

Check for updates

1 Progress, challenges and prospects of protein purification technology

Yan Xuan^{1†}, Zhuru Hou^{2†}, Ling Liu^{1,3†*}, Yan Xuan¹,
Yong Chen⁴, Lei Fan⁴, Zhuoxi Li⁴ and Benjin Xu^{1,3*}

1 School of Biotechnology, Beijing University of Aeronautics and Astronautics, Beijing 100191, China; *2* School of Chemical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China; *3* Beijing Key Laboratory of Protein Purification and Purification Technology, Beijing 100191, China; *4* Beijing Key Laboratory of Protein Purification and Purification Technology, Beijing 100191, China

Protein purification is a key step in the production of recombinant proteins. The development of protein purification technology has been rapid in recent years. This review summarizes the progress, challenges and prospects of protein purification technology. The current status of protein purification technology is reviewed, including the development of various purification methods, such as affinity chromatography, ion exchange chromatography, size exclusion chromatography, and hydrophobic interaction chromatography. The challenges of protein purification technology are discussed, including the low yield and high cost of protein purification. The prospects of protein purification technology are discussed, including the development of novel purification methods, such as membrane chromatography and simulated moving bed chromatography.

Keywords: protein purification, chromatography, membrane chromatography, simulated moving bed chromatography, recombinant proteins

1 Introduction

Protein plays a key role in cell life activities and is the most abundant macromolecule in living organisms. It participates in a series of biological events, such as maintaining the structures and properties of organisms, catalyzing certain chemical reactions, transporting nutrients and metabolic wastes, providing a material basis for the body's

† Yan Xuan, Zhuru Hou, and Ling Liu contributed equally to this work. *Correspondence: lingliu@buaa.edu.cn (L.L.), xuanyan@buaa.edu.cn (Y.X.)

Expression system	Advantages	Disadvantages	Applications
<i>E. coli</i>	<i>E. coli</i> is small and grows rapidly. It is comparatively easier genetic manipulation and fermentation process (Tripathi and Shrivastava, 2019), well genetically characterized (McElwain et al., 2022), simple cultivation conditions (Rosano and Ceccarelli, 2014), and a clear genetic background (Nagai et al., 2018). It has the potential of highly expressed proteins (Kesidis et al., 2020)	Target proteins are easy to form inclusion body and codon bias, which lacks the post-translational modification system (Liu et al., 2019) and endotoxin issues (Mamat et al., 2015), interfering with the expression of target proteins	The most commonly used prokaryotic expression system, widely used in the development of the vaccine (Shirley and Taha, 2018), the production of the hormone (Zielinski et al., 2019; Keshavarz et al., 2021), the synthesis of the important enzyme (Lu et al., 2018). The expressed protein is used for structural insights (Muller et al., 2022)
<i>Saccharomyces cerevisiae</i>	It has high expression level, rapid growth, easy culture, high tolerance to the environment (Tesfaw and Assefa, 2014), low culture costs (Duman-Scheel, 2019), and has a post-translational modification system (Baghban et al., 2019)	Recombinant proteins may accumulate excessively in cells, resulting in reducing the subsequent yield (Tomás-Gamisans, 2020). Intracellular post-translational modifications can lead to the production of hyperglycosylated proteins, which are immunogenic to the human body (Rasala and Mayfield, 2015; Xu et al., 2016). The replication plasmid YRp of <i>Saccharomyces cerevisiae</i> is unstable (Xie et al., 2018)	It can be used as an expression system for the inhibition of human dipeptidyl peptidase IV/CD26 (hDPPIV) (Zhang et al., 2016). It can produce hormones (Chigira et al., 2008), functional foods (Lahue et al., 2020), and biofuels (Boonchuay et al., 2021; Hong and Nielsen, 2012), and itself can also be used as probiotics
<i>Pichia pastoris</i>	As methanol is used as the sole carbon source and energy source, there are a methanol-regulated alcohol oxidase promoter (PAOX1) and a post-translational modification system, which has an efficient secretion mechanism (Juturu and Wu, 2018). It is highly similar to the mammalian cell expression system, but has lower culture costs and rapid expression than the mammalian cell expression system, more advantageous than <i>S. cerevisiae</i> (Karbalaie et al., 2020), and less endogenous secreted protein production to facilitate subsequent isolation and purification (Tachioka et al., 2016)	Glucose, glycerol, and ethanol severely inhibited promoter PAOX1 activity and affected exogenous protein expression (Turkanoglu Ozcelik et al., 2019). Methanol is flammable and toxic, which is detrimental to cell growth (Santoso et al., 2012)	It can express membrane proteins very well (Bill and Hedfalk, 2021). The production of anticancer drug (Liu et al., 2015), viral surface antigens (Gupta et al., 2020), antiviral proteins, important enzymes as well as viral subunit recombinant protein vaccines (Ardizzone et al., 2011; J342.6738-11.

immune defense, participating in intracellular redox reaction, electron transfer, learning and memory.

The first person engaged in the purification of proteins was Edwin Joseph Cohn, an American biochemist, who purified serum protein (Labrou, 2021). In recent years, with the completion of genome sequencing of many species (Shendure and Lieberman Aiden, 2012; Beigh, 2016; Park and Kim, 2016; Ing-Simmons and Vaquerizas, 2019), researches in protein related fields have gradually become the focus of the biotechnology industry (Labrou, 2021), and many purification technologies have been gradually applied from protein and enzyme separation to drug synthesis, vaccine research and development, clinical detection, environmental analysis and biophysical measurement (Pfaunmiller et al., 2013). Therefore, high-quality protein samples are of great significance for the subsequent structure and function research and related products development (Tang et al., 2021). However, proteins have the characteristics of complex structure, easy to be affected by internal and external factors, easy to be degraded, etc. It is difficult and challenging to purify them without being contaminated by the host, maintaining structural integrity and biological activity.

As a downstream technology of the biological industry, the key to protein purification is to define the purification strategy and finally establish an optimal scheme through continuous optimization of purification conditions, that is, using the least steps to achieve the purity, concentration, the activity and the yield we need, which is crucial for subsequent researches (Lojewska et al., 2016; Owczarek et al., 2019; Schiermeyer, 2020). The purification strategy is mainly based on the unique physical and chemical properties and the three-dimensional structures of the target proteins, including the sequence and number of amino acids, the charge, polarity, and hydrophilicity/hydrophobicity of the polypeptide chains, the shape of the proteins, and the

susceptibility to oxidation, and potential post-translational modification sites, etc. Should we choose prokaryotic expression, eukaryotic cell expression, or insect expression system to express the target proteins? Can the correct post-translational modification of the target protein be ensured after expression? Which tag should be selected to facilitate overexpression and purification when constructing the vector? Which method should be used for cell disruption after expression, and whether the disruption process will

possible to achieve the best effect. In addition, the size of the chromatographic column, the concentration of the protein obtained, whether it is necessary to maintain its activity and avoid unnecessary contaminants should also be considered. The additives should be used as little as possible, the impurities that damage the sample should be removed as early as possible, and the additional purification steps should be reduced.

2.3 Identify detection and analysis techniques

For protein solutions containing different components, the detection methods for target proteins are usually different. Different detection methods should be selected in different purification steps according to their sensitivity, precision, accuracy, etc., so as to evaluate the activity, purity and recovery of samples quickly and accurately.

2.4 Determine protein storage conditions

In order to improve the stability of the target protein, prevent microbial growth and maintain the protein activity, various surfactants, metal chelators, protease inhibitors, preservatives and reducing agents (sh-mercaptoethanol or DTT) can usually be added to the buffer solution. In addition, under the premise of ensuring the activity of the target protein, it should be stored at -70°C , and repeated freezing and thawing should be avoided.

cause the denaturation and structural change of the target protein? The purity, concentration, activity and yield of the final product should be determined to avoid over purification or failure to achieve the desired purity due to insufficient purification steps or resolution. All the above issues need to be considered comprehensively.

2.2 Reasonable selection of purification method

Select suitable purification methods according to the characteristics of the target protein and impurities, for example, choose to salt out, organic solvent precipitation, or choose centrifugation, electrophoresis, chromatography and other methods (Table 2). The purification scale should be determined according to the purpose of the target proteins, and the simple purification scheme should be used as far as

3 Cell culture

Recombinant protein synthesis is the process of producing proteins in prokaryotic microorganisms or eukaryotic cells using recombinant DNA technology, and it is an important branch of the emerging synthetic biology. It is becaun. emdentlopmt

gene was determined and suitable vectors (Table 3) and competent cells (Table 4) were selected, then the target gene was linked to an empty vector to construct a recombinant expression vector, and then the recombinant vector was transformed into host cells for culture. When OD_{600} reaches 0.4–0.6, an appropriate proportion of inducer was added to induce protein expression at a certain temperature, and finally the thalli were collected for purification (Figure 2). As an expression host, *E. coli* has become one of the best hosts for the production of recombinant proteins due to its rapid reproduction, low costs, and rapid expression of a large amount of target proteins (Rosano and Ceccarelli, 2014). In order to avoid the influence of lack of post-translational modification systems, inclusion body formation, frameshift mutations and endotoxin production on subsequent purification, several means such as changing host, vector and target gene sequence constantly, optimizing culture conditions, and co-expressing target proteins with molecular chaperones are adopted by researchers to increase protein solubility and expression (Gopal and Kumar, 2013). For the problems encountered in the expression process, a series of effective solutions have been developed. For example, T7RNAP expression activity (Li Z. J. et al., 2022; Chee et al., 2022) or growth decoupling system (Li et al., 2016; Darlington et al., 2018) can be specifically regulated when expressing toxic proteins, but this method is not universally applicable. Zhang et al. (2022) developed a dynamic equilibrium system that can achieve the overexpression of basic growth-related genes (rRNA, RNAP core enzyme, sigma factor), accurately predict and express key proteins using an *ec_iECBD_1354* enzyme constraint model, and dynamically regulate the expression intensity of key growth-related proteins based on a load-driven promoter. This system alleviates the host burden effect, improves the production of recombinant proteins, and is helpful to efficiently develop expression hosts based on the properties of target proteins. However, the *E. coli* expression system lacks post-

translational modifications, which is not conducive to the expression of eukaryotic proteins and many enzymes (Gupta et al., 2019), and increases the subsequent purification processes. Currently, the modified *E. coli* expression system can achieve the simple glycosylation modification of proteins (Gupta and Shukla, 2016), but it still needs to be further optimized. Some rare codons exist in protein-coding genes of higher animals, but these rare codons are not common in *E. coli*. Therefore, the expression of these proteins by *E. coli* expression system may lead to the reduction of target protein expression or premature termination of protein translation (Owczarek et al., 2019). In view of this, several online analysis softwares have been developed to detect the presence of rare codons in the gene sequence, and the expression of the target proteins can be improved through codon optimization (Gupta et al., 2019; Rosano et al., 2019). In addition, introducing short-term heat shock before induction (Oganesyan et al., 2007), adding D-sorbitol, glycerol, ethanol, NaCl (Diamant et al., 2001; Bowden et al., 1991; Oganesyan et al., 2007) and other chemical additives and/or target protein cofactors (Bt0.9(.6725ence,)-406.6(and)-401.7(t)-0.1e6.9(7s0TJ0or)78214-222E(al.,33

import the target proteins into the periplasmic region (Dow et al., 2015; Vargas-Cortez et al., 2017) are usually adopted.

4 Pretreatment

Before purification, the sample should be pretreated to release the total proteins from the tissue cells while maintaining the activity of the target proteins. Since different tissues and cells have different structures, different materials need to be pretreated by different methods. Commonly used pretreatment methods include mechanical method and non-mechanical method, the former includes high-speed bead milling, high-pressure homogenization and ultrasonic crushing, etc., while the latter includes osmotic pressure impact, freeze-thaw crushing, enzymatic hydrolysis and chemical crushing, etc. For example, connective tissue and adipose tissue should be removed by grinding, ultrasonic crushing (Tang et al., 2015) and centrifugation before the purification of animal materials. The seed materials should be shelled or even peeled to avoid the pollution of tannins and other substances, especially the oil seeds should be degreased with low boiling organic solvents such as ether. Plant tissues contain a wide variety of compounds, such as phenols, lipids, pigments, organic acids, carbohydrates, etc., which greatly interfere with protein extraction and proteomic analysis (Wang et al., 2008). In the purification of natural plant proteins, appropriate parts should be selected first for sampling, and then the cell wall and other components should be broken by mechanical or non-mechanical methods (Kielkopf et al., 2021) to obtain more cell lysates (Hunt, 2005), so as to ensure the smooth progress of the subsequent purification process. The purification of recombinant proteins produced by genetic engineering technology also requires pretreatment. The *E. coli* expression hosts are often collected by centrifugation at low temperature and broken by ultrasound. In addition, due to the mild and highly efficient action conditions of the enzyme, selecting a suitable enzyme can effectively decompose the plant or bacterial cell wall to achieve a better purification effect. Therefore, enzymatic hydrolysis has also been widely used in the process of sample pretreatment in recent years. When extracting oil crop proteins, the use of Viscozyme can fully degrade the cell wall

structure and promote the release of oil and protein in peanut cells (Liu et al., 2020a). Penha et al. (2020) using Viscozyme L under the condition of 50°C can improve the recovery rate of soybean protein from 42% to 83%, the recovery rate of isoflavones from 59% to 93%, and also reduce the residue of soybean dregs by 85%, so the utilization rate of raw materials was better improved. Enzymatic hydrolysis can also reduce the sensitization of food. Liang et al. (2021) found that Alcalase, Protamex and Flavourzyme could decompose the sensitizing proteins such as casein, β -lactoglobulin and α -lactoferrin in milk, which not only reduced the sensitization of milk but also improved the content of free amino acids and nutritional quality. Therefore, the enzymatic hydrolysis can also be used to produce hypoallergenic dairy

2016), which is a suitable choice for rough fractionation. The commonly used precipitation methods include salting out, isoelectric point precipitation and organic solvent fractionation.

