

Advancements in Inducer Systems for Recombinant Protein Production in *E. coli*



Jaya A. Gupta, Kathiresan Pandi, and Anurag S. Rathore

Contents

1	Tuning Recombinant Protein Production in <i>E. coli</i>	2
2	Promoter Systems of <i>E. coli</i>	3
3	Inducer Systems in <i>E. coli</i>	4
3.1	Lactose Induction System	4
3.2	Arabinose Induction System	7
3.3	Rhamnose Induction System	11
3.4	pNEW Induction System	14
4	Conclusions and Future Trends	16
	References	17

Abstract *Escherichia coli* (*E. coli*) is the predominant industrial microorganism and cost-effective microbial cell “factory” for producing commercial products, such as metabolites, enzymes, biochemicals, and high-value biotherapeutics. It is the primary workhorse for non-glycosylated recombinant proteins production as it is the production organism used to produce around 30% of biopharmaceutical products. Heterologous protein expression in *E. coli* is often a preferred choice for many reasons: it is economical, provides a fast growth rate, and can reach high cell density and high product titers. Being widely studied, an abundance of biochemical and physiological knowledge is also available. Also, various genetic tools are available to manipulate *E. coli*, including a significantly extensive catalog of expression plasmids and engineered strains and several cultivation strategies. Selecting appropriate expression plasmids and promoter strength is crucial for efficient protein production. Promoter strength depends on the promoter components that can be either constitutive or inducible. Constitutive promoters allow continuous gene expression, which can cause a metabolic burden affecting cell growth, whereas inducible promoters control gene expression through specific inducers. The existing expression systems are diverse and complex, and its selection depends on the nature

J. A. Gupta · K. Pandi · A. S. Rathore (✉)
Department of Chemical Engineering, Indian Institute of Technology, Hauz Khas, India
e-mail: asrathore@biotechcmz.com

of target protein that requires in particular cases to be produced at high concentrations, or at different cell growth phases, or at particular location whether in the cytoplasmic or periplasmic space or in some cases, secretion into medium.

1 Tuning Recombinant Protein Production in *E. coli*

E. coli served as the workhorse of modern biotechnology. It is the predominant industrial microorganism and most frequently used prokaryotes to produce commercial products, such as metabolites, enzymes, biochemicals, and high-value biotherapeutics (Gupta and Shukla 2017; Walsh 2018). A substantial amount of knowledge has been generated over the past 30 years about developing *E. coli* strains for recombinant proteins expression in cytoplasm, periplasm, or secretion into medium (Rosano et al. 2019). Moreover, many expression vectors containing regulated promoters, signal sequences, antibiotic selection and tags for efficient protein purification have been developed (Rosano et al. 2019). Despite all these advancements, process optimization is required to achieve a high titer of recombinant proteins. Due to variations in vectors, gene expression and its products, promoter strength, plasmid copy number, and host–vector interactions, process optimization is often a tedious and time-consuming task (Sahdev et al. 2008). Moreover, there are recombinant proteins that are toxic to the cell. In *E. coli*, it has been confirmed that the expression of heterologous proteins in uncontrolled fashion results in the initiation of metabolic stress responses, undesired product loss, and eventually cell death (Bentley et al. 1990; Bonomo and Gill 2005). Therefore, the complex interplay of media processing, the nature of the product, and the expression system determines productivity, expression level, and product quality. A high level of foreign gene expression that results in 25–35% of total cellular protein may sometimes lead to significant genetic instability (Bentley et al. 1990; Hoffmann and Rinas 2004). When biomass is high, promoters in the plasmid should be tightly regulated (Saïda et al. 2006). Multiple different transcriptional regulations of the promoter are found in *E. coli*. Perhaps, the most common control mechanism is the binding of regulatory proteins to the specific region of the promoter. However, there exist other mechanisms of regulation, such as transcriptional attenuation, anti-termination, anti-sense RNA, variation in sigma factors, and anti-sigma-factors (Tropp 2008). In practice, based on carbon source availability, promoter activity can be controlled by regulatory proteins (Tropp 2008). Transcriptional proteins control the activity of the promoter in two ways. Negatively controlled systems prevent RNA polymerase binding to the promoter using regulatory elements, thereby repressing transcription. On the other hand, regulatory elements in positively controlled systems allow the binding of RNA polymerase to the promoter through activator proteins, thus permitting transcription (Tropp 2008). Either inducer or repressor can control the negatively and positively regulated promoters (Rosano and Ceccarelli 2014). In addition, a negatively regulated inducer system is activated by inhibiting the binding of the repressor to the operator. In contrast, a positively

regulated inducer system is activated in the presence of an effector by inducer binding to the activator (Tropp 2008). Both inducer systems have been successfully used to produce various heterologous proteins from *E. coli* (Sørensen and Mortensen 2005; Rosano and Ceccarelli 2014).

This chapter reviews the most popular and industrially useful promoter systems, such as lactose, arabinose, and rhamnose.

2 Promoter Systems of *E. coli*

A commercially useful promoter must be strong, tightly regulated, have a low basal expression level, must be strain independent, require simple and cost-effective induction method for gene expression, and should be independent on the commonly used ingredients of culturing media. Among the various promoter systems available for protein expression in *E. coli*, only a few of them are used. One example is constitutive promoters that have been opted for heterologous gene expression in *E. coli*. As gene expression in *E. coli* is based on multicopy plasmids, maintaining high-level gene expression constitutively along with cellular growth would cause a metabolic burden to the cells leading to the premature arrest of the system. Therefore, the constitutive promoters would not be ideal for heterologous protein production.

Another example is auto-inducible promoters, which do not require any external inducer for gene expression. Auto-inducible promoters aim toward large-scale protein production as they get induced either at the late log phase or at the stationary phase (i.e., a growth stage when cells are not dividing but remain metabolically active, and the genes required for the cell survival are expressed during this phase). One category of auto-inducible promoters is stationary-phase promoters. It has been reported that nearly 20% of the *E. coli* genes are expressed at an increased level in this phase (Rava et al. 1999). Stationary phase promoters show enhanced activity in the stationary phase, but little or no activity in the exponential phase. Nonetheless, many of these promoters exhibit low strength and, therefore are not popular for constructing expression vectors (Jaishankar and Srivastava 2017).

Starvation promoters constitute another category that has been widely used in *E. coli* (Shin and Seo 1990). These promoters are induced by starvation of a particular molecule/metabolite promoting the induction of heterologous genes. Alkaline phosphate promoter (*phoA*) is an example of starvation promoters, where phosphate starvation can be used to control the expression of heterologous genes during the non-growth phase (Keasling 1999). It is the starvation of external inorganic phosphate (P_i) in growth medium which induces the phosphate starvation response in the cell. This causes a release of alkaline phosphatase in periplasmic space to generate P_i from organophosphates because the physiological role of this enzyme is to perform hydrolysis of organic phosphate compounds present in growth medium to fulfill the cell's requirement of inorganic phosphate (Wanner 1996). Also, there is no known internal sensor for phosphate, so internal phosphate stores

cannot be sensed by the cells. It is the external phosphate starvation that drives the heterologous protein production from these promoters. Starvation promoters are particularly anticipated in bioremediation-related applications, where removal of environmental contaminants can control induction from these promoters (Keasling 1999).

A controlled expression system that could be induced at a desired time and condition is advantageous for efficient recombinant production (Nieto et al. 2000).

3 Inducer Systems in *E. coli*

Repressors, activators, and inducers are three important regulatory proteins that control gene expression in operons. Repressor protein inhibits gene transcription in response to an external stimulus. Activator protein increases gene transcription in response to an external signal. Inducers activate or deactivate transcription depending on the sugar availability and requirement in the cells (Tropp 2008). In *E. coli*, many genes are mostly expressed in a constitutive manner, meaning that the gene expression is always switched “on.” In other cases, when their transcribed proteins are required by the cells, their expression will be controlled and hence referred as inducible expression. The most studied and popular promoters in bacteria are those regulating operons for sugar metabolism, e.g., lacZYA operon, araBAD operon, and rhaBAD operon (Rosano and Ceccarelli 2014). These promoters are inducible promoter system which is beneficial for recombinant protein production.

