Protein J
DOI 10.1007/s10930-013-9502-5

Strategies for the Production of Recombinant Protein in *Escherichia coli*

Gopal Jee Gopal · Awanish Kumar

© Springer Science+Business Media New York 2013

Abstract In the recent past years, a large number of proteins have been expressed in *Escherichia coli* with high productivity due to rapid development of genetic engineering technologies. There are many hosts used for the production of recombinant protein but the preferred choice is *E. coli* due to its easier culture, short life cycle, well-known genetics, and easy genetic manipulation. We often face a problem in the expression of foreign genes in *E. coli*. Soluble recombinant protein is a prerequisite for structural, functional and biochemical studies of a protein. Researchers often face problems producing soluble recombinant proteins for over-expression, mainly the expression and solubility of heterologous proteins. There is no universal strategy to solve these problems but there are a few methods that can improve the level of expression, non-expression, or less expression of the gene of interest in *E. coli*. This review addresses these issues properly. Five levels of strategies can be used to increase the expression and solubility of over-expressed protein; (1) changing the vector, (2) changing the host, (3) changing the culture parameters of the recombinant host strain, (4) co-expression of other genes and (5) changing the gene sequences, which may help increase expression and the proper folding of desired protein. Here we present the resources available for the expression of a gene in *E. coli* to get a substantial amount of good quality recombinant protein. The resources include different strains of *E. coli*, different *E. coli* expression vectors, different physical and chemical agents and the co expression of chaperone interacting proteins. Perhaps it would be the solutions to such problems that will finally lead to the maturity of the application of recombinant proteins. The proposed solutions to such problems will finally lead to the maturity of the application of recombinant proteins.

Keywords *E. coli* · Recombinant protein expression · Vector · Host cell

Abbreviation

aa	Amino acid
bp	Base pairs
GST	Glutathione S-transferase
HAT	Histone acetyltransferase
IPTG	Isopropylthio-β-galactoside
MBP	Maltose binding protein
NEB	New England Biolabs
MCS	Multiple cloning site
Ni-NTA	Nickel nitrilotriacetic acid
ORF	Open reading frame
SUMO	Small ubiquitin modifier
Trx	Thioredoxin

1 Introduction

There are many hosts used for the production of recombinant protein, but the preferred choice is *Escherichia coli* because it is easy to culture, has a very short life cycle and is easily manipulated genetically due to its well-known genetics. *E. coli* was the first host used to produce a recombinant DNA (rDNA) biopharmaceutical, enabling

G. J. Gopal
School of Biotechnology, Jawaharlal Nehru University,
New Delhi 110067, India

A. Kumar (✉)
Department of Biotechnology, National Institute of Technology,
Raipur 492010, Chhattisgarh, India
e-mail: drawanishkr@gmail.com

Published online: 30 July 2013

the approval of Eli Lilly's rDNA human insulin in 1982. We also use *E. coli* host often to produce therapeutic protein to minimize the cost of production. In *E. coli*, we face two main problems; (1) is difficult or little expression of a foreign gene and (2) is solubility of recombinant proteins for over expression. Structural, functional and biochemical studies of protein requires an ample amount of good quality protein. There is no universal strategy to solve this problem but there are a few methods that can improve the level of expression or non-expression or decreased expression of a gene of interest in *E. coli*. Factors such as mRNA's secondary structure, a protein's intrinsic ability to fold, its solubility, preferential codon use, its toxicity, and its need for post translational modification have influence on expression of foreign gene in *E. coli* [27]. At least four strategies can useful for increasing the expression and solubility of over-expressed protein; Changing the vector, changing the host strain, and adding of some chemicals during the induction or co-expression of other genes may help in the proper folding of the desired protein. If the above strategies are unsuccessful, then one can alter the gene sequence without changing the functional domain. Gene sequence alteration include the removal of a signal peptide coding sequence, optimisation of codon according to *E. coli*, deletion of hydrophobic patch coding sequences, and prevention of stable mRNA secondary structure formation by changing the sequence which forms secondary structure etc. The main focus of this review is to describe various strategies for the expression of foreign genes/proteins in *E. coli* with minimum difficulty. This review highlights the problem encountered problems and knowledge gained during widespread research on proteins expression and production.

