Microbial Production of Industrial Proteins and Enzymes Using Metabolic Engineering

12

Pooja Doshi, Manju Shri, Poonam Bhargava, Chaitanya G. Joshi, and Madhvi Joshi

Abstract

Metabolic engineering is a field of science, which takes advantage of previously gathered information about a particular pathway in a living organism and utilizes this for the improvement of product that could be either metabolite, enzyme, or any protein. Advances in various field of science specifically r-DNA technology, bioinformatics, synthetic biology, molecular genetics as well as other protein engineering technologies had given wings to metabolic engineering. Metabolic engineering has the capacity to mold the flux of a completely enzymatic pathway to a very newly designed pathway. It allows the modulation and production of either previously working metabolite or the production of a new novel enzyme in a different microbial strain. In the present era, there is huge demand of microbial enzymes and proteins for various purposes such as medication, oil and gas industry, dairy industry, baking industry, etc. Microbial strains are utilized as micro factories for the production of microbial enzymes and proteins via metabolic engineering. Therefore, in this book chapter we are dealing with the various criteria that are utilized for the selection of the strains, various approaches that are routinely utilized for the higher expression of genes, as well as various metabolic engineering strategies.

P. Doshi · M. Shri
Gujarat Biotechnology Research Center, Department of Science and Technology, Government of Gujarat, Mahanagar Seva Sadan Building (MS Building), Gandhinagar, Gujarat, India

P. Bhargava · C. G. Joshi · M. Joshi (✉)
Gujarat Biotechnology Research Center, Department of Science and Technology, Gandhinagar, Gujarat, India
e-mail: jd1-gbrc@gujarat.gov.in

12.1   Introduction

Microbes have been traditionally used in fermentation processes to obtain desired products. Scientific studies of the fermentation processes have led to the great advancements in this area. Many significant products, which were previously synthesized by chemical processes, are now being produced by fermentation using microbial cell factories and biotransformation. Microbes have various features such as to the ease of their mass cultivation, fast growth, use of cheap substrates and the diversity of the potential products produced by them, which makes them suitable for the production of value-added proteins. In the recent years, the application of microbial proteins has been increased enormously. The industries involving the use of microbial proteins include detergents, starch processing, animal feed, paper and pulp, fruit/vegetable processing, oil and gas, dairy, brewery, textile, baking, tanning, and so on. There could be further increase in the production as well as supply of the microbial products by improving our understanding of the recombination in microbes, metagenome mining, various fermentation processes, and recovery methods.

   Metabolic engineering is a new emerging field with huge impact on human socioeconomic level, for example, its applications in the production of fuels, materials, chemicals, medicine, and pharmaceuticals where modifications are made at the genetic level. It can be defined as enhancement of formation of desired product and/or cellular properties via editing-specific biochemical reaction(s) involved in its formation or the complete designing and introduction of new reactions with the use of r-DNA technology. The targeted biochemical reaction needs to be very specific to obtain maximum yield of the desired product. The procedure involves identification of the target reactions followed by the use of r-DNA techniques to augment, knockdown or delete, transfer, or deregulate the corresponding genes or their protein products. In this book chapter, we are dealing with the various criteria that are utilized for the selection of the strains, various approaches that are routinely utilized for the higher expression of genes, as well as various metabolic engineering strategies.

12.2   Strain Selection Criteria

There is a vast diversity of microbes available in nature. Many microbes especially those isolated from extreme environments are being explored as a source for production of industrially important enzymes since their optimum growth conditions help avoid contamination and produce more stable enzymes. The microbial strain to be used for industrial production of desired protein varies greatly and is usually selected on the basis of the following:

1. The input of efforts required in engineering and the toolsets as well as the availability of resources.
2. The nature and compatibility of the product in the selected host.

3. Metabolic requirements of the selected microbe for the production of the desired protein, viz. synthesis pathways, substrates/precursors, enzymes, and cofactors.

There are various factors which must be considered, such as metabolic resources, secretion of products, proper folding of proteins when selecting a microbe as a host for metabolic engineering. In the coming section, these factors will be discussed briefly.

12.2.1 Metabolic Resources

For the production of protein of interest by a microbe, it requires the precursors and various organic and inorganic cofactors. Cofactors may either act as redox carriers for the biosynthetic reactions or mediate energy transfer for the cell. Manipulation of cofactors include change in culture conditions, modification of pathway for increased availability or introduction of novel cofactors for enhancement of biochemical processes and overproduction of protein of interest (Wang et al. 2013). The advancements in the tools of bioinformatics and the availability of whole genome sequences of microbes helped a lot in exploring their biodiversity and evaluating the diverse potential of the microbial hosts thus enabling to identify the best potential host that can accumulate the pathway for production of desired protein. The availability of databases and computerized methods of model building like ModelSEED (Henry et al. 2010) and Path2models (Buchel et al. 2013) aid in the development of models for potential hosts for production of desired protein. There are various other tools also available like MetaNetX that can facilitate the direct assimilation of the new synthetic pathways available in other genome-scale models, where in their interactions and effects on the host metabolic network can be understood.

