



Review

Engineering Biology to Construct Microbial Chassis for the Production of Difficult-to-Express Proteins

Kangsan Kim ^{1,2}, Donghui Choe ^{1,2} , Dae-Hee Lee ³ and Byung-Kwan Cho ^{1,2,4,*}

¹ Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon 34141, Korea; kskim2474@kaist.ac.kr (K.K.); robinald@kaist.ac.kr (D.C.)

² KAIST Institute for the BioCentury, Korea Advanced Institute of Science and Technology, Daejeon 34141, Korea

³ Synthetic Biology & Bioengineering Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 34141, Korea; dhlee@kribb.re.kr

⁴ Intelligent Synthetic Biology Center, Daejeon 34141, Korea

* Correspondence: bcho@kaist.ac.kr

Received: 6 December 2019; Accepted: 31 January 2020; Published: 2 February 2020



Abstract: A large proportion of the recombinant proteins manufactured today rely on microbe-based expression systems owing to their relatively simple and cost-effective production schemes. However, several issues in microbial protein expression, including formation of insoluble aggregates, low protein yield, and cell death are still highly recursive and tricky to optimize. These obstacles are usually rooted in the metabolic capacity of the expression host, limitation of cellular translational machineries, or genetic instability. To this end, several microbial strains having precisely designed genomes have been suggested as a way around the recurrent problems in recombinant protein expression. Already, a growing number of prokaryotic chassis strains have been genome-streamlined to attain superior cellular fitness, recombinant protein yield, and stability of the exogenous expression pathways. In this review, we outline challenges associated with heterologous protein expression, some examples of microbial chassis engineered for the production of recombinant proteins, and emerging tools to optimize the expression of heterologous proteins. In particular, we discuss the synthetic biology approaches to design and build and test genome-reduced microbial chassis that carry desirable characteristics for heterologous protein expression.

Keywords: heterologous protein expression; difficult-to-express proteins; synthetic biology; systems biology; genome reduction; genome synthesis.

1. Introduction

Heterologous protein expression by means of the genetically engineered prokaryotic host has made available a wide spectrum of recombinant proteins that are otherwise confined to limited natural origin, in a scalable and cost-effective manner. Marked with the expression of an active human protein somatostatin in *Escherichia coli* back in 1977 [1] followed by monumental success in the expression of human insulin [2] and its marketization [3], much focus has been shed on microbe-based protein expression system for its versatile nature and potential for large-scale production. Demands for commercial proteins have also become increasingly diverse—characterized by extensive variations in biochemical and structural properties. These include, but are not limited to, therapeutics for clinical treatment [4,5], antibodies for diagnostics [6], and enzymes for industrial use [7]. Such divergence in biochemical, structural, and functional aspects of recombinant proteins means that there are a vast number of factors to consider before achieving functional expression of the recombinant proteins in bacteria.

While bacterial systems are capable of expressing a wide spectrum of heterologous proteins in their functional forms [4], there remain imminent challenges and limitations in utilizing this system to its fullest. Every so often, heterologous expression of proteins alien to the host system poses significant difficulties and requires extensive optimization steps to tame. For instance, problems highly recursive in heterologous protein expression include improper folding of target proteins, especially those of higher eukaryotes that render the protein to lose its native function. This is largely attributed to host factors such as differences in cytoplasmic redox potential that interfere with disulfide bond formation [8], differences in codon usage, and repetitive DNA sequences that affect protein translation and subsequent protein folding [9]. Furthermore, the functional expression is largely implicated with the size and characteristics of the heterologous protein. It was reported that proteins with large molecular weight or those that harbor several membrane domains pose a far greater tendency to form insoluble aggregates and are prone to proteolysis [10]. Moreover, additional challenges include mimicking eukaryotic post-translational modifications (such as glycosylation) [11], production of harmful endotoxins (*E. coli*) [12], reduced cell viability resulting from unwanted by-products [13], and low protein yields [14]. To tackle these obstacles, different recombinant DNA technologies, including manipulation of gene expression control [15–17], directing proper bond formation and protein folding [18,19], interfering host metabolic pathways [20,21], random or directed evolutions of bacterial strains or enzymes [22], and series of other methods have been systematically employed (summarized in [23]). While these engineering strategies already provide practical solutions for the expression of various protein forms, the engineered expression systems can further be optimized to achieve a better performance.

Functional protein expression has met a dramatic shift paralleled with the advances in systems and synthetic biology; that is, data-driven understanding in systems biology has provisioned a knowledge-based framework on which to rationally design, build, and test biological entities in an increasingly sophisticated manner [24]. For instance, the following have altogether enabled characterization of the bacterial host and recombinant proteins at the systems level: a large influx of high-throughput information on DNA, RNA, proteins, and metabolites that enabled visualization of the global landscape of transcriptional and translational processes [25,26]; availability of robust computational tools and in silico models streamlining detailed analysis and prediction of the cellular networks; and rapid advances in DNA synthesis, assembly techniques that facilitated the process to build and test the designed biological systems [27,28]. Such new schemes provide one possible way around biological uncertainties; that is, by systematically eliminating genetic redundancies, it is expected to minimize the confounding effects of unknown genetic elements and to reduce metabolic interference from native metabolic pathways while endorsing more predictable regulation of cellular functions [24]. This, in essence, would help to further improve the performance of the extant microbial cell factories. In addition, the design and synthesis of minimal genomes can be paralleled with re-engineering efforts to repurpose a cell to execute certain functions, including genome recoding to expedite production of unnatural polypeptides [29], and refactoring genes to make transcriptional and translational regulation more controllable [30]. Hosts of different microbial synthetic minimal genomes have been published to date (reviewed in [31]), with many harboring genetic and phenotypic traits to better house and support biosynthesis of heterologous proteins and commodity biomolecules.

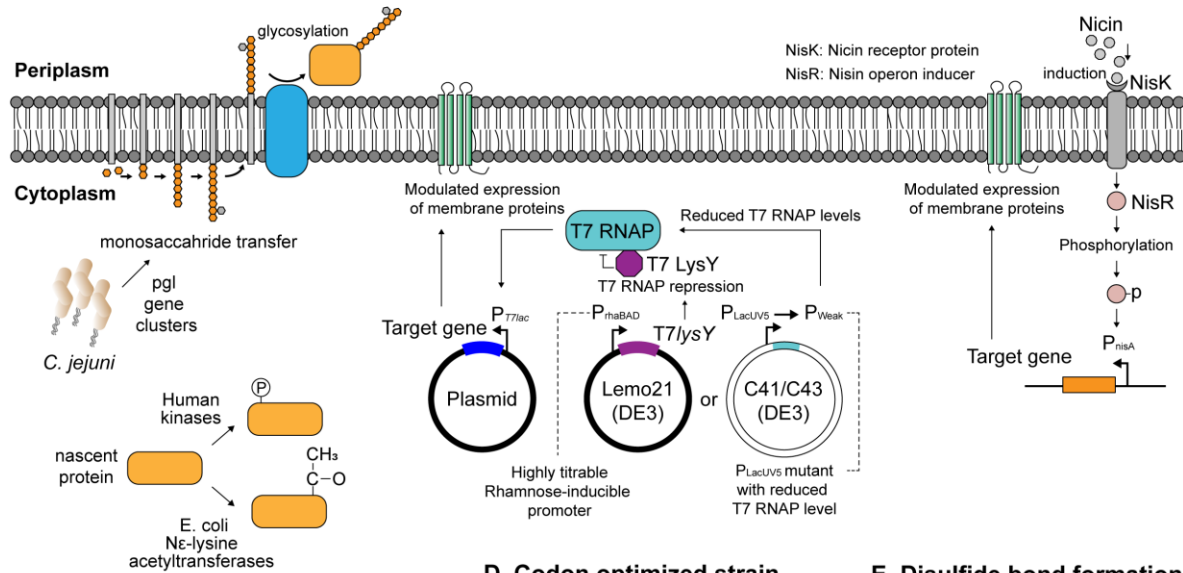
2. Heterologous Expression of Biologically Functional Proteins by Bacteria

Production of recombinant proteins into a biologically functional form necessitates robust expression systems that can reliably execute proper folding, precise post-translational modifications, and efficient translocations while maintaining cellular viability (Figure 1). In recent years, prokaryotic expression systems have been an effective means to deliver functional proteins at high yields. Genomic, genetic, and biochemical aspects of the microbial systems have been extensively explored, culminating in series of genetically engineered microbial chassis with features desirable for heterologous protein expression. Notwithstanding these efforts, no single expression system can support all existing

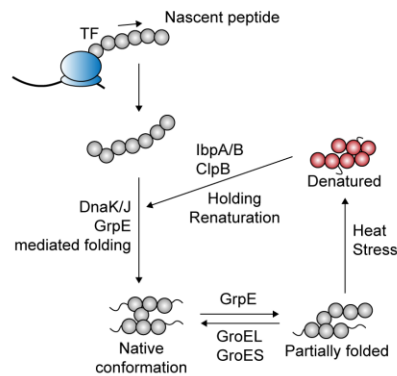
recombinant proteins, and this is especially the case for difficult-to-express proteins, which often require highly specialized strain engineering.