3.1 Lactose Induction System

Inducible sugars can control both the negatively and positively regulated promoters. Lac operon is one of the well-studied systems in *E. coli*. Regulation of lac operon is directed by the lac repressor gene, lacI (Kercher et al. 1997). In the absence of lactose, lacI binds to the promoter and prevents the transcription of the lactose operon genes (Fig. 1). In the presence of inducer, allolactose produced by galactosidase (lacZ) using lactose, the repressor weakens its binding affinity to the operator (Wheatley et al. 2013) (Fig. 2). Galactosidase also hydrolyzes lactose to galactose and glucose, and metabolizes them into the central carbon metabolism. Lactose permease (lacY) of lac operon allows lactose transport into the cell through membranes. The galactoside acetyltransferase (lacA) enzyme metabolizes acetyl-CoA to galactoside by transferring the acetyl group (Juers et al. 2012). Catabolite activator protein (CAP, also known as cyclic adenosine monophosphate (cAMP) receptor protein or CRP) positively controls its transcription (Wheatley et al. 2013). Efficient transcription is possible only when the CAP-cAMP complex binds to the -35 region upstream of the promoter (Simpson 1980; Malan and McClure 1984). A low concentration of glucose facilitates the increase of cAMP levels in the cells and

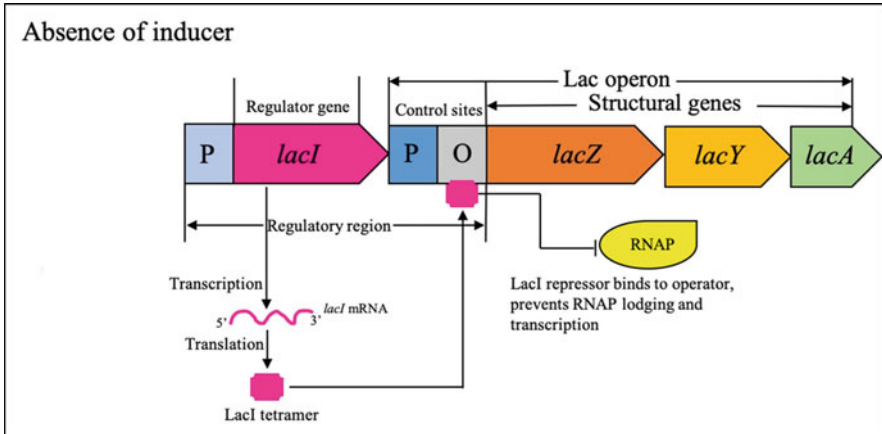


Fig. 1 Repressed lactose operon. This shows the repressed state of lactose operon, where absence of inducer allows binding of LacI repressor to the operator, thus inhibiting RNA polymerase mediated transcription of structural genes

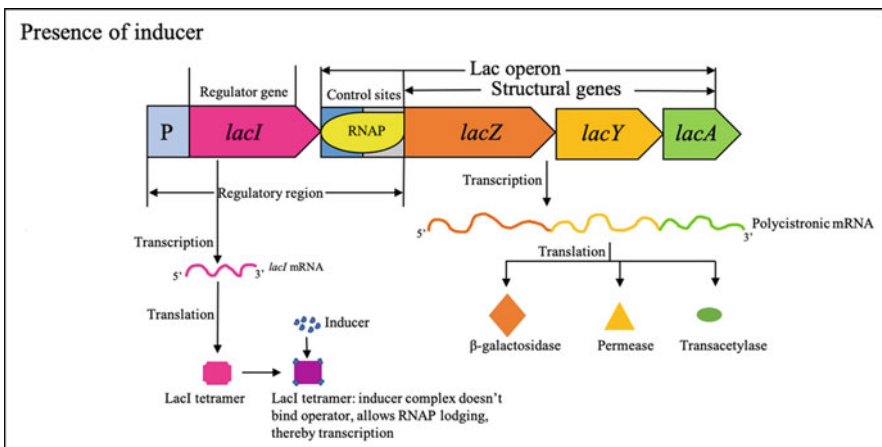


Fig. 2 Induced lactose operon. This shows the induced state of lactose operon, where presence of inducer forms a complex with LacI repressor, which induces conformational changes in repressor and repressor no longer be able to bind the operator, thus promoting RNA polymerase mediated transcription of structural genes

allows transcription of lac operon genes (Inada et al. 1996). Modified lac promoters are routinely used in the *E. coli* vectors for heterologous gene expression. The strength of the lactose promoter was modified by fusing the -35 region of the tryptophan promoter from *trp* operon with the -10 region of the lactose promoter (De Boer et al. 1983; Neubauer et al. 1991). This modified promoter is referred to as tac promoter, which is tenfold more efficient than lac promoter, particularly on

multicopy plasmids. Therefore, *tac* promoter is used for recombinant proteins production at a commercial scale (De Boer et al. 1983).

The *lacUV5* promoter is another modified promoter derived from the *lac* operon to guide the expression of heterologous genes on a plasmid. It is similar to the *lac* promoter, comprising just two base-pair mutations in the -10 hexamer region. It independently works regardless of activators or other cis-regulatory elements (Brosius et al. 1983; Shibui et al. 1988). *LacI* repressor alone can control *lacUV5* promoter expression. The *lacI* gene found in the chromosome leads to very few *lac* repressor molecules (Brosius et al. 1983). Repressors will not occupy the *lac* operators due to low concentration, resulting in basal level expression from *lac*-based promoters. To control basal level expression, three operators were inserted into a plasmid with the correct spacing leads to complete inhibition of the promoter (Oehler et al. 1990, 1994). Several studies have been performed to enhance the strength of *lac* repression in *lac* promoter system. The first study was the isolation of the *lacIq* mutants, which resulted in a tenfold increase in *LacI* expression (Calos and Miller 1981). When plasmids carry the *lacIq* gene, they are less dependent on *E. coli* (Brosius et al. 1983). That *E. coli* strains with a high concentration of *LacI* are still not sufficient for full lactose induction but are still fully inducible with non-metabolizable IPTG. Nevertheless, IPTG is not recommended for large-scale recombinant proteins production owing to its high cost, cellular toxicity and regulatory issues.

Another modified *lac* promoter system with widespread use is pET vector-based *lac* promoter. The pET vectors contain a strong T7 promoter controlled by T7 RNA polymerase (Studier 2014). The T7 RNA polymerase is chromosomally integrated using the prophage DE3 and controlled by *lacUV5* promoter in the genome (Tabor and Richardson 1985). However, leakiness is a significant issue in the pET vectors that can be improved by adding either a copy of *lacI* gene in the expression vector, a *lac* operator downstream of the T7 promoter, or T7 lysozyme that controls the expression of T7 RNA polymerase independently (Studier 1991, 2014). These modifications lead to the development of commercially successful promoters to produce many heterologous proteins in *Escherichia coli* (Samuelson 2011). The lactose-based promoters are predominantly used to produce high expression levels of recombinant proteins in *E. coli*, and many lactose- or IPTG-based vectors have been designed to optimize recombinant proteins expression (Samuelson 2011). Hundreds of studies on lactose-based medium and process optimization have been performed to improve recombinant protein expression.