5.1.1 Salting out precipitation

The solubility of protein is susceptible to the influence of ionic strength in solution. At low salt concentrations, the solubility of proteins will increase with the increase of salt concentrations in the solvent. This phenomenon is called salting in, and the proteins at this time still maintains the folded conformation and stability. If the salt concentration continues to increase, the hydration force of the solution will be enhanced, leading to the destruction of the hydration film on the surface of the proteins, resulting in the reduction of protein solubility, and then aggregation or precipitation. This phenomenon that the solubility of proteins decreases with the increase of salt concentrations is called salting out (Duong-Ly and Gabelli, 2014). The salting out method has the advantages of safety, mild action conditions, maintaining the biological activity of the target protein, low cost and simple operation. However, a large amount of salting-out agent is often left in the sediment, and some contaminants will also precipitate along with the target proteins. Therefore, desalination treatment is also an important part of the salting out precipitation method. Commonly used neutral inorganic salts are ammonium sulfate, sodium sulfate, magnesium sulfate and potassium phosphate. Ammonium sulfate has the advantages of good solubility, low price, easy preparation of high purity products, stability of protein structure and so on. It is the most commonly used neutral inorganic salt in salting out, especially for the purification of acidic proteins. In addition, ammonium sulfate can also be used with other precipitants to achieve better precipitation effect. Ovotransferrin has the functions of anti-oxidation, antibacterial and promoting iron absorption (Rathnapala et al., 2021). Abeyrathne et al. (2013) separated ovotransferrin from egg white by combining ammonium sulfate precipitation with citric acid. This process does not introduce highly polluting chemical reagents, so the ovotransferrin isolated by this method can be used in food and drug production after ultrafiltration desalination. In practical research, the optimal salt concentration of protein precipitation is usually determined by gradually increasing the salt concentration to separate the target proteins one by one under certain temperature and pH conditions. Hazim Abdul Hameed and Hussein Ali (2021) reported that 60% ammonium sulfate was the optimal ratio for rough fractionation when purifying the extracellular L-glutamate oxidase of *Streptomyces*, and the specific activity of the target protein could reach 8.25 U/mg. In addition, the ammonium sulfate salting out precipitation method can also be used for the isolation and purification of special

DNA structures, such as the purification of DNA origami nanostructures reported by Hanke et al. (2022).

5.1.2 Organic solvent precipitation

Organic solvent precipitation is another commonly used method for rough fractionation of proteins, the principle of organic solvent precipitation is that multifold organic solvent is added to the extract of biological macromolecules such as proteins to reduce the dielectric constant of the solution and increase the interaction between protein molecules, so that the solubility of proteins is significantly reduced and eventually aggregation or precipitation occurs. This method is conducive to the precipitation of high molecular weight proteins, while low molecular weight proteins or peptides are not easy to aggregate in organic solvents and remain in the supernatant (Baghalabadi and Doucette, 2020). Different organic solvents usually differ in their abilities to precipitate proteins, and common organic precipitants include acetone, isopropanol, ethanol, and methanol.

Organic solvents are usually volatile, so the residual organic solvents in the target proteins can be easily and quickly removed, and some organic solvents themselves can also be used as protein bactericides. As a commonly used organic solvent, acetone can greatly reduce protein degradation, have little impact on protein activity, and avoid contamination of impurities such as salt and polyphenol (Shaw and Riederer, 2003). When using the organic solvent precipitation to purify proteins on a large scale, it is necessary to first determine the optimal solvent concentration of the target proteins and the optimal volume ratio of protein extract to precipitant. The organic solvent selected as the precipitant should not have the ability to dissolve the target proteins, so the precipitant can be used to dissolve the soluble substances first, and then select an appropriate way to further extract the target proteins and other insoluble substances remaining in the precipitates (Wang et al., 2008; Wu et al., 2014). However, some organic solvents can destroy the hydrogen bonds in proteins and change the spatial structure of target proteins. For example, organic solvents such as ethanol release heat when mixed with water, which is easy to cause protein denaturation and inactivation. Furthermore, the precipitation process needs to consume a large amount of organic solvents and most of them are toxic and flammable, the operation is complex and the costs are high, the process needs to be carried out at low temperature, and the recovery rate is also lower than that of the salting-out method, all these make this method has certain limitations. However, in proteomic analysis, organic solvent precipitation can selectively deplete high molecular weight proteins or enrich low molecular weight proteins in soluble components, increase the detection coverage of small molecular weight proteins, and help improve the accuracy

of proteomic analysis (Periasamy et al., 2021; Nickerson et al., 2022).

5.1.3 Isoelectric point precipitation

The principle of isoelectric point precipitation is based on the fact that different proteins have different isoelectric points. When the pH value of proteins extract reaches the isoelectric point of the target protein, the net charge of the target protein is zero, the solubility is minimum, and the conformation is in the most compact state (Matulis, 2016), and no swimming occurs in the electric field, so as to achieve the purpose of separation. The isoelectric point precipitation is easy to operate and has various pH adjustment methods. However, this method also has defects. For example, when the pH reaches the isoelectric point of the target protein, it still has a certain solubility, resulting in incomplete precipitation. In addition, the inorganic acid/base (hydrochloric acid and sodium hydroxide) introduced in the adjustment of pH is easy to cause irreversible denaturation of the target proteins, and the isoelectric points of many proteins are relatively close, so the effect of isoelectric point precipitation alone is not ideal, with poor resolution and low purity, and generally only used for the preliminary purification of proteins. In order to solve the above problems, isoelectric point precipitation is often used together with other methods. For example, Watanabe et al. (2015) used isoelectric point precipitation combined with electrolytic water treatment technology to purify proteins from rice bran, which can not only significantly improve the purity and extraction efficiency of proteins, but also recover phosphorus containing compounds. Li et al. (2021) used ultrasonic assisted alkaline electrolyzed water to extract proteins from *Euphausia superba*, which can not only reduce the consumption of NaOH, but also maintain the activity of proteins and improve the yield of krill proteins.

5.1.4 Other precipitation methods

Nonionic polymer precipitation was first used to extract immunoglobulin. In recent years, it has been widely used in the isolation and purification of proteins, polypeptides, nucleic acids and enzymes (Wang et al., 2008). The polymers include dextran, NPEP, and polyethylene glycol (PEG) with different molecular weights.

The commonly used polymer PEG is an inert substance with high stability to proteins and can be stored at -20°C for a long time. PEG is generally harmless to the human body and is widely used in cosmetics and medical products (Geng et al., 2019), but some people may develop systemic allergic reactions when using these products (Sellaturay et al., 2021). PEG combined with other forms of precipitant can improve the purification efficiency and reduce the consumption of some substances. Aqueous two-phase systems (ATPS), for example, have emerged as an alternative technology for protein recovery and concentration. Menegotto et al. (2021) established the maximum recovery condition of *A. platensis* protein using 16% sodium citrate and 18% PEG

(1500 Da), and achieved a purification factor of 1.02 and a protein recovery rate of 75%. Geng et al. (2019) used 15% PEG to separate ovalbumin, ovomucin and ovotransferrin at 10°C , pH6.5 and 100 mM NaCl, then obtained a purity of 95.1% ovalbumin by isoelectric precipitation (pH4.5, 4°C) combined with HPLC purification, and the recovery rate was 46.4%. Ovalbumin can be extracted from several kilograms of egg white within 2–3 h. Furthermore, PEG is easier to extract LDL from egg yolk than ammonium sulfate (Wang et al., 2018).

5.2 Centrifugation method

The principle of centrifugation is that when an object moves in a circle around a central axis, the moving object is subjected to centrifugal force, under the action of centrifugal force to achieve the purpose of protein separation. Centrifugation is also a suitable choice for rough fractionation, which can be used as a rough fractionation scheme alone or as a key step in multi-step purification. For example, after precipitation or mechanical crushing, more precipitates can be gathered together by centrifugation to achieve a better separation effect. The commonly used centrifugation methods in protein purification are differential centrifugation and density gradient centrifugation.

5.2.1 Differential centrifugation

Differential centrifugation is based on the difference in sedimentation velocity of proteins with different sizes. The method achieves the effect of separation by continuously increasing the relative centrifugal force, controlling the centrifugation time, and conducting multiple centrifugations (Figure 3). Muthunayake et al. (2020) used differential centrifugation to enrich Bacterial Ribonucleoprotein Bodies (BR bodies) from *Caulobacter crescentus*. Ma et al. (2020) isolated the apoptotic bodies by differential centrifugation after inducing the apoptosis of osteoclasts. Differential centrifugation can obtain different cells more quickly than the conventional washing method during the isolation of mouse bone marrow cells, and it does not affect the cell viability and the distribution of hematopoietic cell populations (Heib et al., 2021). The advantage of differential centrifugation is that it is relatively convenient to separate multiple samples simultaneously. However, this method requires repeated centrifugation, which is more complex. In addition, the precipitate needs to be washed, dissolved and re precipitated for many times during the purification process, which is easy to cause sample loss and yield reduction.

5.2.2 Density gradient centrifugation

Density gradient centrifugation is also called zonal centrifugation. The protein samples to be separated were placed on the surface of density gradient formed by the medium (sucrose,

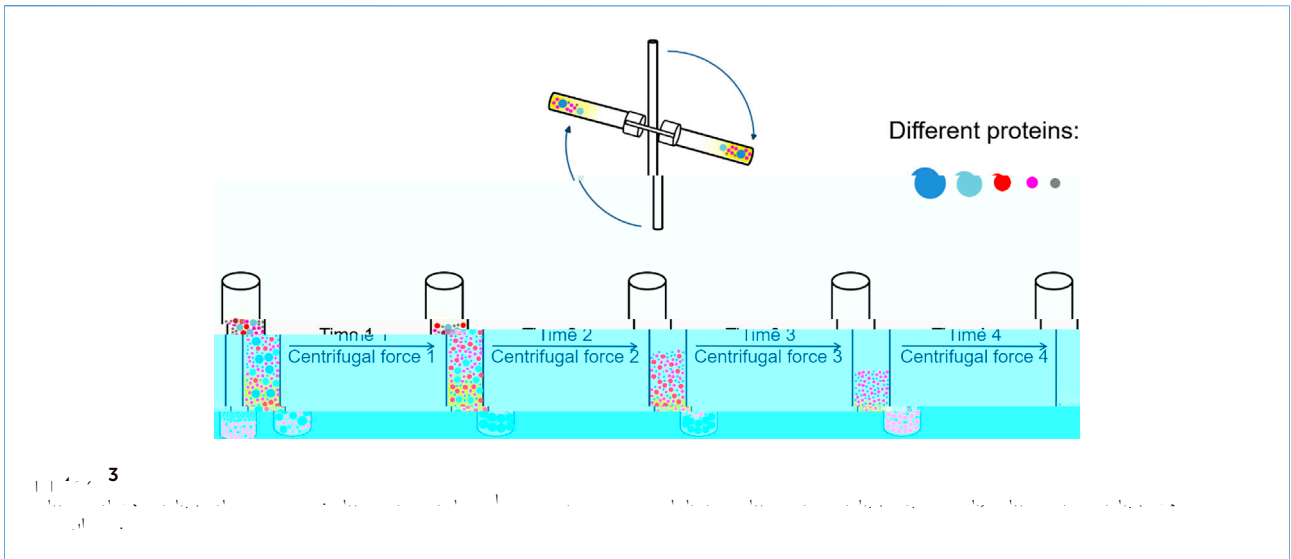


FIGURE 3 Schematic diagram of differential centrifugation. (Adapted from: <https://www.researchgate.net/publication/332211114>)