A. Post-translational modifications

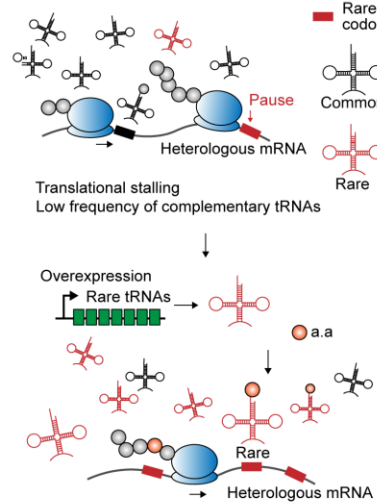
B. Membrane Protein Expression



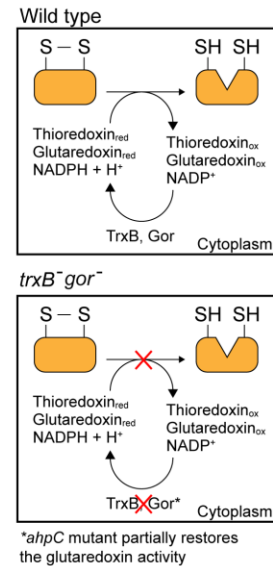
C. Chaperone overexpression



D. Codon optimized strain



E. Disulfide bond formation



F. Elimination of extracellular protease activity

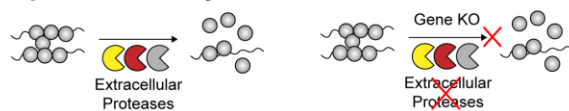


Figure 1. Notable characteristics of microbial chassis. (A) Post-translational modification systems involving glycosylation, phosphorylation, and acetylation; (B) various protein expression systems catered for the expression of membrane proteins; (C) simplified mechanism of the peptide folding mediated by the molecular chaperones; (D) expression system optimized for the expression of the heterologous proteins with high rare codon frequencies; (E) genetically engineered *E. coli* strain (*trxB⁻*, *gor⁻*) enabling the formation of disulfide bridges within the cytoplasm; (F) increasing recombinant protein yield by genetically deleting extracellular protease genes.

2.1. Engineered *E. coli* for Wide Array of Recombinant Proteins

Over 50% of the recombinant proteins registered to date (Protein Data Bank [32]) are of eukaryotic origin, with more than 90% of them being produced in prokaryotic expression systems [32,33]. Among all recombinant expression hosts at work, *E. coli* is by far the most favored strain, owing to its

outstanding genetic tractability, relative ease of cultivation, and innate capacity to accommodate and express exogenous proteins. Specifically, it is the model microbe with the most extensively characterized genome, transcriptome and translome architectures, and underlying regulations [25,34]. In addition, the availability of a vast repertoire of synthetic biology tools such as libraries of promoters, ribosome-binding sites (RBS) and 5'-untranslated regions (5'-UTRs), expression vectors, and synthetic circuits effectively streamline its genetic manipulation [35–39]. In aspects of protein productivity, *E. coli* was reported to dedicate nearly 40% of its dry cell weight entirely for recombinant protein in fed-batch culture conditions [40], and are able to functionally express a wide spectrum of non-glycosylated proteins. Such inherent features of *E. coli* have much been exploited in the industry for mass-production of many commodity proteins [41].

Even so, securing high-purity, high-yield recombinant proteins, especially those of eukaryotic origin, has remained challenging in *E. coli* due to constraints inherent to cellular physiology and translational regulations. For instance, just as many prokaryotic expression systems, *E. coli* lacks post-translational modification machineries required for functional expression of proteins in eukaryotic origins, such as glycosylation [11]. This could present a serious drawback in prokaryote-based expression systems, considering that more than 50% of eukaryotic proteins are predicted to be glycosylated [42]. Other classes of post-translational modifications highly prevalent across eukaryotic proteins include phosphorylation and acetylation [43]. These modifications contribute as much, if not more, to proper protein folding [44], and endows proteins with important functionalities (Table 1).

Table 1. Examples of prokaryote-based heterologous expression systems.

Strain	Defining Features	Reference
<i>E. coli</i> (<i>pgl</i> cluster)	Expression of <i>N</i> -glycosylation pathway of <i>C. jejuni</i> origin, catalyzing glycosylation of recombinant proteins where appropriate.	[45]
<i>E. coli</i> (human JNK1)	Coexpression of human Jun N-terminal kinase 1 (JNK1) effectively catalyzes recombinant protein phosphorylation	[46]
<i>E. coli</i> (<i>rimJ</i>)	Use of the native acetylation machinery in <i>E. coli</i> to acetylate human proteins.	[47]
<i>E. coli</i> (yeast NatA NatB)	Coexpression of yeast-derived NatA NatB acetylation enzymes for amino-terminal acetylation	[48]
<i>E. coli</i> (N ϵ -acetyllysine)	Site-directed incorporation of N ϵ -acetyllysine using a three plasmid system expressing recoded target genes, suppressor tRNA, and evolutionarily engineered aminoacyl-tRNA synthetase.	[49]
<i>E. coli</i> (Chaperone overexpression)	Coordinated co-overexpression of <i>E. coli</i> native molecular chaperones—GroEL/ES, DnaK/J/GrpE, IbaA/B, and ClpB—significantly improves the solubility of the recombinant proteins.	[50]
<i>E. coli</i> Origami	Facilitates formation of disulfide bonds within the cytoplasmic compartment, through inactivation of thioredoxin and glutathione reductase pathways (Δ <i>trxB</i> , Δ <i>gor</i> , <i>aphC</i>)	[51]
<i>E. coli</i> CyDisCo	Introduction of eukaryotic thiol oxidase and disulfide isomerase encourages formation of disulfide bonds within the <i>E. coli</i> cytoplasm.	[52]
<i>E. coli</i> C41(DE3), C43(DE3)	BL21(DE3) derivative with mutations that confer increased tolerance to toxic membrane proteins.	[53]
<i>E. coli</i> Lemo21(DE3)	Harbors a gene expression system that allows fine-tuning of overexpression intensity. Suitable for membrane protein production.	[54]
<i>E. coli</i> Rosetta	Alleviates codon-bias by overexpression of tRNA species orthogonal to rare codons in <i>E. coli</i> —AUA, AGG, AGA, CGG, CUA, CCC, and GGA.	[55]

Table 1. Cont.

Strain	Defining Features	Reference
<i>B. subtilis</i> WB600	Strain that lacks six out of seven extracellular proteases to circumvent host-mediated proteolysis	[56]
<i>B. subtilis</i> (<i>dlt</i> ⁻)	Inactivation of D-anlanylation in <i>dlt</i> ⁻ <i>B. subtilis</i> increases the availability of folding factors Mg ²⁺ , Ca ²⁺ and Fe ³⁺ around the cell membrane microenvironment, alleviating protease-mediated recombinant protein degradation.	[57]
<i>L. lactis</i> NZ9000	Microbial expression system that effectively supports the production of various prokaryotic and eukaryotic membrane proteins.	[58]
<i>Haloferax volcanii</i>	Halophilic archeon that stably overexpresses seven transmembrane helix proteins such as bacterioiopsins. The transmembrane protein expression machinery can be exploited to express eukaryotic proteins with similar protein topology.	[59,60]
<i>Geobacillus kaustophilus</i>	Thermophilic bacteria with an array of heat-stable, sugar-inducible promoters demonstrated soluble expression of heterologous enzymes otherwise insoluble in mesophilic host. With maximal protein yield of 59 mg/L	[61]

For instance, the finding that a gram-negative bacterium *Campylobacter jejuni* harbors a functional N-linked glycosylation pathway encoded by *pgl* gene cluster, which can be reconstituted and expressed in *E. coli*, has widened the scope of prokaryotic-based heterologous protein expression to various glycoproteins (Figure 1A and Table 1) [45,62]. Subsequently, several technical limitations arising from the differences in bacterial and eukaryotic N-glycosylation systems, such as different glycan specificity [63], were circumvented with novel engineering attempts, hence further refining the applicability of the glycosylation-competent *E. coli* [63,64]. On the other hand, phosphorylation-competent prokaryotic expression system has been developed by introducing mammalian kinases into the host cell. Several human kinases coexpressed with recombinant protein in *E. coli* were shown to successfully phosphorylate the target proteins in their native patterns (Figure 1A and Table 1) [46,65]. Furthermore, a number of expression systems have been devised in *E. coli* to deliver proteins with intact acetylation patterns. For example, *E. coli* is known to harbor multiple endogenous N^ε-lysine acetyltransferases (KATs), which catalyzes acetylation of lysine side chains embedded within proteins [66]. It was demonstrated that the native acetylation machineries in *E. coli* can be exploited to effectively acetylate human thymosin α1, suggesting that production of acetylated proteins is likely achievable in recombinant microbes (Figure 1A and Table 1) [47]. Similar to the method used in protein phosphorylation, coexpression of eukaryotic acetylation enzymes, such as NatA or NatB complex from yeast, was shown to provide an effective means for amino-terminal acetylation (Table 1) [48]. Other, more complicated approaches involved site-directed incorporation of acetyllysine, an acetyl-derivative of lysine, on target recombinant protein using a three-plasmid expression system (Table 1) [49]. Each of the three plasmids expresses (i) a target gene with lysine codons replaced with amber codons; (ii) a suppressor tRNA that recognizes the amber codon; (iii) an aminoacyl-tRNA synthetase evolutionarily engineered to charge N^ε-acetyllysine to the suppressor tRNA. By suppressing the action of native deacetylase, the recombinant *E. coli* managed to synthesize manganese superoxide dismutase with site-specific acetylation.