Recently, auto-induction has been gaining attention to produce recombinant proteins in *E. coli*. Auto-induction works when cells shift from an un-induced to an induced gene expression under metabolic control of the cell. The auto-inducing method comprises an optimal combination of glucose, lactose, glycerol and other essential nutrients in the medium (Chen et al. 2014). After glucose is exhausted, lactose is converted to allolactose, thus inducing the expression of recombinant protein through lactose-based promoter systems (Nie et al. 2013). Many proteins have been successfully produced using auto-induction in *E. coli*. For instance, auto-induction method has been demonstrated with enhanced expression of recombinant

tissue plasminogen activator (tPA) (Fathi-Roudsari et al. 2018). By optimizing parameters like temperature and medium composition, recombinant protein expression was significantly improved. This study found that a decrease in temperature and auto-induction medium highly influenced the yield of active soluble tissue plasminogen activator in *E. coli*. Using a highly enriched auto-induction medium, biologically active tPA was improved by 30%. Another example is the recombinant production of pullulanase which functionally participates in specific hydrolysis of starch processing. Recombinant pullulanase was successfully expressed in *E. coli* with 14 U/mL pullulanase activity when induced with IPTG (Nie et al. 2013).

Basal expression of recombinant proteins is prevalent in the lactose-based expression system, which might be disadvantageous to the host and ultimately lead to system instability issues and negatively affect the accumulation of recombinant protein. Therefore, repression of lactose operator in plasmid and genome of *E. coli* was constructed, and lactose operator was bound with lactose repressor to prevent T7 RNA polymerase activity and expression of target protein before induction. Auto-induction strategy has been used to considerably improve the recombinant pullulanase activity and yields (i.e., up to 580 U/mL) (Nie et al. 2013). Nitrile hydratase has been used to produce valuable chemicals such as acrylamide, nicotinamide, and 5-cyanovaleramide (Gupta et al. 2010). Recombinant nitrile hydratase was expressed and produced in *E. coli* using auto-induction. A glycerol limited fed-batch process produced 2170 U/mL of nitrile hydratase and biotransformation with recombinant nitrile hydratase exhibited a productivity of 187 g of nicotinamide/g dry cell weight/h. Auto-induction has been also used to produce industrial enzyme nitrile hydratase. *E. coli* has been considered the industrial workhorse for producing N-glycosylated proteins since the breakthrough work of engineering the glycosylation pathway from *Campylobacter jejuni* (Wacker et al. 2002). It is well known that N-glycosylation decides the biological activity of glycoproteins. Therefore, efficient glycosylation is a prerequisite, and the developing process is critically important. Different glycosylation sequons in *E. coli* carrying the heterologous *pgl* locus from *C. jejuni* have been used. Compared to IPTG induction, auto-induction method produced glycoproteins with 100% glycosylation efficiency (Ding et al. 2017).

3.2 *Arabinose Induction System*

Positively regulated systems are defined by transcriptional regulatory elements and activators, coupled with RNA polymerase binding to the specific region in the promoter, thus guiding transcription. Positively regulated systems are functionally characterized by a slower induction response.

pBAD is a classic model of a positively regulated promoter system induced by L-arabinose (Schleif 2000). *araE* and *araFGH* are two arabinose transport systems present in the arabinose operon and function to convert arabinose into xylulose-5-phosphate by the action of enzymes encoded by ribulokinase (*araB*), isomerase (*araA*), and epimerase (*araD*) (Schleif 2000). The *araBAD* genes are located close

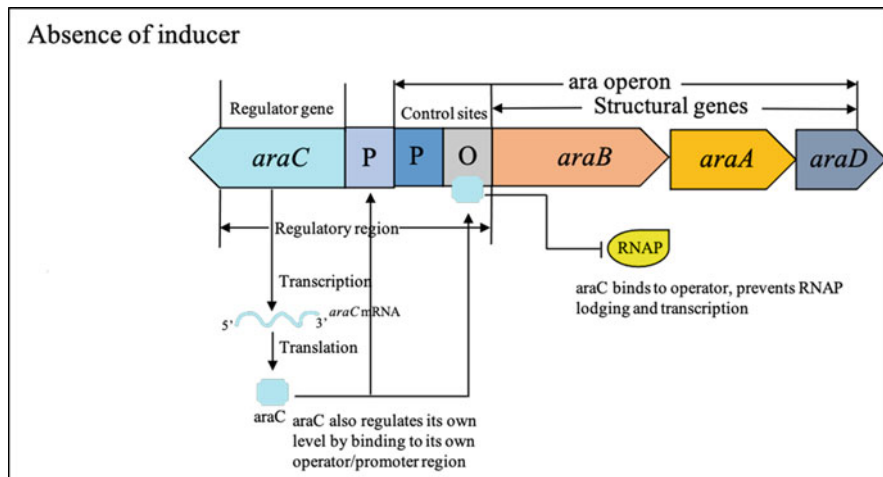


Fig. 3 Repressed arabinose operon. This shows the repressed state of arabinose operon, where the absence of inducer allows binding of *araC* repressor to the operator, thus inhibiting RNA polymerase mediated transcription of structural genes

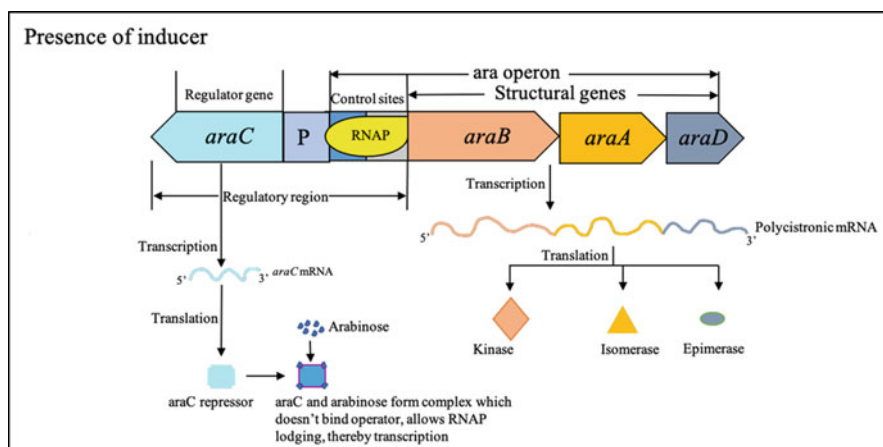


Fig. 4 Induced arabinose operon. This shows the induced state of arabinose operon, where the presence of inducer forms a complex with *araC* repressor, which induces conformational changes in repressor and repressor no longer be able to bind the operator, thus promoting RNA polymerase mediated transcription of structural genes

to *araFGH* and *araE* in the operon. *AraC* protein regulates its own expression and *ara* operon genes (Figs. 3 and 4). The *araC* gene is located close to the *araBAD* operon but in opposite orientation (Johnson and Schleif 1995). It belongs to the *araC/xylS* family, which belongs to a positively regulated system (Mari et al. 1997). The *araBAD* gene products are induced approximately 300 times by *AraC* compared to the un-induced level. Arabinose promoter has the lowest basal expression level, but

the efficiency of repression is gene-dependent, and the repression is not tightly regulated. Induction of the genes also reflects catabolic repression, and induction is delayed due to the presence of glucose and some other sugars. Repression is controlled by the cAMP, which in turn affects the CRP activity. Studies of the araBAD promoter confirm that to activate transcription araC protein binding site must overlap the -35 region of the promoter by 4 bp. Moreover, the two half transcription sites recognized by AraC protein must be in the same direct repeat orientation to activate transcription. This is due to specific contacts made between RNA polymerase and AraC protein at pBAD promoter. In vivo studies in arabinose promoters found that araFGH expression, another arabinose promoter, is more sensitive to catabolite repression but not to arabinose concentration in comparison with araE and araBAD promoters. The relative levels of inducibility in wild-type cells of araBAD, araFGH, and araE have been reported to be 6.5, 5, and 1, respectively (Johnson and Schleif 1995). Researchers have concluded that all promoter systems are rapidly responsive and inducible by 0.53 mM of arabinose, and the arrangement of araC binding site is different among araBAD, araFGH, and araE promoters (Johnson and Schleif 1995). It is mechanistically known that CRP-cAMP and araC bind to specific promoter sites. AraC protein is known to bind to three sites within promoters such as araI, araO1, and araO2 (Ogden et al. 1980; Lee et al. 1981). However, initial studies found that only the transcriptionally active AraC protein could bind to the araI site (Ogden et al. 1980; Lee et al. 1981; Miyada et al. 1984). Later studies have proven that when arabinose is present, AraC protein can bind to all three sites (Hahn et al. 1984). It is known that CRP can bind to pBAD promoter region in *E. coli* and cause repression. Few mutational studies show that even in the absence of CRP, arabinose promoter is still threefold repressed when AraC protein is present (Hahn et al. 1984). Results suggest that some other mechanism may be involved in the repression in the absence of CRP-cAMP. When cells lack adenylyl cyclase (*cya*), pBAD promoter expression is poor upon arabinose addition. However, when pBAD repression is abolished by deleting the operator (*areO2*) site in the promoter, elevated expression of the promoter in *cya* deleted strain is observed. Also, when araC expression was increased, it proportionally increased the expression of the pBAD promoter (Hahn et al. 1984). This observation is likely due to stimulation by CRP. Therefore, Hahn et al. (1984) concluded that pBAD promoter expression is reduced in the absence of CRP-cAMP upon the addition of arabinose. In the absence of CRP-cAMP repression, araC-arabinose is prevented from functioning at araI site to induce the expression.