2 Strategies for Over Expression of Foreign Gene and Production of Recombinant Protein

2.1 Changing the Vector

Changing the vector refers to either changing the promoter with which the gene of interest is to be cloned or changing the fusion 'tag' which influence solubilization of recombinant protein in *E. coli* host (Table 1). The pET vector with 6-Histidine tag (His-tag), is usually the first choice to get recombinant protein because His-tag is a smaller affinity tag with high levels of expression due to the presence of a strong promoter and the availability of robust purification system like Ni-NTA. The pET vector is provided by 'Novagen'. Consecutive 6-Histidine can possibly decrease the solubility of a fused protein. If consecutive Histidine poses a problem in the solubility of a protein, pHAT vector can be used. The pET vector is commercially provided by 'Clontech', and encodes a novel polyhistidine epitope tag. This tag is 19 amino acids long and has a non adjacent 6-Histidine. This sequence of nonadjacent histidine residues has a lower overall charge than tags with consecutive histidine residues. As a result, HAT fusion proteins exhibit solubility close to wild-type proteins without compromising the affinity for metal.

Table 1 Some common tags used in enhancing protein solubility

Tag	Protein	Size
GST	Glutathione S-transferase	211 aa
MBP	Maltose-binding protein	396 aa
NusA	N-utilization substance	495 aa
Sumo	Small ubiquitin modifier	100 aa
Trx	Thioredoxin	109 aa

When a histidine tagged gene does not express the protein or recombinant protein forms inclusion bodies, the gene of interest should be cloned in pGEX system of 'GE healthcare' or GST tag, or in pMAL system of 'New England Biolabs'. It has been well established that maltose binding protein (MBP) tag [3, 7, 12] and GST tag [31] increases solubility of fused proteins and expression of genes. The drawback of pMAL system and pGEX vector is their large sizes of fusion tag (MBP: 44 kDa; GST: 25 kDa). The large size of fusion tags may interfere with the activity of recombinant proteins if the tag is not cleaved properly. Different enzyme sites have been incorporated in this vector so that cleavage would be precise and no extra amino acid remain fused with recombinant protein. Tags may interfere with the structure and function of the target protein, therefore provision tags must be removed after the expression and purification of the protein. Multiple cleavage sites can be engineered into the expression construct to remove tags. Enterokinase (cleavage site DDDDK↑), Factor Xa (cleavage site IEGR↑), SUMO protease (recognizes SUMO tertiary structure and cleaves at the C-terminal end of the conserved Gly–Gly sequence in SUMO), Thrombin (cleavage site LVPR↑GS) are the common proteases used to remove tags often from recombinant protein (Table 1). Three variations of pMAL vector are available with NEB based on enzyme cleavage site: factor Xa, genenase, enterokinase classified as P/C5X, P/C5G, and P/C5E respectively. C and P indicate cytoplasmic or periplasmic expression respectively. SUMO tag recombinant protein has a tendency to be soluble in the *E. coli* host [2]. pSUMO vector with kanamycine and ampiciline resistance marker available. pSUMO vector with dual expression (*E. coli*/yeast and *E. coli*/mammalian cell) and only bacterial expression is also commercially available from 'LifeSensors Inc., 271, Great Valley Parkway'. One big advantage of using pSUMO vector is the ability to get recombinant

protein without any extra amino acid after cleavage of the tag by SUMO protease. N-utilization substance A (NusA) tag also increases solubility of the tagged protein. Solubility modeling of *E. coli* protein shows that NusA tag has the highest potential for solubility [4, 11]. pET-43 and pET-44 vectors supplied by 'Novagen' can be used for NusA tagged recombinant protein. Fusion of 109-aa long Trx.tag thioredoxin protein also increases the solubility of conjugated protein. The pET-Trx plasmid vector can be used for this tag [16, 24]. pET-39b and pET-40b plasmid are used to conjugate 208 aa DsbA and 236 aa DsbC proteins involved in periplasmic disulphide bond formation respectively. DsbA helps in the formation of disulphide bond and DsbC assists in the isomerization of disulphide bond. The recombinant protein(s) conjugated with this protein tag can form disulphide bond in *E. coli* [10, 11, 14, 23, 35]. Fusion tags like NusA, Trx and SUMO needs extra affinity tag for purification. The main disadvantage of NusA, Trx, and SUMO is their requirement for extra affinity tag for protein purification [23].