Inside a cell, newly synthesized peptides are highly susceptible to be misfolded, either as the result of aberrant translational termination or due to abrupt changes in protein conformation in response to environmental stresses. The molecular chaperones play a crucial role in maintaining cellular proteomic integrity by guiding proper folding of nascent proteins into their native conformations and directing refolding of misfolded or aggregated proteins in the cytoplasmic milieu [19]. The chaperone systems most relevant for de novo protein folding include trigger factor (TF), GroEL system, DnaK system, GrpE, and ClpB, each with a distinct but partially overlapping mode of actions. For instance, de novo folding of polypeptides emerging from the ribosome is aided by TF, DnaK system, and subsequently with GroESL system to reach a native conformation (Figure 1C) [19]. In a synthetic biology perspective, this ubiquitous protein folding machinery is an important molecular asset to tackle the problem in heterologous expression of recombinant proteins—protein misfolding. In *E. coli* and other prokaryotic systems, accumulation of overexpressed recombinant proteins in the form of misfolded, insoluble aggregates is rather common [8]. The high incidence of misfolding among the overexpressed recombinant proteins may reflect the inherent upper limit of the host chaperone capacity. Efforts have been made to overexpress single and multiple combinations of chaperones native to the host, which often resulted in meaningful improvements in the net solubility of recombinant proteins [67,68]. Likewise, the co-overexpression of heat-shock chaperones and ClpB chaperone that serve to refold the insoluble aggregates also enhanced the solubility of numbers of recombinant proteins (Table 1) [50,68]. In addition, it was shown that the protein solubility can further be augmented by allowing a prolonged incubation time for in vivo protein–chaperone interaction, which increased the solubility of an array of heterologous proteins for up to 42-folds [50].

On the other hand, reducing conditions in the prokaryotic cytoplasm may inflict biochemical perturbations on nascent polypeptides, interfering with proper disulfide bond formation and rendering the afflicted proteins insoluble [8,69]. Disulfide linkage represents an abundant group of a structural scaffold in protein (with more than one-third of eukaryotic proteins predicted to harbor disulfide bond [70]) and is implicated in diverse aspects of protein functionalities. However, disulfide bond formation in the prokaryotic cytoplasm, such as that of *E. coli*, is disturbed by the presence of numerous reductases and other reductants [69]. Instead, native proteins that necessitate disulfide bonds are subjected to periplasmic secretion with N-terminal signal peptides directing their translocation. The oxidizing periplasmic space provides an amenable medium for oxidation of cysteine residues, and disulfide bond formation is further expedited by the presence of thiol–disulfide oxidoreductases enzymes, such as DsbABCD [70]. Hence, one approach to express disulfide-positive recombinant proteins in the microbial periplasm involves fusion with bacterial signal peptides, which is covered in detail elsewhere [71]. An alternative approach is to genetically engineer a bacterial host to reconfigure the cytoplasmic environment conducive for disulfide bond formation. For example, two pathways—glutaredoxin and thioredoxin pathways—are responsible for maintaining the reducing cytoplasm in *E. coli* (Figure 1E) [72]. Several collective findings have led to an *E. coli* expression system *E. coli* FA113, catered for disulfide bond formation (Table 1) [51,72,73] (widely known as Origami™). More recently, CyDisCo (cytoplasmic disulfide bond formation in *E. coli*) system revealed that disulfide bond formation can be reliably achieved in an otherwise wild-type *E. coli* cytoplasm (Table 1) [52]. CyDisCo strain harbors genes that encode mitochondrial thiol oxidase Ervp1 of yeast origin, and the human protein disulfide isomerase (PDI). In addition, this system can be brought up to a large-scale, fed-batch culture yielding considerable titer for human antibody, interleukins, and growth hormones [74].

Membrane proteins present yet another challenge for heterologous expression. Arising in complex multiple membrane-spanning domains, membrane proteins exhibit extensive variations in terms of structural, chemical, and functional representations. Problems recurrent in heterologous expression include (i) inclusion body formation and cell death resulting from saturation of membrane protein synthesis pathway and (ii) low-level expression [53]. Several expression systems more tolerant to membrane protein expression have been developed. For instance, *E. coli* C41(DE3) and C43(DE3) isolated from BL21(DE3) mutants are able to express membrane proteins more efficiently than the wild-type BL21(DE3) (Figure 1B and Table 1) [53]. Subsequent assessment of individual mutations in C41(DE3) revealed that increased tolerance toward the toxic expression is ascribed to the mutations in the P_{lacUV5} promoter which reduced T7 RNA polymerase levels upon isopropyl β -D-1-thiogalactopyranoside (IPTG) induction in C41(DE3) [75]. This is consistent with the fact that the overexpression of heterologous membranous proteins is especially toxic in the prokaryotic hosts and often suffers low protein yield. In this sense, the reduced expression strength of the P_{lacUV5} promoter in C41(DE3) strain may have helped reach a compromise between protein productivity and cellular fitness. The case of C41(DE3) sheds light on the important rule of thumb that tight control of gene expression is imperative to obtain membrane proteins while preventing undesirable outcomes such as cell death. In this respect, *E. coli* Lemo21(DE3) is a preferred host of choice for membrane protein expression which allows fine-tuning of overexpression intensity (Figure 1B and Table 1) [54]. Specifically, Lemo21(DE3) is isogenic to BL21(DE3), except that it harbors an expression vector that contains T7 lysozyme (T7 RNA polymerase inhibitor) tightly regulated by a rhamnose promoter which confers a highly titratable, broad range of expression control [76].

2.2. *Bacillus Subtilis* as A Versatile Host with Highly Efficient Protein Secretion Systems

Being an efficient producer of many desirable enzymes, *Bacillus* species have long been used as an efficient workhorse for the production of various commodity molecules, including industrial enzymes [77] and pharmaceuticals [78]. Excellent genetic tractability, robust growth in laboratory settings, and the vast availability of genetic tools and expression vectors comparable to those of *E. coli* made them highly pliable for genetic modifications and culture [79]. Notably, *Bacillus* species are known for their excellent protein secretion systems. As gram-positive bacteria with a single sheath of membrane separating between cell and outer environment, proteins can be readily secreted into the culture medium in biologically active forms via specialized secretion machineries. This confers a great advantage compared with *E. coli* expression system, where the overexpressed proteins tend to accumulate inside the cellular compartments, often in the form of insoluble aggregates. This defining trait in *Bacillus* considerably reduces the time and cost associated with downstream purification or extraction processes. As a side note, lack of apparent pathogenicity (viewed as Generally Recognized As Safe organism) gives an extra perk in its utilization, obviating the cumbersome and costly processes to remove toxic impurities such as lipopolysaccharide (LPS) or endotoxin in *E. coli* [80]. Importantly, several *Bacillus* strains, including *B. subtilis*, harbor large-scale fermentation capacity and are used for bulk manufacturing of industrially used proteins. These characteristics, along with several others, make *Bacillus* a robust production platform for heterologous proteins.

As it is said, the choice of *Bacillus* is almost exclusively ascribed to their highly efficient secretion system, which, given the right condition, yields as high as tens of grams per liter of proteins [81]. It is now understood that most of the protein exports are mediated via the general secretory (Sec) pathway, with the remaining few transported by a set of more specialized transport systems including twin-arginine translocation (TAT) pathway [82]. These existing modalities of protein export can be genetically engineered to facilitate secretion of heterologous proteins. For instance, the absence of compatible secretory proteins required for secretion of heterologous proteins in *B. subtilis* can be complemented by coexpressing a secretory protein SecB, along with genetic modifications in the host translocase component SecA to effectively bind SecB [83]. In another instance, engineering the SecA-dependent secretion pathway alone can double the productivity of recombinant proteins [84].

However, one drawback of such signal-peptide-mediated protein export is that the process requires complex interplay of multiple cellular components, which requires an extensive optimization for a large-scale protein secretion. Alternatively, several reports have shed light on nonclassical protein secretion systems that mediate translocation of proteins that lack signal peptides as a more straightforward means to direct protein secretion [85]. Specifically, all routes by which the cytoplasmic proteins without apparent secretion motifs travel out into the extracellular space are termed as *nonclassical secretion pathway* due to their largely elusive mechanism of transport [86]. Notably, in analogy to fusion tags, physically linking such nonclassically secreted proteins to heterologous expression target was shown to functionally mimic the role of signal peptides, enhancing the secretion rates of fusion proteins without compromising their biological functions [85]. This suggests that nonclassically secreted proteins may serve as a novel molecular tool for effective secretion of target proteins.

