International Journal of Biochemistry Research & Review
12(4): 1-10, 2016, Article no.IJBCRR.26639
ISSN: 2231-086X, NLM ID: 101654445

SCIENCEDOMAIN *international*
www.sciencedomain.org

Strategies for the Engineering of Recombinant Protein and its High Level Expression

Fayqa Komal[1*] and Faiza Asghar[1]

[1]School of Biological Sciences, University of Punjab, Qaid-e-Azam Campus, Lahore, Pakistan.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/IJBCRR/2016/26639
Editor(s):
(1) Cheorl-Ho Kim, Molecular and Cellular Glycobiology Unit, Department of Biological Science, Sungkyunkwan University, South Korea.
Reviewers:
(1) Anonymous, Iib-Intech, Conicet, Argentina.
(2) Kian Mau Goh, Universiti Teknologi Malaysia, Malaysia.
(3) Eduardo Mere Del Aguila, Universidade Federal do Rio de Janeiro, Brazil.
(4) Xudong Tang, Jiangsu University of Science and Technology, China.
Complete Peer review History: http://sciencedomain.org/review-history/15432

Review Article

Received 26th April 2016
Accepted 24th June 2016
Published 19th July 2016

ABSTRACT

A large number of proteins are efficiently produced by microbes. Protein engineering especially working with enzymes has become a very promising section of the biotechnology industry. Due to increasing demands of proteins, bioengineering strategies have been gaining importance to modify natural enzymes. Commercially, hundreds of proteins are produced, the production of recombinant proteins still constitutes a challenge in many cases. Most common protein engineering techniques include i) Directed evolution, ii) Site-directed mutagenesis, iii) Truncation, iv) Terminal fusion. After engineering the desired protein, there is another great challenge to get the high level expression and solubility of the proteins. Four levels of strategies can be used to increase the expression and solubility of recombinant proteins; (1) vector selection, (2) host selection, (3) fermentation optimization and (4) codon optimization. Here we present the latest methods of protein engineering and molecular expression of industrially important enzymes and to get good quality of recombinant proteins. In this article, we have reviewed the different approaches, common problems, their solutions and also covered pros and cons of many of the latest used techniques in this ever-growing field.

**Corresponding author: E-mail: komal.faiqa1@gmail.com;*

Keywords: Recombinant proteins; directed evolution; site directed mutagenesis; truncation; terminal fusion; vector selection; host selection; codon optimization.

1. INTRODUCTION

Protein engineering is emerging and remarkable technique and different strategies have been developed in a very short time to produce cellulases with improved qualities [1]. Development in field of biological sciences and recombinant DNA technology is emerging very rapidly [2]. It gives driving force to biologists and biochemists to produce more effective and new methods for screening of good mutants [3]. They can produce purified protein immediately by this technique instead of using kilograms of plant and animal tissues.

In former studies, this topic is studied many times with great detail under different headings. That's why this review is all about the most recent advances in protein engineering and its expression. We also discussed pros and cons of different modern options and approaches, currently considered very popular for expressing a great number of proteins, for young researchers in the field of making heterologous proteins.

2. STRATEGIES FOR PROTEIN ENGINEERING

Usually some proteins in their natural form are not up to the mark in their properties under inconsiderate industrial conditions, and thus optimization is essential to achieve an appropriate protein variant for production needs [4]. For enhancing protein characteristics these four experimental routes can be followed. When there is complete information available about protein structure, the prior approach would be rational design [5]. While in the absence of this information directed evolution, truncation and CBM fusion are considered as the best options.

2.1 Directed Evolution

In routine lab work, we have limited knowledge about structure and mechanisms of the protein of interest [6]. So, as an alternative measure, evolutionary methods are preferred. All these techniques are collectively termed as directed evolution [7]. It mimics darwinian optimization process, whereby the rightest subjects are selected from a group of variants [8]. Here we summarized the basic steps of directed evolution as: 1) generation of library of mutants, 2) screening/selection and 3) gene amplification. For the successful directed evolution experiment the utmost and vital step is selecting the suitable techniques for creating a library and screening or selection [9].

There are three best strategies so far for creating a library of rightest variants: 1) random mutagenesis, 2) semi-rational design, and 3) gene shuffling. In first approach, commonly used mutations can either be point mutations, in-del mutations, inversions, or frame-shift mutations, whereas the combination of first approach with site-directed mutagenesis is known as Semi-rational design [10]. In last technique, gene shuffling, there is an exchange of fragments between genes for creating a library of chimeric descendants. In directed evolution, gene shuffling is normally done by homologous recombination (HR) or non-homologous recombination (NHR).