pQE system of Qiagen (containing T5 promoter) has been shown to increase expression of gene but it does not express with T7 promoter. T5 promoter is recognised by *E. coli* polymerase and it has two *lac* operator sequences. The presence of these two *lac* operator sequence increase the *lac* repressor binding and ensure for the sufficient repression of proteins. This vector has 6-His tag at the –N or –C terminal and is suitable for the expression in *E. coli* host strains like, M15[pREP4] and SG13009[pREP4]. These strains contain pREP4 plasmid. This plasmid encodes *lac* repressor which ensures tight regulation of expression. Sometimes periplasmic translocation of gene products can be required for proper folding and/or disulphide bond formation or for easy purification. To fulfil these requirement, there are also plasmid variants are available with *E. coli* signal peptide coding sequences at the 5′ of MCS like p5 series (c5 series is for cytoplasmic expression) of pMAL vector, pET-22 vector etc. Solubility of recombinant protein in *E. coli* can be determined or checked with freely available software by conjugating the sequence of the fusion tag with the desired protein sequence. This exercise may be useful before cloning the gene of interest. It is also important to search common cloning sites for different vectors before designing the primers. This allows the desired gene fragment to be subcloned from one vector to other without purchasing new primer pairs and without PCR amplification of insert using new primer sets. A vector should be chosen in such a way with consideration that the tag increase solubility but does not disrupt the activity of the recombinant protein or the cleavage of the tag. It has also been observed (personal observation) that non expression of gene is due to frame shift in the reading frame during cloning which causes stop codon incorporation. New researchers could prevent this during cloning at the NcoI restriction endonuclease site of pET-28a plasmid. Cloning at NcoI site in this vector requires two extra base pair in the primer sequence at 5′ to amplify open reading frame (ORF).

2.2 Changing the Host Strain

Nowadays so many biotech companies provide different types of genetically altered *E. coli* strain as per suitability of expression of foreign genes (Table 2). Genes cloned with 'tac' promoter (pMAL system, pGEX system) can be expressed in the cloning host itself but the BL21 strain is the preferred choice for expression because of absence of two main proteases genes cloned in pET system with T7 promoter should be expressed in BL21(DE3). *E. coli* strain BL21(DE3) has a T7 polymerase encoding gene introduced in its genome as well as *Lon* and *OmpT* protease deleted. *Lon* and *OmpT* protease deficient strain of *E. coli* are unable to degrade foreign protein. Leaky expression of the desired gene in BL21(DE3) cell is possible. To minimise leaky expression of toxic gene, the BL21 host strain was improved. The improved strain is BL21(DE3)pLys S and BL21(DE3)pLys E. Both the strain have lysozyme coding plasmid. Lysozyme is inhibitor of T7 polymerases which inhibits residual T7 polymerase and thus prevents leaky expression. Leaky expression of toxic gene is detrimental to the host cell—the host cell will not survive or it will change its mechanism so that even in presence of inducer it will not allow production of toxic protein. Many times expression does not occur in *E. coli* due to a difference in codons present in gene of interest and preferential codon use by the host *E. coli* strain. To solve this codon bias in host cell machinery, a few companies provide modified *E. coli* hosts that have extra tRNA coding genes (AUA, AGG, AGA, CUA, CCC and GGA) compensating for the scarcity of rare tRNA (Fig. 1). The host strain is designed as CodonPlus *E. coli* strain of 'Stratagene' (RIL or RP) and Rosetta strain (Novagene). Rosetta host strains are BL21(DE3) derivatives designed to enhance the expression of heterologous proteins by containing codons rarely used in *E. coli*. CodonPlus-RIL strain is used to overcome the bias of AT-rich genome while CodonPlus-RP strain is used to overcome GC rich genome bias. Unlike CodonPlus of Stratagene, the Rosetta strain contains all the genes coding rare tRNA in the same strain so separate strain for AT and GC rich gene expression would need to be used. Origami™ *E. coli* strain of 'Novagene' may be used if disulphide bond is required for proper folding of protein. The Origami strain is an oxidising *E. coli* strain with mutation in thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes that allows disulphide bond formation within cytoplasm. 'SHuffle' *E. coli* strains from 'NEB'