12.2.2 Minimum Metabolic Adjustment

Metabolic engineering of bacteria many a times forces it to pass through the evolutionary pressure to overcome its wild-type form and function and attain optimality to produce protein of interest (Fong and Palsson 2004). Therefore, it might be always advantageous to use a microbial host requiring marginal metabolic adaptations and minimum adjustment through progression (Fisher et al. 2014). Advancements in synthetic biology is making use of different bioinformatic tools for selection of microbial host for different de novo biosynthetic pathways for the production of protein of interest. Toolsets used for modeling and substrate-based analysis make use of global sensitivity analysis and agent-based modeling for screening significant components. Along with components, various other parameters such as genome-scale metabolic flux modeling have been done for identification of metabolic networks and medium formulations so that the expression of a biosynthetic pathway can be maximized (Apte et al. 2014).

12.2.3 Secretion of Proteins

The kinetics of protein production and purification vary greatly from lab scale to industrial scale. When going for industrial production, it is an added advantage if the protein of interest is secreted extracellularly. It aids in down streaming as well as the yield could be higher in comparison to the intracellular expression. *Bacillus* spp. are favored as the host for secretion of protein in comparison to *E. coli*. However, research is still ongoing for the identification of other protein secretion hosts (Ferrer-Miralles and Villaverde 2013a), such as Streptomyces (Okesli et al. 2011) halophiles that enhance solubility (Tokunaga et al. 2010) and yeast (Mattanovich et al. 2012). Yeast has some other advantages also such as posttranslational modifications. The final protein of interest can influence the microbial host as well.

12.2.4 Genomic Toolsets

Although *E. coli* is a common and preferred host for expression and production of recombinant protein due to its wide exploration both genetically and physiologically, other microbes are also being investigated in the same direction for protein production by metabolic engineering on an industrial scale. Toolsets like exonucleotide-based Gibson assemblies (Zhang et al. 2012) are of great help that increase the reliability of the assembly to be used for the protein synthesis. This assembly method is useful for the construction of natural and synthetic genes, pathways, and entire genomes (Fisher et al. 2014).

12.2.5 Proper Folding and Functionality of Protein of Interest

A critical step in recombinant protein production is to obtain properly folded and functional protein in the host used for expression. In case of use of metabolic engineering especially when expression of eukaryotic proteins is required, various factors like availability of the necessary chaperons and chapronins as well as transcript reading for synthesis and folding and posttranslational modifications to yield a functional protein need to be taken into consideration (Hartl et al. 2011; Bernal et al. 2014; Osterlehner et al. 2011). Eukaryotic protein expression in microbes where posttranslational modifications are required usually made use of yeasts like *Pichia pastoris* as a host besides the human and baculovirus-based insect cell lines. There is also significant interest being shown in transferring the posttranslational capabilities of eukaryotic cells to other microbes as well (Fisher et al. 2014). Heterologous expression, proper folding and N-linked glycosylation of eukaryotic proteins AcrA and IgG have been achieved in *E. coli* using combinatorial libraries, codon optimization, and shotgun proteomics (Pandhal et al. 2013). Earlier metabolic engineering experiments used to fail at the stage of translation because of lack of tools to check if the heterologous protein will be expressed properly or not. A translational coupling cassette is now available to quickly determine if the

heterologous mRNA will be translated in the chosen host along with its level of expression, even when it expresses large multidomain enzymes. It also helps to isolate the translation problems to the C-terminal domains, and to optimize conditions for expressing a codon-optimized sequence variant (Mendez-Perez et al. 2012).

12.3   Approaches to Attain Higher Expression of Industrial Enzymes

12.3.1 E. coli as a Host

E. coli is still one of the favorite organisms for the expression of heterologous proteins. Its usage as a microbial cell chassis is very highly acknowledged and well explored. Multiple studies carried out on its expression and regulation have made it the most popularly used expression platform. For *E. coli* system, various molecular biology tools and protocols are available which give higher expression and production of heterologous proteins. There is a huge number of expression plasmids, and many engineered strains and diverse cultivation strategies are also available for *E. coli* system (Rosano and Ceccarelli 2014). The different levels at which the E. coli expression system can be regulated are described in the following.

12.3.1.1 Transcriptional Regulation

A strong and controlled promoter as well as tight regulation of expression of protein is a requisite for efficient and high-level recombinant protein expression in E. coli chassis. In E. coli, large-scale production of protein makes use of chemical or thermal inducers for expression (Chao et al. 2004). Promoters that are tightly regulated help in designing many novel and highly repressible or inducible expression systems. Studies on regulation of promoters help in providing vital tools as information on regulation for gene expression.

Besides transcription promoters, transcription terminator elements are also crucial in controlling the expression and stability of heterologous gene expression, by enhancing the mRNA stability, which further increases protein expression (Newbury et al. 1987).