Once library of mutants is generated, next step is to select the desired mutant. This selection is done either by screening or selection process [11]. Each mutant in the library is evaluated individually in screening process while in selection methods the entire library is evaluated at once[12]. For instance, use of antibiotics to bacteria on an agar plate and only resistant clones will grow. Practically selection methods are considered preferable.

The only difficulty in these directed evolution experiments, is finding the desirable variant from the library [13]. As size of the library is typically very large ($> 10^{4-6}$), and close evaluation of each variant is not feasible. Though, once an evolved enzyme with improved characteristics is found, there are whole lot chances of success for you to be part of in biochemical industry [14,15].

2.2 Site Directed Mutagenesis

Site-directed mutagenesis is one of the most important *in-vitro* method [16] that is used for the modification of enzyme and to introduce desired properties into protein of interest. It is commonly used in protein chemistry to analyze the function of specific amino acids and to enhance their properties, such as pH, thermal stability and specific activity of different proteins. This

technique is as famous as rational design but requires huge data on the protein structure [17]. To be successful, modification of the desired region usually involves knowledge of structure and function of the existing region and the desired one. Sometimes if the structure and catalytic mechanism of target enzyme are well known, molecular modification for the desired function may not be achieved [18]. The effect of specific amino acids and their role in enzymatic properties can be easily achieved by this technique and it can allow major modifications of an enzyme [19].

In rational design, molecular engineering demands the deep understanding of composition and configuration of protein by single and combinational mutation. Mutation at particular amino acid can be made by overlap PCR method [20] and then the resultant mutant protein is expressed and purified. After purification, native and mutant proteins are evaluated to access weather the required characteristics have been incorporated in protein or not. Second and third round of mutations followed by expression, purification can also be done for the further improvement in proteins characteristics [21].

For example, Hsieh and Vaisvila described single site-directed mutagenesis and multiple mutations as simplified methods for mutagenesis [22]. When cysteine residue at position 22 in oryza sativa is replaced with alanine in Phi-class glutathionine S-transferase F3, there was a 2.2-fold increase in km value than that of the wild type [23]. When site-directed mutagenesis is used with the combination of other techniques can considerably increase the properties of the enzymes. For instance, the ability of maltose-binding of protein-fused Hepl from recombinant *Escherichia coli* was reported to be increased rapidly by using site-directed mutagenesis in combination with calcium [24].

2.3 Truncation

Truncation allows generation of a large number of enzyme variants by relatively simple and fast methods [25]. Extensive screening of the library generated can be employed to identify and isolate the enzyme variants that are more active and stable under a range of conditions as compared to their native counterparts. Production of highly active and thermostable enzymes can make a major contribution in reducing the costs for the production of second-generation bioethanol, in addition to other important applications. Truncation is an incredible approach for improving enzyme performance without knowing the protein structure and enzyme-substrate interactions [26]. Different methods and techniques are available to generate molecular diversity but truncation would be best option for the Creation of hybrid enzymes. This technique is for combining two genes randomly. The primary advantage of truncation is that there is no requirement of sequence similarity for the two genes [27].

Truncation not only improves the properties of industrial enzymes but also makes the process very cost effective and easy to perform as compared to other conventional procedures. Cost effective enzymes with improved thermostability, increased catalytic efficiency and enhanced solubility are highly demanded for the industrial purposes [25]. For instance celluloses and hemicelluloses are first hydrolyzed by enzymatic treatment followed by bacterial or yeast fermentation in lignocellulose bioethanol technology. The conventional acid hydrolysis requires expensive reactors and produces reversion compounds at the end of reaction. Enzymatic digestion is quite specific, and environment friendly as compared to the acid hydrolysis. The production cost of these enzymes can be decreased by adopting truncation techniques [28]. Generation of mutants via truncation and direct transformation of desired cells for soluble expression of protein bypasses the costly and time consuming steps of conventional cloning [29].