The pBAD promoter from *E. coli* is used to express many recombinant proteins. Most recombinant plasmids contain araC gene due to low araC expression from a single chromosomal copy in *E. coli*, considering that many operators present on multicopy plasmids would require more araC repressor. Arabinose promoter varies between 250- and 1300-fold induction ratio in expression vectors and the basal level is very low due to carbon catabolite repression when glucose is present (Guzman et al. 1995). Although pBAD promoter is around 2.5- to 4.5-fold weaker than the tac promoter. Despite their low strength, recombinant proteins can accumulate up to 30% of total cell protein, depending on translation and kinetics. The induction can be

modulated by adding suboptimal concentrations of arabinose. Arabinose-based induction system has been successfully used to produce many recombinant products. Researchers have developed a high cell density fed-batch process to produce 5-hydroxymethylfurfural oxidase (HMFO) and eugenol oxidase (EUGO) (Román et al. 2020). They achieved high cell density culture (HCDC) using two stages of fed-batch process cultivated with glucose for improving biomass. Once glucose was exhausted, arabinose was added for induction. After inducing with arabinose, glycerol was used as an additional carbon source, and this resulted in an eightfold improvement in protein yield compared to IPTG-based inducer system (Román et al. 2020). Arabinose-based promoter expression has been tested for difficult to express proteins such as antibody fragments, membrane proteins, and vaccines. In one study, the expression of human recombinant tetanus toxoid and antibody fragment has been expressed under the control of pBAD and compared with the Lac promoter (Clark et al. 1997). Recombinant proteins were accumulated in a soluble fraction of periplasmic space, which is desirable for downstream processing. Compared with lac promoter induction using IPTG, production of Fab could be more strictly repressed under the control of arabinose promoter and low concentrations of arabinose were required, which is a significant advantage where production of a highly expressed Fab is toxic to the *E. coli* host.

E. coli NEB10 β strain (Miret et al. 2020) has been used for the expression of the complex fusion protein phosphite dehydrogenase-cyclohexanone monooxygenase (PTDH-CHMO) using arabinose promoter. The fed-batch process was developed using a chemically defined medium supplemented with amino acids and glycerol. It resulted in a 9.2-fold improvement of the recombinant protein yields than in a complex medium and an accumulation of up to 2 g/L of PTDH-CHMO fusion protein after 6 h of induction. Arabinose-based expression was also demonstrated for the production of the alcohol dehydrogenase (ADH) enzyme using the same NEB10 β strain (Miret et al. 2020). Arabinose expression system (Eberhardt et al. 2017) was tested along with T7, T5, tac, lactose expression system to produce thermostable steryl glucosidase in *Escherichia coli*. Based on comparative analysis, arabinose expression system offered a 40% improvement in shake flasks. Further process optimization in bioreactor enhanced the production of steryl glucosidase to 200-fold with a maximum activity of 260 U/mL after 6 h of arabinose induction in high cell density culture. Arabinose expression system has also been successfully demonstrated to produce high titer of inclusion bodies in *E. coli*. Human antimicrobial peptide LL-37 has been produced as a fusion protein using arabinose. The results show that active LL-37 can be produced at 1 g/L (Krahulec et al. 2010).

Many commercial vectors are constructed to induce recombinant proteins using arabinose. Apart from gene expression vectors, *E. coli* BL21-AI strain has been specifically developed to utilize the arabinose expression system. This strain carries a genomic insertion of a cassette containing the T7 RNA polymerase gene in the araB locus, allowing the expression of T7 RNA polymerase to be regulated by the araBAD promoter. BL21-AI strain is suitable for high-level recombinant protein expression from any T7-based expression vector. Levels of T7 RNA polymerase can be tightly regulated under arabinose control (Narayanan et al. 2011). High levels of

heterologous gene expression from T7 promoter led to low basal expression and makes it possible to express recombinant genes whose products are toxic for the host cells (Muntari et al. 2012). Another merit of arabinose promoter is its ability to fine-tune the expression level of recombinant proteins depending on its concentration in the medium (Guzman et al. 1995). The efficiency of the arabinose expression system is further increased by the higher stability of the plasmid DNA. These characteristic features of arabinose induction could be potentially used to produce recombinant proteins, which tend to form inclusion bodies when produced at high levels in *E. coli*.

3.3 Rhamnose Induction System

Rhamnose induction is another model for a positive regulation system. Rhamnose induction is a two-step tight control induction process. Its transporter (*rhaT*) controls L-rhamnose uptake and rhamnose is first converted to l-rhamnulose by an isomerase (*rhaB*), and then metabolized to rhamnulose-5-phosphate by a kinase (*rhaA*), and hydrolyzed by an aldolase (*rhaD*). The products dihydroxyacetone phosphate and L-lactaldehyde are consumed by other metabolic pathways (Tobin and Schleif 1987; Moralejo et al. 1993). The gene *rhaBAD* forms an operon located close to the *rhaT* gene (Fig. 5). Two activator proteins, *rhaR* and *rhaS*, are located upstream to *rhaBAD* which are arranged in opposite orientations. When rhamnose is consumed by the cell, *rhaR* protein binds with the rhamnose and becomes activated (Fig. 6). *rhaR* induces its own operon *rhaSR*. Once *rhaS* and *rhaR* accumulate high levels in the cell, then *rhaS* begins to activate the *rhaBAD* expression (Giacalone et al. 2006).

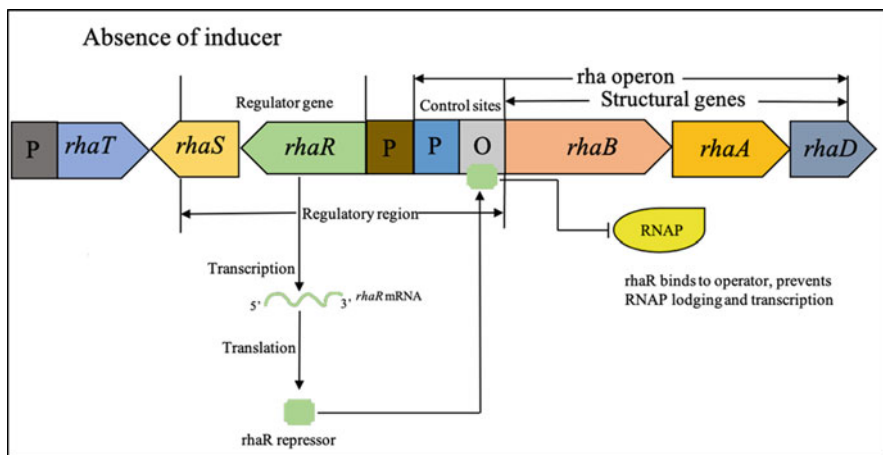


Fig. 5 Repressed rhamnose operon. This shows the repressed state of rhamnose operon, where the absence of inducer allows binding of *rhaR* repressor to the operator, thus inhibiting RNA polymerase mediated transcription of structural genes