12.3.1.2 Translational Regulation

Translational regulation involves controlling and optimizing the factors involved in translation. Regulation can be done at the level of initiation codon, secondary structures at the site of initiation, or even stability of translated mRNA (Sprengart et al. 1996). The initiation codon AUG has been shown to be much more efficient as compared to GUG or UUG. The residues at $3'$ end of initiation codon especially the second one have been shown to influence the translation rate. The Shine-Dalgarno sequence upstream to the initiation codon that initiates translation is also shown to be more proficient than other translational initiators. The efficacy also varies directly to the distance of the translation initiation sequence (Ringquist et al. 1992). The bases

458–466 of the 16S rRNA of *E. coli* when placed upstream of the ribosomal binding site (RBS) has also shown to increase the translation efficiency by 110-fold (Olins and Rangwala 1989). All these factors can be regulated and can be taken into consideration during strain development.

12.3.1.3 Enhancement by Formation of Additional Proteins

Attempts for overexpression of proteins many times lead to accumulation of inclusion bodies within the cell. Purification and refolding to obtain active proteins from inclusion bodies is many a times a tedious job and leads to losses in downstreaming. Methods like use of molecular chaperones and low temperatures has been shown to be efficient in reducing the inclusion body formation and proper refolding of recombinant proteins. Proper folding of proteins is influenced by the oxidative environment of the periplasm and proper signal peptide cleavage during translocation. Various strategies have been used for efficient transport and protein folding. These include overproduction of signal peptidase I (Zhang et al. 1997), simultaneous expression of the secEand prlA4 genes (Pérez-Pérez et al. 1994), addition of Golgi retention or endoplasmic reticulum sequences (Zhan et al. 1998), and mutations in secY (Brinkworth et al. 2011). Uthandi et al. 2012 carried out deletion of the twin-arginine translocation motif, whereas Brinkworth et al. 2011 has described use of type III secretion chaperone to facilitate translocation.

The extracellular secretion of proteins is always favored over the intracellular expression due to several advantages like high expression levels and simplified downstream purification and protein folding. Since E. coli secretes very few proteins, use of signal peptide mutations compatible to E. coli membrane (Ismail et al. 2011) and limited leakage of outer membrane by synergistic use of EDTA and lysozyme (Liu et al. 2012) have also been employed for extracellular protein secretion.

12.3.1.4 Use of Fusion Proteins or Molecular Chaperones

A fusion protein is the product of two or more genes that are translated together with no stop codon in between them. In protein overexpression systems, fusion proteins serve to increase protein yield due to various modes of action like increased solubility of expressed protein, improved folding, efficient mRNA translation (Rosano and Ceccarelli 2014). Translational fusion of trpE gene fragments has positive effect of the expression heterologous genes in E. coli (Makoff et al. 1989). Molecular chaperones are very well known for their assistance in protein folding under normal and various stress conditions like heat or temperature stress (Hartl 1996) as well as they are also capable of providing correctly folded, biologically active proteins that were found to be difficult to be produced in E. coli. It is not necessary to use target protein as well as chaperones from the same organisms. In a two-vector system in *E. coli*, soluble gp37 protein was reported to be effectively produced by co-expression of two bacteriophage T4 chaperones (Bartual et al. 2010).

12.3.1.5 Codon Optimization

The presence of biased codons or codons requiring the rare tRNAs for their expression results in alteration of the proficiency of expression of heterologous protein in *E. coli*. Expression of such genes (without codon optimization) in *E. coli* displays a nonrandom utilization of identical codons which further affects the expression of host genes, or it may lead to an adversely rigorous response (Burgess-Brown et al. 2008). For example, AGA and AGG the arginine codons are mainly infrequent in *E. coli*, and they lead to lower protein expression and mistranslational errors (Calderone et al. 1996). In such cases, there is a need for codon optimization strategies to improve the fidelity of transformation as well as expression of enzymes (Hutterer et al. 2012). The issue can be overcome by conversion of rare codons to commonly used codons by site-directed mutagenesis. Another approach for rescue involves the co-expression of the genes encoding rare tRNA along with the protein of interest (Kleber-Janke and Becker 2000). Kim and Lee (2006) and Gustafsson et al. (2004) have reported use of synthetic DNA with optimized and commonly used codons for successful expression of desired enzymes.

12.3.2 Other Bacteria as a Host

The exploration of microbes, their physiology, and metabolism is attracting other bacterial hosts as microbial chassis. The diversity and biosynthetic potential which is the result of the adaptation due to exposure to varied environments makes other bacteria as a useful host (Ferrer-Miralles and Villaverde 2013b). Besides E. coli, *Bacillus* species are also regularly used as a host for production of recombinant protein due to their high capacity of secretion of protein and ability to export proteins directly into extracellular media. Different *Bacilli* systems that have been explored for protein production include *B. licheniformis,* B. subtilis*, B. megaterium*, B. brevis, and B. amyloliquefaciens. Another group of potential and promising microbes used for recombinant protein expression is lactic acid bacteria. They are safer expression hosts since they do not produce endotoxins like *Bacillus* spp. Besides these, proteobacteria and actinobacteria have also been used as host for protein expression. Table 12.1 highlights significant bacterial groups used as cell factories for recombinant protein production (Ferrer-Miralles and Villaverde 2013b).