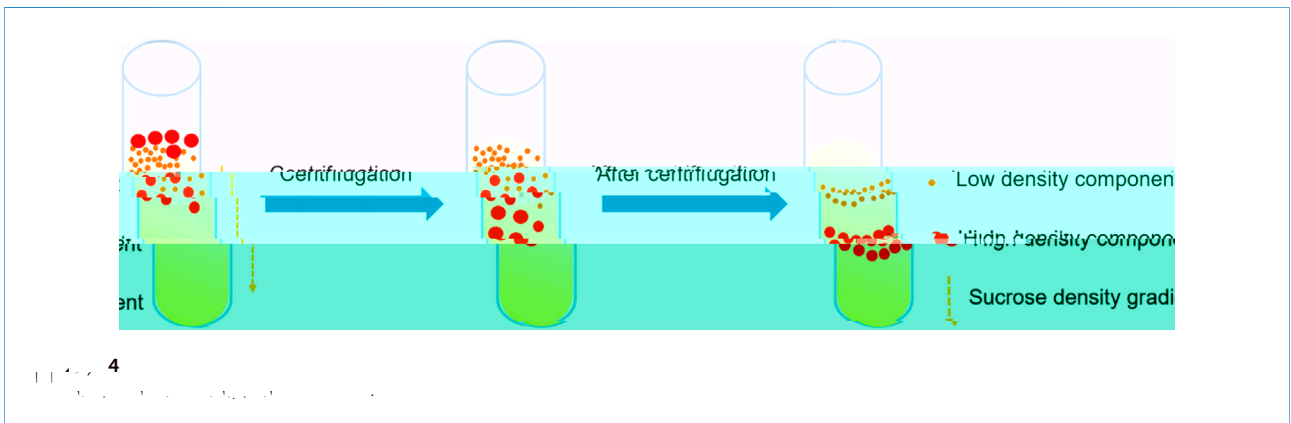


FIGURE 4 Schematic diagram of sucrose density gradient centrifugation. (Adapted from: <https://www.researchgate.net/publication/332211114>)

cesium chloride, etc.) with gradually increasing density and high solubility, and the proteins of different shapes and sizes were separated by centrifugal force to form different settling zones at different settling speeds (Figure 4). This method can separate multiple components in mixed samples with high resolution and low cost. At present, it has been widely used for the isolation and purification of cells, organelles, viruses, bacteria, nucleic acids, proteins, etc. For example, the inner and outer membrane of liver tissue can be separated by sucrose density gradient centrifugation (Lin et al., 2010), and mitochondria after differential centrifugation can also be further refined purified (Clayton and Shadel, 2014). Szelechowski et al. (2013) purified adenoviruses and bacteriophages by cesium chloride density gradient centrifugation. Nasukawa et al. (2017) reduced the centrifugal force from 100,000g to 40,000 g and increased the centrifugation time from 1 to 2 h under the condition of 4°C, and completed the phage purification by non ultracentrifugation, realizing the virus purification without ultracentrifuge. In the final

step of bacterial ribosome purification, polysomes were purified for *in vitro* cell-free translation systems by centrifugation of samples containing 70S monomers, ribosomal subunits, and polysomes through a linear sucrose gradient of 7%–30% (w/w) (Rivera et al., 2015). Wichmann et al. (2021) used OptiPrep density gradient centrifugation to isolate a higher proportion of bacteria from the complex matrix with low microbial load, confirming the compatibility of this method with Raman spectroscopy, and the combination of the two methods can improve the accuracy of bacterial infection diagnosis.

5.3 Dialysis method

Dialysis is a method to separate colloidal substances such as proteins from other small molecular substances including inorganic salt ions and reducing agents based on the principle that protein and other macromolecular particles are large, have

TABLE 4 | *E. coli* strains used in this study.

Strain name	Important features	Application
BL21	<i>E. coli</i> polymerase but no T7 RNA polymerase	It is mainly suitable for the expression of non-toxic proteins, and it can be used for the protein expression of plasmids (such as pGEX and pMAL)
BL21 (DE3)	The λ phage DE3 region containing T7 phage RNA polymerase is integrated on chromosome BL21 allowed the simultaneous expression of both T7 RNA polymerase and <i>E. coli</i> RNA polymerase. (Ivanov et al., 1993)	For the efficient expression of the genes cloned in the expression vectors containing the phage T7 promoter (such as the pET series), and for the protein expression of the plasmids (such as pGEX and pMAL)
BL21 (DE3) pLysS	Carrying the pLysS plasmid, having chloramphenicol resistance, and containing genes expressing T7 lysozyme. T7 lysozyme can reduce the background expression level of the target gene, but does not interfere with IPTG induced expression	Suitable for the expression of toxic proteins and non-toxic proteins
BL21 (AI)	It has tetracycline resistance (Bhawsinghka et al., 2020), lacks Lon protease and OmpT extramembrane protease. It can effectively prevent the degradation of heterologous proteins in <i>E. coli in vivo</i>	Suitable for any T7 promoter-based expression vectors. Widely used for the high-level expression of toxic recombinant proteins (Bhawsinghka et al., 2020)
BL21 Star (DE3)	It contains <i>E. coli</i> RNA polymerase, has chloramphenicol resistance, and contains the rne131 mutation, which reduces the accumulation of endogenous RNase and enhances mRNA stability of intracellular mRNA in the strain (Jiang et al., 2021), thereby it increases the expression level of heterologous proteins	Suitable for the high-level protein expression of the T7 promoter expression vector (such as the pET series) and the non-T7 promoter expression vector (such as pGEX and pMAL), it is not suitable for the expression of toxic proteins (Jiang et al., 2021)
BL21 Star (DE3) pLysS	Containing the <i>E. coli</i> RNA polymerase, carrying the pLysS plasmid, having chloramphenicol resistance, and containing the rne131 mutation, enhances the cellular stability of the mRNA within the strain, thus increasing the expression level of the heterologous protein. Lower the background expression level of the target gene, but did not interfere with the IPTG-induced expression	It is suitable for high-level expression of protein expression in T7 promoter expression vectors (such as the pET series) and non-T7 promoter expression vectors (such as pGEX and pMAL)
BL21-CodonPlus (DE3)-RIPL	It is absent of the Lon protease and the OmpT protease, thus reducing the degradation of the recombinant proteins and enhancing the expression levels of foreign genes. In particular, the expression levels of AT-or GC-rich eukaryotic genes in the prokaryotic system. It can simultaneously express T7 RNA polymerase and <i>E. coli</i> RNA polymerase, and has resistance to tetracycline, chloramphenicol, streptomycin, and spectacular mycin	It can be used for the protein expression of plasmids (such as the pET series, pGEX, and pMAL)

colloidal properties, and cannot freely pass through the semi-permeable membrane (Echave et al., 2021). The activity of purified protease can be improved by attaching specific substances targeting the target protease to the semi permeable membrane (Labus et al., 2020). In order to improve the recovery rate of the target proteins during dialysis, factors such as buffer exchange time, the design of dialysis system, and the chemical and morphological characteristics of dialysis membrane should be considered. In terms of buffer exchange time, the increasing temperature can accelerate the molecular movement and increase the intrinsic diffusion of proteins across the semi-permeable membrane to reduce dialysis time, but increasing temperature may also have adverse effects on protein activity. Therefore, the maximum temperature at which protein activity can be maintained should be determined in advance when using dialysis for protein purification (Burgess, 2016). Dialysis is also commonly used for desalting treatment after the end of salting out. When the protein solution is in a high concentration and high salinity state, dialysis is usually the best method to remove ammonium sulfate from the sample (Duong-Ly and Gabelli, 2014). Phycocyanin has antioxidant, anticancer, liver and kidney protection effects (Piovan et al., 2022), and has been widely used in the production of drugs and nutritional products in recent

years. Khazi et al. (2020) pretreated *Cyanobacteria* to obtain the crude extract of phycocyanin. Ammonium sulfate was added into the crude extract in sections, and the supernatant was removed by centrifugation after full stirring, then the phycocyanin was resuspended and dialyzed overnight to achieve further purification effect.

5.4 Ultrafiltration method

Ultrafiltration is also a method based on membrane separation, which can achieve the effect of concentration and is also a common way of desalination (Thammasena et al., 2020). The performance of ultrafiltration membrane is the key to the success of protein separation. Most ultrafiltration membranes consist of a sturdy scaffold structure with a very thin polymer layer attached to it. Cellulose acetate can be used for filtration, gas separation, adsorption and ion exchange (Vatanpour et al., 2022) due to its advantages of film-forming, good chemical activity (Ma et al., 2022), high stability (Muhmed et al., 2019) and hydrophilicity, environmental protection, and appropriate costs (Sharma et al., 2020; Syamani, 2020). Although the production technology of membranes used for ultrafiltration

has not changed in the past few decades, the control technology of thin layer pore size distribution, membrane morphology and membrane modification has been significantly improved. It is this thin layer that provides the properties required for a selective permeability membrane and determines the flow resistance (Burgess, 2016). The new dopamine modified cellulose acetate ultrafiltration membrane is a relatively advanced filter membrane. The use of Dopamine-modified cellulose acetate overcome the trade-off between permeability and selectivity of conventional cellulose acetate membranes to some extent, possess good antifouling capability and long-term stability, and intercept effectively the target protein as well as improve the water permeability (Ma et al., 2022). In order to avoid the accumulation of intercepted proteins on the membrane surface, tangential flow filtration can be used (Busatto et al., 2018). The principle is that the protein solution is driven by the pump to flow along the direction tangent to the membrane surface. The pressure difference formed on the membrane makes part of the liquid pass through the membrane, while the other part of the liquid flows tangentially through the membrane surface to wash away the intercepted protein molecules. The currently developed one-way tangential flow ultrafiltration technique enables the continuous ultrafiltration of mAb by increasing the surface area and the residence time (Thakur et al., 2022). Ultrafiltration is also applied to the treatment of diseases such as acute heart failure (Costanzo, 2019), the removal of pathogens from seawater (Cordier et al., 2020), the purification of the vaccine product (Emami et al., 2019) and recombinant ferritin (Palombarini et al., 2019), and the treatment of galvanized wastewater (Oztel et al., 2020).

To sum up, the rough fractionation methods are relatively simple and have a large processing flux. They can not only remove a large number of impurities, but also concentrate the protein solution to achieve partial purification, while maintaining a high recovery rate (Mazi et al., 2016). Among the above rough fractionation techniques, PEG precipitation, ultrafiltration, differential centrifugation and other methods can achieve the concentration of viruses in wastewater (Dumke et al., 2021). During the SARS-CoV-2 pandemic, many countries used these methods to enrich and detect viruses in wastewater, thereby boosting epidemiological surveillance and playing an early warning effect (Ahmed et al., 2020; Medema et al., 2020; Ai et al., 2021; Hillary et al., 2021). However, the effect of rough fractionation is poor, so in some important structural and functional analysis studies, it is necessary to combine more refined fractionation methods to obtain the target proteins with higher purity.

6 Fine fractionation

After a series of pretreatment and rough fractionation, the volume of the sample was reduced and most of the impurity

proteins were removed. The purpose of fine fractionation is to separate the target proteins from some proteins of similar size and physical and chemical properties to obtain higher purity, so as to meet the needs of different research fields. At the same time, higher resolution and stronger specificity are required in the process of fine fractionation.

6.1 Chromatography

Chromatography is based on the difference in physical and chemical properties of different substances (Scopes, 2001). It has been used to extract plant pigments as early as the beginning of the 20th century. At present, chromatography has been widely used in the separation of proteins, nucleic acids, polysaccharides, peptides and other biological macromolecules. Chromatography can be divided into adsorption chromatography and non-adsorption chromatography according to whether the sample is bound to the filler. The former mainly includes ion exchange chromatography, hydrophobic chromatography, affinity chromatography and reverse phase chromatography, while the latter mainly includes gel filtration chromatography. Several commonly used chromatographic techniques are introduced as follows. (Table 5).

6.1.1 Gel filtration chromatography

Gel filtration chromatography, also known as molecular sieve chromatography, is one of the most effective methods developed in the 1960s to separate and purify different proteins based on molecular size, and is also the gold standard for separating protein polymers from their monomers (London et al., 2014). In the elution process, proteins with high molecular weight first flow out along the gap between the gel particles, while proteins with low molecular weight can enter the gel mesh, resulting in flow obstruction and slow outflow.