2.4 Fusion

Protein stability can be enhanced by two or more genes or cDNA which can be from same or different source and fused together to make new chimeric nucleotide sequence [30]. This nucleotide sequence coded the protein which is known as chimeric protein (or chimera). Chimeras can be made by ligating different cDNA or by overlap extension PCR. To express chimeric gene, they 1st have to be cloned into a plasmid. Sometimes these chimeric proteins may have better properties as compared to original ones. In human beings, an example of chimeric protein is Philadelphia chromosomal mutation leads to myelogenous leukemia [31]. Chimeric proteins are also used to study disease development in humans, different toxins and antibodies can be engineered to proteins and their effects in the body can be observed. By

domain rearrangements chimeric variants of cellulase can also be made [32].

CBM's (carbohydrate binding modules) are most distinctive, diverse and robust. It is the important component of both free cellulases and cellulosome, bind tightly to crystalline cellulose and thus play a key role in cellulose degradation through their substrate targeting capacity [33,34]. One of the important types of CBM's ,known as CBM3c, is fused to the catalytic module of family 9 glycoside hydrolases (GH9s), and serve to alter the enzymatic characteristic of the parent protein from a standard endoglucanase to a processive enzyme [35,36].Sumo fusion is also important for the soluble expression of many therapeutic proteins [37]. In short, production of new enzymes with required characteristics can be expressed in suitable hosts to produce these enzymes in large quantity [38]. Bulk production of these enzymes can find many applications in a number of industries.

3. STRATEGIES FOR HIGH LEVEL EXPRESSION OF FOREIGN GENE

Due to increased industrial applications demand of protein engineering is increasing immensely [39,40]. Theoretically it looks very simple as in such cases you take your desired gene, clone it in expression vector, and after transformation in specific host, your protein is ready for purification and characterization. Practically, protein engineering is not an easy task because; there are many factors which can go wrong. For example inappropriate usage of the cloning vector [41], metabolic burden on expression host [42], inclusion body (IB) formation [43], protein inactivity [44], codon optimization [45], fermentation optimization [46] and even not getting any protein expression at all are some of the considerable difficulties. So far, there is no single solution exists for successfully the production and expression of all recombinant proteins. As a substitute, it is always beneficiary to have access to a wide collection of expression techniques.

3.1 Vector Selection

Vectors are also known as replicons, as they undergo replication as autonomous units [47]. An expression vector must have features such as strong promoter, replicons, selection markers and multiple cloning sites containing unique restriction sites for the accurate insertion of the gene. It is possible to clone the desired gene directly in expression vector but cloning vectors may also use for this purpose. A large number of cloning vectors are available and it is very easy to get lost in the catalogue when choosing the suitable vector. For this reason, all the characteristics have to be carefully evaluated according to the desired gene. Choosing the vector may depend on copy number of the vector. Replicon is considered as the control center of the copy number [48]. It is logical to think that high copy number is useful as it produces more yield of recombinant protein as many expression entities reside in the cell [49].

Most commonly used expression plasmid is the pET vector with 6-Histidine tag. Generally His-Tag is very useful and smaller affinity tag for the proteins primarily expressed as inclusion bodies (IB). These are formed by an unequal equilibrium between protein aggregation and solubilization [50]. Under fully denaturing environment, His-tag is used to achieve affinity purification to solubulize the protein [51]. But if it fails to prevent the formation of 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' [52,53].

The pET vector is provided by 'Novagen'. A continuous stretch of consecutive 6-Histidine residues may consider as the main possibility for the decrease in solubility of the fused protein. If consecutive Histidine creates hinderence in solubilty, pEt vector can be replaced by pHAT. Dual and triple expression of recombinant proteins can be achieved by p15A ori which is present in pACYC and pBAD series of plasmids and pSC101 plasmids respectively (pACYC and pBAD series of plasmids,10–12 copies per cell; [54,55] (pSC101 <5 copies per cell; [56]. Alternatively, Duet vectors (Novagen) streamlines co-expression of two genes in the single plasmid. The Duet plasmids carries two multiple cloning sites, each preceded by a T7 promoter, a *lac* operator and a ribosome binding site.

3.2 Organism Selection

Expression systems are mainly used in research [57], medicine [58], life sciences [59] and biotechnological processes [60]. They are genetic constructs that are referred to as host and designed to make proteins, or RNA, either inside or outside a cell.