12.3.3 Yeasts as a Host

Yeasts are also exceptional microbial hosts used for recombinant proteins expression since they have a dual advantage of being unicellular, fast growth and easy genetic manipulation like prokaryotes as well as secretory pathway, protein processing, and ability of posttranslational modifications as in eukaryotes. Hence they are of significant interest as microbial chassis especially when the source of protein to be produced is of eukaryotic origin. Yeasts hosts mainly used as expression systems

Table 12.1 Significant bacterial groups used as cell factories for recombinant protein production

S. no	Group	Advantage	Host	Protein product
1	Caulobacteria	Easy purification of secreted RSaA fusions	*Caulobacter crescentus*	Hematopoietic necrosis virus capsid proteins
2	Phototrophic bacteria	High production of membrane proteins	*Rodhobacter sphaeroides*	Membrane proteins
3	Cold adapted bacteria	Improved protein folding	*Pseudoalteromonas haloplanktis*	3H6 Fab, human nerve growth factor
			Shewanella sp. *strain Ac10*	β-Lactamase, peptidases, glucosidase
4	Pseudomonads	Efficient secretion	*Pseudomonas fluorescens*	Human granulocyte colony-stimulating factor
			Pseudomonas putida	Single-chain Fv fragments
			Pseudomonas aeruginosa	Penicillin G acylase
5	Halophilic bacteria	Solubility favored	*Halomonas elongate*	β-Lactamase
6	Streptomycetes	Efficient secretion	*Streptomyces lividans*	M. tuberculosis antigens
			Streptomyces griseus	Trypsin
7	Nocardia	Efficient secretion	*Nocardia lactamdurans*	Lysine-6-aminotransferase
8	Mycobacteria	Posttranslational modifications	*Mycobacterium smegmatis*	Hsp65-hIL-2 fusion protein
9	Coryneform bacteria	High-level production and secretion; GRAS	*Corynebacterium glutamicum*	Protein-glutaminase
			Corynebacterium ammoniagenes	Pro-transglutaminase
			Brevibacterium lactofermentum	Cellulases
10	Bacilli	High-level production and secretion	*Bacillus subtilis*	β-Galactosidase
			Bacillus brevis	Disulfide isomerase
			Bacillus megaterium	Antibodies
			Bacillus licheniformis	Subtilisin
			Bacillus amyloliquefaciens	Amylases

(continued)

Table 12.1 (continued)

S. no	Group	Advantage	Host	Protein product
11	Lactic acid bacteria	Secretion; GRAS	*Lactococcus lactis*	Fibronectin-binding protein A, internalin A, GroEL
			Lactobacillus plantarum	β-Galactosidase
			Lactobacillus casei	VP2-VP3 fusion protein of infectious pancreatic necrosis virus
			Lactobacillus reuteri	Pediocin PA-1
			Lactobacillus gasseri	CC chemokines

include *Pichia pastoris, Arxula adeninivorans,* Saccharomyces cerevisiae, *Schizosaccharomyces pombe, Hansenula polymorpha*, and *Yarrowia lipolytica.*

Details of the different yeasts that have been used as microbial host for recombinant proteins production, their advantages, and shortcomings are mentioned in Table 12.2. Besides use of yeast vectors, integration of gene encoding the protein product of interest into yeast genome has also been done for more expression and added stability of the insert.

12.3.4 Filamentous Fungi as a Host

Besides bacteria and yeasts, the research efforts for the production of industrially important proteins have also focused on filamentous fungi as a host. For production from these fungi, their already known best regulatory, expression, and secretory machinery is employed in conjunction with heterologous gene producing recombinant protein for commercial production. Table 12.3 highlights the fungal strains employed for recombinant protein production. The use of fungi as industrial hosts for protein production is still under exploration phase since fungi encompasses a very large and diverse group of microbes, and many of them are yet to be well characterized for their biosynthetic potential. Filamentous fungi also have disadvantages like their ability to produce and secrete homologous proteases, which might degrade the recombinant product. Also, the process of glycosylation of proteins is different in mammalian and fungal cells. This can also affect the final product synthesized using fungal machinery (Ward 2012).