Table 2 Characteristics of *E. coli* strains used in recombinant protein expression

E. coli strain	Used for	Characteristics at genetic level
ArcticExpress	Grow at low temperature	Low temperature adapted chaperone encoded gene
BL21	Less protease degradation of recombinant protein	*Lon* and *OmpT* protease deficient
BL21-Codonplus (RIL)	Overcome the effect of codon biasness (AT rich gene)	tRNA encoding gene (*argU, ileY* and *leuW*)
BL21(DE3)	Expression under T7 promoter	T7 pol encoded
BL21(DE3)pLys S/E	Controlled expression	Lysozyme encoded plasmid
BL21 Star	Increase stability of mRNA	*rne31* gene mutated
C41(DE3)	Overexpressing toxic proteins	Carry the lambda DE3 lysogen which expresses T7 RNA polymerase from the lacUV5 promoter by IPTG induction
C43(DE3)	Membrane and globular protein	-Same-
Codon plus (RP)	Overcome the effect of codon biasness (GC rich gene)	tRNA encoding gene (*argU* and *proL*)
Lemo21(DE3)	Membrane, globular and toxic protein expression	Lysozyme encoded under rhamnose inducible promoter
M15	Gene under T5 promoter	Constitutively expresses *lac* repressor at high levels
Origami	Protein required disulphide bond formation	*gor* and *trxB* genes mutated
Rossetta	Both the AT and GC rich gene	All the rare tRNA coding gene
SG13009	Enabling *trans* repression of protein expression prior to IPTG induction	Carry the repressor plasmid pREP4, which constitutively expresses *lac* repressor at high levels
Shuffle	Proper disulphide bond formation	*DsbC* gene encoded

are better than 'Origami' strain for the expression of putative disulphide bond forming protein. In addition to *trxB* and *gor* mutation, SHuffle strains express *DsbC* within the cytoplasm, which directs correct disulfide bond formation and also acts as a general chaperone for protein folding.

For the expression of gene which coding globular or membrane protein, C43(DE3) strain of *E. coli* [1] provided by Lucigen can be used. C41(DE3) and C43(DE3) strain (walker strain) are also useful in overcoming the problem of plasmid instability during toxic recombinant protein expression [15]. Walker strain is derivative of BL21(DE3). Membrane protein overexpression characteristics of these strains are due to a mutation in *lac*UV5 promoter which causing weaker expression of T7 RNA polymersae. Lemo21(DE3) *E. coli* strain from 'NEB' is more improved host for the expression of gene cloned under T7 promoter. The concept of Lemo21(DE3) strain came after the discloser of C43(DE3) and C41(DE3) strain's genetic makeup [29]. Lemo21 (DE3) is derivative of BL21(DE3) strain and are genetically engineered in such manner that it can act as C43(DE3) self mutant strain [28]. T7 polymerase is cloned with the well-titrable rhamnose promoter and its activity can be precisely controlled by its natural inhibitor T7 lysozyme found in this strain [34]. Using this Lemo21(DE3) strain, it has also been shown that Sec-translocon saturation is the main bottleneck for heterologous membrane protein production [28]. This strain is also used for the expression of difficult proteins such as membrane protein, toxic protein or the protein that aggregates in *E. coli* without compromising their stability and function [28, 29]. The flexibility of this strain makes it potent, time-saving and cost-effective to identify the optimal conditions for the overexpression of protein of interest.

High level expression usually causes formation of an inclusion body. One way to avoid or decrease the inclusion body formation is to reduce cultivation temperature of the host after induction. However at low the temperature, the expression of chaperonine, which fold newly synthesized or misfolded protein, also reduce drastically. ArcticExpress strains of *E. coli* overcome this problem because these genetically engineered strain co-express cold-adapted chaperonins *Cpn10* and *Cpn60* from the psychrophilic bacterium, *Oleispira antarctica*. Chaperonine has high protein folding activity at 4–10 °C. This strain is commercially available from Agilent technology, Genomics. Acetylated recombinant protein has also been produced in genetically engineered *E. coli* strain but these strains are not commercially available. A high yield of protein also depends on the stability of corresponding mRNA. A BL21 derivative having an *rne131* mutation confers higher stability of mRNA and ultimately increased gene expression. This strain is available from Invitrogen under the name BL21 Star [10, 13, 18, 19]. Glycosylated recombinant protein has been produced from genetically engineered

