chain and light chain synthesis. *Protein Expr. Purif.* **2002**, *26*, 309–320. [\[CrossRef\]](#)
39. Jalalirad, R. Efficient Chromatographic Processes for Elevated Purification of Antibody Fragment (Fab D1.3) from Crude *Escherichia coli* Culture. *Biotechnology* **2013**, *12*, 74. [\[CrossRef\]](#)
40. Nascimento, A.; Pinto, I.F.; Chu, V.; Aires-Barros, M.R.; Conde, J.P.; Azevedo, A.M. Studies on the purification of antibody fragments. *Sep. Purif. Technol.* **2018**, *195*, 388–397. [\[CrossRef\]](#)
41. Rodrigo, G.; Gruvegård, M.; Van Alstine, J.M. Antibody Fragments and Their Purification by Protein L Affinity Chromatography. *Antibodies* **2015**, *4*, 259–277. [\[CrossRef\]](#)
42. Roque, A.C.A.; Lowe, C.R. Rationally Designed Ligands for Use in Affinity Chromatography. In *Affinity Chromatography*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 93–110. [\[CrossRef\]](#)
43. Balasundaram, B.; Chhatre, S.; Arora, D.; Titchener-Hooker, N.J.; Bracewell, D.G. Study of the conditions for multi-modal chromatographic capture of Fab' from dual-salt precipitated *E. coli* homogenate. *J. Chem. Technol. Biotechnol.* **2013**, *88*, 372–377. [\[CrossRef\]](#)
44. Balasundaram, B.; Sachdeva, S.; Bracewell, D.G. Dual salt precipitation for the recovery of a recombinant protein from *Escherichia coli*. *Biotechnol. Prog.* **2011**, *27*, 1306–1314. [\[CrossRef\]](#) [\[PubMed\]](#)

-
45. Bhambure, R.S.; Gani, K.M. A Process for the Purification of Recombinant Antibody Fragments. US20200095277A1, 26 March 2020.
 46. Eifler, N.; Medaglia, G.; Anderka, O.; Laurin, L.; Hermans, P. Development of a novel affinity chromatography resin for platform purification of lambda fabs. *Biotechnol. Prog.* **2014**, *30*, 1311–1318. [[CrossRef](#)] [[PubMed](#)]