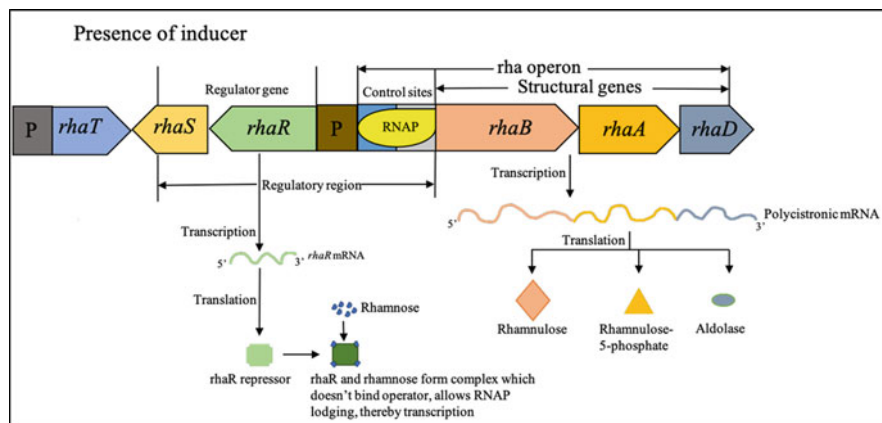


Fig. 6 Induced rhamnose operon. This shows the induced state of rhamnose operon, where the presence of inducer forms a complex with rhaR repressor, which induces conformational changes in repressor and repressor no longer be able to bind the operator, thus promoting RNA polymerase mediated transcription of structural genes

Preliminary studies found that rhamnose induction can induce up to 30,000-fold. *rhaS* was produced at high levels even in the absence of rhamnose. Complementation analysis showed that rhamnose (*rhaBAD*) induction requires RhaS, not RhaR. Gene deletion and *in vitro* transcription-translation assay studies found that RhaS protein is an essential regulator in *rhaBAD* induction. Promoter deleted strains found that two *cis*-acting elements are involved in *rhaBAD* induction. Deletion of upstream element resulted in 60-fold decrease in *rhaBAD* induction. DNA mobility shift assays concluded that CRP protein binds to the DNA region in *rhaBAD* operon. Thus, complete induction of *rhaBAD* expression requires CRP-cAMP complex (Badía et al. 1989; Via et al. 1996). When compared to *araBAD* promoter, *rhaBAD* promoter is tightly regulated. The basal level of *rhaBAD* is about tenfold lower than the lactose promoter but *rhaBAD* and *araBAD* promoter induction strengths are similar (Haldimann et al. 1998). In contrast to the case with the *araBAD* promoter, the regulatory genes *rhaRS* are critical for the function of *rhaBAD* and the chromosomal copy seems to produce enough regulatory proteins to saturate the binding sites on the expression vector. The induction times for attaining a high level of recombinant protein expression varied between 4 and 12 h. Therefore, prolonged response rates of *rhaBAD* promoter were found to be highly beneficial for proteins that have to be transferred into the periplasm or secretion into the medium (Giacalone et al. 2006). The *rhaBAD* regulatory system has been successfully used in HCD fermentation due to its low basal level of expression. A high cell-density fed-batch process for producing heterologous proteins in *E. coli* has been developed using rhamnose. The optimized process with rhamnose resulted in the production of 100 g/L cell dry weight and 3.8 g/L of recombinant L-N-carbamoylase from *E. coli* (Wilms et al. 2001). Production rates of membrane and secretory recombinant proteins need to be controlled in *E. coli* for obtaining high yields (Schlegel et al. 2013). *rhaBAD* based

promoter has been used to produce difficult to express proteins in *E. coli*, since the rhamnose promoter system allows precise tuning of the expression levels in a concentration-dependent fashion (Schlegel et al. 2013). In this study, sfGFP as a model protein was used for monitoring the in vivo kinetics of rhaBAD induction. Moreover, the researchers monitored the production of both the membrane and the secretory proteins in wild-type *E. coli* with the RhaT-mediated L-rhamnose uptake deficient single and double mutants (Hjelm et al. 2017). rhaB rhaT deleted strain improved production yields of all membrane and secretory proteins tested (Hjelm et al. 2017). Penicillin G amidase (PGA) from *Alcaligenes faecalis* was produced using rhamnose in *E. coli*. Production of PGA was induced by repeated addition of the inducer rhamnose. A biomass yield of 13.5 g/L of dry weight and PGA yield of 4500 units per liter were obtained (Deak et al. 2003).

Rhamnose induction system has been successfully used to manufacture recombinant therapeutic proteins. High cell density cultivation was developed in *E. coli* using the rhaBAD expression system. Single-chain antibody fragment was produced in fed-batch cultivation, and a specific product concentration of up to 20 mg/g was obtained. Slow and tight induction of rhamnose produces periplasmic 700 mg/L of antibody fragment in *E. coli* within 4 h (Lindner et al. 2014). Therefore, rhaBAD promoter has some characteristic features that make it exceptionally well suited for expressing the recombinant proteins. These include tight control of expression, low basal activity, and production rates controlled by rhamnose concentration, and it functions independent of any *E. coli* strain (Kelly et al. 2016). However, rhamnose suffers from one major limitation and that is its consumption as sugar by a specific pathway in *E. coli*. Due to its consumption in the central carbon metabolism, a decrease in inducer concentration is observed over time, resulting in lower expression and eventually leading to a decrease in product yield (Kelly et al. 2016). The problem of transient expression caused by inducer degradation has been addressed for some other promoter systems by using nonmetabolized inducer analogs such as IPTG and anhydrotetracycline (ATC). These inducer analogs will not be metabolized and have proven widely valuable for the production of recombinant products. Sugars resembling rhamnose in structure have been studied as potential inducers of the rhaBAD promoter system. Fluorescence studies revealed that L-mannose is the potential inducer among 35 sugars tested, as it provides a broader window of expression levels, good graded response to inducer concentration, and prolong induction (Kelly et al. 2016). L-Lyxose sugar may also be useful at lower expression levels, where it may perform better than L-mannose with respect to induction regulation. Due to commercial availability, L-Mannose and L-lyxose might be useful for bioprocess applications in the future (Kelly et al. 2016).

Most recombinant protein-producing genetic modules and expression systems were constructed in the last decade. Large amounts of soluble recombinant proteins are obtained by a high level of expression using different induction systems. However, the expression of difficult protein targets such as antibody fragments, membrane proteins, large proteins with many disulfide bonds using the expression systems continues to be a challenge with significant losses incurred during protein refolding (Sandomenico et al. 2020). Many strategies can be employed to slow down

the rate of recombinant protein expression and increase the expression level in the soluble fraction, e.g. lowering the growth temperature, optimizing inducer concentration, using an auto-induction medium, or employing a weaker promoter (Sandomenico et al. 2020). Although such tinkering can be very successful, determining the optimal strategy can be laborious as it is dependent on the particular target protein in question (Sandomenico et al. 2020).