At the same time, *Bacillus* strains are also known for their profuse secretion of intracellular and extracellular proteases, with at least seven different extracellular proteases reported to date in *B. subtilis* [87]. While secretory proteins native to *Bacillus* exhibit resistance against the proteolytic activity of the proteases, most recombinant proteins are highly prone to proteolysis and are often rendered nonfunctional [88]. Protease-deficient *Bacillus* strains, such as *B. subtilis* derivative WB600, have thence been genetically engineered to circumvent host-mediated proteolysis (Figure 1F and Table 1) [56], though it only remained a partial success due to the persisting proteolytic activity [57]. Nonetheless, WB600 yielded higher productivity of recombinant proteins including β -lactamase [56], streptokinases [89], and several others. Subsequent removal of the remaining protease-encoding genes *vpr* and *wprA* in WB600 has led to an increased stability of the expressed protein molecules [90]. Several proteins still suffered from proteolytic degradation despite near-complete absence of the proteolytic activity in the protease-deficient *B. subtilis* strains. It was suggested that proteins in their prefolded state are highly susceptible to proteolytic activity, and finding a way to expedite folding of the nascent proteins would significantly improve their half-life post-secretion. One way to facilitate protein folding is by increasing the availability of metal cations such as Mg^{2+} , Ca^{2+} , and Fe^{3+} around the cell membrane microenvironment. These ions are also found in several protein species, including recombinant protective antigen (rPA), and serve as essential components to promote protein folding (Table 1) [57].

The development of novel genetic engineering tools and techniques are further refining the capacity and applicability of the existing *Bacillus* heterologous expression systems in recombinant protein production [91]. Several studies provide guidelines to utilize the existing techniques to their best efficiency and compatibility. One such report describes a combinatorial optimization scheme involving fine-tuning of promoter strength, translation, and folding efficiency in a protease-deficient *B. subtilis*, which altogether led to a ninefold increase in the production of human fibroblast growth factor 21, whose soluble expression is difficult due to protein aggregation [14]. Other recent technical aspects of recombinant protein expression in *Bacillus* have been reviewed elsewhere [92].

2.3. *Lactococcus Lactis* for the Expression of Recombinant Membrane Proteins

Having been widely used in dietary industries as a fermenting bacterium, *Lactococcus lactis* is a gram-positive, lactic-acid-producing bacterium with an emerging role in heterologous protein expression. It earned its recognition following the functional secretion of the heterologously expressed cytokine interleukin-10, which alleviated the symptoms of the inflammatory bowel disease (IBD) in the IBD model mouse in vivo [93]. Several characteristics of *L. lactis* expression system were deemed desirable for development and production of recombinant proteins and therapeutics; these include (i) single-membraned prokaryote with efficient protein secretion system; (ii) negligible extracellular proteolysis activity; (iii) generally recognized as safe (GRAS) strain free of endotoxins; (iv) scalable fermentation capacity; and (v) the availability of both inducible and constitutive genetic control systems such as the nisin-inducible expression (NICE) system.

The NICE system is a popular expression system of choice among the *Lactococcal* strains for the expression of a range of proteins, including hard-to-obtain membrane proteins (Figure 1B and Table 1). This system constitutes a two-component signal transduction system of *nisK* and *nisR*, which together activate the promoters in *cis* in response to subinhibitory concentrations of nisin [58]. In particular, an *L. lactis* NZ9000 strain that incorporates *nisK* and *nisR* has been developed [58] and along with its derivatives, is in active use for recombinant protein expression [94]. This system, along with others, has been used to produce various prokaryotic as well as eukaryotic proteins of animal and plant origin [95–97], with disulfide bonds intact [98]. Of particular note, *L. lactis* expression system has been proven to be highly efficient in membrane protein expression, where the expressed membrane proteins are localized to the cytoplasmic membrane and the use of mild detergents can readily solubilize these membrane proteins [94]. It is also worth noting that the quality of protein conformation and solubility of recombinant proteins in *L. lactis* are closely coupled to culture conditions including temperature and production time [99]. While synthetic biology tools available for *Lactococcus* are not as diverse as those of *E. coli* or *Bacillus*, the engineered *L. lactis* strain and its derivatives nonetheless exhibited an excellent potential for the production of difficult-to-express proteins. With growing availability of genetic tools and techniques [100], it is expected to serve as a robust microbial chassis for recombinant protein expression.

2.4. Extremophiles as Alternative Protein Expression Systems

Extremophiles refers to microbial cells that function optimally in environments that are extreme for common life forms, as characterized by harsh temperature, salinity, and pH. Some of the adaptive features harbored by these extremophiles hold potentially desirable biotechnological values. For instance, the cytoplasmic compartments of moderate halophiles are densely packed with organic compounds or osmolytes as part of an adaptive response to cope with high-salt conditions [101]. This led to speculation that the moderate halophilic cytoplasm could confer higher protein folding efficiency compared to other mesophiles like *E. coli* [102,103]. It was empirically demonstrated that several halophiles successfully accommodate mammalian proteins in a soluble state. These include functional expression of human β_2 -adrenoceptor and mammalian olfactory receptors in *Haloferax volcanii* [59,60] (See Table 1) and human brain serine racemase from *Halomonas* sp. and *Chromohalobacter salexigens* [104]. Furthermore, the expression of proteins with a high content of acidic amino acids in halophilic bacteria are under vigorous investigation as potential protein fusion partners to expedite solubilization of recombinant proteins [103].

Thermophiles are increasingly viewed as an alternative prokaryotic expression system that would complement the inherent shortcomings of mesophilic bacteria. Thermophilic microbes represent species of bacteria and archaea that thrive at high temperatures, usually more than 45 °C [105]. The biotechnological importance of thermophiles is appraised in their capacity for high-temperature bioprocesses and functional expression of thermostable enzymes including DNA polymerases and many other biocatalysts with industrial significance [106]. As for the thermostable enzyme production, although several thermostable proteins have been successfully expressed in *E. coli* hosts [107,108], they are largely limited in their extent to simple monomeric proteins that come without strict requirements for cofactors, additional modifications, or processing [106]. In fact, a significant portion of thermophilic proteins, especially those of hyperthermophilic origin, require high temperature and the presence of specific chaperones or other cofactors for proper protein folding [109]. This signifies that the expression of thermostable enzymes, especially heteromultimeric enzymes, would require a cell factory closely related to native thermophiles.

Despite its significance, heterologous expression of thermostable proteins in thermophilic hosts is still an underexplored field of expertise. One limitation is that genetically tractable thermophiles are relatively scarce compared to other mesophilic bacteria. This is largely ascribed to harsh laboratory culture conditions—high temperature, pH, and salinity—that render most of the temperature-sensitive molecular markers obsolete [105]. Regardless, significant efforts have been put forward to develop

genetic systems orthogonal to thermophilic bacteria and archaea in the past decade (summarized in [105,110]), enabling overexpression of homologous and heterologous proteins in the thermophiles. For instance, heat- and arabinose-inducible promoters in the hyperthermophilic archaeon *Sulfolobus solfataricus* have been used for the soluble expression of proteins by up to 1 mg per liter [109]. In another instance, discovery of a novel inducible promoter system in a thermophilic bacterium *Geobacillus kaustophilus* HTA426 led to functional expression of recombinant enzymes, which were otherwise insoluble in mesophilic hosts, for up to 59 mg per liter [61] (Table 1). However, the current status of thermophilic expression hosts is far from being ideal compared to that of highly optimized mesophiles which yield as much as tens of grams per liter of recombinant proteins [81]. Yet, with steady improvements in the molecular genetics of thermophiles, they are highly anticipated as emerging hosts for the overexpression of much-coveted thermostable proteins.

3. Heterologous Protein Expression by Systematically Engineered Bacteria

3.1. Concept and Overview of Synthetic Minimal Genome

Some of the defining characteristics of heterologous expression host can be described in two distinct, but not mutually exclusive, perspectives. First, contemporary ideas on expression hosts mainly describe them as versatile biological systems capable of accommodating and supporting the functional expression of exogenous genetic elements in a controllable manner [111]. Following the rise of multi-omics technology, an emerging view on an ideal expression host is characterized by microbial strains that harbor simplified genetic and metabolic networks amenable to prediction and control, through which an efficient biosynthesis of desired products can be achieved [112]. The former and the latter statements represent the popular ideas in the field of synthetic and systems biology in regard to microbial chassis strains tailored for heterologous expressions. The shared perspectives of synthetic and systems biology together brought into realization the concept of minimal genomes along with their proposed advantages across diverse fields of applications, including expression of heterologous proteins. A growing number of reports already highlight the advantages of genome-reduced microbial strains outperforming their wild-type counterparts in terms of productivity of target proteins [113] and desired biomolecules [114], which result from the elimination of unfavorable features for protein expression and rewired metabolic networks redirected to produce heterologous proteins. In this section, we attempt to provide a brief overview of the concept, design, and mechanism underlying synthesis of the minimal genome.