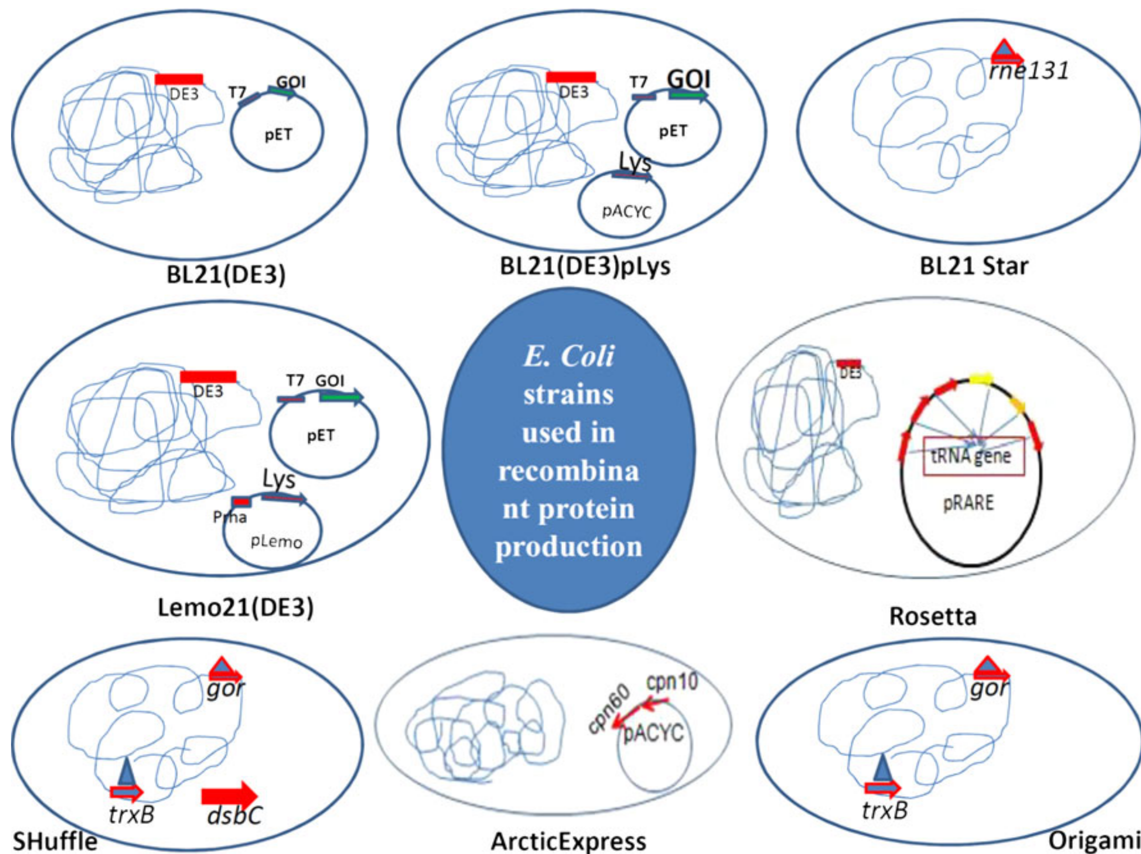

Fig. 1 *E. coli* strains frequently used in recombinant protein production having different vector, promoter, multiple cloning sites, tRNA and other genes

E. coli with the *pgl* operon of *Campylobacter jejuni* [22]. This strain is not commercially available and the glycosylation pattern is different from that of eukaryote so it is not frequently used. *N*-linked protein glycosylation is the most common post-translational modification in eukaryotes. Recently, Valderrama-Rincon et al. [33] successfully engineered an *E. coli* strain and got N-glycosylation of eukaryotic recombinant protein. Strains of *E. coli* frequently used in recombinant protein production are shown and summarized in Fig. 1.

