**ORIGINAL PAPER**

# Factors involved in heterologous expression of proteins in *E. coli* host

Mahin Pouresmaeil[1] · Shahnam Azizi-Dargahlou[1]

## Abstract

The production of recombinant proteins is one of the most significant achievements of biotechnology in the last century. These proteins are produced in the eukaryotic or prokaryotic heterologous hosts. By increasing the omics data especially related to different heterologous hosts as well as the presence of new amenable genetic engineering tools, we can artificially engineer heterologous hosts to produce recombinant proteins in sufficient quantities. Numerous recombinant proteins have been produced and applied in various industries, and the global recombinant proteins market size is expected to be cast to reach USD 2.4 billion by 2027. Therefore, identifying the weakness and strengths of heterologous hosts is critical to optimize the large-scale biosynthesis of recombinant proteins. *E. coli* is one of the popular hosts to produce recombinant proteins. Scientists reported some bottlenecks in this host, and due to the increasing demand for the production of recombinant proteins, there is an urgent need to improve this host. In this review, we first provide general information about the *E. coli* host and compare it with other hosts. In the next step, we describe the factors involved in the expression of the recombinant proteins in *E. coli*. Successful expression of recombinant proteins in *E. coli* requires a complete elucidation of these factors. Here, the characteristics of each factor will be fully described, and this information can help to improve the heterologous expression of recombinant proteins in *E. coli*.

**Keywords** Fusion tag · Inclusion body · Recombinant protein · Heterologous hosts · Recombinant DNA

## Introduction

A protein encoded in recombinant DNA and expressed in a heterologous host is a recombinant protein (Young et al. 2012). In 1973, Herbert Boyer and Stanley Cohen pioneered the use of recombinant DNA in cloning and gene expression processes. In the 1970s, the use of recombinant DNA expression for the production of recombinant proteins increased, and today the production of recombinant proteins has dramatically increased, and the techniques used in this science are available in all countries. Nowadays, some vaccines applied against COVID-19 disease are based on recombinant DNA technology (Francis et al. 2022; Ghiasi et al. 2021) and this technology will be a promising method to fight against cancers and incurable diseases.

Choosing a suitable heterologous host for the expression of heterologous genes is an important step. Researchers have used different prokaryotic and eukaryotic organisms as heterologous hosts, such as *Escherichia coli* (Rosano et al. 2014), *Bacillus subtillis* (Schumann 2007), *Pseudomonas fluorescens* (Zhang et al. 2018), filamentous fungi (Ward 2012), yeasts (Çelik and Çalık 2012), insect cells (Ikonomou et al. 2003), mammalian cells (Lai et al. 2013), and plant cells (Hellwig et al. 2004). Each of these hosts has its own merits and demerits (Shanmugaraj et al. 2020). The selection of suitable expression systems depends on many criteria such as the facilities of the laboratory and local expertise (Ma et al. 2013), final production cost (Gifre et al. 2017), the characteristics of the recombinant protein (Demain et al. 2009), the intrinsic feature of the expression host (Palomares et al. 2004), the source of the heterologous gene (Deng et al. 2017), easy optimization of heterologous expression (Kaur et al. 2018), regulatory considerations (Hellwig et al. 2004), the purpose of recombinant protein production (Puetz and Wurm 2019), etc. Therefore, before the production of recombinant protein in the heterologous host, the ideal host for the expression of the gene of interest should be selected

---

✉ Shahnam Azizi-Dargahlou
   shahnambio@gmail.com

1   Agricultural Biotechnology, Department of Biotechnology, Azarbaijan Shahid Madani University, Tabriz, Iran

according to the purpose of the experiment and the characteristics of the host (Aune 2008; Khan 2013).

According to the marketsandmarkets website https://www.marketsandmarkets.com/ report (Karyolaimos et al. 2019), the production of recombinant proteins in mammalian cells had the highest income among other systems in 2021, and the income from this system was equal to 41.70% of total revenue from recombinant proteins. Also, expression based on the bacterial systems was in second place in terms of income (Fig. 1). However, disadvantages such as high cost and relatively low protein yield are more common in mammalian expression systems than in prokaryotic expression systems (Portolano et al. 2014). Prokaryotic expression systems are suitable hosts for the expression of recombinant proteins, and these organisms are known as the affordable and simplest host. The *E. coli* expression system is the most extensively used bacterial expression system and is the first choice for laboratory research (Koehn and Hunt 2009). Considering the vital importance of recombinant protein production in this era, this review aims to study the optimal expression conditions for recombinant proteins and the effect of various factors on recombinant protein expression in *E. coli*. Here, we also review the bottlenecks and workarounds that researchers have reported for this host.

## Recombinant protein expression in *E. coli*

*E. coli* is one of the important hosts for the production of recombinant proteins. It was first isolated by Theodor Escherich in the late nineteenth century. It is a gram-negative rod-shaped proteobacterium and mammalian intestinal pathogen (Overman et al. 2014). The genetics of *E. coli* has been studied more than any other gram-negative bacteria. Due to recent advances in the understanding of transcription and translation processes, the mechanism of protein folding,

and the availability of improved genetic tools, the use of *E. coli* to express recombinant proteins has increased (Demain and Vaishnav 2009).

Bacteria are commonly used to produce recombinant proteins less than 30 kDa (Demain and Vaishnav 2009) and production of larger sizes is problematic, but the production of far larger proteins has also been reported in this host. Chloroplast is a promising heterologous host for the production of recombinant proteins in the future. Comparing the potential of *E. coli* and plant chloroplast to produce recombinant protein can be a good comparison. In the chloroplast, the integration of heterologous genes into the genome is achieved by means of homologous recombination (Burnett et al. 2020), and polycistronic mRNA is produced in this organelle. Introduced and integrated heterologous genes into the chloroplast genome can have 10,000 polyploidy in a young leaf cell (Verma and Daniell 2007), which is unmatched by any other expression system. Also, some plant-derived recombinant therapeutic proteins can be used as oral drugs (Kurup and Thomas 2020) without the need for a purification process, which sharply decreases the cost of the final product. In addition, plants can grow in the presence of soil, water, and light, which is cost-effective compared to the growth media of other expression systems (Burnett and Burnett 2020). This expression system has some drawbacks. One of the main problems is the process of integration of heterologous genes into the chloroplast genome, which is mostly done using a gene gun device (Burnett and Burnett 2020), which is expensive, as well as optimization of chloroplast transformation is a time-consuming and ultimately production of the homoplasmic transplastomic cell lines from produced primary heteroplasmic cells lines is necessary. Moreover, the metabolic burden is another reported disadvantage of this host which can lead to retardation of plant growth followed by biomass reduction and low recombinant protein yield (Oey et al. 2009).

Fig. 1 Comparison of heterologous hosts by income: In total, revenue from recombinant proteins which are produced in all heterologous hosts is very considerable as in 2020 and 2021 this amount is 358.4 and 556.6 million dollars, respectively. Also, the bacteria-based expression system is the second most profitable after mammalian cells (Karyolaimos et al. 2019)
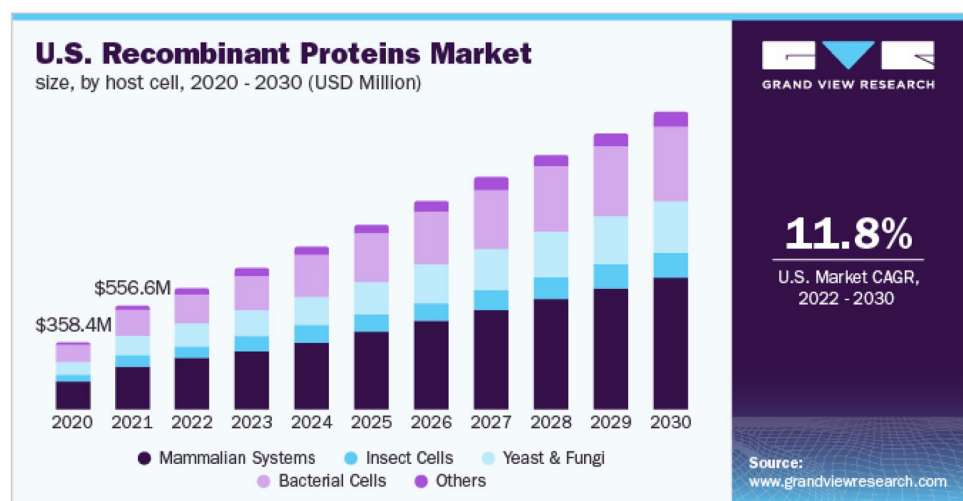
Easy manipulation of the genome of unicellular hosts such as bacteria is their main advantage over multicellular organisms. In addition, unlike multicellular organisms, the transformation of bacteria is simple, highly efficient, and cost-effective (Baneyx 1999). In general, the cost of production, the complexity of the work path, the different patterns of post-translation modifications, the high protease property of hosts, the inability to secrete recombinant proteins, the contamination with pathogens and toxins, the solubility of recombinant proteins, and the unrelated codon preference can be a major bottleneck in the expression of recombinant proteins in various sorts of hosts (Bock 2015). To improve recombinant protein expression in the *E. coli* system we can apply some strategies such as using different promoters to regulate the over-expression of genes, utilizing different strains of *E. coli*, reducing the growth temperature, secretion of protein to extracellular space, changing the culture medium, using tags, and simultaneous expression of recombinant proteins with chaperones or foldases to help the correct folding of the proteins (Baeshen et al. 2015; Wang et al. 2007).

## Essential components of *E. coli*-based expression systems

Several factors are involved in the expression of recombinant proteins in *E. coli*. To optimize and increase the expression of recombinant proteins, some of these factors can be applied or modified, and here we describe each of these factors.

## Vector

To call a DNA molecule as a vector, it must demonstrate specific characteristics. Each vector must be able to independently replicate inside the host cell so that multiple copies of the recombinant DNA can be produced and transferred to the daughter cells. Also, a cloning vector should be relatively small (less than 10 kb in size) because larger molecules are susceptible to breaking during extraction and are more difficult to be manipulated. Two types of DNA molecules with these characteristics can be found in bacterial cells: plasmids and bacteriophage chromosomes (Brown 2020). For the expression of recombinant proteins in *E. coli*, various expression plasmids and different strains have been introduced. In recent years, various expression plasmids have been designed to bypass various obstacles in the recombinant protein expression pathway. The most common expression plasmids have multiple vital elements such as the origin of replication, promoter, an affinity tag, a coding sequence for removal of tags, a multiple cloning site (MCS), a selectable marker gene (to screen recombinant bacteria), and a terminator sequence (Georgiou and Segatori 2005) (Fig. 2). Today, there are many expression plasmids, and the choice is made according to the purpose of the experiment (Rosano and Ceccarelli 2014). In most cases, a derivative of pBR322 plasmid is used as an expression plasmid in *E. coli* (Klumpp 2011). Elements of expression plasmids have a great influence on various aspects of recombinant proteins that are expressed in *E.coli*. Therefore, a complete and accurate understanding of these elements is essential. Here we state most of these elements.