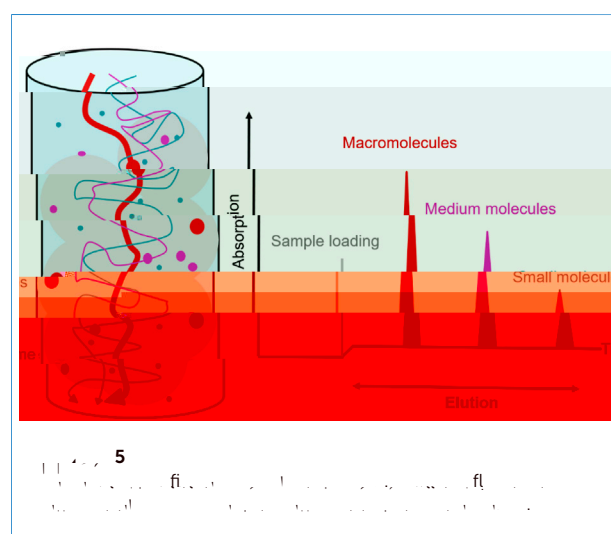


Figure 5

Figure 5: Gel filtration chromatography. The diagram shows a column with gel particles and a detector. The column is divided into three zones: Macromolecules (top), Medium molecules (middle), and Small molecule (bottom). The detector shows three peaks corresponding to these zones. The process is labeled 'Absorption' and 'Elution'.

FIGURE 5 | Comparison of chromatography methods.

Types	Gel filtration chromatography	Ion exchange chromatography	Hydrophobic chromatography	Affinity chromatography
Characteristic				
Separation mechanism	Size	Charge	Hydrophobicity	Specific affinity
Selectivity	Secondary	High	Medium/high	Very high
Load	Low	High	High	High
Purification rate	Medium/low	High	High	High
Biocompatibility	Very nice	Good	Medium/good	Good
Purified protein yield	High	High	Medium/high	High
Capture concentration phase	+	+++	+++	+++
Intermediate purification stage	+	+++	+++	+++
Refining and purification stage	+++	+++	+	++

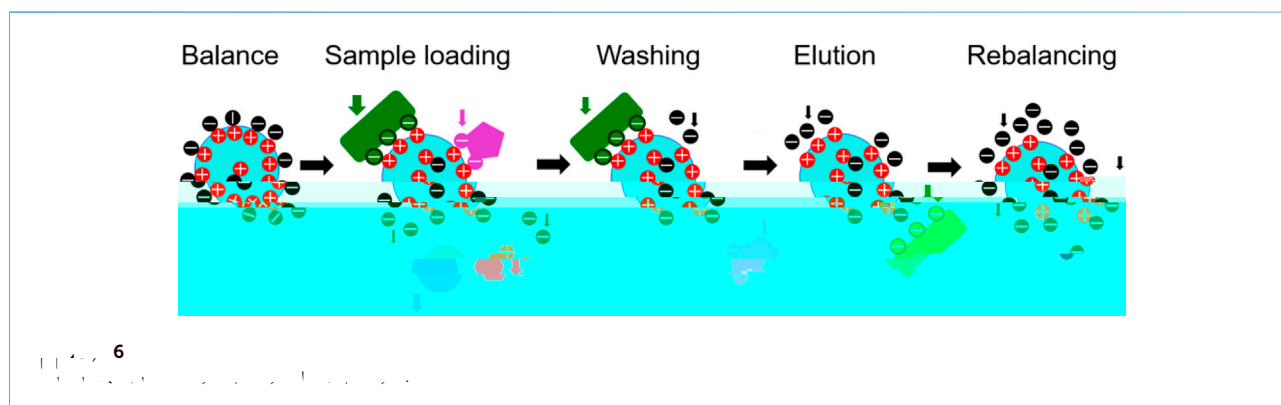


FIGURE 6 | Comparison of chromatography methods.

The required proteins can be collected according to the time sequence of elution (Figure 5). The matrix for gel filtration can be composed of a variety of materials, including simple substances such as dextran (Sephadex™ Series), agarose (Sepharose™ Series) and polyacrylamide (bio-Gel Series) or mixtures composed of dextran polyacrylamide (Sephacryl™ Series) or dextran-agarose (Superdex™ Series). Gel filtration chromatography, which has the advantages of simple operation, rapid separation without affecting biological activity, has been widely used for the isolation and purification of proteins or peptides (Gao et al., 2016). At the same time, the dextran gel can also be used for desalination after salting out. However, the gel filtration chromatography also has certain defects. The chromatography column is longer than other separation columns, so the flow rate is slower, the elution time is longer, and higher column pressure and more fillers are required. Gel filtration chromatography is often used in combination with other chromatography methods (Li Z. et al., 2022). To further investigate the effect of D614G substitution on the structure of SARS-CoV-2 spike protein, Zhang et al. (2021) purified

the SARS- CoV-2 spike protein carrying D614 or G614 with detergent DDM, further purified it through gel filtration chromatography and used it for structural analysis, providing a structural basis for the development of new coronavirus vaccine. Ye et al. (2022) used 70% methanol as elution buffer, and obtained high concentrations of isoorientin and four flavone C-glycosides from bamboo leaf flavonoid by gel filtration chromatography. Gel filtration chromatography can also be used in the last step of purification of low molecular weight hyaluronic acid (Karami et al., 2021). Cao et al. (2022) used TALON IMAC metal chelation chromatography and gel filtration chromatography to obtain pure serotonin 2A receptor (5-HT2AR) for research on the treatment of depression. Zhang Y. et al. (2021) obtained high-purity NMDA receptors by combined use of Strep-Tactin affinity chromatography and gel filtration chromatography (Superose six Increase), and the analysis of its three-dimensional structure by cryo-EM provided new insights for the development of antidepressants. Moreover, gel filtration chromatography also plays an important role in medical

examination and is the gold standard for the detection of macroprolactin (Vilar et al., 2019). When the recovery rate of thyroid stimulating hormone (TSH) does not decrease, the gel filtration chromatography can be used to detect the presence of macro-TSH as the evaluation index of normal thyroid function (Fukushita et al., 2021).

Gel filtration chromatography is suitable for the separation of proteins with larger or smaller molecular weight. When the molecular weight of proteins in the sample is moderate, the purity of the target proteins obtained is relatively low. When the molecular weight of proteins is 25% different from each other, they can be completely separated by a single gel column. In addition, the viscosity of the sample should not be too high, otherwise the mass transfer resistance will increase, and the gel particles will sometimes have non-specific adsorption, which is easy to block the column.

6.1.2 Ion exchange chromatography

Ion exchange chromatography uses ion exchanger as stationary phase to separate and purify according to the difference in reversible binding strength between component ions in the mobile phase and equilibrium ions on the exchanger. Ion exchangers are made by introducing several dissociable groups (active groups) into an insoluble polymer substance (the parent body). It is the many covalently bound charged groups and convertible ions fixed to the parent body that play a key role in the chromatography process. The interaction strength between the proteins and ion exchangers changes with the buffer salt concentration or pH, and the protein is eluted according to the binding strength, so as to achieve the purpose of separation and purification. According to the properties of active groups, ion exchangers can be divided into cation exchangers and anion exchangers. According to the difference of parent, ion exchangers can be divided into

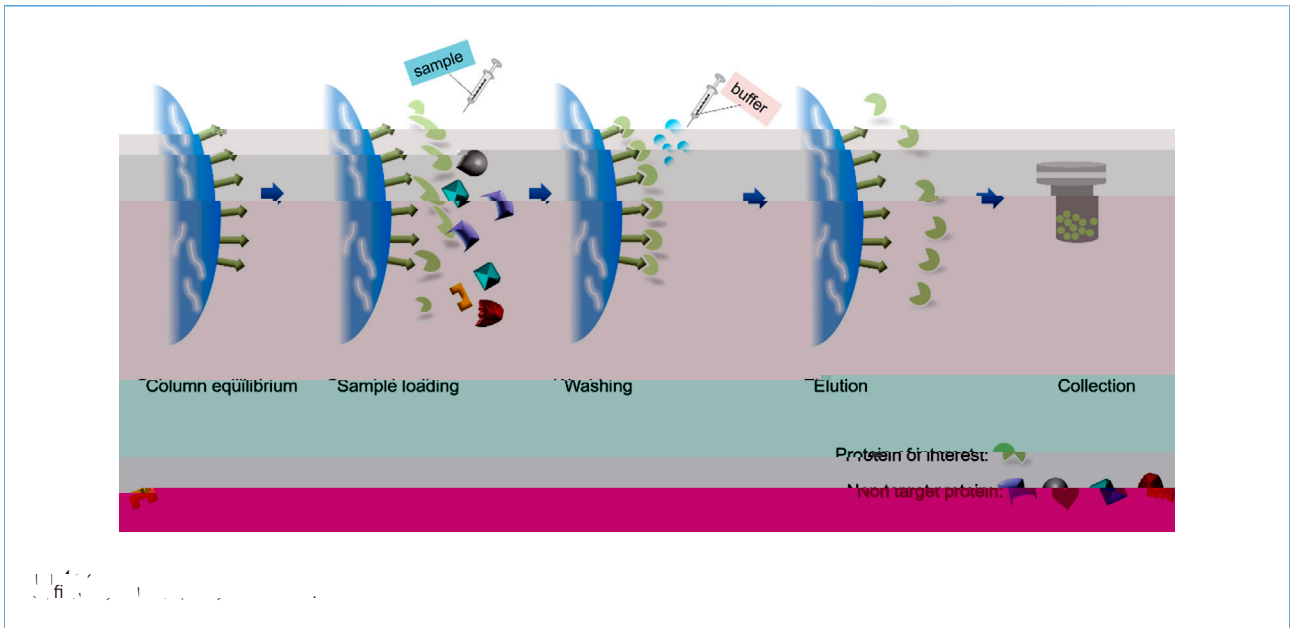


FIGURE 1 | Affinity chromatography.

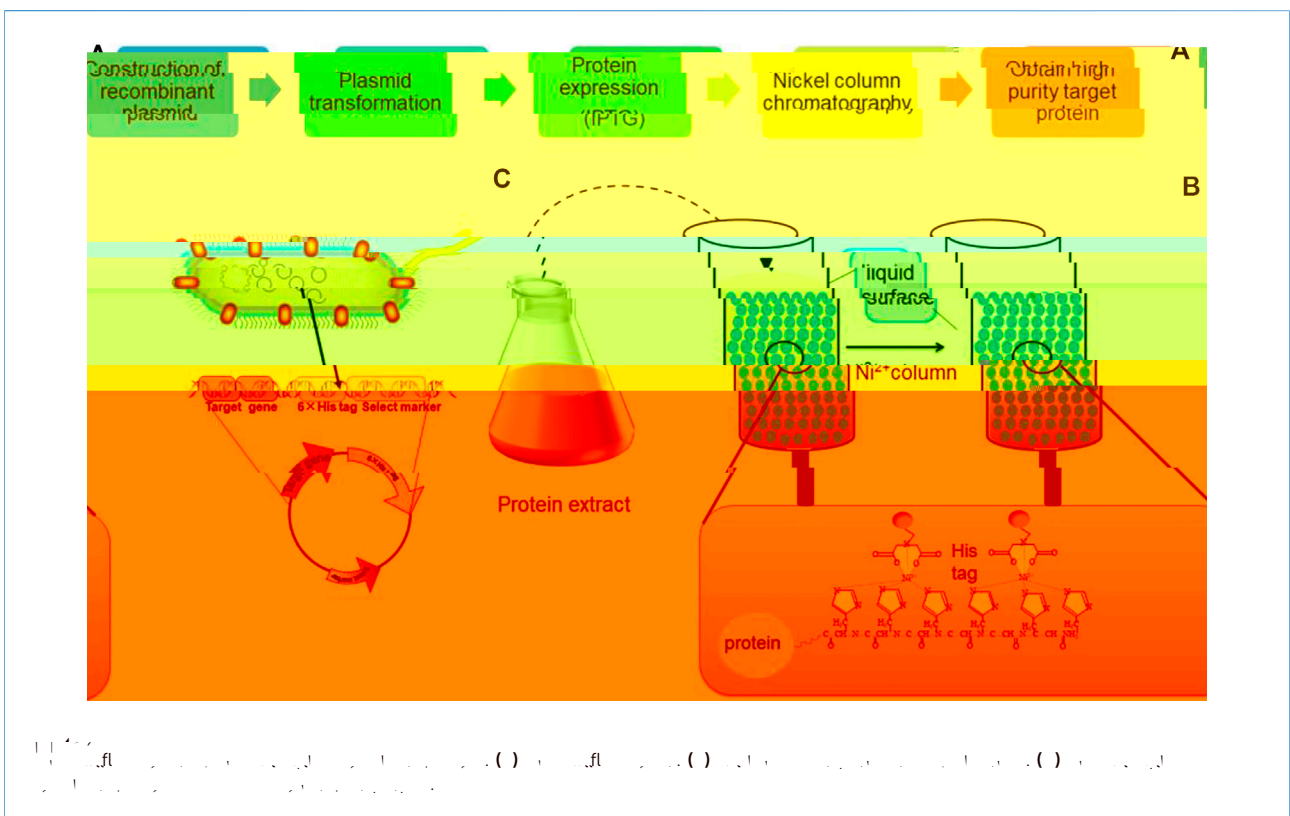


FIGURE 2 | Construction of recombinant plasmid, plasmid transformation, protein expression (IPTG), nickel column chromatography, and obtain high purity target protein.

binding conditions. Affinity chromatography is fast, simple, and highly efficient (Lacki and Riske, 2020), which is often used to isolate compounds with specific tags or to study interactions between biological macromolecules. The

interactions between proteins and ligands are based on non-covalent interactions such as electrostatic gravity, molecular hydrophobicity, van der Waals forces and hydrogen bonding forces.