Table 12.2 Yeast strains used as microbial host for recombinant proteins production, their advantages, and shortcomings

S. no.	Name	Advantages	Products	Shortcomings	Strains used
1	*Saccharomyces cerevisiae*	Well-known expression system, generally recognized as safe, tolerance to wide pH range, resistance to elevated osmotic pressures, intra- and extracellular expression of proteins	Hepatitis B surface antigen, hirudin, insulin, glucagon, urate oxidase, macrophage colony-stimulating factor, and platelet-derived growth factor	The hyperglycosylation of proteins, low protein yield, and plasmid instability (26)	BY4716, W303, S288c, A634A, CEN.PK, BJ5464, ∑1278b,SK1, BY4741, BY4742, BY4743
2	Yarrowia lipolytica	Produce high levels of heterologous proteins and metabolites, grow on n-paraffin, utilize acetate, alcohols, and hydrophobic substrates including oils, alkanes, and fatty acids	Single cell protein (SCP), citric acid, single cell oil (SCO)	Can use only limited range of C6 sugars such as glucose, fructose, and mannose	W29, E150, E129, YB423, CX161-1B
3	Pichia pastoris	Powerful methanol-regulated alcohol oxidase promoter (PAOX1), highly efficient secretion mechanism, posttranslational modification capabilities, and high cell density growing on defined medium	Human erythropoietin, phospholipase C, phytase, human superoxide dismutase, trypsin, human serum albumin, collagen, and human monoclonal antibody 3H6 Fab fragment	The high concentration of proteases, difficulties in the systematic study due to the product-specific effects and risks related to storage of large amounts of methanol	Y-11430, X-33, Mut+ (AOX1+, AOX2+), Muts (AOX1−, AOX2+), Mut− (AOX1−, AOX2) SMD1163 strain (his4pep4prb1), SMD1165 strain (his4prb1), SMD1168 strain (his4pep4), GS115 (his4), KM71 (Δaox1::SARG4 his4 arg4), SMD1168 (His4, pep4)
4	Hansenula polymorpha	Use methanol as the only carbon and energy source, multicopy integration system and the powerful inducible promoters, thermotolerant (48–50 °C), resistant to heat, oxidative stress, and heavy metals	Hepatitis B vaccine, insulin, interferon alpha-2a, hirudin, phytase, hexose oxidase, and lipase.	Accessibility of the multicopy integration system, powerful inducible promoters, thermotolerant (48–50 °C), resistanceto heat, oxidative stress and heavy metals, showed reduced hypermannosylation	CBS4732 (CCY38–22-2; ATCC34438, NRRL-Y-5445), DL-1 (NRRL-Y-7560; ATCC26012) and NCYC495 (CBS1976; ATAA14754, NRLLY-1798)

Table 12.3 Fungal strains used as microbial host for recombinant proteins production, their advantages, and shortcomings

S. no.	Genus	Proposed hosts	Advantages	Shortcomings	Some potential products
1	Aspergillus	A. niger, A. nidulans, A. oryzae, A. awamori, A. terreus	Generally recognized as safe (GRAS) list of the Food and Drug Administration (FDA), genetic characterization	Produces aflatoxins (A. niger)	Alpha1-proteinase inhibitor, antigen-binding (Fab') fragment, corticosteroid binding globulin, epithelial growth factor, granulocyte macrophage colony-stimulating factor, growth hormone, humanized IgG1 (kappa) antibodies, interferon-alpha-2, interleukin-6, lactoferrin, lysozyme, mucus proteinase inhibitor, parathyroid hormone, single-chain variable region fragment (scFv) anti-lysozyme construct, superoxide dismutase, and tissue plasminogen activator.
2	Trichoderma	T. reesei, T. altoviride, T. vireus	Genome sequence available, extraordinary ability to secrete proteins	–	Cellulase, hemicellulose, endoglucanase I, endoglucanase II and cellobiohydrolase I, cinnanoyl esterase from an unsuitable host, the anaerobic fungus Piromyces equi
3	Penicillium	P. purpurogenum, P. funiculosum, P. emersonii, P. chrysogenum, P. Camemberti, P. roqueforti	Efficient promoters	Limited genomic characterization	Arabinofuranosidases, beta-glucosidase, xylanases, acetyl esterases and ferulyl esterases, pectin lyase
4	Rhizopus	Rhizopus oryzae, Rhizomucor miehei, and Rhizomucor pusillus	Characterized annotated expanded gene families of secreted proteases, source of important extracellular industrial and medical enzymes	Pathogenic	Aspartic proteases and subtilases
5	Neurospora	N. crassa	Biochemically and genetically characterized, grown in simple media with high growth rates producing high amounts of proteins	Not yet recognized as an important industrial host	Influenza hemagglutinin (HA) and neuraminidase antigens (NA)

12.4 Metabolic Engineering Strategies

12.4.1 Directed Evolution

The use of enzymes on an industrial scale many times needs properties like thermo-stability and resistance to osmotic pressure under actual process conditions that are not found in naturally occurring enzymes. Directed evolution can be defined as the tuning of the natural enzyme in a lab by a process similar to natural evolution by random mutagenesis and recombination followed by efficient screening and selection of the mutants so obtained for desired activity.