**Fig. 2** Important elements of expression plasmid in bacteria. The distance between different elements is not to scale
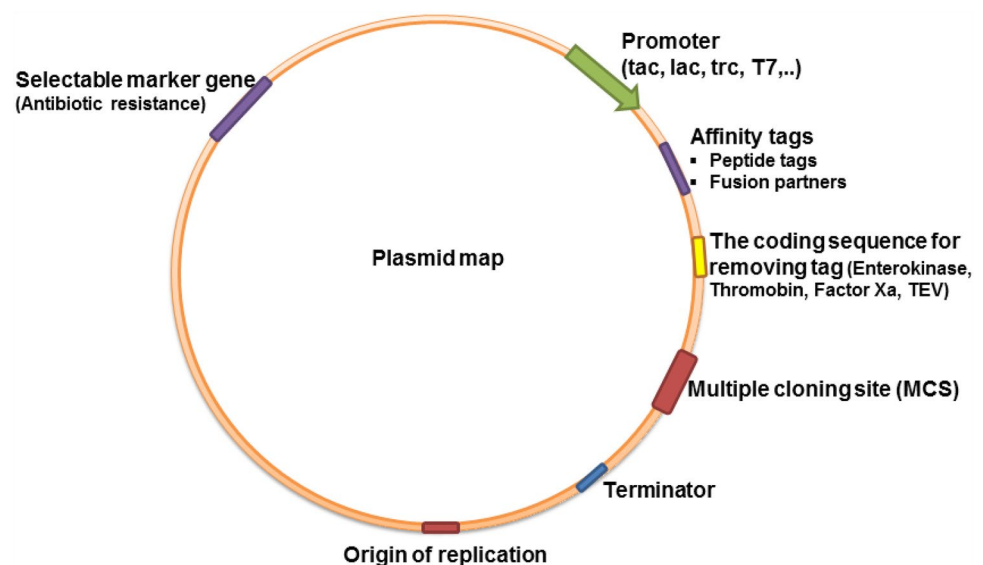
## Replicon

A replicon is a complete segment of DNA that can replicate independently through a single origin of replication. A replicon contains an origin of replication with its associated cis-control elements (Ekundayo and Bleichert 2019). One of the factors for selecting a vector for protein expression is the copy number of plasmids, which is controlled by the replicon. The copy number of plasmids that are commonly used to construct expression vectors in *E.coli* are shown in Table 1. It should be noted that increasing the copy number of a plasmid does not always mean increasing the efficiency of recombinant protein production (Lozano Terol et al. 2021). The relationship between recombinant protein production and plasmid copy number is a delicate balance, as it has been identified that the metabolic burden increases by 0.063% for each additional plasmid, and maintaining a balance between host growth and recombinant protein yield is necessary (Rouches et al. 2022). In addition to the copy number of plasmids, their compatibility with each other is another important factor to be considered. This compatibility is related to the replication mechanism of plasmids. Overall, plasmids that are closely related to each other are more competent to be incompatible (Thomas 2021). It is impossible to maintain two different plasmids that use the same mechanism for their replication, in the same cell. When two plasmids in the same group are incompatible, it means that they cannot replicate together in the same cell, but compete with each other to recruit replication factors (Velappan et al., 2007). For example, pMB1 and colE1 plasmids are in the same group and are incompatible (Gao 1986; Minton 1984; O'brien et al. 2006).

## Promoter

Promoter and enhancer elements are regulatory sequences. Enhancers are short nucleotide sequences that increase the transcription rate (Palstra and Grosveld 2012). Promoters control the binding of RNA polymerase to DNA, so a promoter controls when and where a gene is expressed in an organism (Blazeck and Alper 2013). An efficient expression vector requires a strong promoter with a high affinity for the RNA polymerase to increase the transcription of the heterologous gene [16]. A strong and adjustable promoter can control the transcription of the target gene in a certain period. These promoters are regulated using inhibitors or activators. A promoter used in an expression vector should have some specific features for the ideal expression of recombinant proteins. They must be strong and have a high affinity for polymerase and produce recombinant proteins at 10–30% or even more of the total cellular proteins, as well as promoter induction should be easy and affordable (Makrides 1996). The characteristics of some promoters used in *E. coli* are listed in Table 2. In general, promoter strength and plasmid copy number are important factors in the yield of recombinant proteins, and regulation of these two factors can bypass the problem of metabolic burden in *E.coli*, which reduces the yield of recombinant proteins (Lozano Terol et al. 2021).

## Fusion tags

Tags are molecules composed of amino acids chain which can have multiple effects on the target proteins (Lichty et al. 2005). Because of their critical function in the pathway of recombinant protein production, we discuss them thoroughly. Affinity tags are unique proteins/peptides which are added to the N- or C-terminus or both ends of recombinant proteins to primarily help the purification of target proteins. It is usually more useful to place tags at the N-terminal than the C-terminal. When tags are placed at the N-terminal, it is easier to purify the protein and also causes effective translation initiation (Costa et al. 2014). Several factors should be considered when designing and placing tags in the N or C terminus of the target proteins. For example, the added tag should not disturb the function of the target protein, also during protein purification these tags should not be out of reach (Booth et al. 2018; Wingfield 2015). In general, tags have numerous important functions in the expression of heterologous proteins. These functions include helping to optimize protein expression, increasing the solubility and

**Table 1** Copy number of plasmids that are generally utilized for constructing expression vectors in *E. coli*

| Plasmid | Replicon | Copy number | An element that controls copy number of plasmid | Mechanism of controlling plasmid copy number | References |
|---------|----------|-------------|------------------------------------------------|---------------------------------------------|------------|
| pBR322 | pMB1 | 15–60 | RNAI (different from RNAi) | Controls the processing of pre-RNAII into primer | (Parada and Marians 1991) |
| colE1 | colE1 | 15–20 | RNAI | Controls the processing of pre-RNAII into primer | (Camps 2010) |
| pUC | mod pMB1 | 500–700 | RNAI | Controls the processing of pre-RNAII into primer | (Chen et al. 2015) |
| pACYC | p15A | 18–22 | Directly repeated sequences (iterons) | Sequesters of the RepA protein | (Stüber and Bujard 1981) |
| pSC101 | pSC101 | <5 | Directly repeated sequences (iterons) | Sequesters of the RepA protein | (Ingmer et al. 2001) |

**Table 2** Common promoters used to drive the recombinant protein expression in *E. coli*

| Promoter | Inducer | Characteristic | References |
|---|---|---|---|
| Lac | IPTG | The binding site for RNA polymerase | (Giordano et al. 1989) |
| trc and tac | IPTG, thermal | Control and increase the expression level of genes of interest | (Brosius et al. 1985) |
| Trp | Indoleacrylic acid, Tryptophan starvation | High-level production of proteins on small and large scales | (Hu and Eftink 1993; Yang and Yanofsky 2005) |
| T7 | Lactose or IPTG | It is very popular for recombinant protein expression | (Rosano and Ceccarelli 2014) |
| Phage promoter pL | Thermal | A synthetic translation initiation region (Kopanic et al.), and convenient cloning sites | (Cheng and Patterson 1992) |
| araP$_{BAD}$ | Arabinose | Suitable for toxic protein expression | (Giacalone et al. 2006) |
| rhaPBAD | L-rhamnose | Regulated by two activators, RhaS and RhaR | (Wegerer et al. 2008) |
| proU | Osmolarity | Encodes a high-affinity glycine betaine transport system which is osmotically inducible | (Mellies et al. 1994) |
| recA | Nalidixic acid | RecA can stimulate the relaxation activity of topoisomerase | (Keller et al. 1990; Reckinger et al. 2007) |
| phoA | Phosphate starvation | It is an inducible promoter of the alkaline phosphatase gene (phoA) derived from *E. coli* | (Wang et al. 2005) |
| cadA | pH | A cadmium-sensitive promoter | (Yoon and Silver 1991) |
| Nar | Anaerobic conditions, presence of nitrate | A dissolved oxygen (DO)-dependent promoter | (Hwang et al. 2017) |

stability of proteins, helping protein purification, (Rosano and Ceccarelli 2014) and reducing the toxicity of recombinant proteins on host cells (Arnau et al. 2006; Costa et al. 2014). In addition, they can be used to confirm and quantify the expression of recombinant genes in heterologous hosts.

Affinity tags are in two forms, peptide tags (FLAG, poly-His, poly-Arg, Stre-taq, c-Myc, S-taq), which interfere less with protein of interest (POI) due to their small size (Bozarth et al. 2009). Sometimes they may cause negative effects on the tertiary structure of the POI (Khlebnikov and Keasling 2002). It has been discovered that poly-His tag can reduce the solubility of the POI (Woestenenk et al. 2004). The other one is in the form of larger proteins which, in addition to helping to purify the POIs, may also increase their solubility (Tochio et al. 2015). Tags that play a role in increasing protein solubility include maltose-binding protein (MBP), N-Utilization protein A, thioredoxin, glutathione s-transferase (GST), ubiquitin, Fh8, and SUMO (Costa et al. 2014; Rosano and Ceccarelli 2014). Recent studies show that the Fh8 fusion tag is the best solubility enhancer among the tags (Costa et al. 2014). In addition, various researchers have successfully applied several non-commercial and uncommon tags (Han et al. 2018; Morris et al. 2016; Nguyen et al. 2019; Ojima-Kato et al. 2017).

Removing fusion tags is usually necessary to achieve intact and active POIs. Enzymatic or chemical cleavage can be used to remove tags. To enzymatically remove tags, the coding sequence for tag removal is inserted into the expression vector that encodes a protease cleavage site. In choosing a suitable protease for removing tags, factors such as the number of unwanted amino acid residues remaining after cleavage, the ease of removing the protease after enzymatic cleavage, and the cost and characteristics of the protease should be considered. Note, the buffer required for protease activity should not affect the properties of the POI. These proteases include tobacco etch virus (TEV) protease, thrombin, enterokinase, and factor Xa (Blommel and Fox 2007; Jenny et al. 2003). In chemical cleavage, a series of chemical reagents are used to remove fusion tags. Cheapness is the advantage of this method, but the reaction conditions are harsh. One of the chemicals used to remove tags is Cyanogen bromide (CNBr). Methionine is an infrequent amino acid in the structure of proteins. CNBr is used to make cleavage at the C-terminal of the methionine residue. We can then use this reagent to release the target proteins from the tags, where the methionine is only present just before the N-terminus of target proteins (Andreev et al. 2010). The presence of methionine in the structure of tags is not important in this respect, but it must not be present in the structure of the target proteins except at their beginning. As proteins increase in size, the frequency of methionines increases throughout the protein sequence, then the application of this method to

smaller target proteins such as antimicrobial peptides will be more feasible. Recently, the cleavage of peptide bonds using metals, especially nickel or palladium, which could be used to cleave the tags, has been established (Andreev et al. 2010; Hwang et al. 2014).

In the above-mentioned affinity tags, we need some instruments and specific solutions to separate fusion proteins from the total proteins of hosts. Purification of recombinant proteins based on cost-effective physico-chemical methods does not require special tools such as chromatography to separate recombinant proteins from other cellular extracted substances (Wang et al. 2020). In these methods, purification is based on thermal, salinity, and pH changes or a combination of these factors. The elastin-like polypeptide (EST) (Gong et al. 2022; Pereira et al. 2021), four-helix bundle protein DAMP4 (Wang et al. 2017) and the cell surface protein B (CspB) (Nonaka et al. 2019) tags fall into this category. In the future, the introduction of new purification methods based on physico-chemical methods is highly felt. Overall, non-chromatographic-based purification methods plus non-enzymatic elimination of tags would be double cost-effective strategies for large-scale production of recombinant proteins. In one experiment, scientists used a non-chromatographic and self-cleavable ELP-intein tag for the purification of specific antimicrobial peptides (AMPs) (Camps) that heterologously expressed in *E. coli* strain BL21. In this tag, the ELP has been applied with the goal of promoting aggregation of target protein in the special condition followed by rapid and cost-effective isolation of AMP and the intein has been applied for releasing this tag from AMP (Sousa et al. 2016). In another study, researchers utilized a system called the TSGIT to solve the problems associated with the cleavages of tags (unwanted amino acid residues) and purifying intact recombinant proteins without any contamination with truncated recombinant proteins. In this system, three types of tags include purification, solubilization, and cleavable tags were added to both termini of the POI. Purification tags were two hexa-histidine tags in a row in the N terminal and two consecutive Strep tags in the C-terminus of the POI. The solubilization tag in both termini of the protein was thioredoxin. They used the SUMO cleavage tag in the N terminal and the intein cleavage tag in the C terminus. These tags released the POI without any unwanted extra N- or C-terminal amino acids and in intact form. The limitation of this system was the relatively high concentration of thiol-containing reductant which was required for the cleavage of the C-terminal cleavage tag (Raducanu et al. 2021).