Tags	Sequence/Size	Features	Application
Epitope Tags			
His	6–10 His/0.84 kDa (Mishra, 2020)	It has a small molecular weight, and no impact on the function and downstream application of the target proteins. It has low immunogenicity so that purified proteins were directly immunized to animals and antibodies were prepared (Freitas et al., 2022)	Bromberg et al. (2022) found that His tag can be used as a decoy modulating preferred orientation in cryo-EM, providing a new scheme for improving cryo-EM sample processing Williams et al. (2021) used the optimized His tag the integrate with recombinant proteins, which is not only convenient for affinity chromatography, but also provides an effective binding site for the radiolabeling of the recombinant protein. Therefore, the optimized His tag are used in the process of cancer molecular imaging and radionuclide therapy with more economic advantages
FLAG	DYKDDDDK/1.01 kDa (Mishra, 2020)	It has a small molecular weight, high hydrophilicity (Schmidt et al., 2012), The tag has high sensitivity and are easy to remove after use	Han et al. (2017) purified functional reprogramming factors from HEK293 cells using FLAG tag. Imagawa et al. (2021) purified virus-like particles with FLAG-tagged envelope protein through one-step affinity purification with FLAG tag, which could be used as a tetravalent dengue vaccine candidate
HA	YPYDVPDYA/1.1 kDa	The tag is derived from the influenza virus hemagglutinin (Zhou et al., 2012). The tag has little influence on the spatial structure of the target protein, and it is easy to be constructed into tag proteins fused to the N or C terminus	Vyas et al. (2020) CRISPR-Cas9 gene editing technology was used to connect HA tag as locators to the C-terminus of mouse nicotinic acetylcholine proteins (nAChRs) $\alpha 9$ or $\alpha 10$. The expression of $\alpha 9$ and $\alpha 10$ proteins can be detected
Strep	Including Strep-tag II (WSHPQFEK) (Tremante et al., 2021), Twin-Strep-tag (WSHPQFEK-GGGSGGGSG-SA-WSHPQFEK) (Schmidt et al., 2013)	Strep-tag II is small, which does not affect the folding ability of protein, and has strong specific binding ability with affinity resin, which effectively avoids non-specific binding. Therefore, the purity of the target proteins obtained is high. The target proteins can be purified under mild conditions to maintain its activity (Johar et al., 2017)	Kawaguchi et al. (2019)

6 (Continued)

Tags	Sequence/Size	Features	Application
Myc	EQKLISEEDL/1.2 kDa (Mishra, 2020)	The tag is derived from peptide 410–419 of human c-Myc protein, which can be fused to the N-terminal or C-terminal of the target protein	Myc tag can be used as epitope tag, widely used in the detection of anti-c-Myc hybridoma E10 antibody Myc1-9E10 (Schuchner et al., 2020)
Protein/domain tags			
Glutathione S-transferase (GST)	26 kDa (Williams et al., 2020)	The tag was derived from <i>Schistosoma japonicum</i> , which could be expressed in different hosts and was widely applicable. It can be expressed in different hosts, and has a wide range of applications. Different proteases can be easily removed (Kim et al., 2021). It can increase the solubility and improve the expression and stability of foreign proteins. It has good specificity and helps to maintain the antigenicity and the activity of the proteins after eluting under mild and non-denaturing conditions. It has a large molecular weight and may affect protein functions and downstream experiments	Zhou et al. (2020) fused melittin (MET) with the GST tag and expressed melittin using the <i>E. coli</i> expression system. After affinity chromatography, digested with prescission protease, gel filtration chromatography, melittin with a purity of >90% was purified, which also showed strong antibacterial activity. Such a scheme could be used for large-scale industrial production of melittin
MBP	40 kDa (Greenfield et al., 2020)	It is one of the most commonly used crystallization chaperones for the target proteins (Fox et al., 2001), and can increase the solubility of the fusion proteins, especially the eukaryotic proteins. MBP tags can be easily detected by immunoassay	Nosaki et al. (2021) have successfully revealed the crystal structure of AtBIL1/BZR1 DBD and target DNA complex by the combination of MBP-mediated crystallization and MD simulation. This is shown to be capable of deciphering the protein-DNA recognition code of interest Li et al. (2019) found that MBP tag can assist in the crystallization of the CAMP factor of <i>Streptococcus agalaciae</i> , facilitating the subsequent study of the structure of the CAMP factor Guo et al. (2018) used MBP tag to prepare high purity and high activity MBP-MLIF in the <i>E. coli</i> expression system. On the whole, this method has advantages such as simplicity and application, which promotes the scientific research of stem cells
SUMO	11.2 kDa (Heinrich et al., 2019)	It can increase the expression amount of the fusion proteins (Mohammadi et al., 2022), and has the functions such as anti-protease hydrolysis, promoting the correct folding of the target protein, enhancing the solubility, and shielding the toxic proteins	Kim et al. (2019) used SUMO tag and His tag to construct a new vector pKSEC1 for the production of antimicrobial peptide-Abaecin, providing a new scheme for the treatment of bacterial infection Mohammadi et al. (2022) used SUMO tag to express a large number of recombinant human angiotensin-converting enzyme 2 (rhACE2) in <i>E. coli</i> , and the solubility of recombinant human angiotensin-converting enzyme 2 (rhACE2) could be improved in combination with freeze-thaw method, providing a reasonable choice for subsequent large-scale purification
Halo	33 kDa (Freidel et al., 2016)	Halo Tag monomeric proteins can be fused to either the N or C terminus of the recombinant proteins and express in prokaryotic and eukaryotic systems. Halo Tag can rapidly bind to ligands with high specificity and affinity to form stable covalent compounds	Deo et al. (2021) used the self-labeling Halo tag protein as a fluorescent indicator sensor scaffold to detect a single action potential in cultured neurons. Liu et al. (2018) connected the Halo tag and tobacco Etching virus (TEV) protease cleavage site at the n-terminal in the <i>E. coli</i> expression system. After chromatography and enzyme digestion, high purity Halo-FGF7 and rhFGF7 were respectively determined by <i>in vitro</i> and <i>in vivo</i> activity, MTT, the

(Continued on following page)

Tags	Sequence/Size	Features	Application
SNAP	19.4 kDa (Freidel et al., 2016)	The tag is derived from human O6-alkylguanine-DNA-alkyltransferase (Padayachee et al., 2019). It can covalently bind with the substrate to tag proteins with biotin or fluorescent groups; highly specialized, stable and suitable for protein detection and purification in various environments	protective effects of rhFGF7 and Halo-FGF7 on acute liver injury were further investigated by Western blot. Akin et al. (2021) used Halo tag to mark Nav1.7 channels in nerve cells to study the effect of paclitaxel (PTX) increase axonal localization and vesicular trafficking of Nav1.7 Fu et al. (2021) used SNAP as the tag to construct Snap-tagged EGFR/CMC in combination with HPLC-IT-TOF-MS to screen EGFR-targeted active compounds in traditional Chinese medicine, efficiently isolate and accurately identify them, and screen out important candidates for disease treatment. Moinpour et al. (2022) used the SNAP-tag biocoupling strategy to form stable covalent bonds between benzyl guanine (BG) -modified phospholipids and SNAP-tag fusion proteins, thus realizing programmable control of protein capture. It is also proved that the SNAP-tag biological coupling strategy can directly isolate proteins in the presence of complex biochemical mechanisms, which will facilitate the development of advanced lipid-based artificial organelles
Calmodulin-binding peptide (CBP)	RWKKNFIAVSAANRFKKIS (Mukherjee et al., 2015)	The tag is derived from the 26 amino acid sequences of myosin light streptokinase of	

The carrier of affinity chromatography should have the following characteristics (Chen et al., 2003): 1) Insoluble in water, but highly hydrophilic; 2) Inert substance with good specificity; 3) There are enough chemical groups for activation; 4) Stable physical and chemical properties; 5) Good mechanical properties, with a certain particle shape to maintain a certain flow rate; 6) Good permeability, with porous network structure for free passage of macromolecular substances; 7) It can resist the damage to microorganisms and alcohols. The commonly used stationary phase carriers include alumina,

polyacrylamide gel, dextran gel, cellulose, agarose, metal chelate and so on.

Affinity tags are important recognition structures for affinity chromatography, located inside the N terminus of the target protein, whose essence is a protein or a peptide. Some affinity tags also help the protein fold (Kou et al., 2007), increase solubility (Hammarstrom et al., 2002; Chen et al., 2005; Nallamsetty and Waugh, 2006), and in turn increase the yield of the target proteins (Sun et al., 2005). Currently, multiple types of affinity tags have been developed. The most commonly used His tag

(Kinrade et al., 2020) is composed of six or more tandem histidine residues, and it is small, cheap, and has little impact on the structure and function of the target proteins (Figure 8) (Lichty et al., 2005). Abis et al. (2019) purified human soluble cyclooxygenase-2 by nickel ion chelation chromatography combined with and BTS (Benzylthio-Sepharose) affinity chromatography, and the target proteins showed a high purity, as determined by SDS-PAGE analysis. Ge et al. (2022) purified the recombinant SARS-CoV-2 S1 protein by immobilized metal affinity chromatography and used it as an antigen for immunizing hens. Öztürk and Demir (2021) proposed to use CTS-p (HEMA) -Cu²⁺ as an IMAC adsorbent, which can quickly, efficiently and repeatedly detect the amount of residues of melamine in different complex substrates. Wang et al. (2022) used titanium (IV) ion-fixed metal affinity chromatography (Ti⁴⁺ -IMAC) to jointly extract DNA binding proteins (DBPs) and RNA binding proteins (RBPs), which amount of

6.1.4 Hydrophobic chromatography

There are some amino acids with hydrophobic side chains on the surface of most proteins, such as phenylalanine, tryptophan, methionine, etc. The number, size and distribution of these amino acids determine the properties of proteins (O'Connor and Cummins, 2017). Hydrophobic chromatography is a purification technology developed based on the hydrophobicity difference of proteins. The balance of hydrophobicity is mainly controlled by salt (Figure 10). Generally, controlling the concentration of salt ions can effectively remove the self-aggregation or self-interaction caused by hydrophobic interaction within the sample. Hydrophobic chromatography can also be used to remove impurities and monitor the purity of the target proteins. Weigel et al. (2019) used hydrophobic chromatography to reversibly combine virus particles to remove residual contaminated DNA and proteins when purifying influenza A and B viruses.

7 Challenges and future perspectives

Protein is the material basis of life activities. With the continuous development of molecular biology, structural biology, genomics and bioinformatics, people gradually realize that it is far from enough to clarify the

Author contributions

All authors listed have made a substantial, direct and intellectual contribution to this work, and approved it for publication. Conception: LL and BX; Writing—original draft: MD, ZH, LL, and BX; Methodology, investigation and revising: ZH, YX, XC, LF, and ZL; Writing—Reviewing and Editing: MD, ZH, LL, and BX.