3.4 *pNEW Induction System*

To broaden the applicability of the expression system across a wide range of *E. coli* strains a cumate (p-isopropylbenzoate)-regulated expression system has been developed by researchers (Choi et al. 2010). It was first developed for mammalian cells and later for methylotrophic bacteria (Mullick et al. 2006; Choi et al. 2006). A cumate-inducible expression system adapted for *E. coli* is designated as pNEW. It carries a synthetic operator and the regulator (*cymR*) of the *Pseudomonas putida* F1 *cmt* operon (Choi et al. 2006; Eaton 1996; Mullick et al. 2006). The target gene expression is controlled transcriptionally with these regulatory elements through chemical inducer cumate. Generally, the cumate switch requires four major components: a strong promoter, a repressor-binding DNA sequence or operator, expression of a repressor-encoding gene (*cymR*), and the chemical inducer cumate (Choi et al. 2006; Mullick et al. 2006). The absence of cumate keeps the system in an off state (Fig. 7) while the addition of cumate quickly changes the binding of the repressor CymR to the operator which further releases the repressor from the operator, resulting in the formation of the CymR-cumate complex and the expression of the

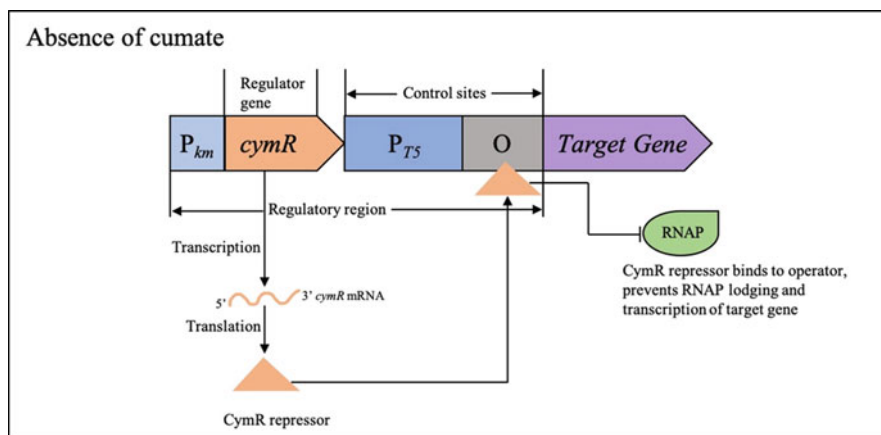


Fig. 7 Repressed cumate gene switch. This shows switched-off state of cumate system, *cymR* gene is under the control of weak constitutive promoter P_{km} . The levels of CymR produced is sufficient to inhibit transcription of target gene in absence of cumate

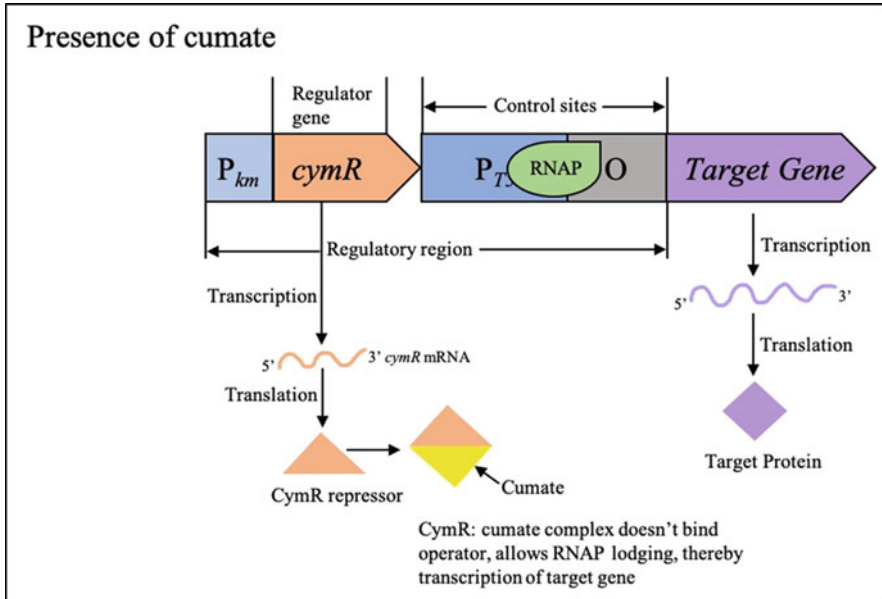


Fig. 8 Induced cumate gene switch. This shows switched on state of cumate system, where the presence of cumate, forms a complex with CymR repressor, induces conformational changes and repressor can no longer be able to bind with operator, thus allowing transcription of target gene

downstream target gene (Fig. 8). In pNEW plasmid, a 26 bp truncated operator fragment of the *cmt* operon was inserted downstream of the strong T5 promoter (PT5), which is mostly recognized by RNA polymerases in *E. coli* strains (Bujard et al. 1987). To minimize transcriptional read-through from PT5, the T7 terminator was provided just downstream of MCS, and for the selection of recombinant clones, kanamycin resistance gene (*Kmr*) was used. It was demonstrated that pNEW system is tightly regulated and operational across all *E. coli* strains, recommending its broad applicability (Choi et al. 2010). The rheostat (dose-dependent) control through cumate switch has also been determined in *E. coli* where authors measured the specific GFP yields of cells harboring the pNEW-gfp construct in which recombinant culture was induced with a range of cumate concentrations (0–122 mM). A linear, exponential specific GFP yield response (rheostat mechanism) was observed in the range of 0–10 mM cumate (subsaturating concentrations), which culminated in a wide range of specific GFP yields (0–90 mg/g dry weight of cells). The GFP yield of 90 mg/g was achieved at the higher limit of the linear response range, which signifies almost 50% of maximal achievable target protein during high-cell density fermentations induced with 100 mM (Choi et al. 2010). Expression of target proteins directed by pLac and pBAD promoters can also be controlled by changing the inducer concentrations, but the rheostat control functions only at much higher concentrations ($\geq 100 \mu\text{M}$) of arabinose and IPTG (Giacalone et al. 2006; Hashemzadeh-Bonehi et al. 1998). Using the cumate switch, the rheostat zone is

feasible with more economical and lower concentrations of chemical inducer (cumate). If required, it also holds the potential to achieve high heterologous protein yields (50% of theoretical maximum) at higher inducer concentrations of the linear rheostat range (Choi et al. 2010). This tunability of production yield can be beneficial for process development (Kaur et al. 2009; Otto et al. 1995) and developing desired production strategies. Another important characteristic of the cumate induction system is a homogenous expression of microbial population post-induction with the inducer cumate, mainly when the purpose is to work through metabolically altered pathways. However, under similar culture conditions, pET-gfp resulted in lesser GFP yield than pNEW-gfp, which may be due to homogenous and consistent induction trait of cumate inducible system. Lower leakiness compared to pET system is the next important trait of cumate switch, which is mainly important when the expressed target protein is toxic. Therefore, cumate switch offers the margin to induce the recombinant cells in the desired growth phase and time, which results in optimum recombinant protein yield and optimum growth of the production host. The relatively quick response of cumate-based induction as compared to IPTG-based induction using pET-system is the next important characteristic that results in optimum protein yield. The plausible reason for relatively quick induction by cumate switch is attributed to its direct and simpler induction which only requires displacing the repressor CymR from the operator upon cumate addition thereby promoting the downstream transcription of target genes by existing *E. coli* RNA polymerases. However, IPTG-based induction is indirect, as it first induces the transcription of T7 RNA polymerase, which then further induces the transcription of heterologous target genes from the T7 promoter of pET-system (Choi et al. 2006). The specific target protein yield and biomass yield in high cell density fermentation of the recombinant *E. coli* resulted in higher recombinant protein (GFP) yield and biomass induced with similar concentrations of cumate and IPTG at the end of fermentation. The demonstration of cumate switch by Choi and co-workers showed that cumate offers several advantages over IPTG-based induction such as lower cost, non-toxic, non-leaky, ability to modulate induction using different inducer concentrations and it can be used across all *E. coli* hosts targeted for recombinant protein production. Hence, cumate based induction system can offer cost-effective benefits, particularly for large-scale bioprocesses compared to an IPTG-based induction system (Choi et al. 2006).

4 Conclusions and Future Trends

This chapter describes the different induction systems and genetically modified promoters used to facilitate the expression of recombinant proteins in *E. coli*. Problems concomitant with high-level overexpression have also been addressed. Several categories of promoter-inducer systems for heterologous production have been presented, each of them offering a variety of advantages and limitations. While

homogenous and long-term induction of recombinant proteins continues to be a challenge, significant advancements have been made in the past few decades.