Fusion tags sometimes cause structural changes in protein function, so when the goal of recombinant protein production is biopharmaceutical applications, removing the fusion tag without any amino acid residue in the goal protein is necessary to ensure the safety and efficiency of protein therapy. In 2020, researchers introduced a Caspase-2 cleavage site (a pentapeptide site that includes VDVAD amino acids) as a specific site for cleavage by the cpCasp2 enzyme which lead to an authentic N terminal in the recombinant protein (Cserjan-Puschmann et al. 2020). Introducing new types of tags with novel attractive features to circumvent obstacles in heterologous protein synthesis in *E. coli* is an ongoing process, and it is recommended to stay up-to-date to select the appropriate tags based on your projects.

## Codon preference

Many factors are involved to increase recombinant protein expression in heterologous hosts. The relative abundance of different tRNAs for a particular amino acid can vary greatly between species and results in species-specific preferences (Terpe 2006). The expression of heterologous proteins in *E. coli* is affected by codon usage bias (Rosano and Ceccarelli 2009). Codon preference is a common phenomenon among organisms, and differences in the pattern of codons for a specific amino acid or differences in the pattern of tRNAs abundance can lead to it. Two strategies are used to avoid this problem. In the first scenario, researchers substitute the gene of interest codons sequence with the preferred codons in the host organism, and finally, the modified gene is ordered to be artificially synthesized. In the second option, researchers use a genetically modified strain of *E. coli* that can translate all codons well. Modification of codon usage can have a distinct effect on protein expression levels. Some studies have shown that codon optimization increases protein expression even more than 1000-folds (Mauro 2018). Codon optimization increases the rate of translation speed by overcoming limitations related to species-specific differences in codon usage and tRNA abundance (Lipinszki et al. 2018; Portolano et al. 2014). In one study, a high level of production of human growth hormone was achieved in *E. coli* through the codon optimization strategy, and it was found that codon optimization is an important solution to avoid the problems forced by rare codons or unwanted secondary structure formation of mRNA (Ghavim et al. 2017).

## *E. coli* strains have a fundamental role in heterologous protein expression

Several strains of *E. coli* are used for recombinant protein production depending on the intrinsic properties of the POIs. *E. coli* K12 and BL21 (DE3) and many of their derivatives are the most commonly used hosts for heterologous expression (Sahdev et al. 2008; Tegel et al. 2010). The genome sequence of *E. coli* K12 was published in 1997, (Blattner et al. 1997) and the reduction of its genome with the removal of unnecessary regions improved its properties (Posfai et al. 2006). *E. coli* strains differ based on specific characteristics acquired through

the genetic engineering of their genomes. In general, they were engineered to acquire features such as higher mRNA stability, ability to form disulfide bonds in the cytoplasm, the absence of proteases, suitability for eukaryotic gene expression, suitability for toxic gene expression, having stable plasmids, stringent control on heterologous overexpression, etc. Characteristics of *E.coli* strains are demonstrated in Table 3 (Baeshen et al. 2015; Hayat et al. 2018; Terpe 2006). In the *E. coli* expression system, one of the problems is the lack of obtaining recombinant protein with high yield, which is due to the metabolic burden. Metabolic burden causes growth decrement and cell density reduction in *E. coli*, which leads to a decrease in recombinant protein yield. Studies have shown that this imbalance can be controlled by regulating heterologous gene expression by considering the copy number of a plasmid and the adjusting promoter strength (Lozano Terol et al. 2021). pET-based expression plasmids were widely used for research and commercial production (Zhang et al. 2022). This plasmid has a T7 promoter that drives the heterologous expression of recombinant proteins. The T7 promoter was derived from λ prophage, as well as the T7

RNA polymerase gene of this virus was integrated into the genome of *E.coli* BL21 (DE3) strain (Studier and Moffatt 1986). In addition, the gene of this RNA polymerase exists in the genome of other *E. coli* strains (Robichon et al. 2011; Vikström et al. xxxx). This RNA polymerase transcribes genes eight times more than the native RNA polymerases of *E.coli* (Iost et al. 1992). This RNA polymerase sometimes reduces the level of recombinant proteins through the deleterious effects that it causes, such as metabolic burden, toxic issues, and low host cell density. To overcome these issues, researchers developed strains of *E. coli* which have adjustable promoters for the RNA polymerase gene that can tightly control its expression. Scientists by redesigning some parts of the pET28a vector increased its efficacy in the production of recombinant proteins (Shilling et al. 2020). In another study, researchers knocked down the genes related to blocking stress responses in *E. coli*, and these mutated *E. coli* revealed a great potential to express recombinant proteins in higher amounts in comparison to non-mutated counterparts (Sharma et al. 2020).

**Table 3** *E. coli* strains used for recombinant protein production

| Strain | Derivation | Characteristics | References |
|---|---|---|---|
| BL21 | B834 | High-level protein expression and easy induction | (Manual xxxx) |
| BL21* | BL21 | Mutation in the rne131 gene for increasing mRNA stability | (Makino et al. 2011) |
| BL21-pLys | BL21 | Suitable for the expression of toxic proteins and is resistant to chloramphenicol | (Manual xxxx) |
| Origami | K-12 | Mutation in *trxB* & *gor* genes for improving disulfide bond formation in the cytosol | (Makino et al. 2011) |
| Origami B | K-12 | Mutations in trxB and gor genes plus mutation in lacZY gene | (Kaplan et al. 2016) |
| Rosetta | BL21 | Elevate the expression of eukaryotic genes that contain codons which rarely used in *E. coli* | (Fathi-Roudsari et al. 2016) |
| Rosetta-gami | BL21 | Elevate the expression of eukaryotic proteins as well as enhance disulfide bond formation in the target protein | (Zarkar et al. 2020) |
| C41 | BL21 | Natural mutant strain with the ability to overexpress membrane and toxic proteins | (Dumon-Seignovert et al. 2004; Schlegel et al. 2015) |
| C43 | BL21 | Natural mutant strain with the ability to overexpress membrane and toxic proteins | (Dumon-Seignovert et al. 2004; Tian et al. 2009) |
| Rosetta 2- pLys | BL21 | Enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli* | (Kopanic et al. 2013) |
| Shuffle | *E. coli* B (C3029) & *E. coli* K-12 (C3026) | Assisting in the correct folding of proteins containing disulfide bonds in their structures in the cytoplasm | (Lobstein et al. 2012) |
| Tuner | BL21 | lactose permease LacY-deficient | (Binder et al. 2016) |
| Lemo21 | BL21 | Suitable for proteins with problems in expression (membrane proteins, toxic proteins, proteins with solubility issues) | (Schlegel et al. 2012) |
| SoluBL21 | BL21 | Soluble mammalian protein expression | (Hata et al. 2013) |
| W3110 | K12 | Wild-type laboratory strain | (Khankal et al. 2008) |
| HMS174 | K-12 | *recA* mutant for increasing plasmid stability | (Striedner et al. 2010) |
| BLR | BL21 | recA-deficient derivative of BL21(DE3) | (Goffin and Dehottay 2017) |

## Culture conditions

Another effective factor in the expression of recombinant proteins is the conditions of bacterial cultures. Researchers use shake flasks for the optimization of growth conditions. These shake flasks do not require any special instruments (Ladner et al. 2019). However, unlike bioreactors, standard cultivation of *E. coli* in shake flasks is done in batches. Batch growth in flasks has several limitations such as lack of pH control, low aeration, and nutrient shortage (Weuster-Botz et al. 2001). In aerobic cultures, oxygen supply is vital. Under oxygen-restricted conditions, the metabolism of a microorganism is substantially altered, and unwanted by-products like organic acids such as acetate and ethanol may be composed (Losen et al. 2004). In *E.coli*, high cell density followed by the formation of acetate leads to toxicity during the culture period. Controlling oxygen levels can greatly control toxicity (Johnson et al. 2010). Acetate production and its toxicity can be solved by continuously feeding the culture with glucose and maintaining the growth rate, especially before entering the acetate production stage (Millard et al. 2021).

Different types of bioreactor systems have been introduced for large-scale production of recombinant proteins in *E.coli* (Blunt et al. 2018; Lopes et al. 2014; Muhamad Ali et al. 2015). In one experiment, researchers used a new device (gas-vortex bioreactor) for large-scale production of recombinant proteins in *E. coli* and claimed their method outperformed the stirred bioreactor (Savelyeva et al. 2017). Culture conditions (such as temperature, oxygen supply, and pH) were confirmed to have a significant effect on the large-scale production of recombinant proteins in *E. coli* (Zare et al. 2019). In one research, nerve growth factor (β-NGF) protein was heterologously produced in *E. coli* DE3 strain, and it was reported that the 30% dissolved oxygen and post-induction temperature of 28.5 °C were the best values to produce this protein in the bioreactor (Hajihassan et al. 2019). In another study, improved growth performance of *E. coli* was reported by manipulating the expression of small non-coding RNAs. These small RNAs play a role in regulating stress conditions such as fluctuations in temperature, pH, dissolved oxygen, glucose, and salt concentration. These stress conditions can occur especially in large-volume bioreactors (Negrete and Shiloach 2017). Different amounts of recombinant proteins have been produced in *E. coli*. During a batch bioreactor culture, the concentration of Cas9 protein that heterologously was expressed in *E. coli* was approximately 420.1 mg/L, obtained 6 h after the induction time (Carmignotto and Azzoni 2019). In another study, the produced recombinant TGase enzyme in bioreactor cultivation was 17.5 g/L (Duarte et al. 2021). One of the problems in the high-volume production of recombinant proteins is the risk of bioreactor contamination. The heatinducible-based

expression system is generally used at the industrial level for the synthesis of various recombinant proteins, which does not require the addition of chemical inducers and thus minimizes the risks of contamination (Restrepo-Pineda et al. 2021).

The introduction of auto-induction of heterologous gene expression methods is a great advance in heterologous gene expression. Various auto-induction of promoters based on nutrient manipulation (Lu et al. 2020; Tahara et al. 2020) and genetic modification of expression vectors (Shang et al. 2021) has been reported. These auto-induction systems have several advantages over classic ones. They do not need the inducers such as the isopropyl-β-D-thiogalactoside (IPTG) (which is toxic to *E. coli* (Dvorak et al. 2015; Kosinski et al. 1992; Malakar and Venkatesh 2012) and more expensive than its natural analog, lactose (Khani and Bagheri 2020)) and periodic measurements of the $OD_{600}$ to add inducers at proper times. Moreover, they are cost-effective and convenient methods for heterologous protein expression. Researchers have reported that by applying the auto-induction method, the amount of recombinant proteins and *E. coli* biomass has increased in comparison to using an inducer to over-expression of the recombinant proteins (Gao et al. 2010; Lu et al. 2020; Sarduy et al. 2012). Moreover, this strategy was introduced as a suitable method for the production of soluble recombinant proteins (Fathi-Roudsari et al. 2018). In the auto-induction procedure based on nutrient manipulation, researchers utilized a pretty defined medium such as the ZYM-5052 (Sarduy et al. 2012) which lactose in these media functions as a heterologous expression inducer. In auto-induction media based on nutrient manipulation, researchers use a balanced mixture of glucose, lactose, and glycerol as carbon substrates (Li et al. 2011). *E.coli* first uses glucose as a carbon source, and around the mid-to-late log phase, which is the best time for induction of expression, glucose (the repressor of the lac promoter) is depleted and lactose is used as a carbon source (Greicius et al. 2022). Lactose also acts as a lac promoter inducer, which drives the recombinant protein expression. Glycerol is consumed along with lactose to provide adequate carbon source support (Ukkonen et al. 2013). Adjusting the appropriate proportions of these elements in auto-induction media can be a critical step to produce the desired amount of recombinant proteins.