Availability of gene-essentiality information facilitates the genome reduction process in that it provides a shortlist of genetic elements that could be targeted for removal, circumventing the laborious trial-and-error gene deletion approaches. With gene essentiality information in hand, genome reduction can proceed in either a bottom-up or top-down manner (Figure 2). The bottom-up approach entails designing and building an artificially synthesized genome from scratch. Laboratory-made genetic fragments that constitute a section of the target genome are assembled into larger segments by means of enzymatic assembly [27] and yeast homologous recombination into the partial or whole genome [115]. This way, the minimal genome of *Mycoplasma genitalium* and *Mycoplasma mycoides* were de novo assembled into biologically functional forms [115,116]. More recently, the genome-reduced *E. coli* MDS42 was de novo assembled and synthesized into a genome-recoded version [117] using newly developed homologous recombination methods that facilitate the iterative replacement of large stretches of DNA strands in a highly efficient manner [118]. Implications of the synthetic genome in heterologous protein expression are centered on genome-scale codon replacements, where incorporation of synthetic or rare codons may help facilitate synthesis of unnatural polypeptides [119] and heterologous proteins of eukaryotic origin that grossly differ in their codon content. Details on heterologous protein synthesis using synthetic genome will be discussed in subsequent sections.

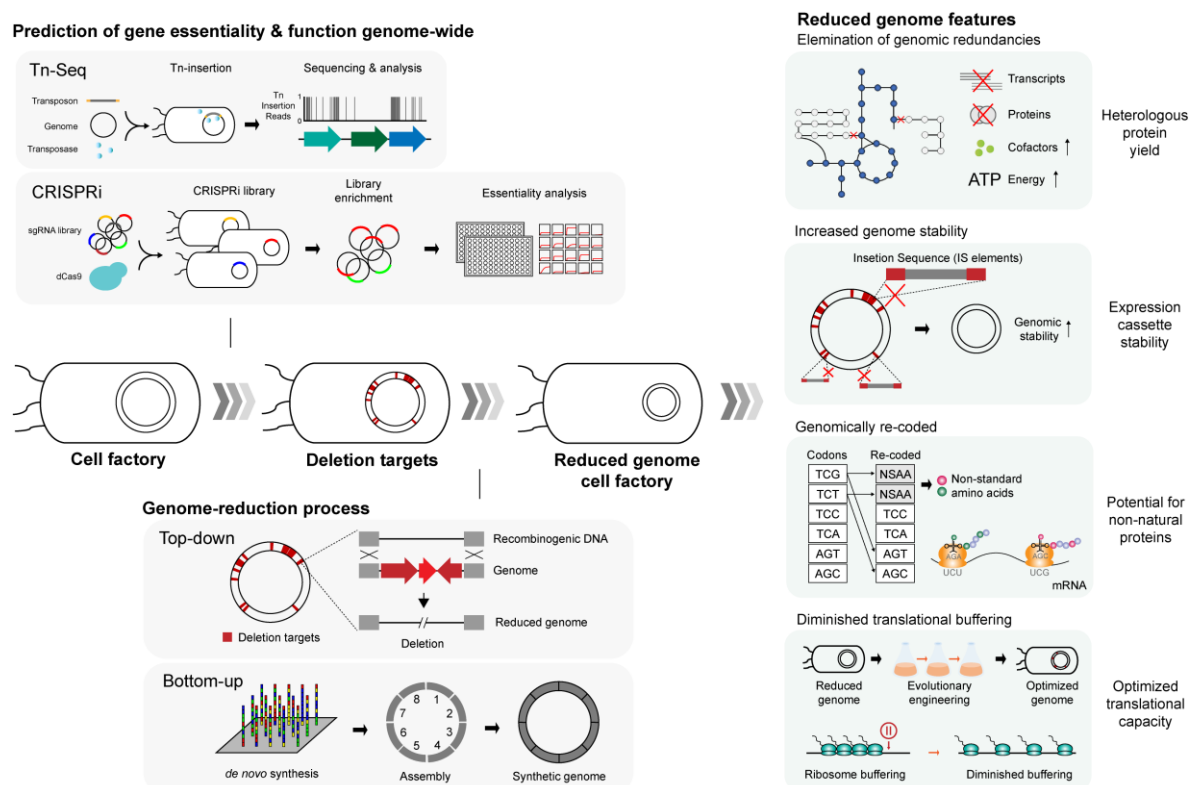


Figure 2. Illustration of the process behind genome streamlining and several notable traits of the genome-reduced strains with relevance to heterologous protein expression.

Conversely, top-down genome-reduction proceeds with systematic removal of nonessential genetic elements starting from an intact genome of the target organism. Typically, homologous recombination mediated by λ red [120] or site-specific Cre-*loxP* [121] and Flp-FLT [122] recombination methods are employed in the process of genomic deletions. Due to relative ease of underlying procedures and cheaper costs, a larger number of top-down genome reduction projects have been published to date as compared with that of bottom-up synthesis (Table 2). Strains subjected to top-down genome-reductions to date include *E. coli* [123–127], *Streptomyces* strains [114,128–133], *B. subtilis* [92,134–138], *L. lactis* [139], *Pseudomonas putida* [140,141], *Corynebacterium glutamicum* [142], and a yeast strain [143]. This strategy enabled the removal of unnecessary proportions of genome at large scale without causing palpable defects in the target organisms (except in a few instances [135,144]) and sometimes yielded improved cellular performances in terms of growth rates, cell density, transformation efficiency, and protein productivity compared with their wild-type counterparts. With an unintended discovery of the desirable properties that synthetic minimal genomes may convey in microbial production, several studies have explored their potential for industrial applications for recombinant protein production [145–148]. The following sections have been outlined to elaborate in detail on cellular mechanisms underlying improved performance in genome-reduced strains and some of the practical examples of the minimal genome expression system, and also to discuss potential applications of genome-reduced strains in heterologous protein expressions.

Table 2. Examples of synthetic genome-reduced strains with defining characteristics beneficial for heterologous protein expression.

Strain	Designation	Genome Reduction	Notable Characteristics	Reference
<i>E. coli</i> MG1655	MDS42	663 kbp (14.3%)	IS-free strain with increased stability of exogenous genetic construct. Increased yield of chimeric fusion protein. Improved efficiency of electroporation comparable to commercial DH10B strain. Additional engineering yielded improved recombinant protein productivity	[124]
<i>E. coli</i> MG1655	MS56	1068 kbp (23.0%)	IS-free strain with increased stability of exogenous genetic construct. Higher electroporation efficiency.	[127]
<i>E. coli</i> MG1655	eMS57	1089 kbp (23.5%)	Evolutionary engineering of MS56 restored the growth on minimal medium. Diminished translational buffering predicted to increase production of recombinant proteins.	[148]
<i>E. coli</i> W3110	MGF-01	1030 kbp (22.2%)	Increased growth density.	[125]
<i>E. coli</i> W3110	DGF-298	1670 kbp (35.9%)	Higher genome stability, increased growth rate.	[149]
<i>B. subtilis</i> 168	MGIM	991 kbp (23.5%)	Small reduction in growth rate, comparable enzyme production.	[135]
<i>B. subtilis</i> 168	MBG874	874 kbp (20.7%)	Protein productivity increased up to 2.5-fold. Enhanced nutrient utilization.	[113]
<i>B. subtilis</i> 168	PG10	1460 kbp (~36%)	Improved secretory protein production, including that of some of the difficult-to-produce proteins.	[137]
<i>L. lactis</i> NZ9000	9k-4	72 kbp (2.8%)	2.2- to 2.5-fold increase in recombinant protein activities. Higher final cell density and growth rates.	[139]
<i>P. putida</i> KT2440	EM383	266 kbp (4.3%)	Higher growth rate and final cell density. Showed as much as 41% increase in recombinant protein yield depending on the carbon source used.	[141]
<i>C. glutamicum</i> ATCC 13032	MB001	205 kbp (6.2%)	Increased recombinant protein activity and transformation efficiency.	[142]

3.2. Applications of the Minimal Genome: from Gene Essentiality to Protein Production

3.2.1. Construction of Genome-Reduced Microbial Strains

The essence of genome reduction lies in identifying genes that are indispensable for sustaining cell viability. Previous attempts to determine gene essentiality involved a priori predictions derived

from in silico comparative genomics in search of homologs and paralogs among remotely related species [150] or systematic inactivation of individual genes from the bacterial genome [151,152]. However, these line of efforts to define essential genes only remained a partial success, as it turned out that evolutionary gene conservation is a poor indicator of gene essentiality [150], and that the nature of gene essentially is highly condition-specific, rendering it almost infeasible to determine using the earlier experimental designs that are extremely time-consuming and laborious [151,152]. Today, the construction of minimal genome is usually accompanied by high-throughput techniques such as transposon mutagenesis sequencing (Tn-Seq) or CRISPR-dCas9, which present a reliable systems-level reference on genome-wide gene essentiality information (Figure 2) [153,154], allowing rational selection of genomic regions for targeted removal. Genome reduction then serves as a testbed to further evaluate the essentiality of the targeted genes or genomic regions and classify them into subcategories in accordance with resulting cellular viability and changes in growth rates ensuing from the targeted deletion (Figure 2) [116].