Funding

This study was supported by Fundamental Research Program of Shanxi Province (Grant no. 202103021223397), Science and Technology Innovation Project of Colleges and Universities in Shanxi Province (Grant no. 2020L0749), Key R&D Projects of Introducing High-Level Scientific and Technological Talents in Lvliang City (Grant no. 2021RC-1-4), the Project of Luliang City Science and Technology Program (Grant no. 2020SHFZ29), the National College Students' Innovation and Entrepreneurship Training Program (Grant no. 202117114001), the Key Projects of Innovation and Entrepreneurship Training for College

Students in Shanxi Province (Grant no. S202117114008; Grant no. 20221577), and Projects of Innovation and entrepreneurship training program for college students of Fenyang College of Shanxi Medical University (Grant no. FDC202111).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the

dynamic gene regulation. *ACS Synth. Biol.* 11 (2), 921–937. doi:10.1021/acssynbio.1c00545

Chen, H. Q., Xu, Z. N., Xu, N. Z., and Cen, P. L. (2005). Efficient production of a soluble fusion protein containing human β -defensin-2 in *E. coli* cell-free system. *J. Biotechnol.* 115 (3), 307–315. doi:10.1016/j.jbiotec.2004.08.012

Chen, J., Pan, Z. D., Han, L., Zhou, Y. X., Zong, H. F., Wang, L., et al. (2021). A novel bispecific antibody targeting CD3 and Lewis Y with potent therapeutic efficacy against gastric cancer. *Biomedicine* 9 (15), 1059. doi:10.3390/biomedicine9081059

Chen, J., Xu, W., Li, L. T., Yi, L. C., Jiang, Y. H., Hao, P. F., et al. (2022). Immunogenicity and protective potential of chimeric virus-like particles containing SARS-CoV-2 spike and H5N1 matrix 1 proteins. *Front. Cell. Infect. Microbiol.* 12, 967493. doi:10.3389/fcimb.2022.967493

Chen, L., Yang, X. Y., Luo, D., and Yu, W. C. (2016). Efficient production of a bioactive bevacizumab monoclonal antibody using the 2A self-cleavage peptide in transgenic rice callus. *Front. Plant Sci.* 7, 1156. doi:10.3389/fpls.2016.01156

Chen, T., Liu, Y., and Pan, J. Q. (2003). A new technology of isolation and purification: Affinity chromatography. *Mod. Food Technol.* 19 (2), 98–101. doi:10.3969/j.issn.1673-9078.2003.02.038

Chen, W. H., Chag, S. M., Poongavanam, M. V., Biter, A. B., Ewera, E. A., Rezende, W., et al. (2017). Optimization of the production process and characterization of the yeast-expressed SARS-CoV recombinant receptor-binding domain (RBD219-N1), a SARS vaccine candidate. *J. Pharm. Sci.* 106 (8), 1961–1970. doi:10.1016/j.xphs.2017.04.037

Chigira, Y., Oka, T., Okajima, T., and Jigami, Y. (2008). Engineering of a mammalian O-glycosylation pathway in the yeast *Saccharomyces cerevisiae*: Production of O-fucosylated epidermal growth factor domains. *Glycobiology* 18 (4), 303–314. doi:10.1093/glycob/cwn008

Clayton, D. A., and Shadel, G. S. (2014). Purification of mitochondria by sucrose step density gradient centrifugation. *Cold Spring Harb. Protoc.* 2014 (10), prot080028. doi:10.1101/pdb.prot080028, pdb

Cordier, C., Stavrakakis, C., Morga, B., Degremont, L., Voulgaris, A., Bacchi, A., et al. (2020). Removal of pathogens by ultrafiltration from sea water. *Environ. Int.* 142, 105809. doi:10.1016/j.envint.2020.105809

Costanzo, M. R. (2019). Ultrafiltration in acute heart failure. *Card. Fail. Rev.* 5 (1), 9–18. doi:10.15420/cfr.2018.29.2

Darlington, A. P. S., Kim, J., Jimenez, J. I., and Bates, D. G. (2018). Dynamic allocation of orthogonal ribosomes facilitates uncoupling of co-expressed genes. *Nat. Commun.* 9 (1), 695. doi:10.1038/s41467-018-02898-6

Das, T. K., Narhi, L. O., Sreedhara, A., Menzen, T., Grapentin, C., Chou, D. K., et al. (2020). Stress factors in mAb drug substance production processes: Critical assessment of impact on product quality and control strategy. *J. Pharm. Sci.* 109 (1), 116–133. doi:10.1016/j.xphs.2019.09.023

de Marco, A., Deuerling, E., Mogk, A., Tomoyasu, T., and Bukau, B. (2007). Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *E. coli*. *BMC Biotechnol.* 7, 32. doi:10.1186/1472-6750-7-32

Deo, C., Abdelfattah, A. S., Bhargava, H. K., Berro, A. J., Falco, N., Farrants, H., et al. (2021). The HaloTag as a general scaffold for far-red tunable chemigenetic indicators. *Nat. Chem. Biol.* 17 (6), 718–723. doi:10.1038/s41589-021-00775-w

Diamant, S., Eliahu, N., Rosenthal, D., and Goloubinoff, P. (2001). Chemical chaperones regulate molecular chaperones *in vitro* and in cells under combined salt and heat stresses. *J. Biol. Chem.* 276 (43), 39586–39591. doi:10.1074/jbc.M103081200

Dirisala, V. R., Nair, R. R., Srirama, K., Reddy, P. N., Rao, K. R. S. S., Satya Sampath Kumar, N., et al. (2017). Recombinant pharmaceutical protein production in plants: Unraveling the therapeutic potential of molecular pharming. *Acta Physiol. Plant.* 39 (1), 1–9. doi:10.1007/s11738-016-2315-3

Donovan, R. S., Robinson, C. W., and Glick, B. R. (1996). Review: Optimizing inducer and culture conditions for expression of foreign proteins under the control of the lac promoter. *J. Ind. Microbiol.* 16 (3), 145–154. doi:10.1007/BF01569997

Dow, B. A., Tatulian, S. A., and Davidson, V. L. (2015). Use of the amicyanin signal sequence for efficient periplasmic expression in *E. coli* of a human antibody light chain variable domain. *Protein Expr. Purif.* 108, 9–12. doi:10.1016/j.pep.2014.12.017

Dudenhoeffer, B. R., Borggraefe, J., Schweimer, K., and Knauer, S. H. (2020). NusA directly interacts with antitermination factor Q from phage λ . *Sci. Rep.* 10 (1), 6607. doi:10.1038/s41598-020-63523-5

Duman-Scheel, M. (2019). *Saccharomyces cerevisiae* (Baker's Yeast) as an interfering RNA expression and delivery system. *Curr. Drug Targets* 20 (9), 942–952. doi:10.2174/1389450120666181126123538

Dumke, R., de la Cruz Barron, M., Oertel, R., Helm, B., Kallies, R., Berendonk, T. U., et al. (2021). Evaluation of two methods to concentrate SARS-CoV-2 from untreated wastewater. *Pathogens* 10 (2), 195. doi:10.3390/pathogens10020195

Duong-Ly, K. C., and Gabelli, S. B. (2014). Salting out of proteins using ammonium sulfate precipitation. *Methods Enzymol.* 541, 85–94. doi:10.1016/B978-0-12-420119-4.00007-0

Echave, J., Fraga-Corral, M., Garcia-Perez, P., Popovic-Djordjevic, J., H Avdovic, E., Radulovic, M., et al. (2021). Seaweed protein hydrolysates and bioactive peptides: Extraction, purification, and applications. *Mar. Drugs* 19 (9), 500. doi:10.3390/md19090500

Emami, P., Motevalian, S. P., Pepin, E., and Zydney, A. L. (2019). Purification of a conjugated polysaccharide vaccine using tangential flow diafiltration. *Biotechnol. Bioeng.* 116 (3), 591–597. doi:10.1002/bit.26867

Evert, K., and Irwin, H. (2011). Ion exchange chromatography. *Methods Biochem. Anal.* 54, 93–133. doi:10.1002/9780470939932.ch4

Fox, J. D., Kapust, R. B., and Waugh, D. S. (2001). Single amino acid substitutions on the surface of *Escherichia coli* maltose-binding protein can have a profound impact on the solubility of fusion proteins. *Protein Sci.* 10 (3), 622–630. doi:10.1110/ps.45201

Freidel, C., Kaloyanova, S., and Peneva, K. (2016). Chemical tags for site-specific fluorescent labeling of biomolecules. *Amino Acids* 48 (6), 1357–1372. doi:10.1007/s00726-016-2204-5

Freitas, A. I., Domingues, L., and Aguiar, T. Q. (2022). Tag-mediated single-step purification and immobilization of recombinant proteins toward protein-engineered advanced materials. *J. Adv. Res.* 36, 249–264. doi:10.1016/j.jare.2021.06.010

Fu, J., Jia, Q. Q., Liang, P. D., Wang, S. S., Zhou, H. X., Zhang, L. Y., et al. (2021). Targeting and covalently immobilizing the EGFR through SNAP-tag technology for screening drug leads. *Anal. Chem.* 93 (34), 11719–11728. doi:10.1021/acs.analchem.1c01664

Fukushita, M., Watanabe, N., Noh, J. Y., Yoshihara, A., Matsumoto, M., Suzuki, N., et al. (2021). A case of macro-TSH consisting of IgA-bound TSH. *Endocr. J.* 68 (10), 1241–1246. doi:10.1507/endocrj.EJ21-0021

Gao, P., Li, J., Li, Z., Hao, J., and Zan, L. (2016). Establishment and application of milk fingerprint by gel filtration chromatography. *J. Dairy Sci.* 99 (12), 9493–9501. doi:10.3168/jds.2015-10655

Ge, S. K., Wu, R., Zhou, T. T., Liu, X., Zhu, J., and Zhang, X. Y. (2022). Specific anti-SARS-CoV-2 S1 IgY-scFv is a promising tool for recognition of the virus. *AMB Expr.* 12 (1), 18. doi:10.1186/s13568-022-01355-4

Geng, F., Xie, Y. X., Wang, J. Q., Li, S. G., Jin, Y. G., and Ma, M. H. (2019). Large-scale purification of ovalbumin using polyethylene glycol precipitation and isoelectric precipitation. *Poult. Sci.* 98 (3), 1545–1550. doi:10.3382/ps/pey402

Goh, J. B., and Ng, S. K. (2018). Impact of host cell line choice on glycan profile. *Crit. Rev. Biotechnol.* 38 (6), 851–867. doi:10.1080/07388551.2017.1416577

Gonzalez-Lamothe, R., Mitchell, G., Gattuso, M., Diarra, M. S., Malouin, F., and Bouarab, K. (2009). Plant antimicrobial agents and their effects on plant and human pathogens. *Int. J. Mol. Sci.* 10 (8), 3400–3419. doi:10.3390/ijms10083400

Gopal, G. J., and Kumar, A. (2013). Strategies for the production of recombinant protein in *Escherichia coli*. *Protein J.* 32 (6), 419–425. doi:10.1007/s10930-013-9502-5

Greenfield, E. A., DeCaprio, J., and Brahmandam, M. (2020). Preparing GST-His or MBP-fusion proteins from bacteria. *Cold Spring Harb. Protoc.* 2020 (9), 100024. doi:10.1101/pdb.prot100024

Guo, Y., Yu, M., Jing, N., and Zhang, S. (2018). Production of soluble bioactive mouse leukemia inhibitory factor from *Escherichia coli* using MBP tag. *Protein Expr. Purif.* 150, 86–91. doi:10.1016/j.pep.2018.05.006

Gupta, J., Kaul, S., Srivastava, A., Kaushik, N., Ghosh, S., Sharma, C., et al. (2020). Expression, purification and characterization of the hepatitis E virus like-particles in the *Pichia pastoris*. *Front. Microbiol.* 11, 141. doi:10.3389/fmicb.2020.00141

Gupta, S. K., Dang, A. K., Smita, M., Dwivedi, S., and Shukla, P. (2019). Chapter 11-Effectual bioprocess development for protein production. *Appl. Microbiol. Bioeng.* 203–227. doi:10.1016/b978-0-12-815407-6.00011-3

Gupta, S. K., and Shukla, P. (2016). Advanced technologies for improved expression of recombinant proteins in bacteria: perspectives and applications. *Crit. Rev. Biotechnol.* 36 (6), 1089–1098. doi:10.3109/07388551.2015.1084264

Hammarstrom, M., Hellgren, N., van Den Berg, S., Berglund, H., a220.1(angl.)-260326.3(0042)(M)