An expression vector must have features such as the system must be easy to culture and maintain, grow rapidly, and produce large amounts of protein. Many different host systems may be used for expression for example bacteria [58], yeast [61,62], unicellular algae [63] and filamentous fungi [64]— each expression system has distinct strengths and weaknesses [12,65]. For example, a prokaryotic expression system may not be used for the eukaryotic proteins in which post-translation modifications such as protein glycosylation are required [66].

The initiation of a whole process depends on the host cell machinery making the selection of an appropriate host a vital step. For large scale production of enzyme, *Escherichia coli* expression system is the easiest and quickest method. Another widely used expression system is use of recombinant baculoviruses (insect viruses). For the enhanced production of biologically active mammalian proteins, baculovirus insect cell expression system is far better than yeast and bacterial expression systems [67]. Insect cells are eukaryotic, resembles closely that of the mammalian cells, thus capable of performing protein folding, protein oligomerization and post-translational modifications (e.g. palmitolation, glycosylation, myristolation, amidation and fatty acid acylation) [68]. In this review, we will precisely focus on *Escherichia coli* as in the past, many reviews have covered other expression systems in detail [69-72].

*E. coli h*as highly developed genetic system and has many advantages as the expression host. (i) It has unequaled fast growth kinetics. Its generation time is about 20 minutes under optimized conditions [73]. (ii) As *E. coli* is known as work horse organism for its simple and easy physiology makes easy to get high cell density cultures and produce large quantities of proteins (iii) low cost culture methods (iv) Foreign DNA is fast and easy to transform. Transformation of plasmid in *E. coli* can be achieved in as few as 5min [74]. (v) flexible system – can carry plasmids with multiple promoters, tags and restriction sites.

3.3 Fermentation Conditions

Just like protein engineering, techniques for protein expression are also very important for enhanced protein production. Fermentation conditions are one of the important factors required for the optimum protein production in *E. coli* in labs [75]. The considerable fermentation conditions include culture systems, nutrient composition, temperature, pH and duration of the experiment. Continuous culture system, batch culture system and fed batch culture system [76] are the three types of the high-cell-density culture systems for the maximum production of recombinant proteins (100 g dry cell weight/liter). These culture systems may have substantial metabolic effects on protein production and on cells as well, whole experiment must be monitored carefully then. For example, temperature and composition of the culture system are the important variables for the translation of different mRNAs.

It has been suggested that for the maximum release of periplasmic proteins into the medium, avoiding considerable cell lysis, addition of glycine in the growth medium is a very good option [77,78]. In a same way, addition of sorbitol and glycyl betaine in the growth medium can enhance the solubility and activity of a protein by>400-fold [58]. Major challenges faced in the production of recombinant protein at high cell density culture batches includean inadequate supply of oxygen and high levels of carbon dioxide [79], the ultimate reasons for the decreased growth rates and acetate saturation [80] causing increased output of heat and decline in the mixing efficacy.

3.4 Codon Usage

'Synonymous codon bias' is seen in Genes of both prokaryotes and eukaryotes [81]. Some of the observations based on the methodical study of codon usage patterns in *E. coli* are enlisted here. (i) In all degenerate codon families there is a preference for one or two codons. (ii) Some of the codons are common amongst different genes regardless of the profusion of the proteins; for instance, commonly used codon for proline is CCG. (iii) Frequency of codon bias system is in direct relation with the expression rates of genes, genes with higher expression rates exhibit a greater degree of codon bias vice versa [82]. (iv) The rate of similar codon usage might reveal the profusion of their related tRNAs [83].

From these interpretations we can say that in *E. coli,* less efficient genes are the ones with increased rare codons and thus by replacing rare codons with their synonymous common ones can enhance protein production [45]. For example in the expression of several mammalian genes tRNA Arg (AGG/AGA) is considered as a limiting

factor [84] , because these are the rarely used codons in *E. coli*. Likewise, Goldman et al. [85] testified about the strong translational blocking of mRNA in presence of leucine and arginine, when there was a continuous stretch of rare codons close to the 59 end of the mRNA.

So far, though, it's been a tough job to design some general and definite "rules" to predict a relation of the content of rare codons in a specific gene with its efficacy of expression in *E. coli* [86]. For the final documentation of accurate results, a long series of variables are monitored closely e.g; occurrence of rare codons, position of their relative tRNA [87], secondary structure of the mRNA, and some other effects. From all this discussion it would not be wrong to say that quantity and quality of a protein is directly affected by its codon context [88].