Apart from its use in assessing genome-scale essentiality, potential uses of minimal genome strains extend to metabolic engineering and protein production. Although there exists a wealth of information on many cellular processes at different subsystems, it was showcased to be extremely challenging to understand the regulatory networks integrated across the subsystems. Synthetic lethality presents a prominent example of unknown regulatory effects, where two independent genes or genomic regions that are otherwise nonlethal when deleted individually are lethal upon simultaneous deletion [126]. Similarly, the deletion of gene candidates that are thought to be redundant or remotely related to certain physiological responses can result in unintended consequences in bioproduction [155]. Hence, the scheme behind minimal-genome-driven metabolic engineering revolves around the idea that simplification of the genome would lead to bacterial chassis with fewer unknown regulations and lesser interference on heterologous pathways, thus allowing more predictable modeling of the chassis of interest [156].

3.2.2. Effect of Genomic Stability on Protein Production

Several genome-reduced strains, by both chance and intent, showcased some of the phenotypes highly favored for the production of heterologous proteins. First, high-fidelity expression of heterologous proteins is ensured by the stability of the genome or expression plasmids that harbor target genes or pathways. Many microbes of industrial use, including *E. coli*, are often plagued by the activation of mobile DNA elements that could randomly inactivate the heterologous expression system [157]. Typically, the expression of heterologous proteins demands a significant portion of metabolic resources from the recombinant host, diverting cellular resources such as ATP, precursors for biosynthesis, and translational machineries away from the host's innate metabolism, hence imposing a metabolic burden on the host microbe [28,158]. In this context, insertion sequence (IS)-mediated mutagenesis serves as an adaptive response to alleviate the metabolic burden by disrupting the expression of foreign proteins, which, in turn, adversely affects the target protein production [157]. The genomes of several *E. coli* genome-reduced strains were shown to be far more stable compared with their deletion-free parental strains (Figure 2) [124,127,149]. This is largely owing to the complete or partial excision of mobile DNA elements, in particular, insertion sequences (ISes) from the genome, ameliorating the possibility of IS-mediated mutagenesis. For instance, it was empirically demonstrated that the first IS-free *E. coli* strain MDS42 with 14% genome reduction restored the yield of a chimeric fusion protein and stably propagated adeno-associated viral vector plasmids that are otherwise frequently inactivated and destabilized in the IS-positive parental MG1655 strain (Table 2) [124]. Similarly, another IS-free genome-reduced *E. coli* strain MS56 exhibited similar genomic and plasmid stability while exhibiting cellular fitness comparable to that of the parental strain, despite a staggering 23% reduction in the genomic content (Table 2) [127]. Succeeding studies have focused on practical applications of the IS-free, genome-reduced strain for the production of commodity biomolecules. *E. coli* MDS42 was re-engineered with additional genetic manipulations to stably overexpress toxic

recombinant protein methyltransferase in its functional form [146] and boosted production of essential amino acid, L-threonine, by 83% [147].

3.2.3. Increased Availability of Cellular Resources

Along with the genetic fidelity of the heterologous expression system imparted by the removal of mutation-causing cellular machineries, many other factors may account for the improved recombinant protein yield in the genome-reduced expression system. In particular, enhanced biosynthesis of desired protein is largely attributable to the removal of innate but redundant metabolic pathways, which prompt to redirect energy, precursor pools, cofactors, and transcriptional and translational machineries toward heterologous metabolic pathways [24,28,128,158]. The metabolic burden imposed by expressing heterologous genes arises from competition for shared but limited cellular resources between exogenous genes and host metabolism. Thus, removal of host genes that are nonessential translates to an overall decrease in the resource cost spent on now-absent host metabolism, which then reduces the net metabolic load [159] and allows for efficient resource utilization and protein translation (Figure 2) [24,137,145]. In addition, reduced genomic complexity in the minimal cells may help to curtail regulatory interference with the targeted or exogenous metabolic pathways [114], reducing the number of compounding variables that pose unknown effects on heterologous expression. Such shared metabolic predispositions of the genome-reduced strains offer a fascinating platform for the bioproduction of numerous valuable commodity molecules, including heterologous proteins.

Empirically, *B. subtilis* MBG874 strain that underwent a 20% genome reduction exhibited increased productivity of episomally expressed proteins up to 2.5-fold higher than that of the wild-type *B. subtilis* 168 (Table 2) [113]. In another instance, mini-*Bacillus* strains that retained only two-thirds of the *B. subtilis* 168 genome, with extra re-engineering effort, were able to stably express four individual staphylococcal antigens that are recalcitrant for heterologous protein expression in the existing *B. subtilis* chassis (Table 2) [145]. Genome-streamlined *L. lactis* chassis NZ9000 specializing in the production of many membranous and disulfide-containing proteins also exhibited superior growth rate, biomass yield, and nutrition utilization capacity [139]. Importantly, 2.8% of genome reduction targeting the prophages and their cognate genes in *L. lactis* significantly improved the production of the model recombinant proteins, leucocin C and the red fluorescence protein (RFP). Phenotype screenings of the target proteins showed up to 2.5-fold increase in the inhibiting zone (bacteriocin leucocin C) and a 2.2-fold increase in the RFP intensity, compared with wild-type strain [139]. Other industrially relevant strains in the context of heterologous protein expressions, such as *Pseudomonas putida* and *Corynebacterium*, showed equally encouraging outcomes in terms of protein productivity [141] and genetic engineering [142] (Table 2).

Together, this evidence highlights the advantages of genome-reduced strains in heterologous expression. Generally, overexpression of heterologous genes force-feeds metabolic fluxes and gene expression machineries that are otherwise dedicated to host metabolism to the production of target proteins [28,158]. Subsequent depletion of metabolic precursors and shared translational resources invoke cellular adaptive responses that resemble starvation and amino acid depletion, often manifested in the form of reduced cell biomass, growth retardation, and loss of cell viability [28,158]. Systematic deletion of redundant genes and associated metabolic pathways seems to alleviate the genetic load in the genome-reduced strain, as evident through the increased growth and target protein yield. Second, targeted removal of mutation-causing insertion elements from the genome dramatically increases the genetic fidelity of recombinant expression systems, enabling a long-term, stable recombinant protein expression.

3.2.4. Optimization of Codon Usage

Another frequent issue in recombinant protein production lies in processing rare codons in the target heterologous genes. Differences in the usage of synonymous codons on the coding regions of the genome have been reported to influence transcriptional and translational features, including translation

elongation rates [160], mRNA folding stability [161], protein production rate [161], and protein folding [160]. The effect of codon usage is especially compelling in heterologous protein expression, where incompatibility of codon usage between recombinant genes and expression host (also known as codon bias) has often hampered target protein expression due in large part to a sheer shortage of rare tRNAs or lack thereof [9,162]. It was suggested that codon abundance is positively correlated with the availability of complementary tRNA species, which is then proportional to protein production [9]. The paucity of host tRNA molecules orthogonal to a cluster of codons highly abundant in the target protein could inflict a host of translational errors during heterologous expression, including premature termination (ribosomal stalling) [163], amino acid misincorporation [164], translational frameshifts [165], and decreased protein quality and yield [9,55], which is further aggravated upon overexpression of the target heterologous protein. Efforts to thwart codon-bias-derived expression problems have been computationally optimizing the codon profile of a target gene and selecting for the synonymous codons that best fit the codon usage of the expression host [166]; gene sequence optimization [165]; or co-overexpressing rare tRNA genes accompanied with supplementation of cognate amino acids in the culture media [9]. The latter option has been exploited in commercially available *E. coli* Rosetta(DE3) strain, which contains a plasmid to coexpress tRNA species common in eukaryotic proteins but rare in *E. coli* (Figure 1D and Table 1) [167]. Expression of rare tRNAs accompanied with substitution of *E. coli* Rosetta(DE3) could enhance the overall productivity of industrial enzymes and recombinant protein fragments [55,167], except in some instances, productivity of individual proteins was poorer compared with that of the control strain, suggesting the presence of unknown metabolic burden [55].

The recent demonstration of genome-recoding from bottom-up to yield a fully functional *E. coli* [117] opened a way for a new synthetic biology design for the production of heterologous proteins. The synthetic *E. coli* genome now harbors 59 out of 61 sense codons and resembles otherwise isogenic *E. coli* MDS42 (Table 2) [124]. A total of 18,214 codons composed of two serine synonymous codons UCG and UCA, and the amber stop codon UAG underwent substitution with their synonymous counterparts, leaving three unassigned (blank) codons in the genome. Extended application of the availability of blank codons entailed assignment and incorporation of noncanonical amino acids (NCAAs) in desired genes with the help of aminoacyl-tRNA synthetase complementary to the NCAAs (Figure 2) [29,117], opening the possibility of engineering microbes for the production of a wide, heterogeneous spectrum of proteins [29]. Furthermore, with 57-codon genome underway to elucidate many more aspects of translation [168], it is speculated that designing genome-recoded chassis oriented toward expression of heterologous proteins could be synergized by working in tandem with contemporary troubleshooting designs, tRNA overexpression, and gene sequence optimization to alleviate imminent problems of heterologous protein expression originating from the codon bias.