## Localization of recombinant proteins

According to the location of synthesized heterologous proteins in *E. coli*, they can be classified into different groups. Heterologous proteins can be localized in three different spaces including cytoplasm, periplasm, and extracellular space (secretion into the culture medium) (Fig. 3) (Kleiner-Grote et al. 2018). Accumulation of recombinant proteins

in the cytoplasm is the most common strategy for recombinant protein production in *E. coli*, which requires the simplest expression system and demonstrates the highest yield (Makrides 1996). The formation of inclusion bodies (IBs) is one of the bottlenecks in the accumulation of over-expressed recombinant proteins in the cytoplasm (Carrió and Villaverde 2001; Ventura and Villaverde 2006). In this case, there is a need to solubilize and refold the aggregated proteins and as a result reduce the yield and bioactivity of the recombinant proteins (Bhatwa et al. 2021). *E. coli* cytosol has a reducing environment and is not a suitable environment for the expression of recombinant proteins with disulfide bridges. The periplasm of *E.coli* is a suitable space for proteins that contain disulfide bonds in their structures. These bonds are essential for protein folding and activity (Paraskevopoulou et al. 2022b). Recombinant protein secretion into the periplasm or in the culture medium has many advantages over the intracellular production of recombinant proteins in the form of IBs. It helps in the natural folding and stability of proteins and also helps in the production of active and soluble proteins at a reduced cost (Mergulhão et al. 2005). To translocation of the recombinant proteins into the periplasm from the cytosol, various signal peptides have been proposed and utilized (Karyolaimos et al. 2019; Maleki and Hajihassan 2021; Santos et al. 2019). The

presence of a small amount of protease and protein in the periplasm as compared to the cytoplasm facilitates the purification of recombinant proteins. One of the most important aspects of partitioning recombinant proteins in the periplasm is protein purification without needing cell lysis and through osmotic shock (Ramanan et al. 2010). In general, recombinant protein accumulation in the periplasmic space has many advantages over the cytoplasm. The main advantage is the simplicity of purification, but the translocation of recombinant proteins to the periplasmic space requires a more complex expression system, which creates limitations for the production of recombinant proteins (Fig. 3). Moreover, the periplasmic environment is more oxidizing, and hence disulfide bond formation is favored, also the protease activity in the periplasm is lower and as a result the probability of protein stability increases (Makrides 1996; Mergulhão et al. 2005).

## Inclusion bodies are like a double-edged sword

Recombinant proteins may be soluble or insoluble after expression under normal growth conditions. *E. coli* proteins are stable and soluble, but many proteins, especially those
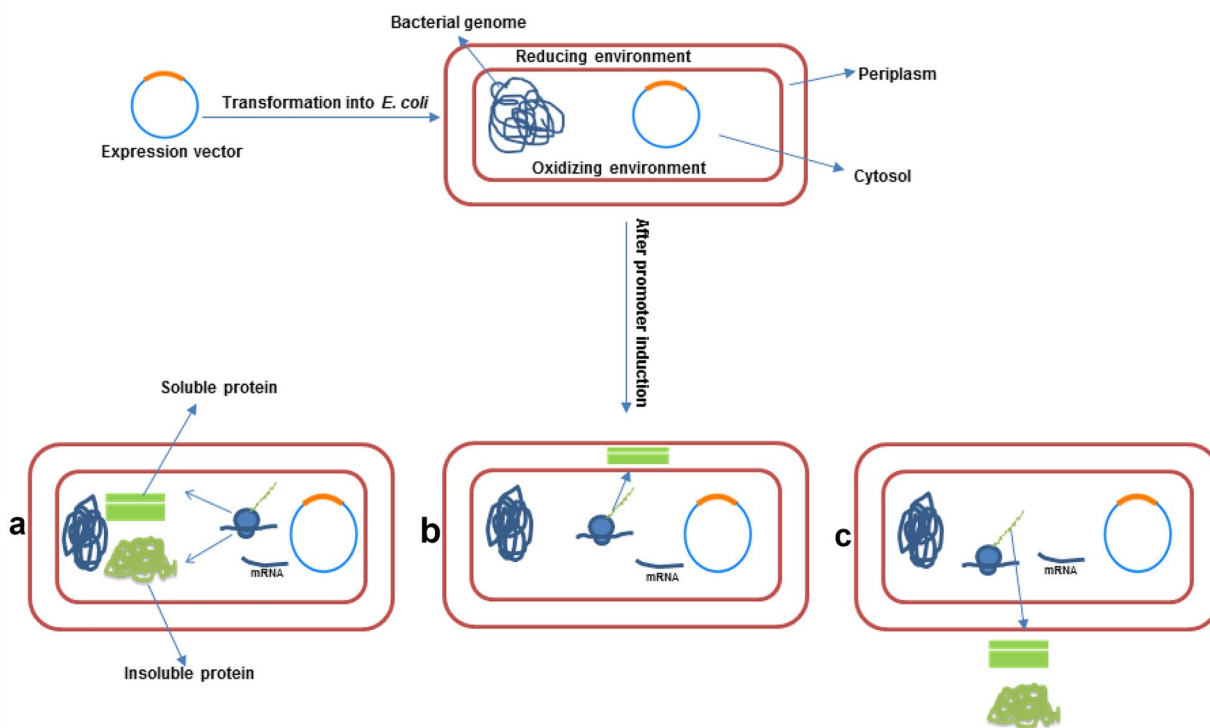


Fig. 3 Translocation of recombinant proteins in *E. coli*. **a** Recombinant proteins in the cytoplasm can be soluble or insoluble **b** Translocation of recombinant proteins into the periplasmic space through signal peptides. **c** Secretion of recombinant proteins into the extracellular space

of eukaryotic origin, often cannot be expressed in soluble form in *E. coli*. In addition, some recombinant proteins may be toxic to *E. coli* cells when express in an active form and therefore produce in the form of IBs (Saïda et al. 2006; Singh et al. 2005). Recombinant proteins that are produced as IBs in *E. coli* are inactive, insoluble, and usually have misplaced intramolecular and intermolecular disulfide bonds and free cysteine molecules are unusual in these particles (Fischer et al. 1993). Moreover, these submicron proteinaceous bodies (typically from 50 to 800 nm) are non-toxic to *E. coli* (Rinas et al. 2017).

IBs are proteins aggregated inside the cytoplasm (Palmer et al. 2012). These bodies are resistant to proteases, and it is easy to separate these bodies from other cellular components. In addition, recombinant proteins that are toxic to the heterologous host can be produced in the form of IBs that are non-toxic to the host cells. The most important problem of IBs is obtaining the active target protein from these bodies, which requires different steps such as denaturation of IBs and refolding and recovery of active target proteins. These steps are very cumbersome and expensive (Singh and Panda 2005) and do not work for all IB types. To obtain active proteins, IBs must be recovered from the cell, which is called the disaggregation process or the solubilization process and involves the disruption or destruction of the secondary and tertiary structures of the proteins. High concentrations of chemical denaturants such as urea and guanidine hydrochloride are used to dissolve IBs (Chew et al. 2020). These materials result in the complete denaturation of secondary structures and often lead to the aggregation of proteins during the refolding process (Singh et al. 2015). Because of these bottlenecks, researchers like to produce recombinant proteins in soluble and active forms rather than IBs. In this section, we describe all the negative and positive features of these bodies.

IBs are insoluble and aggregated forms of recombinant proteins. When researchers want to express heterologous proteins in *E.coli*, the formation of these bodies is a common phenomenon. Studies have shown that usually six factors are involved in the formation of IBs including, 1- Charge average 2- Turn forming residues 3- Proline fraction 4- Cysteine fraction 5- Hydrophilicity index 6- Total number of residues. The first two factors are more fundamental (Diaz et al. 2010; Hirose and Noguchi 2013). The speed of recombinant protein synthesis is another vital factor in the formation of IBs, which has a direct correlation with the formation of IBs.

There are strategies to reduce the formation IBs and to help the proper folding of recombinant proteins. These approaches include bacterial growth at lower temperatures, selection of suitable bacterial strains, co-expression with a chaperone, amino acid substitution, growth and or induction under osmotic stress conditions in the presence of sorbitol and glycyl betaine, growing in an environment containing non-metabolized sugars, adjusting pH and using strains lacking thioredoxin reductase (Gatti-Lafranconi et al. 2011; Malakar and Venkatesh 2012). Here. We discuss some of these factors.

## Molecular chaperones

Several reasons such as lacking post-translational modifications (PTMs), misfolding, lower solubility capacity, and host metabolic burden are involved in creating IBs (Zhang et al. 2022). One of the strategies to overcome the creation of IBs during the over-expression of recombinant proteins is the co-expression of the target proteins with the chaperone molecules. These chaperones can alleviate IBs construction by helping the proper folding of proteins during or post the translation phase (Jo 2022). Chaperones are molecules present in the cell as intermediates of protein folding and also play a role in maintaining the quality and stability of the proteome (Bhatwa et al. 2021). Sometimes, in stressful conditions such as heat shock, the expression of chaperones is increased. According to the function, chaperones can be divided into three groups: folding, holding, and disaggregating chaperones (Sahdev et al. 2008). In the cytoplasm of *E. coli*, trigger factor along with GroEL, GroES, DnaK, DnaJ, and GrpE play a role in protein refolding, and GroEL and DnaK are folding chaperones and work by consuming ATP (Mokry et al. 2015; Wang et al. 2018; Xu et al. 2018). Heat shock proteins are named according to their molecular weight (Li and Srivastava 2004). ClpB is one of the disaggregate chaperones that prevent the aggregates of proteins under stress conditions (Deville et al. 2017).

## Fusion tags

To increase the solubility of recombinant proteins in *E. coli*, they can be co-expressed with fusion tags with high solubility features, but it should be kept in mind that not all tags with high solubility are suitable for increasing target protein solubility (Song et al. 2011). A protein called NusA was used as a tag to increase the solubility and proper folding of recombinant proteins (Li et al. 2013). Antigen-binding fragments (Fabs) for production in *E. coli* must be secreted into the periplasmic space, where oxidizing conditions allow the proper formation of disulfide bonds. So far, various studies have been done to increase the expression of Fabs protein in the periplasmic space in *E. coli* (Gundinger et al. 2022). In 2019, Luo et al. used the alkaline phosphatase (phoA) promoter in the pAT system and concluded that this system is better than the commonly used T7-based pET system for overexpressing Fabs in *E.coli* (Luo et al. 2019). The use of peptide tags to increase the solubility of recombinant

proteins has been reported in numerous studies (Paraskevopoulou and Falcone 2018). In one study, researchers found that adding a peptide tag called a hexalysine tag (6 K) to the C-terminus of the target protein increased the yield, solubility, and integrity of the target protein. This target protein was directed to the periplasmic compartment (Paraskevopoulou et al. 2022a). In another study, these properties of hexalysine were not observed, but it improved the thermal stability of the recombinant protein (Paraskevopoulou et al. 2019). These results show that the addition of peptide tags can sometimes show different responses based on the target protein. We think that *in-silico* analysis of the effects of tags on the target protein should be considered before recombinant DNA construction.

## Decreasing the growth temperature

During the production of recombinant proteins, low temperature reduces the accumulation of proteins and is one of the ways to reduce the production of IBs. Studies have shown that a growth temperature between 15 and 25 °C can be a suitable temperature when the host is facing the problem of IBs formation (San-Miguel et al. 2013). Very low temperatures may slow down the growth rate and thus decrease the concentration of proteins, in addition, chaperones may not function properly and protein folding may be affected (Strocchi et al. 2006). Studies have shown that low temperature is effective in increasing the solubility of some complex proteins. Hydrophobic bonds, which are effective in creating IBs, are highly temperature dependent, so protein expression under low-temperature conditions leads to high stability and proper folding of proteins (Vera et al. 2007).