4. CONCLUSION

Protein engineering, sub discipline of genetic engineering has played the key role in improving commercial enzymes and finding their new applications. Here we have highlighted some of the recent and promising tools and strategies which are being applied nowadays for bioengineering. Directed evolution and In vitro site-directed mutagenesis are considered as invaluable techniques for studying protein structure-function relationships and gene expression, and for carrying out protein modifications. Other techniques of protein engineering which are reported before chiefly include terminal fusion and truncation**.** In order to decide which sites are responsible for the improvement in their catalytic performance, gene sequences of parental proteins are compared with that of positive mutants. Then site-directed mutagenesis is performed for the further amplification in performance. But a plus point of directed evolution over site directed mutagenesis is that there is no need to understand the mechanism of the desired activity or how mutations would affect it. Truncation and terminal fusion not only improves the properties of industrial enzymes but also makes the process very cost effective with improved thermo stability, increased catalytic efficiency and enhanced solubility. These chimeric and truncated proteins are also used to study disease development in humans, different toxins and antibodies, and their effects in the body. After protein engineering, our next target was to improve the expression of recombinant proteins. We have discussed here a complete list of expression systems which are evolving continuously with the emergence of new vectors, various hosts, culture parameters of recombinant host strain and codon optimization. For optimized cloning of desired protein, pET vector series is considered as an efficient prokaryotic expression vector. While for the efficient expression of cloned protein, *E. coli* is considered as the easiest vector to handle. In *E. coli*, expression and solubility of recombinant proteins can also be increased by replacing the rare codons with high-usage codons and by optimizing the culture condition of recombinant *E. coli.* We hope that this review demonstrates the real progress, being made in modifying the commercially important enzymes.

ACKNOWLEDGEMENTS

It is a pleasure to acknowledge Dr Waheed Akhtar who taught us about protein engineering. We are grateful to Razia Tajwar, Sahar Shahid and Shaista Bashir for their critical reading of the manuscript and their thoughtful comments. Any errors are solely our own responsibility. We appreciate the constructive comments of the reviewers and their suggestions on improving the manuscript.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Vazquez E, et al. Modular protein engineering in emerging cancer therapies. Current Pharmaceutical Design. 2009; 15(8):893-916.
2. Glick BR, Pasternak JJ. Principles and applications of recombinant DNA. ASM, Washington DC. 1998;683.
3. Leatherbarrow RJ, Fersht AR. Protein engineering. Protein Engineering. 1986;1(1):7-16.
4. Gotor V, Alfonso I, García-Urdiales E. Asymmetric organic synthesis with enzymes. John Wiley & Sons; 2008.
5. Chica RA, Doucet N, Pelletier JN. Semi-rational approaches to engineering enzyme activity: Combining the benefits of directed evolution and rational design. Current Opinion in Biotechnology. 2005; 16(4):378-384.
6. Bloom JD, et al. Evolving strategies for enzyme engineering. Current Opinion in Structural Biology. 2005;15(4):447-452.