3.2.5. Changes in Translation Efficiency

The abundance of mRNA level ensuing from the overexpression of heterologous gene alone may not necessarily correspond to high protein levels, indicating the presence of post-transcriptional regulations [169,170]. This regulatory mechanism is an extension to codon bias in that the presence of rare codons causes retardation in translation elongation rate [163], and it was shown to be negatively correlated to the abundance of mRNA transcripts and their translation [169]. An investigation into this phenomenon gave rise to the term translation efficiency (TE), which is poised to explain the variation between the abundance of mRNA transcripts and cognate protein molecules [169]. TE provides a measurable means to quantitatively assess how effectively translation is taking place within the mRNA transcripts as presented in the ratio between ribosome-protected mRNA fragment levels (RPF) and the corresponding transcript counts per gene [169]. In the same study, it was demonstrated that mRNA transcripts with variable abundance showed strikingly lesser variations in ribosome occupancy. In other words, translational regulations work to conserve cellular protein abundance by controlling the number of mRNAs actively translated by ribosomes, as featured through ribosome-occupied mRNA transcript counts. Such a mechanism effectively buffers for the variations in mRNA levels,

possibly functioning as part of a compensatory mechanism for aberrant mRNA expression and to stably maintain steady-state protein levels. This phenomenon, often referred to as translational buffering, hinders the proper expression of heterologous proteins in prokaryotes [9,163]. However, the molecular basis underlying translational buffering is still under active investigation [169–171], rendering it highly challenging to establish strategies for translational-level regulation control. It is only out of hindsight that we know that the interplay of multiple intrinsic [163] and extrinsic [172] factors can, each with differing magnitudes [148,173], pose an influence on translational buffering. These include the sequence on 5' untranslated region (5' UTR), action of regulatory proteins, and ribosomal stalling (rare codons). Among the proposed measures to alleviate translational buffering, increasing cellular tRNA availability is regarded as the most feasible solution to mitigate ribosomal pausing [173]. Until recently, a study reported diminished translational buffering in a minimal genome *E. coli* that underwent evolutionary engineering (Figure 2) [148], where the genome-reduced *E. coli* MS56 [127] subjected to adaptive evolution for 800 generations (resulting in eMS57) constantly exhibited higher TE values for mRNAs that share a similar expression level with that of MG1655 control (Table 2). While evolutionary engineering of prokaryotes has been shown to rewire metabolic pathways and reorganize transcriptomes [22], eMS57 showed a rare instance of global-scale remodeling of translome mediated by adaptive evolution. The authors concluded that the diminished translational buffering in eMS57 was indeed communicated by the translome-wide remodeling in the process of adaptation, and in principle, this change was driven by an increased availability of ribosome as per the reduced gene counts and a more spacious spatial proximity between neighboring genes that may buffer the expression of one another [174]. Although actual protein productivity of this strain has not been empirically tested, it can be said that the integration of evolutionary engineering with rational genome reduction provides a potentially promising strategy to optimize translational regulation.

4. Conclusion and Perspectives

Microbial expression systems have been delivering wide varieties of high-quality recombinant proteins in a scalable and cost-effective manner. While incompatibility between microbial translation machineries and eukaryotic protein architecture presents a serious setback to heterologous protein expression, new discoveries and engineering strategies have been put forward to circumvent the problems. Today, catalogs of highly specialized microbial chassis are available for the functional expression of recombinant proteins with specific structural and biochemical demands (Table 1). Despite the efforts, the complex interplay of cellular physiology and recombinant expression pathways still renders the optimization of protein expression a highly challenging task. Recent advances in the systems-level understanding of cellular expression machineries and the wealth of transcriptomic and proteomic data together present an alternative scheme that could possibly mitigate the bottlenecks in recombinant protein expression.

The shared perspectives of systems and synthetic biology fostered an emerging idea that cells can be made simpler and more efficient by eliminating components of the genome deemed unnecessary for survival. Indeed, precisely designed genome reductions have consistently led to enhanced cellular features including genomic stability, growth rate, biomass, protein productivity, and genetic pliability (more amenable to genetic engineering) in the target microbial cells. For the most part, an increasing number of existing microbial chassis have been genome-streamlined to resolve some of the bottlenecks in heterologous protein expression and to further optimize their performance. Two opposing approaches exist for the construction of reduced genome strains: top-down genome reduction and bottom-up genome synthesis. Despite the shared goal, the two opposing approaches essentially differ with respect to their area of expertise. While the top-down genome reduction offers a pragmatic approach to further improve the functionality of existing cell factories, the current status of the bottom-up synthesis is more inclined toward understanding the fundamental aspects of the minimal genome. Although it is said that the bottom-up genome synthesis may provide new ground to synthesize non-natural proteins for medicinal use, the underlying cost and time to build one such

cell do not outweigh the benefits of chemical-based synthesis methods, at least for the near future. Nevertheless, genome-reduction approaches overall represent an ingenious attempt to understand the gene functions and correct for the glitches in genetic makeups that potentially undermine the performance of the cell factories. However, it is to be noted that the genome reduction alone is far from sufficient to tackle every aspect of difficult-to-express protein expression and that the scope of genome reduction in heterologous protein expression is highly specific to certain aspects (i.e., cellular resource reallocation, alleviation of protein toxicity, translation modulation). A carefully designed genome-reduction strategy would open an additional window to further improve the cellular capacity for heterologous protein expression.

Author Contributions: B.-K.C. conceptualized and supervised the project; K.K., D.C., D.-H.L., and B.-K.C., wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study is supported by the Korea Bio Grand Challenge (2018M3A9H3024759 to B.-K.C.) and Bio & Medical Technology Development Program (2018M3A9F3079664 to B.-K.C.) through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Itakura, K.; Hirose, T.; Crea, R.; Riggs, A.D.; Heyneker, H.L.; Bolivar, F.; Boyer, H.W. Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. *Science* **1977**, *198*, 1056–1063. [[CrossRef](#)] [[PubMed](#)]
2. Williams, D.C.; Van Frank, R.M.; Muth, W.L.; Burnett, J.P. Cytoplasmic inclusion bodies in *Escherichia coli* producing biosynthetic human insulin proteins. *Science* **1982**, *215*, 687–689. [[CrossRef](#)] [[PubMed](#)]
3. Carlson, R. Estimating the biotech sector's contribution to the US economy. *Nat. Biotechnol.* **2016**, *34*, 247–255. [[CrossRef](#)]
4. Sanchez-Garcia, L.; Martin, L.; Mangues, R.; Ferrer-Miralles, N.; Vazquez, E.; Villaverde, A. Recombinant pharmaceuticals from microbial cells: A 2015 update. *Microb. Cell Fact.* **2016**, *15*, 33. [[CrossRef](#)] [[PubMed](#)]
5. Tanhaiean, A.; Azghandi, M.; Razmyar, J.; Mohammadi, E.; Sekhavati, M.H. Recombinant production of a chimeric antimicrobial peptide in *E. coli* and assessment of its activity against some avian clinically isolated pathogens. *Microb. Pathog.* **2018**, *122*, 73–78. [[CrossRef](#)] [[PubMed](#)]
6. Robinson, M.P.; Ke, N.; Lobstein, J.; Peterson, C.; Szkodny, A.; Mansell, T.J.; Tuckey, C.; Riggs, P.D.; Colussi, P.A.; Noren, C.J.; et al. Efficient expression of full-length antibodies in the cytoplasm of engineered bacteria. *Nat. Commun.* **2015**, *6*, 8072. [[CrossRef](#)]
7. Jaeger, K.E.; Eggert, T.; Eipper, A.; Reetz, M.T. Directed evolution and the creation of enantioselective biocatalysts. *Appl. Microbiol. Biotechnol.* **2001**, *55*, 519–530. [[CrossRef](#)]
8. Derman, A.I.; Beckwith, J. *Escherichia coli* alkaline phosphatase fails to acquire disulfide bonds when retained in the cytoplasm. *J. Bacteriol.* **1991**, *173*, 7719–7722. [[CrossRef](#)]
9. Kane, J.F. Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Curr. Opin. Biotechnol.* **1995**, *6*, 494–500. [[CrossRef](#)]
10. Vallejo, L.F.; Rinas, U. Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb. Cell Fact.* **2004**, *3*, 11. [[CrossRef](#)]
11. Jenkins, N.; Curling, E.M. Glycosylation of recombinant proteins: Problems and prospects. *Enzym. Microb. Technol.* **1994**, *16*, 354–364. [[CrossRef](#)]
12. Paddon, C.J.; Westfall, P.J.; Pitera, D.J.; Benjamin, K.; Fisher, K.; McPhee, D.; Leavell, M.D.; Tai, A.; Main, A.; Eng, D.; et al. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **2013**, *496*, 528–532. [[CrossRef](#)] [[PubMed](#)]
13. Luli, G.W.; Strohl, W.R. Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations. *Appl. Environ. Microbiol.* **1990**, *56*, 1004–1011. [[CrossRef](#)] [[PubMed](#)]
14. Li, D.; Fu, G.; Tu, R.; Jin, Z.; Zhang, D. High-efficiency expression and secretion of human FGF21 in *Bacillus subtilis* by intercalation of a mini-cistron cassette and combinatorial optimization of cell regulatory components. *Microb. Cell Fact.* **2019**, *18*, 17. [[CrossRef](#)] [[PubMed](#)]