Random mutagenesis for evolution makes use of the DNA libraries generated by techniques like error-prone PCR, combinatorial oligonucleotide mutagenesis, DNA shuffling, and staggered extension process (StEP recombination) for strain development followed by screening to obtain a better host with enhanced efficacy. Recombination target modifications make use of nonhomologous recombination, exon shuffling, or alternative splicing to develop and select folded proteins from the obtained secondary structure elements (Urvoas et al. 2012). Although the method is feasible and proficient, a limitation of this method many times involves a compromise between the targeted gene and other essential properties of host required for proper growth and survival.

12.4.2 Site-Directed Mutagenesis (SDM)

It is a highly versatile and precise molecular biology technique to confer tailored mutation in double-stranded DNA. It is also a widely used technique for protein engineering. The approach can be used in traditional cloning, in mapping and control of regulatory as well as in functional analysis of proteins. It also facilitates genome editing in a defined manner through homology-directed repair. Both single-site-directed mutagenesis (point mutation, insertion, deletion, multiple nearby substitutions) as well as multiple site mutations (100 bp apart) have been reported for modification and regulation of the desired gene/s for protein expression (Hsieh and Vaisvila 2013). Irfan et al. 2018 have reported to improve the thermostability of xylanase from *Geobacillus thermodenitrificans* by using SDM. Thermostability of bacterial chitinase was also shown to improve by 15% by using site-directed mutagenesis to alter the enzyme structure (Emruzi et al. 2018).

12.4.3 Site Saturation Mutagenesis (SSM)

SSM is used for substitution of targeted residue (of a protein) with some other naturally occurring amino acid to improve the expression and efficacy the desired significant proteins. The protocol makes use of single amino acid substitutions achieved by using different sets of primers having degenerate mixture of the four nucleotide bases at the three positions of the codon at the site linked to the

functionality of the protein. The amplified mutated PCR products transformed to competent cells and the activities of all substitutions are checked to determine the substitution having the maximum efficiency (Steffens and Williams 2007). When multisites are iteratively subjected for enzyme optimization by saturation mutagenesis, the procedure is called iterative saturation mutagenesis. SSM has led to the evolution of enzymes with stability, enhanced activity, manipulation of binding properties of antibodies and transcription factors (TFs) as well as stereoselectivity. Besides modifying the gene encoding functional protein, SSM has some applications in the engineering of promoters, transcriptional enhancers, RBS, *trans*-acting factors and *cis*-regulatory (Guazzaroni et al. 2015). ISM for evolution experiments has proved to be the most efficient technique even when compared to error-prone PCR (Yang et al. 2017).

12.4.4 Protein Truncation

Truncation is the random or directed deletion of the protein domains that are not necessary for its activity. It has been reported for enhancement of the desirable properties of the enzymes in some cases and may also lead to reduction of specific activity in others. Truncated versions of endo-dextranases from *Streptococcus mutans* were rendered resistant to degradation by proteases during long-term storage by truncation to remove domains not involved in catalytic activity (Kim et al. 2011). Amylases that are active at a high pH are in demand in textile and detergent industries. The stability and specific activity of the alkaline α-amylase of *Bacillus pseudofirmus* were found to improve to almost 35-fold due to N-terminal truncated mutant (Lu et al. 2016).

12.4.5 Fusion to Generate Chimeric Enzymes

Chimeric proteins are proteins prepared by fusion of the structural genes of the two different proteins/polypeptides having different functions or physico-chemical properties to produce a single protein with higher efficiency. Fusion proteins may be a product of two end-to-end fused sub-units linked by a linker or a product of gene where the amino acids from both sources are interspersed with each other. The product of the end fusion genes usually shows the activity of both the parent genes, whereas the product of the later type often shows a novel activity (Irfan et al. 2018). The fused molecules range from short, synthetic oligonucleotides to full-length structural genes. The increasing amount of publicly available sequenced gene databases provides endless number of fusion partners, thus making this technique a valuable and versatile tool for expression of desired proteins.

12.5   Conclusions

Proteins are present and are synthesized by all living forms. They are a part of cyto-skeleton as well as help body functions to be carried out smoothly by serving as biocatalysts. Use of microbes as cell factories for production of proteins and enzymes is increasing due to their higher efficiency as compared to chemical catalysts and they are ecofriendly and renewable. They usually have simple nutritional requirements and are easy for handling and manipulation as well. Microbes as chassis for production make use of bacteria, fungi, and yeasts. The use of advanced metabolic engineering strategies has led to assimilation of various microorganisms that can be used in microbial cell factories. Metabolic orthogonality is a main objective that is to be achieved in metabolic engineering for the production of different products by microbes. It is preferred because microbes when engineered have to compromise on their natural pathways. The availability of resources in the form of genetic exploration of microbes will help in achieving this goal, and even researchers will have to input less efforts for bioproduction. Amongst bacteria, a major area is still occupied by E. coli. *Bacillus* spp. are employed as microbial chassis due to their secretory properties that help increase the expression level of proteins. Lactic acid bacteria are still under the process of exploration as a host. Yeasts and fungi are preferred hosts due to posttranslational modifications as well as their high secretion capacity. The choice of best microbial skeleton for industrial protein production is a diverse area as still under exploration phase as the information on genetic and metabolic resources is continuously being added. The addition of the more knowledge in terms of databases will also help for selection of metabolic engineering strategy from the above-discussed methods for the production of desired proteins.