7. Doucet N, Pelletier J. Directed evolution of enzymes. Canadian Chemical News. 2004; 56(4):20-21.
8. Turner NJ. Directed evolution of enzymes for applied biocatalysis. Trends in Biotechnology. 2003;21(11):474-478.
9. Bornscheuer UT. Directed evolution of enzymes. Angewandte Chemie International Edition. 1998;37(22):3105-3108.
10. Reetz MT. Combinatorial and evolution-based methods in the creation of enantioselective catalysts. Angewandte Chemie International Edition. 2001;40(2): 284-310.
11. Becker S, et al. Single-Cell high-throughput screening to identify enantioselective hydrolytic enzymes. Angewandte Chemie International Edition. 2008;47(27):5085-5088.
12. Adrio JL, Demain AL. Recombinant organisms for production of industrial products. Bioengineered Bugs. 2010;1(2): 116-131.
13. Jäckel C, Kast P, Hilvert D. Protein design by directed evolution. Annu. Rev. Biophys. 2008;37:153-173.
14. Reetz MT, Jaeger KE. Superior biocatalysts by directed evolution, in biocatalysis-from discovery to application. Springer. 1999;31-57.
15. Binod P, et al. Industrial enzymes—Present status and future perspectives for India. J. Sci. Ind. Res. India. 2013;72:271-286.
16. Ho SN, et al. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene. 1989;77(1):51-59.
17. Kunkel TA, Roberts JD, Zakour RA. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods in Enzymology. 1987;154:367-382.
18. Arnold FH, et al. How enzymes adapt: Lessons from directed evolution. Trends in Biochemical Sciences. 2001;26(2):100-106.
19. Wang J, et al. The molecular nature of the F-actin binding activity of aldolase revealed with site-directed mutants. Journal of Biological Chemistry. 1996;271(12):6861-6865.
20. Aiyar A, Xiang Y, Leis J. Site-directed mutagenesis using overlap extension PCR. In vitro Mutagenesis Protocols. 1996;177-191.
21. Wells JA, Vasser M, Powers DB. Cassette mutagenesis: An efficient method for generation of multiple mutations at defined sites. Gene. 1985;34(2-3):315-323.
22. Hsieh PC, Vaisvila R. Protein engineering: single or multiple site-directed mutagenesis. Enzyme Engineering: Methods and Protocols. 2013;173-186.
23. Jo HJ, et al. Site-directed mutagenesis of cysteine residues in Phi-class glutathione S-transferase F3 from *Oryza sativa*. Bulletin of the Korean Chemical Society. 2012;33(12):4169-4172.
24. Chen S, et al. Combination of site-directed mutagenesis and calcium ion addition for enhanced production of thermostable MBP-fused heparinase I in recombinant *Escherichia coli.* Applied Microbiology and Biotechnology. 2013;97(7):2907-2916.
25. Ostermeier M, Nixon AE, Benkovic SJ. Incremental truncation as a strategy in the engineering of novel biocatalysts. Bioorganic & Medicinal Chemistry. 1999; 7(10):2139-2144.
26. Zhiwei Chen JHP, Hanbin Liu, Huu M. Tran, Nathan SY. Hsu, Dean Dibble, Seema Singh, Paul D. Adams, Rajat Sapra, Masood Z. Hadi, Blake A. Simmons, Kenneth L. Sale. Improved activity of a thermophilic cellulase, Cel5A, from *Thermotoga maritima* on Ionic liquid pretreated switchgrass. PLoS ONE. 2013;8.
27. Patrick WM, Gerth ML. ITCHY: Incremental truncation for the creation of hybrid enzymes. Directed Evolution Library Creation: Methods and Protocols. 2014; 225-244.
28. Hecky J, et al. A general method of terminal truncation, evolution, and re-elongation to generate enzymes of enhanced stability. Protein Engineering Protocols. 2007;275-304.
29. Zhao H, Chockalingam K, Chen Z. Directed evolution of enzymes and pathways for industrial biocatalysis. Current Opinion in Biotechnology. 2002; 13(2):104-110.
30. Li JWH, Vederas JC. Drug discovery and natural products: End of an era or an endless frontier? Science. 2009; 325(5937):161-165.
31. Redaelli S, et al. Activity of bosutinib, dasatinib, and nilotinib against 18 imatinib-resistant BCR/ABL mutants. Journal of Clinical Oncology. 2009;27(3):469-471.
32. Sajjad M, et al. Influence of positioning of carbohydrate binding module on the activity of endoglucanase CelA of