2.3 Changing the Culture Parameters of Recombinant Host Strain

Culture condition of recombinant *E. coli* may also increase expression of gene and solubility of expressed protein. The solubility of recombinant protein is increased by prolonged induction at low temperatures with decreased amounts of IPTG. Sometimes addition of some chemical in culture medium also increases expression of the gene of interest because these chemicals induce chaperone expression. Induced chaperone manages the proper folding of recombinant protein and thus enables it to remain in soluble form. Expressed proteins may be stabilized by the binding with ligands (a small chemical agent) and have found widespread application in recombinant protein productions [9]. This concept is exploited to increase the proportion of recombinant protein expressed in soluble form or to stabilize a protein during purification [10]. Ethanol [8, 14], heat shock i.e. temperature treatment [20], benzyl alcohol [5], osmolytes, ionic strength of buffer [6, 30] etc., are inducers that enhance expression and production of recombinant proteins.

2.4 Co-expression

Few proteins require their counterparts for their stability so the genes coding these protein either do not express or express and then degrade the counterparts immediately. In any case recombinant protein would not be produced. In this situation the interacting protein or protein which stabilizes the desired protein must be identified. The gene coding those proteins must be co-expressed with the protein of interest. For example, chaperone, when co-expressed with gene of interest, increases the expression and solubility of that gene product. Recently deMarco, showed that when chaperone is co-expressed with a gene of interest

nearly all of the gene is expressed [5]. Chaperone coding plasmid vectors are not commercially available but can be obtained by request from deMarco. When two gene product forms a complex, one gene alone is either expressed or not expressed but more importantly, forms inclusion bodies in the foreign host. This is due to exposed hydrophobic patches in the absence of the interacting partner. The interacting partner of a complex protects the hydrophobic patches of both genes' proteins in the native environment. In this case, both genes should be co-expressed in heterologous host using either duet/bicistronic vector or two compatible vectors. Expression of toxin coding gene usually requires the co-expression of an antitoxin coding gene or toxin gene with a mutated toxic region.

2.5 Changing Gene Sequences Without Changing the Functional Domain of Protein

Removal of signal peptide coding sequence also increases the expression and stability of recombinant protein. Signal peptides usually have no role except the export of protein from the site of synthesis to the target site. Therefore, endogenous mature protein does not have signal peptides. Production of recombinant protein without signal sequence in heterologous host would not alter its biochemical characteristics and function. Thus a signal peptide coding sequence can be removed from ORF if it hampers stability and expression of the complete ORF. Many times, the deletion of a transmembrane or hydrophobic patch coding region increases expression of the gene and solubility of the recombinant protein [26]. To avoid codon bias in the heterologous host gene expression, the gene sequence can be optimized as per host preferences of the codon. Gene with optimized codon sequence can be obtained either through site directed mutagenesis (if a few codons have to be changed) or by synthesizing complete ORF. This optimized gene can then be cloned and expressed in the appropriate vector and host respectively. This optimised ORF does not require BL21-codon plus/Rossetta strain for its expression.

The secondary structure of mRNA also hampers the efficient translation of a gene [25]. To avoid the formation of this secondary structure, codon sequences can be manipulated via site directed mutagenesis without changing the amino acid sequence [17]. Gene sequence optimization used to prevent the formation of mRNA secondary structure to smooth translation initiation and prolong processivity. Commercial gene optimization services that optimize gene sequences and then synthesize that gene for efficient expression in a given host are also available. Usually gene optimization software consider following parameters: (a) Elimination of cryptic splice sites and RNA destabilizing sequence elements for increased RNA stability, (b) Addition of RNA stabilizing sequence elements, (c) Codon optimization and G/C content adaptation for your expression system, (dld) Intron removal and avoidance of stable RNA secondary structures.

3 Discussion

The methods for protein production and purification have been reviewed by researchers numerous times in an admirable and comprehensive manner. Here we have compiled the resources available till date for the production of ample and good quality recombinant protein using *E. coli* as a host. Production of recombinant protein in *E. coli* is less costly than using other hosts and the handling is also easier. Therefore, it is beneficial to utilize *E. coli* as a host for the production of desired protein. When recombinant protein is expressed immensely in bacteria, it tends to misfold and accumulate as a soluble and insoluble nonfunctional aggregate. A general strategy to improve the native folding of recombinant proteins is to increase the cellular concentration of viscous organic compounds, osmolytes, or of molecular chaperones that can prevent aggregation. If resources (as discussed above) are known to researchers then they can exploit that resource for their condition. Still there is no guarantee that all of the gene of interest will express in *E. coli* but researcher definitely will get most of the gene product by applying the aforementioned strategy. Strategies for the expression of recombinant proteins in *E. coli* can be prioritized by changing the following factors (1) cultivation parameters, (2) host strain, (3) vector, (4) co-expression, (5) ORF with unaltered functional domain/structure (Fig. 2). Priority can be changed on the basis of known physiochemical properties of proteins. It is suggested to try these strategies in permutation to get overexpression of the gene/protein of interest.