References

Apte AA, Senger RS, Fong SS (2014) Designing novel cellulase systems through agent-based modeling and global sensitivity analysis. Bioengineered 5:1–11

Bartual SG, Garcia-Doval C, Alonso J, Schoehn G, van Raaij MJ (2010) Two-chaperone assisted soluble expression and purification of the bacteriophage T4 long tail fibre protein gp37. Protein Expr Purif 70:116–121

Bernal V, Castaño-Cerezo S, Gallego-Jara J, Écija-Conesa A, de Diego T et al (2014) Regulation of bacterial physiology by lysine acetylation of proteins. New Biotechnol 31:586–595

Brinkworth AJ, Malcolm DS, Pedrosa AT, Roguska K, Shahbazian S, Graham JE et al (2011) *Chlamydia trachomatis* Slc1 is a type III secretion chaperone that enhances the translocation of its invasion effector substrate TARP. Mol Microbiol 82:131–144

Buchel F, Rodriguez N, Swainston N, Wrzodek C, Czauderna T et al (2013) Path 2 models: large-scale generation of computational models from biochemical pathway maps. BMC Syst Biol 7:116

Burgess-Brown NA, Sharma S, Sobott F, Loenarz C, Oppermann U, Gileadi O (2008) Codon optimization can improve expression of human genes in *Escherichia coli:* a multi-gene study. Protein Expr Purif 59:94–102

Calderone TL, Stevens RD, Oas TG (1996) High-level misincorporation of lysine for arginine at AGA codons in a fusion protein expressed in *Escherichia coli*. J Mol Biol 262:407–412

Chao YP, Wen CS, Wang JY (2004) A facile and efficient method to achieve LacZ overproduction by the expression vector carrying the thermoregulated promoter and plasmid copy number. Biotechnol Prog 20:420–425

Emruzi Z, Aminzadeh S, Karkhane AA, Alikhajeh J, Haghbeen K, Gholami D (2018) Improving the thermostability of *Serratia marcescens* B4A chitinase via G191V site-directed mutagenesis. Int J Biol Macromol 116:64–70

Ferrer-Miralles N, Villaverde A (2013a) Bacterial cell factories for recombinant protein production; expanding the catalogue. Microb Cell Factories 12:113

Ferrer-Miralles N, Villaverde A (2013b) Bacterial cell factories for recombinant protein production; expanding the catalogue. Microb Cell Fact 12:113

Fisher AK, Freedman BG, Bevan DR, Senger RS (2014) A review of metabolic and enzymatic engineering strategies for designing and optimizing performance of microbial cell factories. Comput Struct Biotechnol J 11(18):91–99

Fong SS, Palsson BO (2004) Metabolic gene-deletion strains of Escherichia coli evolve to computationally predicted growth phenotypes. Nat Genet 36:1056–1058

Guazzaroni ME, Silva-Rocha R, Ward RJ (2015) Synthetic biology approaches to improve biocatalyst identification in metagenomic library screening. Microb Biotechnol 8(1):52–64

Gustafsson C, Govindarajan S, Minshull J (2004) Codon bias and heterologous protein expression. Trends Biotechnol 22:346–353

Hartl FU (1996) Molecular chaperones in cellular protein folding. Nature 381:571–579

Hartl FU, Bracher A, Hayer-Hartl M (2011) Molecular chaperones in protein folding and proteostasis. Nature 475:324–332

Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B et al (2010) High-throughput generation, optimization and analysis of genome-scale metabolic models. Nat Biotechnol 28:977–982

Hsieh PC, Vaisvila R (2013) Protein engineering: single or multiple site-directed mutagenesis. Methods Mol Biol 978:173–186

Hutterer KM, Zhang Z, Michaels ML, Belouski E, Hong RW, Shah B et al (2012) Targeted codon optimization improves translational fidelity for an Fc fusion protein. Biotechnol Bioeng 109:2770–2777

Irfan M, Gonzalez CF, Raza S, Rafiq M, Hasan F, Khan S, Shah AA (2018) Improvement in thermostability of xylanase from *Geobacillus thermodenitrificans* C5 by site directed mutagenesis. Enzym Microb Technol 111:38–47

Ismail NF, Hamdan S, Mahadi NM, Murad AMA, Rabu A, Bakar FDA et al (2011) A mutant L-asparaginase II signal peptide improves the secretion of recombinant cyclodextrin glucanotransferase and the viability of *Escherichia coli*. Biotechnol Lett 33:999–1005

Kim S, Lee SB (2006) Rare codon clusters at 5′-end influence heterologous expression of archaeal gene in *Escherichia coli*. Protein Expr Purif 50:49–57