Clostridium thermocellum. Journal of Biotechnology. 2012;161(3):206-212.
33. Gunnarsson LC, et al. A carbohydrate binding module as a diversity-carrying scaffold. Protein Engineering Design and Selection. 2004;17(3):213-221.
34. Blake AW, et al. Understanding the biological rationale for the diversity of cellulose-directed carbohydrate-binding modules in prokaryotic enzymes. Journal of Biological Chemistry. 2006;281(39): 29321-29329.
35. Bayer EA, et al. Cellulose, cellulases and cellulosomes. Current Opinion in Structural Biology. 1998;8(5):548-557.
36. Sakon J, et al. Crystal structure of thermostable family 5 endocellulase E1 from *Acidothermus cellulolyticus* in complex with cellotetraose. Biochemistry. 1996;35(33):10648-10660.
37. Peroutka III RJ, et al. SUMO fusion technology for enhanced protein expression and purification in prokaryotes and eukaryotes. Heterologous gene expression in *E. coli*: Methods and Protocols. 2011;15-30.
38. Foo JL, et al. The imminent role of protein engineering in synthetic biology. Biotechnology Advances. 2012;30(3):541-549.
39. Kirk O, Borchert TV, Fuglsang CC. Industrial enzyme applications. Current Opinion in Biotechnology. 2002;13(4):345-351.
40. Gavrilescu M, Chisti Y. Biotechnology—a sustainable alternative for chemical industry. Biotechnology Advances. 2005; 23(7):471-499.
41. Friedman AM, et al. Construction of a broad host range cosmid cloning vector and its use in the genetic analysis of *Rhizobium mutants*. Gene. 1982;18(3): 289-296.
42. Bentley WE, et al. Plasmid-encoded protein: the principal factor in the "metabolic burden" associated with recombinant bacteria. Biotechnology and Bioengineering. 1990;35(7):668-681.
43. Kopito RR. Aggresomes, inclusion bodies and protein aggregation. Trends in Cell Biology. 2000;10(12):524-530.
44. Sørensen HP, Mortensen KK. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli.* Microbial Cell Factories. 2005;4(1):1.
45. Burgess-Brown NA, et al. Codon optimization can improve expression of human genes in *Escherichia coli:* A multi-gene study. Protein Expression and Purification. 2008;59(1):94-102.
46. Weuster-Botz D. Experimental design for fermentation media development: Statistical design or global random search? Journal of Bioscience and Bioengineering. 2000;90(5):473-483.
47. Haase SB, Calos MP. Replication control of autonomously replicating human sequences. Nucleic Acids Research. 1991; 19(18):5053.
48. Del Solar G, Espinosa M. Plasmid copy number control: An ever-growing story. Molecular Microbiology. 2000;37(3):492-500.
49. Rosano GL, Ceccarelli EA. Recombinant protein expression in *Escherichia coli:* Advances and challenges. Recombinant Protein Expression in Microbial Systems. 2014;7.
50. Carrio M, Villaverde A. Construction and deconstruction of bacterial inclusion bodies. Journal of Biotechnology. 2002; 96(1):3-12.
51. Carrio MM, Villaverde A. Protein aggregation as bacterial inclusion bodies is reversible. FEBS Letters. 2001;489(1):29-33.
52. Kimple ME, Sondek J. Overview of affinity tags for protein purification. Current Protocols in Protein Science. 2004;9.9:1-9.9. 19.
53. Costa S, et al. Fusion tags for protein solubility, purification, and immunogenicity in *Escherichia coli:* The novel Fh8 system. Recombinant Protein Expression in Microbial Systems. 2014;24.
54. Chang AC, Cohen SN. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. Journal of Bacteriology. 1978;134(3):1141-1156.
55. Guzman LM, et al. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. Journal of Bacteriology. 1995; 177(14):4121-4130.
56. Venkova-Canova T, Srivastava P, Chattoraj DK. Transcriptional inactivation of a regulatory site for replication of *Vibrio cholerae* chromosome II. Proceedings of the National Academy of Sciences. 2006; 103(32):12051-12056.
57. Macauley-Patrick S, et al. *H*eterologous protein production using the *Pichia pastoris*