It would be beneficial to screen the solubility of recombinant protein through available in silico tools before cloning into particular vectors [32, 36]. To save time and resources it would be better to use an existing vector or host strain but if the existing vector or host strain does not fit to the research design then a host and/or vector can be constructed. A universal host or vector which allows all types of gene expression is required. Host strains with combination of two or more features like 'Rosetta-gami' strain allow disulphide bond formation and also prevent the effect of codon bias. ArcticExpress, RIL/RP strain of *E. coli* has the ability to allow the expression of heterologous gene at low temperature as well as nullify the effect of rare codon in the gene [21]. ArcticExpress codon plus and Rossetta-gami are the variants of ArchticExpress and Rossetta strains that are also commercially available from 'Agilent Technologies' and from 'Novagene' respectively.

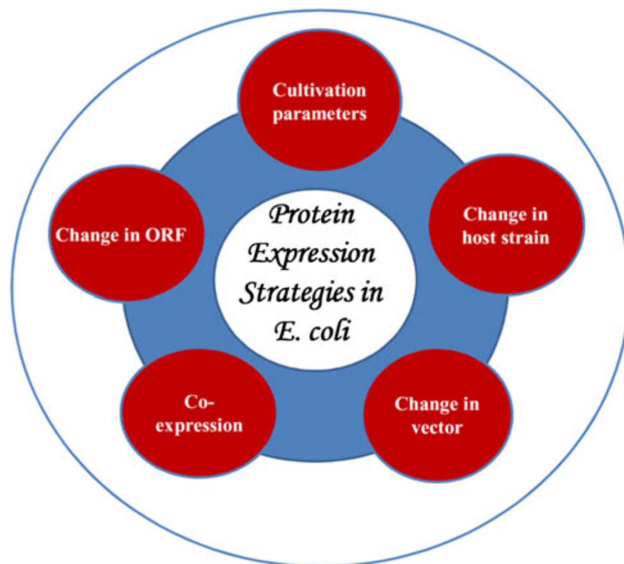

Fig. 2 Strategise for production of recombinant proteins in *E. coli*. Buffers, pH, ionic strength, temperature, time etc. are the important factors and variables for easy and enhanced production of recombinant protein in *E. coli*. By following these expression strategies and manipulating these factors researcher could get rid of troubles during recombinant protein production study

They are suitable for avoiding codon bias as well as allowing disulphide bond formation during gene expression. It would be ideal if one *E. coli* strain could possess all the characteristics required for heterologous gene expression such as, disulphide bond formation, rare t-RNA, low temperature adapted chaperone, slow expression etc. The information described here can subsequently be combined and utilized to express protein in *E. coli*, minimizing the current problems researchers face. Moreover, the detail described in this review regarding protein expression may eliminate barriers and allow for successful implementation and management of multi-parallel expression strategies. Based on the examples presented, it is evident that the discussed strategies have had a significant impact on recombinant protein expression in *E. coli.* This is of utmost importance in the future for enhancing production of recombinant proteins.

Acknowledgments G.J.G. acknowledges the University Grant Commission (UGC), New Delhi, India for endowing him the Dr. D. S. Kothari Post Doctoral Fellowship and Prof. Rakesh Bhatnagar [School of Biotechnology, Jawaharlal Nehru University (JNU), New Delhi, India] for providing the opportunity to work in his lab. The authors are thankful to the Departments of Biotechnology, JNU, New Delhi, India and National Institute of Technology, Raipur (CG), India.