Kim YM, Shimizu R, Nakai H, Mori H, Okuyama M, Kang MS et al (2011) Truncation of N- and C-terminal regions of *Streptococcus mutans* dextranase enhances catalytic activity. Appl Microbiol Biotechnol 91:329–339

Kleber-Janke T, Becker WM (2000) Use of modified BL21 (DE3) *Escherichia coli* cells for high-level expression of recombinant peanut allergens affected by poor codon usage. Protein Expr Purif 19:419–424

Liu S-L, Du K, Chen W-Z, Liu G, Xing M (2012) Effective approach to greatly enhancing selective secretion and expression of three cytoplasmic enzymes in *Escherichia coli* through synergistic effect of EDTA and lysozyme. J Ind Microbiol Biotechnol 39:1301–1307

Lu Z, Wang Q, Jiang S, Zhang G, Ma Y (2016) Truncation of the unique N-terminal domain improved the thermos-stability and specific activity of alkaline α-amylase Amy703. Sci Rep 6:22465

Makoff A, Parry N, Dicken L (1989) Translational fusions with fragments of the trpE gene improve the expression of a poorly expressed heterologous gene in *Escherichia coli*. J Gen Microbiol 135:11–24

Mattanovich D, Branduardi P, Dato L, Gasser B, Sauer M et al (2012) Recombinant protein production in yeasts. Methods Mol Biol 824:329–358

Mendez-Perez D, Gunasekaran S, Orler VJ, Pfleger BF (2012) A translation-coupling DNA cassette for monitoring protein translation in *Escherichia coli*. Metab Eng 14:298–305

Newbury SF, Smith NH, Robinson EC, Hiles ID, Higgins CF (1987) Stabilization of translationally active mRNA by prokaryotic REP sequences. Cell 48:297–310

Okesli A, Cooper LE, Fogle EJ, van der Donk WA (2011) Nine post-translational modifications during the biosynthesis of cinnamycin. J Am Chem Soc 133:13753–13760

Olins PO, Rangwala S (1989) A novel sequence element derived from bacteriophage T7 mRNA acts as an enhancer of translation of the lacZ gene in *Escherichia coli*. J Biol Chem 264:16973–16976

Osterlehner A, Simmeth S, Göpfert U (2011) Promoter methylation and transgene copy numbers predict unstable protein production in recombinant Chinese hamster ovary cell lines. Biotechnol Bioeng 108:2670–2681

Pandhal J, Woodruff LBA, Jaffe S, Desai P, Ow SY et al (2013) Inverse metabolic engineering to improve *Escherichia coli* as an N-glycosylation host. Biotechnol Bioeng 110:2482–2493

Pérez-Pérez J, Márquez G, Barbero JL, Gutiérrez J (1994) Increasing the efficiency of protein export in *Escherichia coli*. Nat Biotechnol 12:178–180

Ringquist S, Shinedling S, Barrick D, Green L, Binkley J, Stormo GD et al (1992) Translation initiation in *Escherichia coli*: sequences within the ribosome-binding site. Mol Microbiol 6:1219–1229

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

Sprengart ML, Fuchs E, Porter A (1996) The downstream box: an efficient and independent translation initiation signal in *Escherichia coli*. EMBO J 15:665

Steffens DL, Williams JG (2007) Efficient site-directed saturation mutagenesis using degenerate oligonucleotides. J Biomol Tech 18(3):147

Tokunaga H, Arakawa T, Tokunaga M (2010) Novel soluble expression technologies derived from unique properties of halophilic proteins. Appl Microbiol Biotechnol 88:1223–1231

Urvoas A, Valerio-Lepiniec M, Minard P (2012) Artificial proteins from combinatorial approaches. Trends Biotechnol 30(10):512–520

Uthandi S, Prunetti L, De Vera IMS, Fanucci GE, Angerhofer A, Maupin-Furlow JA (2012) Enhanced archaeal laccase production in recombinant Escherichia coli by modification of N-terminal propeptide and twin arginine translocation motifs. J Ind Microbiol Biotechnol 39:1523–1532

Wang Y, San KY, Bennett GN (2013) Cofactor engineering for advancing chemical biotechnology. Curr Opin Biotechnol 24:994–999

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

Yang H, Li J, Du G, Liu L (2017) Microbial production and molecular engineering of industrial enzymes: challenges and strategies. In: Biotechnology of microbial enzymes. Academic, Cambridge, pp 151–165

Zhan J, Stayton P, Press OW (1998) Modification of ricin A chain, by addition of endoplasmic reticulum (KDEL) or Golgi (YQRL) retention sequences, enhances its cytotoxicity and translocation. Cancer Immunol Immunother 46:55–60

Zhang YB, Greenberg B, Lacks SA (1997) Analysis of a *Streptococcus pneumoniae* gene encoding signal peptidase I and overproduction of the enzyme. Gene 194:249–255

Zhang F, Carothers JM, Keasling JD (2012) Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. Nat Biotechnol 30:354–359