- Hanke, M., Hansen, N., Chen, R. P., Grundmeier, G., Fahmy, K., and Keller, A. (2022). Salting-out of DNA origami nanostructures by ammonium sulfate. *Int. J. Mol. Sci.* 23 (5), 2817. doi:10.3390/ijms23052817
- Hara, K. Y., Yagi, S., Hirono-Hara, Y., and Kikukawa, H. (2021). A method of solubilizing and concentrating astaxanthin and other carotenoids. *Mar. Drugs* 19 (8), 462. doi:10.3390/md19080462
- Hardy, D., Bill, R. M., Jawhari, A., and Rothnie, A. J. (2019). Functional expression of multidrug resistance protein 4 MRP4/ABCC4. *SLAS Discov.* 24 (10), 1000–1008. doi:10.1177/2472555219867070
- Hazim Abdul Hameed, D., and Hussein Ali, E. (2021). Extraction and purification of extracellular L-glutamate oxidase from streptomyces. *Arch. Razi Inst.* 76 (4), 769–779. doi:10.22092/ari.2021.355928.1738
- Hefferon, K. (2017). Reconceptualizing cancer immunotherapy based on plant production systems. *Future Sci. OA* 3 (3), FSO217. doi:10.4155/fsoa-2017-0018
- Heib, T., Gross, C., Muller, M. L., Stegner, D., and Pleines, I. (2021). Isolation of murine bone marrow by centrifugation or flushing for the analysis of hematopoietic cells - a comparative study. *Platelets* 32 (5), 601–607. doi:10.1080/09537104.2020.1797323
- Heinrich, J., Drewniak, C., Neugebauer, E., Kellner, H., and Wiegert, T. (2019). The Yoaw signal peptide directs efficient secretion of different heterologous proteins fused to a StrepII-SUMO tag in *Bacillus subtilis*. *Microb. Cell Fact.* 18 (1), 31. doi:10.1186/s12934-019-1078-0
- Hillary, L. S., Farkas, K., Maher, K. H., Lucaci, A., Thorpe, J., Distaso, M. A., et al. (2021). Monitoring SARS-CoV-2 in municipal wastewater to evaluate the success of lockdown measures for controlling COVID-19 in the UK. *Metabolic engineering of *Saccharomyces cerevisiae*: A key cell factory platform for future biorefineries.* *Cell. Mol. Life Sci. Cell. Mol. Life Sci.* 20069 (16), 1172142671–1172142690. doi:10.1016/j.watres.2021.117214Hong10.1007/s00018-012-0945-1
- Hong, K. K., and Nielsen, J. (2012). Metabolic engineering of *Saccharomyces cerevisiae*: A key cell factory platform for future biorefineries. *Cell. Mol. Life Sci.* 69 (16), 2671–2690. doi:10.1007/s00018-012-0945-1
- Huang, L. P., Lu, Y. H., Wei, Y. W., Guo, L. J., Wu, H. L., Zhang, F. Y., et al. (2011). Construction and biological characterisation of recombinant porcine circovirus type 2 expressing the V5 epitope tag. *Virus Res.* 161 (2), 115–123. doi:10.1016/j.virusres.2011.05.015
- Hunt, I. (2005). From gene to protein: A review of new and enabling technologies for multi-parallel protein expression. *Protein Expr. Purif.* 40 (1), 1–22. doi:10.1016/j.pep.2004.10.018
- Imagawa, T., Ito, M., Matsuda, M., Nakashima, K., Tokunaga, Y., Ohta, I., et al. (2021). Virus-like particles with FLAG-tagged envelope protein as a tetravalent dengue vaccine candidate. *Sci. Rep.* 11 (1), 17542. doi:10.1038/s41598-021-97038-4
- Ing-Simmons, E., and Vaquerizas, J. M. (2019). Visualising three-dimensional genome organisation in two dimensions. *Development* 146 (19), dev177162. doi:10.1242/dev.177162
- Ivanov, I. G., Saraffova, A., Alexandrova, R., and AbouHaidar, M. G. (1993). Expression of human alpha 1 interferon genes in vectors containing tandemly located promoters recognized by two different RNA polymerases (*Escherichia coli* and T7). *FEMS Microbiol. Lett.* 108 (2), 231–236. doi:10.1111/j.1574-6968.1993.tb06104.x
- Jhamb, K., and Sahoo, D. K. (2012). Production of soluble recombinant proteins in *Escherichia coli*: Effects of process conditions and chaperone co-expression on cell growth and production of xylanase. *Bioresour. Technol.* 123, 135–143. doi:10.1016/j.biortech.2012.07.011
- Jiang, N., Ding, X., and Lu, Y. (2021). Development of a robust *Escherichia coli*-based cell-free protein synthesis application platform. *Biochem. Eng. J.* 165, 107830. doi:10.1016/j.bej.2020.107830
- Jing, S. Y., Shi, C., Leong, H. Y., Yuan, J. J., Gao, D., Wang, H. B., et al. (2021). A novel twin-column continuous chromatography approach for separation and enrichment of monoclonal antibody charge variants. *Eng. Life Sci.* 21 (6), 382–391. doi:10.1002/elsc.202000094
- Johar, S. S., and Talbert, J. N. (2017). Strep-tag II fusion technology for the modification and immobilization of lipase B from *Candida Antarctica* (CALB). *J. Genet. Eng. Biotechnol.* 15 (2), 359–367. doi:10.1016/j.jgeb.2017.06.011
- Juturu, V., and Wu, J. C. (2018). Heterologous protein expression in *Pichia pastoris*: Latest research progress and applications. *ChemBiochem* 19 (1), 7–21. doi:10.1002/cbic.201700460
- Karami, M., Shahraky, M. K., Ranjbar, M., Tabandeh, F., Morshedi, D., and Aminzade, S. (2021). Preparation, purification, and characterization of low-molecular-weight hyaluronic acid. *Biotechnol. Lett.* 43 (1), 133–142. doi:10.1007/s10529-020-03035-4
- Karbalaei, M., Rezaee, S. A., and Farsiani, H. (2020). *Pichia pastoris*: A highly successful expression system for optimal synthesis of heterologous proteins. *J. Cell. Physiol.* 235 (9), 5867–5881. doi:10.1002/jcp.29583
- Kaur, J., Kumar, A., and Kaur, J. (2018). Strategies for optimization of heterologous protein expression in *E. coli*: Roadblocks and reinforcements. *Int. J. Biol. Macromol.* 106, 803–822. doi:10.1016/j.ijbiomac.2017.08.080
- Kawaguchi, N., Ogawa, H., and Date, K. (2019). Preparation and purification of active recombinant human pancreatic lipase in *Escherichia coli*. *Bio. Protoc.* 9 (13), e3286. doi:10.21769/BioProtoc.3286
- Keshavarz, R., Babaeipour, V., Mohammadpour-Aghdam, M., and Deldar, A. A. (2021). Overexpression, overproduction, purification, and characterization of rhGH in *Escherichia coli*. *Biotechnol. Appl. Biochem.* 68 (1), 122–135. doi:10.1002/bab.1902
- Kesidis, A., Depping, P., Lodé, A., Vaitsoyopoulou, A., Bill, M. R., Goddard, D. A., et al. (2020). Expression of eukaryotic membrane proteins in eukaryotic and prokaryotic hosts. *Methods* 180, 3–18. doi:10.1016/j.ymeth.2020.06.006
- Khazi, M. I., Demirel, Z., Liaqat, F., and Dalay, M. C. (2020). Analytical grade purification of phycocyanin from *Cyanobacteria*. *Methods Mol. Biol.* 1980, 173–179. doi:10.1007/7651_2018_202
- Kielkopf, C. L., Bauer, W., and Urbatsch, I. L. (2021). Preparation of cell extracts for purification of proteins expressed in *Pichia pastoris*. *Cold Spring Harb. Protoc.* 2021 (1), pdb.prot102186. doi:10.1101/pdb.prot102186
- Kim, D. S., Kim, S. W., Song, J. M., Kim, S. Y., and Kwon, K. C. (2019). A new prokaryotic expression vector for the expression of antimicrobial peptide abaecin using SUMO fusion tag. *BMC Biotechnol.* 19 (1), 13. doi:10.1186/s12896-019-0506-x
- Kim, H. J., Lee, J. H., Lee, K. B., Shin, J. W., Kwon, M. A., Lee, S., et al. (2021). Transglutaminase 2 crosslinks the glutathione S-transferase tag, impeding protein-protein interactions of the fused protein. *Exp. Mol. Med.* 53 (1), 115–124. doi:10.1038/s12276-020-00549-9
- Kimia, Z., Hosseini, S. N., Ashraf Talesh, S. S., Khatami, M., Kavianpour, A., and Javidanbardan, A. (2019). A novel application of ion exchange chromatography in recombinant Hepatitis B vaccine downstream processing: Improving recombinant HBsAg homogeneity by removing associated aggregates. *J. Chromatogr. B* 1113, 20–29. doi:10.1016/j.jchromb.2019.03.009
- Kinrade, B., Davies, P. L., and Vance, T. D. R. (2020). Bacterial sugar-binding protein as a one-step affinity purification tag on dextran-containing resins. *Protein Expr. Purif.* 168, 105564. doi:10.1016/j.pep.2019.105564
- Kou, G., Shi, S., Wang, H., Tan, M., Xue, J., Zhang, D., et al. (2007). Preparation and characterization of recombinant protein ScFv(CD11c)-TRP2 for tumor therapy from inclusion bodies in *Escherichia coli*. *Protein Expr. Purif.* 52 (1), 131–138. doi:10.1016/j.pep.2006.08.007
- Labrou, N. E. (2021). Protein purification technologies. *Methods Mol. Biol.* 2178, 3–10. doi:10.1007/978-1-0716-0775-6_1
- Labus, K., Wiśniewski, Ł., Cieńska, M., and Bryjak, J. (2020). Effectivity of tyrosinase purification by membrane techniques versus fractionation by salting out. *Chem. Pap.* 74, 2267–2275. doi:10.1007/s11696-020-01060-1
- Lacki, K. M., and Riske, F. J. (2020). Affinity chromatography: An enabling technology for large-scale bioprocessing. *Biotechnol. J.* 15 (1), e1800397. doi:10.1002/biot.201800397
- Lahue, C., Madden, A. A., Dunn, R. R., and Smukowski Heil, C. (2020). History and domestication of *Saccharomyces cerevisiae* in bread baking. *Front. Genet.* 11, 584718. doi:10.3389/fgene.2020.584718
- Lemaitre, R. P., Bogdanova, A., Borgonovo, B., Woodruff, J. B., and Drechsel, D. N. (2019). FlexiBAC: A versatile, open-source baculovirus vector system for protein expression, secretion, and proteolytic processing. *BMC Biotechnol.* 19 (1), 20. doi:10.1186/s12896-019-0512-z
- Li, S., Jendresen, C. B., Grunberger, A., Ronda, C., Jensen, S. I., Noack, S., et al. (2016). Enhanced protein and biochemical production using CRISPRi-based growth switches. *Metab. Eng.* 38, 274–284. doi:10.1016/j.ymben.2016.09.003
- Li, Y. J., Zeng, W. H., Li, Y. L., Fan, W. R., Ma, H., Fan, X. J., et al. (2019). Structure determination of the CAMP factor of *Streptococcus agalactiae* with the aid of an MBP tag and insights into membrane-surface attachment. *Acta Crystallogr. D. Struct. Biol.* 75 (Pt 8), 772–781. doi:10.1107/S205979831901057X
- Li, Y., Zeng, Q. H., Liu, G., Peng, Z., Wang, Y., Zhu, Y., et al. (2021). Effects of ultrasound-assisted basic electrolyzed water (BEW) extraction on structural and functional properties of Antarctic krill (*Euphausia superba*) proteins. *Ultrason. Sonochem.* 71, 105364. doi:10.1016/j.ultsonch.2020.105364
- Li, Z., Huang, X. H., Tang, Q. Y., Ma, M. H., Jin, Y. G., and Sheng, L. (2022a). Functional properties and extraction techniques of chicken egg white proteins. *Foods* 11 (16), 2434. doi:10.3390/foods11162434