## Disulfide bond formation in the cytoplasm

Proteins in the form of IB are mostly inactive and insoluble and have to be refolded. Production of proteins with many disulfide bonds is problematic, and refolding of aggregated proteins is difficult (Berkmen 2012; Palmer and Wingfield 2012). The cytosol of *E. coli* is a reducing environment and triggers the misfolding of proteins that contain intradisulfide bonds and stimulates the construction of IBs (Singh and Panda 2005). The limiting factors in the formation of intracytoplasmic disulfide bonds include the presence of reductase enzymes that inhibit the formation of disulfide bonds as well as the absence of periplasmic oxidase enzymes such as disulfide bond oxidase (DsbA) and disulfide bond isomerase (DsbC) which help the disulfide bonds formation (Hatahet et al. 2014). Recently, engineered strains such as Shuffle and Origami have been released that contain mutations in two genes, thioredoxin reductase (trxB) and glutathione reductase (Gordon et al.

2008). These mutations make the cytoplasm a suitable space for creating disulfide bonds (Hayat et al. 2018).

## The Good face of inclusion bodies

Sometimes the production of recombinant proteins in the form of IBs is the best option. The purification of the IBs is easy, and they are an ideal strategy for the production of recombinant proteins that are toxic to the heterologous hosts. Moreover, the production of recombinant proteins in IBs form is much more efficient in terms of yield. It has been identified that the addition of the ssTorA tag can induce IBs formation even in highly soluble proteins such as MBP and thioredoxin (Jong et al. 2017), so this tag can be applied to produce recombinant proteins in the form of IB.

## Effect of post-translational modification in recombinant protein expression

The lack of the translational modification mechanism and the formation of IBs are significant challenges during recombinant protein production in *E. coli*. The engineering of glycosylation pathway and other post-translational modifications in *E. coli* has great progress in recent years. *E. coli* is unable to glycosylate expressed proteins, but the discovery of N-glycosylation in gram-negative bacteria, *Campylobacter jejuni,* and transferring this mechanism into *E. coli* genome has made it possible to use *E. coli* to express glycoproteins (Feldman et al. 2005). Another issue of *E. coli* system is the lack of an acetylation mechanism and the inability to acetylate eukaryotic proteins. Studies have shown that in eukaryotic cells, nearly 98% of proteins are acetylated at the amino terminus (Johnson et al. 2010). Acetylation affects various cellular processes such as DNA replication, repair and recombination processes, protection of genome stability, cytoskeleton dynamics, metabolism, signal transmission, and protein folding. Therefore, the correct performance of acetylation will be necessary to produce active recombinant eukaryotic proteins (Neumann et al. 2008). Moreover, recent studies have shown that *E. coli* can perform protein phosphorylation through intracellular kinases using adenosine triphosphate as a phosphoryl group donor (Mijakovic et al. 2006).

## Applications of recombinant proteins produced in *E. coli*

Recombinant proteins have many applications. They can be used in vaccine production, drug delivery system, construction of antibodies and enzymes for the treatment of diseases, diagnostic studies, assessment of protein–protein interactions, understanding of biological structures, etc. *E.*

*coli* is the prevalent host for the production of recombinant proteins (Gupta et al. 2017; Werkmeister and Ramshaw 2012). Myriad of heterologous proteins such as insulin, interferon, Preotact (human parathyroid hormone), Nivestim (filgrastim, rhGCSF), and Preos (parathyroid hormone) have been produced in this organism (Baeshen et al. 2015). Developing novel therapies based on recombinant DNA technology and recombinant proteins for many chronic diseases such as cancer and other rare diseases is under investigation. The worldwide heterologous proteins market size was estimated at USD 1.4 billion in 2022 and is forecasted to reach USD 2.4 billion at a compound annual growth rate (CAGR) of 11.4% by 2027 (https://www.marketsandmarkets.com, 26 Jan 2023). In recent years, due to the emergence of antibiotic-resistant microbes, the production of recombinant antimicrobial peptides is one of the promising strategies to overcome this problem. Many studies have been conducted on the expression of these peptides in *E. coli*. It has been reported that the production of antimicrobial peptides in *E. coli* is restricted due to their toxicity to host cells and susceptibility to proteolytic degradation. One of the strategies to circumvent this hindrance in *E. coli* is the co-expression of these peptides with other proteins (Li 2011; Zorko and Jerala 2010). For the expression of antimicrobial peptides that are rich in disulfide bonds, *E. coli* Origami or Rosetta-gami strains are suitable because they have mutations in genes that are responsible for reduction reaction that this type of reaction is a hindrance in the formation of disulfide bridges (Wibowo and Zhao 2019). It is expected that the number of recombinant proteins produced in *E. coli* in the market will be increased with the increase of researchers' knowledge about the effective factors in the expression of recombinant proteins and the introduction of new strains of *E. coli* and various expression vectors.

## Conclusion and future perspective

The production of various sorts of proteins with different applications in heterologous hosts has been one of the most important achievements of biotechnology and genetic engineering in recent years. Recombinant proteins are used in various industries such as pharmaceuticals, cosmetics, food, agriculture, etc. Nowadays, there is a great demand for the production of recombinant proteins to prevent, diagnose and treat human diseases. Overall, to express a recombinant protein, a DNA molecule containing the open reading frame of a particular protein is identified, isolated, cloned into a vector, and finally expressed in a suitable host. A heterologous gene can be expressed in different types of heterologous hosts, therefore, choosing a suitable host is a critical step. Mammalian cells, bacteria, yeasts, plants, and insects are hosts that have been introduced for heterologous expression of a gene or genes of interest. The bacterial host is the most used expression system due to its simple expression method and low cost. *E. coli* strains are the most utilized bacterial host for the expression of recombinant proteins. Known and unknown factors are involved in enhancing the expression of recombinant protein in *E.coli*. These known factors include the selection of suitable expression vectors, the choosing of the adjustable promoter, the picking up of the competent strain of *E. coli*, codon optimization, concentration of the inducer in the induction time, induction period, the translocation of heterologous proteins into the periplasm or extracellular space, application of proper affinity tags, types and condition of the culture medium, etc. Moreover, *E.coli* as a heterologous host has strengths and weaknesses. Identification of every aspect of *E. coli* can be a useful strategy to boost its efficiency in recombinant protein production. Here, we discussed the important elements that are involved in heterologous gene expression in *E. coli*. One of these elements is the IBs formation. Over-expression of recombinant proteins in *E.coli* is very likely to form IBs. These bodies are resistant to proteases and harvesting them from lysed cells can be done through moderate-speed centrifugation as well as in the case of proteins that are toxic to the host, IBs are the best solution to overcome this challenge. These characteristics of the IBs are a shining clue that directs us to exploit their properties. So, the introduction of one or several common, efficient, cost-effective methods for solubilization and refolding of any type of aggregated recombinant proteins would be a major breakthrough in this science. Today, with the sharp progress in omics data and the existence of engineered endonucleases, it is possible to identify the function of genes by heterologous expression or knocking down or knocking out of them in the native host. This information will help us better understand the behavior of the *E. coli* host in different situations and exploit this data to adapt it. One of the limitations in recombinant protein expression in *E. coli* is the lack of an extracellular secretion system. A promising method for this hindrance would be the introduction of genes of non-indigenous extracellular secretion pathways into *E. coli*. In general, progress is limitless and as scientists, we should accelerate it.

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

## References

Andreev YA, Kozlov SA, Vassilevski AA, Grishin EV (2010) Cyanogen bromide cleavage of proteins in salt and buffer solutions. Anal Biochem 407(1):144–146

Arnau J, Lauritzen C, Petersen GE, Pedersen J (2006) Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. Protein Expr Purif 48(1):1–13

Aune, T. E. V. (2008). High level recombinant protein production in Escherichia coli by engineering broad-host-range plasmid vectors containing the Pm/xylS expression cassette.

Baeshen MN, Al-Hejin AM, Bora RS, Ahmed MM, Ramadan HA, Saini KS, Redwan EM (2015) Production of biopharmaceuticals in E. coli: current scenario and future perspectives. J Microbiol Biotechnol 25(7):953–962

Baneyx F (1999) Recombinant protein expression in Escherichia coli. Curr Opin Biotechnol 10(5):411–421

Berkmen M (2012) Production of disulfide-bonded proteins in Escherichia coli. Protein Expr Purif 82(1):240–251

Bhatwa A, Wang W, Hassan YI, Abraham N, Li X-Z, Zhou T (2021) Challenges associated with the formation of recombinant protein inclusion bodies in Escherichia coli and strategies to address them for industrial applications. Front Bioeng Biotechnol 9:630551

Binder D, Probst C, Grünberger A, Hilgers F, Loeschcke A, Jaeger K-E, Drepper T (2016) Comparative single-cell analysis of different E. coli expression systems during microfluidic cultivation. PLoS ONE 11(8):e0160711

Blattner FR, Plunkett G III, Bloch CA, Perna NT, Burland V, Riley M, Mayhew GF (1997) The complete genome sequence of Escherichia coli K-12. Science 277(5331):1453–1462

Blazeck J, Alper HS (2013) Promoter engineering: recent advances in controlling transcription at the most fundamental level. Biotechnol J 8(1):46–58

Blommel PG, Fox BG (2007) A combined approach to improving large-scale production of tobacco etch virus protease. Protein Expr Purif 55(1):53–68

Blunt W, Levin DB, Cicek N (2018) Bioreactor operating strategies for improved polyhydroxyalkanoate (PHA) productivity. Polymers 10(11):1197

Bock R (2015) Engineering plastid genomes: methods, tools, and applications in basic research and biotechnology. Annu Rev Plant Biol 66(1):211–241

Booth WT, Schlachter CR, Pote S, Ussin N, Mank NJ, Klapper V, Chruszcz M (2018) Impact of an N-terminal polyhistidine tag on protein thermal stability. ACS Omega 3(1):760–768

Bozarth A, Maier U-G, Zauner S (2009) Diatoms in biotechnology: modern tools and applications. Appl Microbiol Biotechnol 82(2):195–201

Brosius J, Erfle M, Storella J (1985) Spacing of the-10 and-35 regions in the tac promoter. Effect on its in vivo activity. J Biol Chem 260(6):3539–3541

Brown TA (2020) Gene cloning and DNA analysis: an introduction. Wiley, New York

Burnett MJ, Burnett AC (2020) Therapeutic recombinant protein production in plants: challenges and opportunities. Plants, People, Planet 2(2):121–132

Camps M (2010) Modulation of ColE1-like plasmid replication for recombinant gene expression. Recent Patents DNA Gene Sequ 4(1):58–73

Carmignotto GP, Azzoni AR (2019) On the expression of recombinant Cas9 protein in E. coli BL21 (DE3) and BL21 (DE3) Rosetta strains. J Biotechnol 306:62–70

Carrió MM, Villaverde A (2001) Protein aggregation as bacterial inclusion bodies is reversible. FEBS Lett 489(1):29–33

Çelik E, Çalık P (2012) Production of recombinant proteins by yeast cells. Biotechnol Adv 30(5):1108–1118

Chen, S., Xu, Z., & Zhao, W. (2015). Construction of pCXZ14W, a Novel pUC19-derived Plasmid Encoding the rop Gene.

Cheng X, Patterson TA (1992) Construction and use of λ PL promoter vectors for direct cloning and high level expression of PCR amplified DNA coding sequences. Nucleic Acids Res 20(17):4591–4598

Chew, F., Abidin, N., Yusof, N., Rafi, N., & Chua, G. (2020). *Recovery of inclusion body protein in Escherichia coli: Effects of solubilization methods and process condition.* Paper presented at the IOP Conference Series: Materials Science and Engineering.