IPTG-based induction from T7 expression system remains the gold standard as it ensures strong induction and also because IPTG is not metabolized by the cells. Nonetheless, it is known to cause metabolic burden and toxicity (Dvorak et al. 2015). Lactose-based induction offers several advantages where comparable product titers have been achieved without affecting cellular fitness (Bashir et al. 2016). The selection of an appropriate inducer system for target therapeutic production is a prerequisite. Development in inducer systems (arabinose, rhamnose, and cumate) allows precise regulation of recombinant protein production rates in a concentration-dependent manner. Recently, production using auto-induction is in trend where a combination of appropriate promoter and media/feed is designed in such a way that programmed induction can be achieved.

The prevalence of protein-based approved biopharmaceuticals seems to remain an industry reality in the near future. Amongst biopharmaceuticals approved, antibodies continue to dominate, but new gene-based and nucleic acid-based products and cellular therapies are also being launched. Also, there is a trend toward mammalian-based production, given that many recently approved products belong to the class containing post-translational modifications and therefore require a mammalian system (Walsh 2018). With increasing interest in the expression of more complex proteins, pressure continues for the development of microbial systems that can balance the quantity and quality of the product and its production cost (Castiñeiras et al. 2018). Additionally, along with the improvement in engineered N-glycosylation pathway in *E. coli*, recent developments in systems and synthetic biology arms (Gardner 2013) allow complete characterization of pathways that should accurately predict design models. *E. coli* remains a significant contributor to the biopharmaceutical industry.

References

- Badía J, Baldomà L, Aguilar J, Boronat A (1989) Identification of the rhaA, rhaB and rhaD gene products from *Escherichia coli* K-12. FEMS Microbiol Lett 65:253–257
- Bashir H, Ahmed N, Khan MA, Zafar AU, Tahir S, Khan MI et al (2016) Simple procedure applying lactose induction and one-step purification for high-yield production of rhCIFN. Biotechnol Appl Biochem 63(5):708–714
- Bentley WE, Mirjalili N, Andersen DC et al (1990) Plasmid-encoded protein: the principal factor in the “metabolic burden” associated with recombinant bacteria. Biotechnol Bioeng 35:668–681
- Bonomo J, Gill RT (2005) Amino acid content of recombinant proteins influences the metabolic burden response. Biotechnol Bioeng 90:116–126
- Brosius J, Amann E, Brosius J, Ptaslme M (1983) Vectors bearing a hybrid Trp-lac promoter useful for regulated expression of cloned genes in *Escherichia coli*. Gene 25(2–3):167–178
- Bujard H, Gentz R, Lanzer M et al (1987) A T5 promoter-based transcription-translation system for the analysis of proteins in vitro and in vivo. Methods Enzymol 155:416–433
- Calos MP, Miller JH (1981) The DNA sequence change resulting from the IQ1 mutation, which greatly increases promoter strength. MGG Mol Gen Genet 183:559–560

- Castiñeiras TS, Williams SG, Hitchcock AG, Smith DC (2018) *E. coli* strain engineering for the production of advanced biopharmaceutical products. *FEMS Microbiol Lett* 365:1–10
- Chen WB, Nie Y, Xu Y, Xiao R (2014) Enhancement of extracellular pullulanase production from recombinant *Escherichia coli* by combined strategy involving auto-induction and temperature control. *Bioprocess Biosyst Eng* 37:601–608
- Choi YJ, Morel L, Bourque D et al (2006) Bestowing inducibility on the cloned methanol dehydrogenase promoter (PmxaF) of *Methylobacterium extorquens* by applying regulatory elements of *Pseudomonas putida* F1. *Appl Environ Microbiol* 72(12):7723–7729
- Choi YJ, Morel L, Le FT et al (2010) Novel, versatile, and tightly regulated expression system for *Escherichia coli* strains. *Appl Environ Microbiol* 76(15):5058–5066
- Clark MA, Hammond FR, Papaioannou A et al (1997) Regulation and expression of human fabs under the control of the *Escherichia coli* arabinose promoter, PBAD. *Immunotechnology* 3: 217–226
- De Boer HA, Comstock LJ, Vassert M (1983) The tac promoter: a functional hybrid derived from the trp and lac promoters. *Proc Natl Acad Sci U S A* 80(1):21–25
- Deak PM, Lutz-Wahl S, Bothe H, Fischer L (2003) Bioreactor cultivation of *Escherichia coli* for production of recombinant penicillin G amidase from *Alcaligenes faecalis*. *Biotechnol Lett* 25: 397–400
- Ding N, Yang C, Sun S et al (2017) Increased glycosylation efficiency of recombinant proteins in *Escherichia coli* by auto-induction. *Biochem Biophys Res Commun* 485:138–143
- Dvorak P, Chrast L, Nikel PI, Fedr R, Soucek K, Sedlackova M et al (2015) Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21 (DE3) carrying a synthetic metabolic pathway. *Microb Cell Factories* 14(1):1–15
- Eaton RW (1996) p-Cumate catabolic pathway in *Pseudomonas putida* F1: cloning and characterization of DNA carrying the cmt operon. *J Bacteriol* 178(5):1351–1362
- Eberhardt F, Aguirre A, Menzella HG, Peiru S (2017) Strain engineering and process optimization for enhancing the production of a thermostable steryl glucosidase in *Escherichia coli*. *J Ind Microbiol Biotechnol* 44:141–147
- Fathi-Roudsari M, Maghsoudi N, Maghsoudi A et al (2018) Auto-induction for high level production of biologically active reteplase in *Escherichia coli*. *Protein Expr Purif* 151:18–22
- Gardner TS (2013) Synthetic biology: from hype to impact. *Trends Biotechnol* 31(3):123–125
- Giacalone MJ, Gentile AM, Lovitt BT et al (2006) Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. *BioTechniques* 40:355–364
- Gupta SK, Shukla P (2017) Microbial platform technology for recombinant antibody fragment production: a review. *Crit Rev Microbiol* 43:31–42
- Gupta N, Balomajumder C, Agarwal VK (2010) Enzymatic mechanism and biochemistry for cyanide degradation: a review. *J Hazard Mater* 176(1–3):1–13
- Guzman L-M, Belin D, Carson MJ, Beckwith JON (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* 177:4121–4130
- Hahn S, Dunn T, Schleif R (1984) Upstream repression and CRP stimulation of the *Escherichia coli* l-arabinose operon. *J Mol Biol* 180:61–72
- Haldimann A, Daniels LL, Wanner BL (1998) Use of new methods for construction of tightly regulated arabinose and rhamnose promoter fusions in studies of the *Escherichia coli* phosphate regulon. *J Bacteriol* 180:1277–1286
- Hashemzadeh-Bonehi L, Mehraein-Ghomi F, Mitsopoulos C et al (1998) Importance of using lac rather than ara promoter vectors for modulating the levels of toxic gene products in *Escherichia coli*. *Mol Microbiol* 30(3):676–678
- Hjelm A, Karyolaimos A, Zhang Z et al (2017) Tailoring *Escherichia coli* for the l-rhamnose PBAD promoter-based production of membrane and secretory proteins. *ACS Synth Biol* 6:985–994