- Li, Z. J., Zhang, Z. X., Xu, Y., Shi, T. Q., Ye, C., Sun, X. M., et al. (2022b). CRISPR-based construction of a BL21 (DE3)-derived variant strain library to rapidly improve recombinant protein production. *ACS Synth. Biol.* 11 (1), 343–352. doi:10.1021/acssynbio.1c00463
- Liang, X. N., Qian, G. L., Sun, J., Yang, M., Shi, X. Y., Yang, H., et al. (2021). Evaluation of antigenicity and nutritional properties of enzymatically hydrolyzed cow milk. *Sci. Rep.* 11 (1), 18623. doi:10.1038/s41598-021-98136-z
- Lichty, J. J., Malecki, J. L., Agnew, H. D., Michelson-Horowitz, D. J., and Tan, S. (2005). Comparison of affinity tags for protein purification. *Protein Expr. Purif.* 41 (1), 98–105. doi:10.1016/j.pep.2005.01.019
- Lin, Y., Liu, Y., Li, J. J., Zhao, Y., He, Q. Z., Han, W. J., et al. (2010). Evaluation and optimization of removal of an acid-insoluble surfactant for shotgun analysis of membrane proteome. *Electrophoresis* 31 (16), 2705–2713. doi:10.1002/elps.201000161
- Liu, C., Hao, L. H., Chen, F. S., and Zhu, T. W. (2020a). The mechanism of extraction of peanut protein and oil bodies by enzymatic hydrolysis of the cell wall. *J. Oleo Sci.* 69 (11), 1467–1479. doi:10.5650/jos.ess20148
- Liu, H. N., Dong, W. H., Lin, Y., Zhang, Z. H., and Wang, T. Y. (2022). The effect of microRNA on the production of recombinant protein in CHO cells and its mechanism. *Front. Bioeng. Biotechnol.* 10, 832065. doi:10.3389/fbioe.2022.832065
- Liu, M., Chen, Z. Y., Huang, S. S., Chu, S. H., Issaro, N., Tian, H., et al. (2018). Effective protection on acute liver injury by halo tag-flanked recombinant fibroblast growth factor 7. *Biotechnol. J.* 13 (7), e1700411. doi:10.1002/biot.201700411
- Liu, M., Wang, B., Wang, F., Yang, Z., Gao, D., Zhang, C. Y., et al. (2019). Soluble expression of single-chain variable fragment (scFv) in *Escherichia coli* using superfolder green fluorescent protein as fusion partner. *Appl. Microbiol. Biotechnol.* 103 (15), 6071–6079. doi:10.1007/s00253-019-09925-6
- Liu, X. B., Liu, M., Tao, X. Y., Zhang, Z. X., Wang, F. Q., and Wei, D. Z. (2015). Metabolic engineering of *Pichia pastoris* for the production of dammareniol-II. *J. Biotechnol.* 216, 47–55. doi:10.1016/j.jbiotec.2015.10.005
- Loh, H. S., Green, B. J., and Yusibov, V. (2017). Using transgenic plants and modified plant viruses for the development of treatments for human diseases. *Curr. Opin. Virol.* 26, 81–89. doi:10.1016/j.coviro.2017.07.019
- Lojewska, E., Kowalczyk, T., Olejniczak, S., and Sakowicz, T. (2016). Extraction and purification methods in downstream processing of plant-based recombinant proteins. *Protein Expr. Purif.* 120, 110–117. doi:10.1016/j.pep.2015.12.018
- Lomonosoff, G. P., and D'Aoust, M. A. (2016). Plant-produced biopharmaceuticals: A case of technical developments driving clinical deployment. *Science* 353 (6305), 1237–1240. doi:10.1126/science.aaf6638
- London, A. S., Mackay, K., Lihon, M., He, Y., and Alabi, B. R. (2014). Gel filtration chromatography as a method for removing bacterial endotoxin from antibody preparations. *Biotechnol. Prog.* 30 (6), 1497–1501. doi:10.1002/btpr.1961
- Lu, J., Zhao, Y., and Zhang, J. (2018). High-level expression of *Aerococcus viridans* pyruvate oxidase in *Escherichia coli* by optimization of vectors and induction conditions. *Lett. Appl. Microbiol.* 67 (3), 262–269. doi:10.1111/lam.13013
- Ma, Q. Y., Liang, M. M., Limjunyawong, N., Dan, Y., Xing, J. C., Li, J. M., et al. (2020). Osteoclast-derived apoptotic bodies show extended biological effects of parental cell in promoting bone defect healing. *Theranostics* 10 (15), 6825–6838. doi:10.7150/thno.45170
- Ma, X., Wang, C. Y., Guo, H. X., Wang, Z. F., Sun, N., Huo, P. F., et al. (2022). Novel dopamine-modified cellulose acetate ultrafiltration membranes with improved separation and antifouling performances. *J. Mat. Sci.* 57 (11), 6474–6486. doi:10.1007/s10853-022-07024-y
- Mamat, U., Wilke, K., Bramhill, D., Schromm, A. B., Lindner, B., Kohl, T. A., et al. (2015). Detoxifying *Escherichia coli* for endotoxin-free production of recombinant proteins. *Microb. Cell Fact.* 14, 57. doi:10.1186/s12934-015-0241-5
- Masuda, Y., Tsuda, M., Hashikawa-Muto, C., Takahashi, Y., Nonaka, K., and Wakamatsu, K. (2019). Cation exchange chromatography performed in overloaded mode is effective in removing viruses during the manufacturing of monoclonal antibodies. *Biotechnol. Prog.* 35 (5), e2858. doi:10.1002/btpr.2858
- Matulis, D. (2016). Selective precipitation of proteins. *Curr. Protoc. Protein Sci.* 83, 41–44. doi:10.1002/0471140864.ps0405s83.5.37
- Mazi, B. G., Hamamci, H., Ogrydziak, D. M., and Dungan, S. R. (2016). Single-step partial purification of intracellular β -galactosidase from *Kluyveromyces lactis* using microemulsion droplets. *Appl. Biochem. Biotechnol.* 180 (5), 1000–1015. doi:10.1007/s12010-016-2148-y
- McElwain, L., Phair, K., Kealey, C., and Brady, D. (2022). Current trends in biopharmaceuticals production in *Escherichia coli*. *Biotechnol. Lett.* 44 (8), 917–931. doi:10.1007/s10529-022-03276-5
- Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., and Brouwer, A. (2020). Presence of SARS-Coronavirus-2 RNA in sewage and correlation with reported COVID-19 prevalence in the early stage of the epidemic in The Netherlands. *Environ. Sci. Technol. Lett.* 7 (7), 511–516. doi:10.1021/acsclett.0c00357
- Menegotto, A. L. L., Fernandes, I. A., Steffens, J., and Valduga, E. (2021). Protein purification of *Arthrospira platensis* using aqueous two-phase system composed of polyethylene glycol and potassium phosphate/sodium citrate. *J. Appl. Phycol.* 34, 311–320. doi:10.1007/s10811-021-02652-4
- Mi, S. Y., Tang, Y. J., Dari, G., Shi, Y. Y., Zhang, J. N., Zhang, H. L., et al. (2021). Transcriptome sequencing analysis for the identification of stable lncRNAs associated with bovine *Staphylococcus aureus mastitis*. *J. Anim. Sci. Biotechnol.* 12 (1), 120. doi:10.1186/s40104-021-00639-2
- Mishra, V. (2020). Affinity tags for protein purification. *Curr. Protein Pept. Sci.* 21 (8), 821–830. doi:10.2174/1389203721666200606220109
- Mohammadi, M., Taheri, R. A., Bemani, P., Hashemzadeh, M. S., Farnoosh, G., and Amini, R. (2022). Utilization of SUMO tag and freeze-thawing method for high level expression and solubilization of recombinant human angiotensin-converting enzyme 2 (rhACE2) protein in *E. coli*. *Protein Expr. Lett.* 29 (7), 605–610. doi:10.2174/0929866529666220715101357
- Moinpour, M., Fracassi, A., Brea, R. J., Salvador-Castell, M., Pandey, S., Edwards, M. M., et al. (2022). Controlling protein enrichment in lipid sponge phase droplets using SNAP-tag bioconjugation. *ChemBiochem.* 23 (5), e202100624. doi:10.1002/cbic.202100624
- Muhmed, S. A., Nor, N. A. M., Jaafar, J., Ismail, A. F., Othman, M. H. D., Rahman, M. A., et al. (2019). Emerging chitosan and cellulose green materials for ion exchange membrane fuel cell: A review. *Energy Ecol. Environ.* 5, 85–107. doi:10.1007/s40974-019-00127-4
- Mukherjee, S., Ura, M., Hoey, R. J., and Kossiakoff, A. A. (2015). A new versatile immobilization tag based on the ultra high affinity and reversibility of the calmodulin-calmodulin binding peptide interaction. *J. Mol. Biol.* 427 (16), 2707–2725. doi:10.1016/j.jmb.2015.06.018
- Muller, C., Zhang, L., Zipfel, S., Topitsch, A., Lutz, M., Eckert, J., et al. (2022). Molecular interplay of an assembly machinery for nitrous oxide reductase. *Nature* 608 (7923), 626–631. doi:10.1038/s41586-022-05015-2
- Muthunayake, N. S., Al-Husini, N., and Schrader, J. M. (2020). Differential centrifugation to enrich bacterial ribonucleoprotein bodies (BR bodies) from *Caulobacter crescentus*. *Star. Protoc.* 1 (3), 100205. doi:10.1016/j.xpro.2020.100205
- Nagai, H., Masuda, A., Toya, Y., Matsuda, F., and Shimizu, H. (2018). Metabolic engineering of mevalonate-producing *Escherichia coli* strains based on thermodynamic analysis. *Metab. Eng.* 47, 1–9. doi:10.1016/j.ymben.2018.02.012
- Nallamsetty, S., and Waugh, D. S. (2006). Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners. *Protein Expr. Purif.* 45 (1), 175–182. doi:10.1016/j.pep.2005.06.012
- Nasukawa, T., Uchiyama, J., Taharaguchi, S., Ota, S., Ujihara, T., Matsuzaki, S., et al. (2017). Virus purification by CsCl density gradient using general centrifugation. *Arch. Virol.* 162 (11), 3523–3528. doi:10.1007/s00705-017-3513-z
- Nigi, I., Fairall, L., and Schwabe, J. W. R. (2017). Expression and purification of protein complexes suitable for structural studies using mammalian HEK 293F cells. *Curr. Protoc. Protein Sci.* 90 (5), 1–15. doi:10.1002/cpps.4428.16
- Nickerson, J. L., Baghalabadi, V., Dang, Z. H., Miller, V. A., Little, S. L., and Doucette, A. A. (2022). Organic solvent-based protein precipitation for robust proteome purification ahead of mass spectrometry. *J. Vis. Exp.* 180, e63503. doi:10.3791/63503
- Nishida, T., Kubota, S., and Takigawa, M. (2017). Production of recombinant CCN2 protein by mammalian cells. *Methods Mol. Biol.* 1489, 95–105. doi:10.1007/978-1-4939-6430-7_10
- Nosaki, S., Terada, T., Nakamura, A., Hirabayashi, K., Xu, Y. Q., Bui, T. B. C., et al. (2021). Highlighting the potential utility of MBP crystallization chaperone for *Arabidopsis* BIL1/BZR1 transcription factor-DNA complex. *Sci. Rep.* 11 (1), 3879. doi:10.1038/s41598-021-83532-2
- O'Connor, B. F., and Cummins, P. M. (2017). Hydrophobic interaction chromatography. *Methods Mol. Biol.* 1485, 355–363. doi:10.1007/978-1-4939-6412-3_18
- Oganesyan, N., Ankoudinova, I., Kim, S. H., and Kim, R. (2007). Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallization. *Protein Expr. Purif.* 52 (2), 280–285. doi:10.1016/j.pep.2006.09.015
- Owczarek, B., Gerszberg, A., and Hnatuszko-Konka, K. (2019). A brief reminder of systems of production and chromatography-based recovery of recombinant protein biopharmaceuticals. *Biomed. Res. Int.* 2019, 1–13. doi:10.1155/2019/4216060

Thammasena, R., Fu, C. W., Liu, J. H., and Liu, D. C. (2020). Evaluation of nutrient content, physicochemical and functional properties of desalted duck egg white by ultrafiltration as desalination. *Anim. Sci. J.* 91 (1), e13339. doi:10.1111/asj.13339

Tomás-Gamisans, M., Andrade, C. C. P., Maresca, F., Monforte, S., Ferrer, P., and Albiol, J. (2020). Redox engineering by ectopic overexpression of NADH kinase in recombinant *Pichia pastoris* (*Komagataella phaffii*): Impact on cell physiology and recombinant production of secreted proteins. *Appl. Environ. Microbiol.* 86 (6), e02038–e02019. doi:10.1128/AEM.02038-19

Traore, E. S., Li, J. S., Chiura, T., Geng, J., Sachla, A. J., Yoshimoto, F., et al. (2021). Heme binding to HupZ with a C-terminal tag from group A *Streptococcus*. *Molecules* 26 (3), 549. doi:10.3390/molecules26030549

Tremante, E., Sibilio, L., Centola, F., Knutti, N., Holzapfel, G., Manni, I., et al. (2021). Toolbox: StrepTagged nanoassemblies of antibodydrugconjugates (ADC)

Zhang, Z. W., Wu, X. D., Cao, L. L., Zhong, Z. D., and Zhou, Y. (2016). Generation of glucagon-like peptide-2-expressing *Saccharomyces cerevisiae* and its improvement of the intestinal health of weaned rats. *Microb. Biotechnol.* 9 (6), 846–857. doi:10.1111/1751-7915.12412

Zhang, Z. X., Wang, Y. Z., Nong, F. T., Xu, Y., Ye, C., Gu, Y., et al. (2022). Developing a dynamic equilibrium system in *Escherichia coli* to improve the production of recombinant proteins. *Appl. Microbiol. Biotechnol.* 106 (18), 6125–6137. doi:10.1007/s00253-022-12145-0

Zhao, X. Y., Li, G. S., and Liang, S. F. (2013). Several affinity tags commonly used in chromatographic purification. *J. Anal. Methods Chem.* 2013, 581093. doi:10.1155/2013/581093

Zhou, H. H., Zhang, A. X., Zhang, Y., and Zhu, D. Y. (2012). Cloning, expression, and purification of a recombinant Tat-HA-NR2B9c peptide. *Protein Expr. Purif.* 85 (2), 239–245. doi:10.1016/j.pep.2012.08.011

Zhou, L. X., Liu, Z. Y., Xu, G. Y., Li, L. H., Xuan, K., Xu, Y., et al. (2020). Expression of melittin in fusion with GST in *Escherichia coli* and its purification as a pure peptide with good bacteriostatic efficacy. *ACS Omega* 5 (16), 9251–9258. doi:10.1021/acsomega.0c00085

Zielinski, M., Romanik-Chrusciewska, A., Mikiewicz, D., Lukasiewicz, N., Sokolowska, I., Antosik, J., et al. (2019). Expression and purification of recombinant human insulin from *E. coli* 20 strain *Protein Expr. Purif.* 157, 63–69. doi:10.1016/j.pep.2019.02.002