Costa S, Almeida A, Castro A, Domingues L (2014) Fusion tags for protein solubility, purification and immunogenicity in Escherichia coli: the novel Fh8 system. Front Microbiol 5:63

Cserjan-Puschmann M, Lingg N, Engele P, Kröß C, Loibl J, Fischer A, Brocard C (2020) Production of circularly permuted caspase-2 for affinity fusion-tag removal: cloning, expression in Escherichia coli, purification, and characterization. Biomolecules 10(12):1592

Demain AL, Vaishnav P (2009) Production of recombinant proteins by microbes and higher organisms. Biotechnol Adv 27(3):297–306

Deng T, Ge H, He H, Liu Y, Zhai C, Feng L, Yi L (2017) The heterologous expression strategies of antimicrobial peptides in microbial systems. Protein Expr Purif 140:52–59

Deville C, Carroni M, Franke KB, Topf M, Bukau B, Mogk A, Saibil HR (2017) Structural pathway of regulated substrate transfer and threading through an Hsp100 disaggregase. Sci Adv 3(8):e1701726

Diaz AA, Tomba E, Lennarson R, Richard R, Bagajewicz MJ, Harrison RG (2010) Prediction of protein solubility in Escherichia coli using logistic regression. Biotechnol Bioeng 105(2):374–383

Duarte LS, Matte CR, Dall Cortivo PR, Nunes JES, Barsé LQ, Bizarro CV, Ayub MAZ (2021) Expression of Bacillus amyloliquefaciens transglutaminase in recombinant E. coli under the control of a bicistronic plasmid system in DO-stat fed-batch bioreactor cultivations. Braz J Microbiol 52:1225–1233

Dumon-Seignovert L, Cariot G, Vuillard L (2004) The toxicity of recombinant proteins in Escherichia coli: a comparison of overexpression in BL21 (DE3), C41 (DE3), and C43 (DE3). Protein Expr Purif 37(1):203–206

Dvorak P, Chrast L, Nikel PI, Fedr R, Soucek K, Sedlackova M, Damborsky J (2015) Exacerbation of substrate toxicity by IPTG in Escherichia coli BL21 (DE3) carrying a synthetic metabolic pathway. Microbial Cell Fact 14(1):1–15

Ekundayo B, Bleichert F (2019) Origins of DNA replication. PLoS Genet 15(9):e1008320

Fathi-Roudsari M, Akhavian-Tehrani A, Maghsoudi N (2016) Comparison of three Escherichia coli strains in recombinant production of reteplase. Avicenna J Med Biotechnol 8(1):16

Fathi-Roudsari M, Maghsoudi N, Maghsoudi A, Niazi S, Soleiman M (2018) Auto-induction for high level production of biologically active reteplase in Escherichia coli. Protein Expr Purif 151:18–22

Feldman MF, Wacker M, Hernandez M, Hitchen PG, Marolda CL, Kowarik M, Aebi M (2005) Engineering N-linked protein

glycosylation with diverse O antigen lipopolysaccharide structures in Escherichia coli. Proc Natl Acad Sci 102(8):3016–3021

Fischer B, Sumner I, Goodenough P (1993) Isolation, renaturation, and formation of disulfide bonds of eukaryotic proteins expressed in Escherichia coli as inclusion bodies. Biotechnol Bioeng 41(1):3–13

Francis AI, Ghany S, Gilkes T, Umakanthan S (2022) Review of COVID-19 vaccine subtypes, efficacy and geographical distributions. Postgrad Med J 98(1159):389–394

Gao, L. (1986). The effect of wave reflection of artificial boundaries in viscoelastic media [J]. *J. Hydraulic Engineering, 6.*

Gao M, Ma C, Liu W, Zhu J, Tian H, Gao X, Yao W (2010) Production and purification of an analog of glucagon-like peptide-1 by auto-induction and on-column cleavage in Escherichia coli. World J Microbiol Biotechnol 26(9):1675–1682

Gatti-Lafranconi P, Natalello A, Ami D, Doglia SM, Lotti M (2011) Concepts and tools to exploit the potential of bacterial inclusion bodies in protein science and biotechnology. FEBS J 278(14):2408–2418

Georgiou G, Segatori L (2005) Preparative expression of secreted proteins in bacteria: status report and future prospects. Curr Opin Biotechnol 16(5):538–545

Ghavim M, Abnous K, Arasteh F, Taghavi S, Nabavinia MS, Alibolandi M, Ramezani M (2017) High level expression of recombinant human growth hormone in Escherichia coli: crucial role of translation initiation region. Res Pharm Sci 12(2):168

Ghiasi N, Valizadeh R, Arabsorkhi M, Hoseyni TS, Esfandiari K, Sadighpour T, Jahantigh HR (2021) Efficacy and side effects of Sputnik V, Sinopharm and AstraZeneca vaccines to stop COVID-19; a review and discussion. Immunopathol Persa 7(2):e31–e31

Giacalone MJ, Gentile AM, Lovitt BT, Berkley NL, Gunderson CW, Surber MW (2006) Toxic protein expression in Escherichia coli using a rhamnose-based tightly regulated and tunable promoter system. Biotechniques 40(3):355–364

Gifre L, Arís A, Bach À, Garcia-Fruitós E (2017) Trends in recombinant protein use in animal production. Microb Cell Fact 16:1–17

Giordano TJ, Deuschle U, Bujard H, McAllister WT (1989) Regulation of coliphage T3 and T7 RNA polymerases by the lac represser-operator system. Gene 84(2):209–219

Goffin P, Dehottay P (2017) Complete genome sequence of Escherichia coli BLR (DE3), a recA-deficient derivative of E. coli BL21 (DE3). Genome Announc 5(22):e00441-e1417

Gong L, Yang Z, Zhang F, Gao W (2022) Cytokine conjugates to elastin-like polypeptides. Adv Drug Deliv Rev. https://doi.org/10.1016/j.addr.2022.114541

Gordon E, Horsefield R, Swarts HG, de Pont JJH, Neutze R, Snijder A (2008) Effective high-throughput overproduction of membrane proteins in Escherichia coli. Protein Expr Purif 62(1):1–8

Greicius A, Baliutavicius T, Lastauskiene E, Gudiukaite R (2022) Application of milk permeate as an inducer for the production of microbial recombinant lipolytic enzymes. Fermentation 9(1):27

Gundinger T, Kittler S, Kubicek S, Kopp J, Spadiut O (2022) Recombinant protein production in E. coli using the phoA expression system. Fermentation 8(4):181

Gupta V, Sengupta M, Prakash J, Tripathy BC (2017) Production of recombinant pharmaceutical proteins. Basic and applied aspects of biotechnology. Springer, New York, pp 77–101

Hajihassan Z, Sadat S, Gholami Tilko P (2019) Optimization of recombinant β-NGF production in bioreactor. Modares J Biotechnol 10(1):9–13

Han Y, Guo W, Su B, Guo Y, Wang J, Chu B, Yang G (2018) High-level expression of soluble recombinant proteins in Escherichia coli using an HE-maltotriose-binding protein fusion tag. Protein Expr Purif 142:25–31

Hata S, Kitamura F, Sorimachi H (2013) Efficient expression and purification of recombinant human μ-calpain using an E scherichia coli expression system. Genes Cells 18(9):753–763

Hatahet F, Boyd D, Beckwith J (2014) Disulfide bond formation in prokaryotes: history, diversity and design. Biochim Biophys Acta (BBA) 1844(8):1402–1414

Hayat SM, Farahani N, Golichenari B, Sahebkar A (2018) Recombinant protein expression in Escherichia coli (E. coli): what we need to know. Curr Pharm Des 24(6):718–725

Hellwig S, Drossard J, Twyman RM, Fischer R (2004) Plant cell cultures for the production of recombinant proteins. Nat Biotechnol 22(11):1415–1422

Hirose S, Noguchi T (2013) ESPRESSO: a system for estimating protein expression and solubility in protein expression systems. Proteomics 13(9):1444–1456

https://www.marketsandmarkets.com. (26 Jan 2023). Recombinant Proteins Market worth $2.4 billion by 2027. https://www.marketsandmarkets.com/PressReleases/recombinant-proteins.asp

Hu D, Eftink MR (1993) Interaction of indoleacrylic acid with Trp aporepressor from Escherichia coli. Arch Biochem Biophys 305(2):588–594

Hwang PM, Pan JS, Sykes BD (2014) Targeted expression, purification, and cleavage of fusion proteins from inclusion bodies in Escherichia coli. FEBS Lett 588(2):247–252

Hwang HJ, Kim JW, Ju SY, Park JH, Lee PC (2017) Application of an oxygen-inducible nar promoter system in metabolic engineering for production of biochemicals in Escherichia coli. Biotechnol Bioeng 114(2):468–473

Ikonomou L, Schneider Y-J, Agathos S (2003) Insect cell culture for industrial production of recombinant proteins. Appl Microbiol Biotechnol 62(1):1–20

Ingmer H, Miller C, Cohen SN (2001) The RepA protein of plasmid pSC101 controls Escherichia coli cell division through the SOS response. Mol Microbiol 42(2):519–526

Iost I, Guillerez J, Dreyfus M (1992) Bacteriophage T7 RNA polymerase travels far ahead of ribosomes in vivo. J Bacteriol 174(2):619–622

Jenny RJ, Mann KG, Lundblad RL (2003) A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa. Protein Expr Purif 31(1):1–11

Jo BH (2022) An intrinsically disordered peptide tag that confers an unusual solubility to aggregation-prone proteins. Appl Environ Microbiol 88(7):e00097

Johnson M, Coulton AT, Geeves MA, Mulvihill DP (2010) Targeted amino-terminal acetylation of recombinant proteins in E. coli. PLoS ONE 5(12):e15801

Jong WS, Vikström D, Houben D, van den Berg van Saparoea, H. B., de Gier, J.-W., & Luirink, J. (2017) Application of an E. coli signal sequence as a versatile inclusion body tag. Microb Cell Fact 16:1–13

Kaplan O, Zarubova J, Mikulova B, Filova E, Bártová J, Bačáková L, Brynda E (2016) Enhanced mitogenic activity of recombinant human vascular endothelial growth factor VEGF121 expressed in E. coli origami B (DE3) with molecular chaperones. PLoS ONE 11(10):e0163697

Karyolaimos A, Ampah-Korsah H, Hillenaar T, Mestre Borras A, Dolata KM, Sievers S, de Gier J-W (2019) Enhancing recombinant protein yields in the E. coli periplasm by combining signal peptide and production rate screening. Front Microbiol 10:1511

Kaur J, Kumar A, Kaur J (2018) Strategies for optimization of heterologous protein expression in E. coli: roadblocks and reinforcements. Int J Biol Macromol 106:803–822

Keller J, Gerber R, Del Tito B, Vannicolo D, Arcuri E (1990) Effects of mutations altering SOS regulation on a nalidixic acid-inducible system for the production of heterologous proteins in Escherichia coli. J Ind Microbiol 6:199–206

Khan KH (2013) Gene expression in mammalian cells and its applications. Adv Pharm Bull 3(2):257

Khani M-H, Bagheri M (2020) Skimmed milk as an alternative for IPTG in induction of recombinant protein expression. Protein Expr Purif 170:105593

Khankal R, Luziatelli F, Chin JW, Frei CS, Cirino PC (2008) Comparison between Escherichia coli K-12 strains W3110 and MG1655 and wild-type E. coli B as platforms for xylitol production. Biotechnol Lett 30(9):1645–1653

Khlebnikov A, Keasling JD (2002) Effect of lacY expression on homogeneity of induction from the Ptac and Ptrc promoters by natural and synthetic inducers. Biotechnol Prog 18(3):672–674

Kleiner-Grote GR, Risse JM, Friehs K (2018) Secretion of recombinant proteins from E. coli. Eng Life Sci 18(8):532–550

Klumpp S (2011) Growth-rate dependence reveals design principles of plasmid copy number control. PLoS ONE 6(5):e20403

Koehn J, Hunt I (2009) High-Throughput Protein Production (HTPP): a review of enabling technologies to expedite protein production. High throughput protein expression and purification. Humana Press, Totowa, pp 1–18

Kopanic JL, Al-Mugotir M, Zach S, Das S, Grosely R, Sorgen PL (2013) An Escherichia coli strain for expression of the connexin45 carboxyl terminus attached to the 4th transmembrane domain. Front Pharmacol 4:106

Kosinski MJ, Rinas U, Bailey JE (1992) Isopropyl-β-D-thiogalactopyranoside influences the metabolism of Escherichia coli. Appl Microbiol Biotechnol 36:782–784

Kurup VM, Thomas J (2020) Edible vaccines: promises and challenges. Mol Biotechnol 62:79–90

Ladner T, Flitsch D, Lukacs M, Sieben M, Büchs J (2019) Combined dissolved oxygen tension and online viscosity measurements in shake flask cultivations via infrared fluorescent oxygen-sensitive nanoparticles. Biotechnol Bioeng 116(12):3215–3227

Lai T, Yang Y, Ng SK (2013) Advances in mammalian cell line development technologies for recombinant protein production. Pharmaceuticals 6(5):579–603

Li Y (2011) Recombinant production of antimicrobial peptides in Escherichia coli: a review. Protein Expr Purif 80(2):260–267

Li, Z., & Srivastava, P. (2004). Heat-shock proteins Current protocols in immunology/edited by John E Coligan [et al.]. In: Appendix.