expression system. Yeast. 2005;22(4): 249-270.
58. Terpe K. Overview of bacterial expression systems for heterologous protein production: From molecular and biochemical fundamentals to commercial systems. Applied Microbiology and Biotechnology. 2006;72(2):211-222.
59. Wurm FM. Production of recombinant protein therapeutics in cultivated mammalian cells. Nature Biotechnology. 2004;22(11):1393-1398.
60. Giddings G, et al. Transgenic plants as factories for biopharmaceuticals. Nature Biotechnology. 2000;18(11):1151-1155.
61. Gellissen G. Heterologous protein production in methylotrophic yeasts. Applied Microbiology and Biotechnology. 2000;54(6):741-750.
62. Cereghino JL, Cregg JM. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris.* FEMS Microbiology Reviews. 2000;24(1):45-66.
63. Fischer R, et al. Plant-based production of biopharmaceuticals. Current Opinion in Plant Biology. 2004;7(2):152-158.
64. Punt PJ, et al. *Filamentous* fungi as cell factories for heterologous protein production. Trends in Biotechnology. 2002; 20(5):200-206.
65. Demain AL, Vaishnav P. Production of recombinant proteins by microbes and higher organisms. Biotechnology Advances. 2009;27(3):297-306.
66. Sahdev S, Khattar SK, Saini KS. Production of active eukaryotic proteins through bacterial expression systems: A review of the existing biotechnology strategies. Molecular and Cellular Biochemistry. 2008;307(1-2):249-264.
67. Merrington CL, Bailey MJ, Possee RD. Manipulation of baculovirus vectors. Molecular Biotechnology. 1997;8(3):283-297.
68. Kato T, et al. Silkworm expression system as a platform technology in life science. Applied Microbiology and Biotechnology. 2010;85(3):459-470.
69. Zhang J, et al. Expression of plectasin in Pichia pastoris and its characterization as a new antimicrobial peptide against *Staphyloccocus* and *Streptococcus*. Protein Expression and Purification. 2011; 78(2):189-196.
70. Swiech K, Picanço-Castro V, Covas DT. Human cells: New platform for recombinant therapeutic protein production. Protein Expression and Purification. 2012;84(1):147-153.
71. Kerrigan JJ, et al. Production of protein complexes via co-expression. Protein Expression and Purification. 2011;75(1):1-14.
72. Jarvis DL. Baculovirus–insect cell expression systems. Methods in Enzymology. 2009;463:191-222.
73. Sezonov G, Joseleau-Petit D, D'Ari R. *Escherichia coli* physiology in Luria-Bertani broth. Journal of Bacteriology. 2007; 189(23):8746-8749.
74. Pope B, Kent HM. High efficiency 5 min transformation of *Escherichia coli.* Nucleic Acids Research. 1996;24(3):536-537.
75. Patnaik R, et al. L-Tyrosine production by recombinant *Escherichia coli:* Fermentation optimization and recovery. Biotechnology and Bioengineering. 2008; 99(4):741-752.
76. Kleist S, et al. Optimization of the extracellular production of a bacterial phytase with *Escherichia coli* by using different fed-batch fermentation strategies. Applied Microbiology and Biotechnology. 2003;61(5-6):456-462.
77. Geciova J, Bury D, Jelen P. Methods for disruption of microbial cells for potential use in the dairy industry—a review. International Dairy Journal. 2002;12(6): 541-553.
78. Pierce J, et al. Factors determining more efficient large-scale release of a periplasmic enzyme from *E. coli* using lysozyme. Journal of Biotechnology. 1997; 58(1):1-11.
79. Kumar R, Shimizu K. Transcriptional regulation of main metabolic pathways of cyoA, cydB, fnr, and fur gene knockout *Escherichia coli* in C-limited and N-limited aerobic continuous cultures. Microbial Cell Factories. 2011;10(1):1.
80. Aristidou AA, San KY, Bennett GN. Metabolic engineering of *Escherichia coli* to enhance recombinant protein production through acetate reduction. Biotechnology Progress. 1995;11(4):475-478.
81. Yang H, et al. Comparative analysis of heterologous expression, biochemical characterization optimal production of an alkaline α-amylase from alkaliphilic Alkalimonas amylolytica in *Escherichia coli* and *Pichia pastoris.* Biotechnology Progress. 2013;29(1):39-47.
82. Stenström CM, et al. Codon bias at the 3′-side of the initiation codon is correlated

with translation initiation efficiency in *Escherichia coli.* Gene. 2001;263(1):273-284.

83. Baca AM, Hol WG. Overcoming codon bias: a method for high-level overexpression of *Plasmodium* and other AT-rich parasite genes in *Escherichia coli.* International Journal for Parasitology. 2000;30(2):113-118.
84. Calderone TL, Stevens RD, Oas TG. High-level misincorporation of lysine for arginine at AGA codons in a fusion protein expressed in *Escherichia coli.* Journal of Molecular Biology. 1996;262(4):407-412.
85. Goldman E, et al. Consecutive low-usage leucine codons block translation only when near the 5′ end of a message in *Escherichia coli.* Journal of Molecular Biology. 1995;245(5):467-473.
86. Bailly-Bechet M, et al. Codon usage domains over bacterial chromosomes. PLoS Comput Biol. 2006;2(4):e37.
87. Kleber-Janke T, Becker WM. Use of modified BL21 (DE3) *Escherichia coli* cells for high-level expression of recombinant peanut allergens affected by poor codon usage. Protein Expression and Purification. 2000;19(3):419-424.
88. Chaney JL, Clark PL. Roles for synonymous codon usage in protein biogenesis. Annual Review of Biophysics. 2015;44:143-166.

Peer-review history:
The peer review history for this paper can be accessed here:
http://sciencedomain.org/review-history/15432