References

1. Bruno M, Walker JE (1996) J Mol Biol 260:289–298
2. Butt TR, Edavettal SC, Hall JP, Mattern MR (2005) Protein Expr Purif 43:1–9
3. Collins-Racie LA, McColgan JM, Grant KL, DiBlasio-Smith EA, McCoy JM, LaVallie ER (1995) Biotechnology (NY) 13:982–987
4. Davis GD, Elisee C, Newham DM, Harrison RG (1999) Biotechnol Bioeng 65:382–388
5. deMarco A, Vigh L, Diamant S, Goloubinoff P (2005) Cell Stress Chaperones 10:329–339
6. Diamant S (2003) Mol Microbiol 49:401–410
7. diGuan C, Li P, Riggs PD, Inouye H (1988) Gene 67:21–30
8. Dinnbier U, Limpinse E, Schmid R, Bakker EP (1988) Arch Microbiol 150:348–357
9. Gräslund S, Nordlund P, Weigelt J, Gunsalus KC (2008) Nat Methods 5(2):135–146
10. Grunberg-Manago M (1999) Annu Rev Genet 33:193–227
11. Harrison RG (2000) Innovations 11:4–7
12. Hu J, Qin H, Gao FP, Cross TA (2011) Protein Expr Purif 80(1):34–40
13. Kido M, Yamanaka K, Mitani T, Niki H, Ogura T, Hiraga S (1996) J Bacteriol 178:3917–3925
14. Lars-Nieba L, Honegger A, Krebber C, Pluckthun A (1997) Protein Eng 10:435–444
15. Laurence DS, Cariot G, Vuillard L (2004) Protein Expr Purif 37(1):203–206
16. LaVallie ER, DiBlasio EA, Kovacic S, Grant KL, Schendel PF, McCoy JM (1993) Biotechnology (NY) 11:187–193
17. Lee N, Zhang SQ, Cozzitorto J, Yang JS, Testa D (1987) Gene 58:77–86
18. Lopez PJ, Marchand I, Joyce SA, Dreyfus M (1999) Mol Microbiol 33:188–199
19. Makino T, Skretas G, Georgiou G (2011) Microb Cell Fact 14:10–32
20. Makrides SC (1996) Microbiol Rev 60:516–538
21. Marblestone JG, Edavettal SC, Lim Y, Lim P, Zuo X, Butt TR (2006) Protein Sci 15:182–189
22. Michael W, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A, Wren BW, Aebi M (2002) Science 298:1790–1793
23. Missiakas D (1994) EMBO J 13(8):2013–2020
24. Novy R (1995) Innovations 3–7
25. Prinz WA, Aslund F, Holmgren A, Beckwith J (1997) J Biol Chem 272:15661–15667
26. Raina S, Missiakas D (1997) Annu Rev Microbiol 51:179–202
27. Redwan ELRM (2006) Arab J Biotechnol 9(3):493–510
28. Schlegel S, Löfblom J, Lee C, Hjelm A, Klepsch M, Strous M, Drew D, Slotboom DJ, de Gier JW (2012) J Mol Biol 423(4):648–659
29. Schlegel S, Rujas E, Ytterberg AJ, Zubarev RA, Luirink J, de Gier JW (2013) Microb Cell Fact 12(12):24
30. Schultz T, Liu J, Capasso P, de Marco A (2007) Biochem Biophys Res Commun 355:234–239
31. Smith DB, Johnson KS (1988) Gene 67:31–40
32. Steczko J, Donoho GA, Dixon JE, Sugimoto T, Axelrod B (1991) Protein Expr Purif 2:221–227
33. Valderrama-Rincon JD, Fisher AC, Merritt JH, Fan YY, Reading CA, Chhiba K, Heiss C, Azadi P, Aebi M, DeLisa MP (2012) Nat Chem Biol 8(5):434–436
34. Wagner S, Klepsch MM, Schlegel S, Appel A, Draheim R, Tarry M, Högbom M, Van Wijk KJ, Slotboom DJ, Persson JO, deGier JW (2008) Proc Natl Acad Sci 105(38):14371–14376
35. Zapun A, Missiakas D, Raina S, Creighton TE (1995) Biochemistry 34:5075–5089
36. Zuo X, Li S, Hall J, Mattern MR, Tran H, Shoo J, Tan R, Weiss SR, Butt TR (2005) J Struct Funct Genomics 6:103–111