Li Z, Kessler W, van den Heuvel J, Rinas U (2011) Simple defined autoinduction medium for high-level recombinant protein production using T7-based Escherichia coli expression systems. Appl Microbiol Biotechnol 91(4):1203–1213

Li K, Jiang T, Yu B, Wang L, Gao C, Ma C, Ma Y (2013) Escherichia coli transcription termination factor NusA: heat-induced oligomerization and chaperone activity. Sci Rep 3(1):1–9

Lichty JJ, Malecki JL, Agnew HD, Michelson-Horowitz DJ, Tan S (2005) Comparison of affinity tags for protein purification. Protein Expr Purif 41(1):98–105

Lipinszki Z, Vernyik V, Farago N, Sari T, Puskas LG, Blattner FR, Gyorfy Z (2018) Enhancing the translational capacity of E. coli by resolving the codon bias. ACS Synth Biol 7(11):2656–2664

Lobstein J, Emrich CA, Jeans C, Faulkner M, Riggs P, Berkmen M (2012) SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. Microb Cell Fact 11(1):1–16

Lopes M, Belo I, Mota M (2014) Over-pressurized bioreactors: application to microbial cell cultures. Biotechnol Prog 30(4):767–775

Losen M, Frölich B, Pohl M, Büchs J (2004) Effect of oxygen limitation and medium composition on Escherichia coli fermentation in shake-flask cultures. Biotechnol Prog 20(4):1062–1068

Lozano Terol G, Gallego-Jara J, Sola Martinez RA, Martinez Vivancos A, Canovas Diaz M, de Diego Puente T (2021) Impact of the expression system on recombinant protein production in Escherichia coli BL21. Front Microbiol 12:682001

Lu J, Zhang C, Leong HY, Show PL, Lu F, Lu Z (2020) Overproduction of lipoxygenase from Pseudomonas aeruginosa in Escherichia coli by auto-induction expression and its application in triphenylmethane dyes degradation. J Biosci Bioeng 129(3):327–332

Luo M, Zhao M, Cagliero C, Jiang H, Xie Y, Zhu J, Yuan Y (2019) A general platform for efficient extracellular expression and purification of Fab from Escherichia coli. Appl Microbiol Biotechnol 103(8):3341–3353

Ma JKC, Christou P, Chikwamba R, Haydon H, Paul M, Ferrer MP, Wigdorowitz A (2013) Realising the value of plant molecular pharming to benefit the poor in developing countries and emerging economies. Plant Biotechnol J 11(9):1029–1033

Makino T, Skretas G, Georgiou G (2011) Strain engineering for improved expression of recombinant proteins in bacteria. Microb Cell Fact 10(1):1–10

Makrides SC (1996) Strategies for achieving high-level expression of genes in Escherichia coli. Microbiol Rev 60(3):512–538

Malakar P, Venkatesh K (2012) Effect of substrate and IPTG concentrations on the burden to growth of Escherichia coli on glycerol due to the expression of Lac proteins. Appl Microbiol Biotechnol 93:2543–2549

Maleki M, Hajihassan Z (2021) De novo designing a novel signal peptide for secretion of neurturin to the periplasmic space of Escherichia coli. Appl Biochem Microbiol 57(1):S54–S63

Manual, I. BL21-Gold Competent Cells, BL21-Gold (DE3) Competent Cells, and BL21-Gold (DE3) pLysS Competent Cells.

Manual, I. BL21 (DE3) Competent Cells, BL21 (DE3) pLysS Competent Cells, and BL21 Competent Cells.

Mauro VP (2018) Codon optimization in the production of recombinant biotherapeutics: potential risks and considerations. BioDrugs 32(1):69–81

Mellies J, Brems R, Villarejo M (1994) The Escherichia coli proU promoter element and its contribution to osmotically signaled transcription activation. J Bacteriol 176(12):3638–3645

Mergulhão F, Summers DK, Monteiro GA (2005) Recombinant protein secretion in Escherichia coli. Biotechnol Adv 23(3):177–202

Mijakovic I, Petranovic D, Macek B, Cepo T, Mann M, Davies J, Vujaklija D (2006) Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. Nucleic Acids Res 34(5):1588–1596

Millard P, Enjalbert B, Uttenweiler-Joseph S, Portais J-C, Létisse F (2021) Control and regulation of acetate overflow in Escherichia coli. Elife 10:e63661

Minton NP (1984) Improved plasmid vectors for the isolation of translational lac gene fusions. Gene 31(1–3):269–273

Mokry DZ, Abrahão J, Ramos CH (2015) Disaggregases, molecular chaperones that resolubilize protein aggregates. An Acad Bras Ciênc 87:1273–1292

Morris J, Jayanthi S, Langston R, Daily A, Kight A, McNabb DS, Kumar TKS (2016) Heparin-binding peptide as a novel affinity tag for purification of recombinant proteins. Protein Exp Purif 126:93–103

Muhamad Ali MA, ISMAINI, I., Depamede, S. N., Bagus Dwi Hari Setyono, B., Alis Mukhlis, A. M., Muhamad Amin, M. A., & Mohammad Ashari, A. (2015) Stirred bioreactor for the robustness production of recombinant gst Vp28 in fed-batch cultivation of Escherichia coli. Sci Study Res-Chem Chem Eng Biotechnol Food Indus 16(3):245–252

Negrete A, Shiloach J (2017) Improving E. coli growth performance by manipulating small RNA expression. Microb Cell Fact 16(1):1–12

Neumann H, Peak-Chew SY, Chin JW (2008) Genetically encoding Nε-acetyllysine in recombinant proteins. Nat Chem Biol 4(4):232–234

Nguyen TKM, Ki MR, Son RG, Pack SP (2019) The NT11, a novel fusion tag for enhancing protein expression in Escherichia coli. Appl Microbiol Biotechnol 103(5):2205–2216

Nonaka T, Tsurui N, Mannen T, Kikuchi Y, Shiraki K (2019) Non-chromatographic purification of teriparatide with a pH-responsive CspB tag. Protein Expr Purif 155:66–71

O'brien KJ, Li L (2006) Rightful resistance in rural China. Cambridge University Press, Cambridge

Oey M, Lohse M, Kreikemeyer B, Bock R (2009) Exhaustion of the chloroplast protein synthesis capacity by massive expression of a highly stable protein antibiotic. Plant J 57(3):436–445

Ojima-Kato T, Nagai S, Nakano H (2017) N-terminal SKIK peptide tag markedly improves expression of difficult-to-express proteins in Escherichia coli and Saccharomyces cerevisiae. J Biosci Bioeng 123(5):540–546

Overman RC, Debreczeni JE, Truman CM, McAlister MS, Attwood TK (2014) Completing the structural family portrait of the human EphB tyrosine kinase domains. Protein Sci 23(5):627–638

Palmer I, Wingfield PT (2012) Preparation and extraction of insoluble (inclusion-body) proteins from Escherichia coli. Curr Protoc Protein Sci. https://doi.org/10.1002/0471140864.ps0603s70

Palomares LA, Estrada-Moncada S, Ramírez OT (2004) Production of recombinant proteins: challenges and solutions. In: Balbas P, Lorence A (eds) Recombinant gene expression: reviews and protocols. Humana Press, New Jersey, pp 15–51

Palstra R-J, Grosveld F (2012) Transcription factor binding at enhancers: shaping a genomic regulatory landscape in flux. Front Genet 3:195

Parada C, Marians K (1991) Mechanism of DNA A protein-dependent pBR322 DNA replication. DNA A protein-mediated trans-strand loading of the DNA B protein at the origin of pBR322 DNA. J Biol Chem 266(28):18895–18906

Paraskevopoulou V, Falcone FH (2018) Polyionic tags as enhancers of protein solubility in recombinant protein expression. Microorganisms 6(2):47

Paraskevopoulou V, Artiaga VG, Rowlinson R, Winkler GS, Gellert P, Stolnik S, Falcone FH (2019) Introduction of a C-terminal hexa-lysine tag increases thermal stability of the LacDiNac binding adhesin (LabA) exodomain from Helicobacter pylori. Protein Exp Purif 163:105446

Paraskevopoulou V, Alissa M, Hage N, Falcone FH (2022a) Introduction of a hexalysine (6 K) tag can protect from N-terminal cleavage and increase yield of recombinant proteins expressed in the periplasm of E. coli. In: Fruitós EG, Giralt AA (eds) Insoluble proteins: methods and protocols. Springer, New York, pp 155–167

Paraskevopoulou V, Alissa M, Hage N, Falcone FH (2022b) Introduction of a hexalysine (6 K) tag can protect from N-terminal cleavage and increase yield of recombinant proteins expressed in the periplasm of E. coli. In: Fruitós EG, Giralt AA (eds) Insoluble proteins. Springer, New York, pp 155–167

Pereira AM, da Costa A, Dias SC, Casal M, Machado R (2021) Production and purification of two bioactive antimicrobial peptides using a two-step approach involving an elastin-like fusion tag. Pharmaceuticals 14(10):956

Portolano N, Watson PJ, Fairall L, Millard CJ, Milano CP, Song Y, Schwabe JW (2014) Recombinant protein expression for structural biology in HEK 293F suspension cells: a novel and accessible approach. J vis Exp 92:e51897

Posfai G, Plunkett G III, Fehér T, Frisch D, Keil GM, Umenhoffer K, De Arruda M (2006) Emergent properties of reduced-genome Escherichia coli. Science 312(5776):1044–1046

Puetz J, Wurm FM (2019) Recombinant proteins for industrial versus pharmaceutical purposes: a review of process and pricing. Processes 7(8):476

Raducanu VS, Raducanu DV, Ouyang Y, Tehseen M, Takahashi M, Hamdan SM (2021) TSGIT: an N-and C-terminal tandem tag system for purification of native and intein-mediated ligation-ready proteins. Protein Sci 30(2):497–512

Ramanan RN, Tan JS, Mohamed MS, Ling TC, Tey BT, Ariff AB (2010) Optimization of osmotic shock process variables for enhancement of the release of periplasmic interferon-α2b from Escherichia coli using response surface method. Process Biochem 45(2):196–202

Reckinger AR, Jeong KS, Khodursky AB, Hiasa H (2007) RecA can stimulate the relaxation activity of topoisomerase I: molecular basis of topoisomerase-mediated genome-wide transcriptional responses in Escherichia coli. Nucleic Acids Res 35(1):79–86

Restrepo-Pineda S, Pérez NO, Valdez-Cruz NA, Trujillo-Roldán MA (2021) Thermoinducible expression system for producing recombinant proteins in Escherichia coli: advances and insights. FEMS Microbiol Rev 45(6):023