- Hoffmann F, Rinas U (2004) Stress induced by recombinant protein production in *Escherichia coli*. *Adv Biochem Eng Biotechnol* 89:73–92
- Inada T, Kimata K, Aiba H (1996) Mechanism responsible for glucose-lactose diauxie in *Escherichia coli*: challenge to the cAMP model. *Genes Cells* 1(3):293–301
- Jaishankar J, Srivastava P (2017) Molecular basis of stationary phase survival and applications. *Front Microbiol* 8:2000
- Johnson CM, Schleif RF (1995) In vivo induction kinetics of the arabinose promoters in *Escherichia coli*. *J Bacteriol* 177(12):3438–3442
- Juers DH, Matthews BW, Huber RE (2012) LacZ β -galactosidase: structure and function of an enzyme of historical and molecular biological importance. *Protein Sci* 21(12):1792–1807
- Kaur P, Agarwal S, Datta S (2009) Delineating bacteriostatic and bactericidal targets in mycobacteria using IPTG inducible antisense expression. *PLoS One* 4(6):e5923
- Keasling JD (1999) Gene-expression tools for the metabolic engineering of bacteria. *Trends Biotechnol* 17(11):452–460
- Kelly CL, Liu Z, Yoshihara A et al (2016) Synthetic chemical inducers and genetic decoupling enable orthogonal control of the rhaBAD promoter. *ACS Synth Biol* 5:1136–1145
- Kercher MA, Lu P, Lewis M (1997) Lac repressor-operator complex. *Curr Opin Struct Biol* 7:76–85
- Krahulec J, Hyršová M, Pepeliaev S et al (2010) High level expression and purification of antimicrobial human cathelicidin LL-37 in *Escherichia coli*. *Appl Microbiol Biotechnol* 88:167–175
- Lee NL, Gielow WO, Wallace RG (1981) Mechanism of araC autoregulation and the domains of two overlapping promoters, Pc and PBAD, in the L-arabinose regulatory region of *Escherichia coli*. *Proc Natl Acad Sci* 78:752–756
- Lindner R, Moosmann A, Dietrich A et al (2014) Process development of periplasmatically produced single chain fragment variable against epidermal growth factor receptor in *Escherichia coli*. *J Biotechnol* 192:136–145
- Malan TP, McClure WR (1984) Dual promoter control of the *Escherichia coli* lactose operon. *Cell* 39:173–180
- Mari M, Gallegos M-T, Schleif R et al (1997) AraC/XylS family of transcriptional regulators
- Miret J, Román R, Benito M et al (2020) Development of a highly efficient production process for recombinant protein expression in *Escherichia coli* NEB10 β . *Biochem Eng J* 159:107612
- Miyada CG, Stoltzfus L, Wilcox G (1984) Regulation of the araC gene of *Escherichia coli*: catabolite repression, autoregulation, and effect on araBAD expression. *Proc Natl Acad Sci* 81:4120–4124
- Moralejo P, Egan SM, Hidalgo E, Aguilar J (1993) Sequencing and characterization of a gene cluster encoding the enzymes for L-rhamnose metabolism in *Escherichia coli*. *J Bacteriol* 175:5585–5594
- Mullick A, Xu Y, Warren R et al (2006) The cumate gene-switch: a system for regulated expression in mammalian cells. *BMC Biotechnol* 6:43
- Muntari B, Amid A, Mel M et al (2012) Recombinant bromelain production in *Escherichia coli*: process optimization in shake flask culture by response surface methodology. *AMB Express* 2:12
- Narayanan A, Ridilla M, Yernool DA (2011) Restrained expression, a method to overproduce toxic membrane proteins by exploiting operator–repressor interactions. *Protein Sci* 20:51–61
- Neubauer P, Wolff C, Hecker M et al (1991) Introduction of the tac-promoter by lactose under fermentation conditions. *Acta Biotechnol* 11:23–29
- Nie Y, Yan W, Xu Y et al (2013) High-level expression of *Bacillus naganensis* pullulanase from recombinant *Escherichia coli* with auto-induction: effect of lac operator. *PLoS One* 8(10):e78416
- Nieto C, Fernández De Palencia P, López P, Espinosa M (2000) Construction of a tightly regulated plasmid vector for *Streptococcus pneumoniae*: controlled expression of the green fluorescent protein. *Plasmid* 43(3):205–213

- Oehler S, Amouyal M, Kolkhof P et al (1994) Quality and position of the three lac operators of *E. coli* define efficiency of repression. *EMBO J* 13:3348–3355
- Oehler S, Eismann ER, Kramer H, Muller-Hill B (1990) The three operators of the lac operon cooperate in repression. *EMBO J* 9:973–979
- Ogden S, Haggerty D, Stoner CM et al (1980) The *Escherichia coli* L-arabinose operon: binding sites of the regulatory proteins and a mechanism of positive and negative regulation. *Proc Natl Acad Sci U S A* 77(6):3346–3350
- Otto CM, Niagro F, Su X, Rawlings CA (1995) Expression of recombinant feline tumour necrosis factor is toxic to *Escherichia coli*. *Clin Diagn Lab Immunol* 2(6):740–746
- Rava PS, Somma L, Steinman HM (1999) Identification of a regulator that controls stationary-phase expression of catalase-peroxidase in *Caulobacter crescentus*. *J Bacteriol* 181(19):6152–6159
- Román R, Lončar N, Casablancas A et al (2020) High-level production of industrially relevant oxidases by a two-stage fed-batch approach: overcoming catabolite repression in arabinose-inducible *Escherichia coli* systems. *Appl Microbiol Biotechnol* 104:1–9
- Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol* 5:172
- Rosano GL, Morales ES, Ceccarelli EA (2019) New tools for recombinant protein production in *Escherichia coli*: a 5-year update. *Protein Sci* 28:1412–1422
- 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* 307:249–264
- 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
- Samuelson JC (2011) Recent developments in difficult protein expression: a guide to *E. coli* strains, promoters, and relevant host mutations. *Methods Mol Biol* 705:195–209
- Sandomenico A, Sivaccumar JP, Ruvo M (2020) Evolution of *Escherichia coli* expression system in producing antibody recombinant fragments. *Int J Mol Sci* 21:1–39
- Schlegel S, Rujas E, Ytterberg AJ et al (2013) Optimizing heterologous protein production in the periplasm of *E. coli* by regulating gene expression levels. *Microb Cell Factories* 12:1–12
- Schleif R (2000) Regulation of the L-arabinose operon of *Escherichia coli*. *Trends Genet* 16(12):559–565
- Shibui T, Uchida M, Teranishi Y (1988) A new hybrid promoter and its expression vector in *Escherichia coli*. *Agric Biol Chem* 52:983–988
- Shin PK, Seo J-H (1990) Analysis of *E. coli* phoA-lacZ fusion gene expression inserted into a multicopy plasmid and host cell's chromosome. *Biotechnol Bioeng* 36(11):1097–1104
- Simpson RB (1980) Interaction of the cAMP receptor protein with the lac promoter. *Nucleic Acids Res* 8(4):759–766
- Sørensen HP, Mortensen KK (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J Biotechnol* 115:113–128
- Studier FW (1991) Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J Mol Biol* 219(1):37–44
- Studier FW (2014) Stable cultures and auto-induction for inducible protein production in *E. coli*. *Struct Genomics Methods Protoc* 1091:17–32
- Tabor S, Richardson CC (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci* 82:1074–1078
- Tobin JF, Schleif RF (1987) Positive regulation of the *Escherichia coli* L-rhamnose operon is mediated by the products of tandemly repeated regulatory genes. *J Mol Biol* 196:789–799
- Tropp BE (2008) *Molecular biology: genes to proteins*. Jones and Bartlett Publishers
- Via P, Badía J, Baldomá L et al (1996) Transcriptional regulation of the *Escherichia coli* rhaT gene. *Microbiology* 142:1833–1840
- Wacker M, Linton D, Hitchen PG et al (2002) N-linked glycosylation in *campylobacter jejuni* and its functional transfer into *E. coli*. *Science* 298(5599):1790–1793

- Walsh G (2018) Biopharmaceutical benchmarks 2018. *Nat Biotechnol* 36:1136–1145
- Wanner BL (1996) Phosphorus assimilation and control of the phosphate regulon. In: Neidhardt FC, Curtiss RI, Gross CA et al (eds) *Escherichia coli* and *Salmonella typhimurium* cellular and molecular biology. 2nd ed, chapter 84. Am Soc Microbiol, Washington, DC
- Wheatley RW, Lo S, Jancewicz LJ et al (2013) Structural explanation for allolactose (lac operon inducer) synthesis by lacZ β -galactosidase and the evolutionary relationship between allolactose synthesis and the lac repressor. *J Biol Chem* 288:12993–13005
- Wilms B, Hauck A, Reuss M et al (2001) High-cell-density fermentation for production of L-N-carbamoylase using an expression system based on the *Escherichia coli* rhaBAD promoter. *Biotechnol Bioeng* 73:95–103