Rinas U, Garcia-Fruitós E, Corchero JL, Vázquez E, Seras-Franzoso J, Villaverde A (2017) Bacterial inclusion bodies: discovering their better half. Trends Biochem Sci 42(9):726–737

Robichon C, Luo J, Causey TB, Benner JS, Samuelson JC (2011) Engineering Escherichia coli BL21 (DE3) derivative strains to minimize E. coli protein contamination after purification by immobilized metal affinity chromatography. Applied Environ Microbiol 77(13):4634–4646

Rosano GL, Ceccarelli EA (2009) Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted Escherichia coli strain. Microb Cell Fact 8(1):1–9

Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in Escherichia coli: advances and challenges. Front Microbiol 5:172

Rouches MV, Xu Y, Cortes LBG, Lambert G (2022) A plasmid system with tunable copy number. Nat Commun 13(1):3908

Sahdev S, Khattar SK, Saini KS (2008) Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. Mol Cell Biochem 307(1):249–264

Saïda F, Uzan M, Odaert B, Bontems F (2006) Expression of highly toxic genes in E. coli: special strategies and genetic tools. Curr Protein Pept Sci 7(1):47–56

San-Miguel T, Pérez-Bermúdez P, Gavidia I (2013) Production of soluble eukaryotic recombinant proteins in E. coli is favoured in early log-phase cultures induced at low temperature. Springerplus 2(1):1–4

Santos BD, Morones-Ramirez JR, Balderas-Renteria I, Casillas-Vega NG, Galbraith DW, Zarate X (2019) Optimizing periplasmic expression in Escherichia coli for the production of recombinant proteins tagged with the small metal-binding protein SmbP. Mol Biotechnol 61(6):451–460

Sarduy ES, Muñoz AC, Trejo SA, Planes M, d. l. A. C. (2012) High-level expression of Falcipain-2 in Escherichia coli by codon optimization and auto-induction. Protein Expr Purif 83(1):59–69

Savelyeva AV, Nemudraya AA, Podgornyi VF, Laburkina NV, Ramazanov YA, Repkov AP, Richter VA (2017) Analysis of the efficiency of recombinant Escherichia coli strain cultivation in a gas-vortex bioreactor. Biotechnol Appl Biochem 64(5):712–718

Schlegel S, Löfblom J, Lee C, Hjelm A, Klepsch M, Strous M, de Gier J-W (2012) Optimizing membrane protein overexpression in the Escherichia coli strain Lemo21 (DE3). J Mol Biol 423(4):648–659

Schlegel S, Genevaux P, de Gier J-W (2015) De-convoluting the genetic adaptations of E. coli C41 (DE3) in real time reveals how alleviating protein production stress improves yields. Cell Rep 10(10):1758–1766

Schumann W (2007) Production of recombinant proteins in Bacillus subtilis. Adv Appl Microbiol 62:137–189

Shang F, Wang H, Zhang D, Wang W, Yu J, Xue T (2021) Construction of an AI-2 quorum sensing induced heterologous protein expression system in Escherichia coli. PeerJ 9:e12497

Shanmugaraj B, Malla A, Phoolcharoen W (2020) Emergence of novel coronavirus 2019-nCoV: need for rapid vaccine and biologics development. Pathogens 9(2):148

Sharma AK, Shukla E, Janoti DS, Mukherjee KJ, Shiloach J (2020) A novel knock out strategy to enhance recombinant protein expression in Escherichia coli. Microb Cell Fact 19(1):1–10

Shilling PJ, Mirzadeh K, Cumming AJ, Widesheim M, Köck Z, Daley DO (2020) Improved designs for pET expression plasmids increase protein production yield in Escherichia coli. Commun Biol 3(1):1–8

Singh SM, Panda AK (2005) Solubilization and refolding of bacterial inclusion body proteins. J Biosci Bioeng 99(4):303–310

Singh A, Upadhyay V, Upadhyay AK, Singh SM, Panda AK (2015) Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. Microb Cell Fact 14(1):1–10

Song J-A, Lee D-S, Park J-S, Han K-Y, Lee J (2011) A novel Escherichia coli solubility enhancer protein for fusion expression of aggregation-prone heterologous proteins. Enzyme Microb Technol 49(2):124–130

Sousa DA, Mulder KCL, Nobre KS, Parachin NS, Franco OL (2016) Production of a polar fish antimicrobial peptide in Escherichia coli using an ELP-intein tag. J Biotechnol 234:83–89. https://doi.org/10.1016/j.jbiotec.2016.07.021

Striedner G, Pfaffenzeller I, Markus L, Nemecek S, Grabherr R, Bayer K (2010) Plasmid-free T7-based Escherichia coli expression systems. Biotechnol Bioeng 105(4):786–794

Strocchi M, Ferrer M, Timmis KN, Golyshin PN (2006) Low temperature-induced systems failure in Escherichia coli: insights from rescue by cold-adapted chaperones. Proteomics 6(1):193–206

Stüber D, Bujard H (1981) Organization of transcriptional signals in plasmids pBR322 and pACYC184. Proc Natl Acad Sci 78(1):167–171

Studier FW, Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol 189(1):113–130

Tahara N, Tachibana I, Takeo K, Yamashita S, Shimada A, Hashimoto M, Suzuki F (2020) Boosting auto-induction of recombinant proteins in Escherichia coli with glucose and lactose additives. Protein Pept Lett 28(10):1180

Tegel H, Tourle S, Ottosson J, Persson A (2010) Increased levels of recombinant human proteins with the Escherichia coli strain Rosetta (DE3). Protein Expr Purif 69(2):159–167

Terpe K (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 72(2):211–222

Thomas CM (2021) Plasmid incompatibility. In: Bell E (ed) Molecular life sciences: an encyclopedic reference. Springer, New York, pp 1–3

Tian Z-G, Dong T-T, Yang Y-L, Teng D, Wang J-H (2009) Expression of antimicrobial peptide LH multimers in Escherichia coli C43 (DE3). Appl Microbiol Biotechnol 83:143–149

Tochio N, Umehara T, Nakabayashi K, Yoneyama M, Tsuda K, Shirouzu M, Sasazuki T (2015) Solution structures of the DNA-binding domains of immune-related zinc-finger protein ZFAT. J Struct Funct Genom 16(2):55–65

Ukkonen K, Mayer S, Vasala A, Neubauer P (2013) Use of slow glucose feeding as supporting carbon source in lactose autoinduction medium improves the robustness of protein expression at different aeration conditions. Protein Expr Purif 91(2):147–154

Velappan N, Sblattero D, Chasteen L, Pavlik P, Bradbury AR (2007) Plasmid incompatibility: more compatible than previously thought? Protein Eng, Des Sel 20(7):309–313

Ventura S, Villaverde A (2006) Protein quality in bacterial inclusion bodies. Trends Biotechnol 24(4):179–185

Vera A, González-Montalbán N, Arís A, Villaverde A (2007) The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. Biotechnol Bioeng 96(6):1101–1106

Verma D, Daniell H (2007) Chloroplast vector systems for biotechnology applications. Plant Physiol 145(4):1129–1143

Vikström, D., Klepsch, M., Wagner, S., Samuelson+, J. C., & de Gier, J.-W. E. coli Lemo21 (DE3).

Wang Y, Ding H, Du P, Gan R, Ye Q (2005) Production of phoA promoter-controlled human epidermal growth factor in fed-batch cultures of Escherichia coli YK537 (pAET-8). Process Biochem 40(9):3068–3074

Wang G-X, Tan M-K, Rakshit S, Saitoh H, Terauchi R, Imaizumi T, Tominaga T (2007) Discovery of single-nucleotide mutations in acetolactate synthase genes by Ecotilling. Pestic Biochem Physiol 88(2):143–148

Wang J, Yu H, Tian S, Yang H, Wang J, Zhu W (2017) Recombinant expression insulin-like growth factor 1 in Bacillus subtilis using a low-cost heat-purification technology. Process Biochem 63:49–54

Wang Z, Zhang M, Lv X, Fan J, Zhang J, Sun J, Shen Y (2018) GroEL/ES mediated the in vivo recovery of TRAIL inclusion bodies in Escherichia coli. Sci Rep 8(1):1–10

Wang S, Lin R, Ren Y, Zhang T, Lu H, Wang L, Fan D (2020) Non-chromatographic purification of thermostable endoglucanase from Thermotoga maritima by fusion with a hydrophobic elastin-like polypeptide. Protein Expr Purif 173:105634

Ward OP (2012) Production of recombinant proteins by filamentous fungi. Biotechnol Adv 30(5):1119–1139

Wegerer A, Sun T, Altenbuchner J (2008) Optimization of an *E. coli* L-rhamnose-inducible expression vector: test of various genetic module combinations. BMC Biotechnol 8(1):1–12

Werkmeister JA, Ramshaw JA (2012) Recombinant protein scaffolds for tissue engineering. Biomed Mater 7(1):012002

Weuster-Botz D, Altenbach-Rehm J, Arnold M (2001) Parallel substrate feeding and pH-control in shaking-flasks. Biochem Eng J 7(2):163–170

Wibowo D, Zhao C-X (2019) Recent achievements and perspectives for large-scale recombinant production of antimicrobial peptides. Appl Microbiol Biotechnol 103(2):659–671

Wingfield PT (2015) Overview of the purification of recombinant proteins. Curr Protoc Protein Sci. https://doi.org/10.1002/0471140864.ps0601s80

Woestenenk EA, Hammarström M, van den Berg S, Härd T, Berglund H (2004) His tag effect on solubility of human proteins produced in *Escherichia coli*: a comparison between four expression vectors. J Struct Funct Genom 5(3):217–229

Xu X, Liang K, Niu Y, Shen Y, Wan X, Li H, Yang Y (2018) BAH1 an E3 ligase from arabidopsis thaliana stabilizes heat shock factor σ32 of *Escherichia coli* by interacting with DnaK/DnaJ chaperone team. Curr Microbiol 75(4):450–455

Yang W-J, Yanofsky C (2005) Effects of tryptophan starvation on levels of the trp RNA-binding attenuation protein (TRAP) and anti-TRAP regulatory protein and their influence on trp operon expression in Bacillus subtilis. J Bacteriol 187(6):1884–1891

Yoon K, Silver S (1991) A second gene in the Staphylococcus aureus cadA cadmium resistance determinant of plasmid pI258. J Bacteriol 173(23):7636–7642

Young CL, Britton ZT, Robinson AS (2012) Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. Biotechnol J 7(5):620–634

Zare H, Sadeghi HMM, Akbari V (2019) Optimization of fermentation conditions for reteplase expression by *Escherichia coli*

using response surface methodology. Avicenna J Med Biotechnol 11(2):162

Zarkar N, Nasiri Khalili MA, Khodadadi S, Zeinoddini M, Ahmadpour F (2020) Expression and purification of soluble and functional fusion protein DAB389IL-2 into the *E. coli* strain Rosetta-gami (DE3). Biotechnol Appl Biochem 67(2):206–212

Zhang W, Lu J, Zhang S, Liu L, Pang X, Lv J (2018) Development an effective system to expression recombinant protein in *E. coli* via comparison and optimization of signal peptides: expression of Pseudomonas fluorescens BJ-10 thermostable lipase as case study. Microb Cell Fact 17(1):1–12

Zhang Z-X, Nong F-T, Wang Y-Z, Yan C-X, Gu Y, Song P, Sun X-M (2022) Strategies for efficient production of recombinant proteins in *Escherichia coli*: alleviating the host burden and enhancing protein activity. Microb Cell Fact 21(1):191. https://doi.org/10.1186/s12934-022-01917-y

Zorko M, Jerala R (2010) Production of recombinant antimicrobial peptides in bacteria. Antimicrobial Peptides. Springer, Berlin, pp 61–76