15. Studier, F.W.; Moffatt, B.A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **1986**, *189*, 113–130. [[CrossRef](#)]
16. Guzman, L.M.; Belin, D.; Carson, M.J.; Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.* **1995**, *177*, 4121–4130. [[CrossRef](#)]
17. Wang, R.F.; Kushner, S.R. Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* **1991**, *100*, 195–199. [[CrossRef](#)]
18. Nallamsetty, S.; Waugh, D.S. A generic protocol for the expression and purification of recombinant proteins in *Escherichia coli* using a combinatorial His6-maltose binding protein fusion tag. *Nat. Protoc.* **2007**, *2*, 383–391. [[CrossRef](#)]
19. Baneyx, F.; Mujacic, M. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol.* **2004**, *22*, 1399–1408. [[CrossRef](#)]
20. Dittrich, C.R.; Vadali, R.V.; Bennett, G.N.; San, K.Y. Redistribution of metabolic fluxes in the central aerobic metabolic pathway of *E. coli* mutant strains with deletion of the ackA-pta and poxB pathways for the synthesis of isoamyl acetate. *Biotechnol. Prog.* **2005**, *21*, 627–631. [[CrossRef](#)]
21. Aristidou, A.A.; San, K.Y.; Bennett, G.N. Metabolic engineering of *Escherichia coli* to enhance recombinant protein production through acetate reduction. *Biotechnol. Prog.* **1995**, *11*, 475–478. [[CrossRef](#)] [[PubMed](#)]
22. Dragosits, M.; Mattanovich, D. Adaptive laboratory evolution—principles and applications for biotechnology. *Microb. Cell Fact.* **2013**, *12*, 64. [[CrossRef](#)] [[PubMed](#)]
23. Liponska, A.; Ousalem, F.; Aalberts, D.P.; Hunt, J.F.; Boel, G. The new strategies to overcome challenges in protein production in bacteria. *Microb. Biotechnol.* **2019**, *12*, 44–47. [[CrossRef](#)] [[PubMed](#)]
24. Choe, D.; Cho, S.; Kim, S.C.; Cho, B.K. Minimal genome: Worthwhile or worthless efforts toward being smaller? *Biotechnol. J.* **2016**, *11*, 199–211. [[CrossRef](#)]
25. Cho, B.K.; Zengler, K.; Qiu, Y.; Park, Y.S.; Knight, E.M.; Barrett, C.L.; Gao, Y.; Palsson, B.O. The transcription unit architecture of the *Escherichia coli* genome. *Nat. Biotechnol.* **2009**, *27*, 1043–1049. [[CrossRef](#)]
26. Nakahigashi, K.; Toya, Y.; Ishii, N.; Soga, T.; Hasegawa, M.; Watanabe, H.; Takai, Y.; Honma, M.; Mori, H.; Tomita, M. Systematic phenome analysis of *Escherichia coli* multiple-knockout mutants reveals hidden reactions in central carbon metabolism. *Mol. Syst. Biol.* **2009**, *5*, 306. [[CrossRef](#)]
27. Gibson, D.G.; Young, L.; Chuang, R.Y.; Venter, J.C.; Hutchison, C.A.; Smith, H.O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **2009**, *6*, 343–345. [[CrossRef](#)]
28. Carr, P.A.; Church, G.M. Genome engineering. *Nat. Biotechnol.* **2009**, *27*, 1151–1162. [[CrossRef](#)]
29. Chin, J.W. Molecular biology. Reprogramming the genetic code. *Science* **2012**, *336*, 428–429.
30. Smanski, M.J.; Bhatia, S.; Zhao, D.; Park, Y.; Lauren, B.A.W.; Giannoukos, G.; Ciulla, D.; Busby, M.; Calderon, J.; Nicol, R.; et al. Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol.* **2014**, *32*, 1241–1249. [[CrossRef](#)]
31. Chi, H.; Wang, X.; Shao, Y.; Qin, Y.; Deng, Z.; Wang, L.; Chen, S. Engineering and modification of microbial chassis for systems and synthetic biology. *Synth. Syst. Biotechnol.* **2019**, *4*, 25–33. [[CrossRef](#)] [[PubMed](#)]
32. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids. Res.* **2000**, *28*, 235–242. [[CrossRef](#)] [[PubMed](#)]
33. Saccardo, P.; Corchero, J.L.; Ferrer-Miralles, N. Tools to cope with difficult-to-express proteins. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 4347–4355. [[CrossRef](#)] [[PubMed](#)]
34. Li, G.W.; Oh, E.; Weissman, J.S. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature* **2012**, *484*, 538–541. [[CrossRef](#)] [[PubMed](#)]
35. Salis, H.M.; Mirsky, E.A.; Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat. Biotechnol.* **2009**, *27*, 946–950. [[CrossRef](#)] [[PubMed](#)]
36. Seo, S.W.; Yang, J.S.; Cho, H.S.; Yang, J.; Kim, S.C.; Park, J.M.; Kim, S.; Jung, G.Y. Predictive combinatorial design of mRNA translation initiation regions for systematic optimization of gene expression levels. *Sci. Rep.* **2014**, *4*, 4515. [[CrossRef](#)]
37. Miksch, G.; Bettenworth, F.; Friehs, K.; Flaschel, E.; Saalbach, A.; Twellmann, T.; Nattkemper, T.W. Libraries of synthetic stationary-phase and stress promoters as a tool for fine-tuning of expression of recombinant proteins in *Escherichia coli*. *J. Biotechnol.* **2005**, *120*, 25–37. [[CrossRef](#)]
38. Chen, Y.J.; Liu, P.; Nielsen, A.A.; Brophy, J.A.; Clancy, K.; Peterson, T.; Voigt, C.A. Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat. Methods* **2013**, *10*, 659–664. [[CrossRef](#)]

39. Zaslaver, A.; Bren, A.; Ronen, M.; Itzkovitz, S.; Kikoin, I.; Shavit, S.; Liebermeister, W.; Surette, M.G.; Alon, U. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nat. Methods* **2006**, *3*, 623–628. [[CrossRef](#)]
40. Strandberg, L.; Andersson, L.; Enfors, S.O. The use of fed batch cultivation for achieving high cell densities in the production of a recombinant protein in *Escherichia coli*. *FEMS Microbiol. Rev.* **1994**, *14*, 53–56. [[CrossRef](#)]
41. Terpe, K. Overview of bacterial expression systems for heterologous protein production: From molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* **2006**, *72*, 211–222. [[CrossRef](#)] [[PubMed](#)]
42. Apweiler, R.; Hermjakob, H.; Sharon, N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* **1999**, *1473*, 4–8. [[CrossRef](#)]
43. Brown, J.L.; Roberts, W.K. Evidence that approximately eighty per cent of the soluble proteins from Ehrlich ascites cells are Nalpha-acetylated. *J. Biol. Chem.* **1976**, *251*, 1009–1014. [[PubMed](#)]
44. Holmes, W.M.; Mannakee, B.K.; Gutenkunst, R.N.; Serio, T.R. Loss of amino-terminal acetylation suppresses a prion phenotype by modulating global protein folding. *Nat. Commun.* **2014**, *5*, 4383. [[CrossRef](#)] [[PubMed](#)]
45. Wacker, M.; Linton, D.; Hitchen, P.G.; Nita-Lazar, M.; Haslam, S.M.; North, S.J.; Panico, M.; Morris, H.R.; Dell, A.; Wren, B.W.; et al. N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* **2002**, *298*, 1790–1793. [[CrossRef](#)] [[PubMed](#)]
46. Murata, T.; Shinozuka, Y.; Obata, Y.; Yokoyama, K.K. Phosphorylation of two eukaryotic transcription factors, Jun dimerization protein 2 and activation transcription factor 2, in *Escherichia coli* by Jun N-terminal kinase 1. *Anal. Biochem.* **2008**, *376*, 115–121. [[CrossRef](#)]
47. Ren, Y.; Yao, X.; Dai, H.; Li, S.; Fang, H.; Chen, H.; Zhou, C. Production of Nalpha-acetylated thymosin alpha1 in *Escherichia coli*. *Microb. Cell Fact.* **2011**, *10*, 26. [[CrossRef](#)]
48. Eastwood, T.A.; Baker, K.; Brooker, H.R.; Frank, S.; Mulvihill, D.P. An enhanced recombinant amino-terminal acetylation system and novel in vivo high-throughput screen for molecules affecting alpha-synuclein oligomerisation. *FEBS Lett.* **2017**, *591*, 833–841. [[CrossRef](#)]
49. Neumann, H.; Peak-Chew, S.Y.; Chin, J.W. Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. *Nat. Chem. Biol.* **2008**, *4*, 232–234. [[CrossRef](#)]
50. de Marco, A.; Deuerling, E.; Mogk, A.; Tomoyasu, T.; Bukau, B. Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *E. coli*. *BMC Biotechnol.* **2007**, *7*, 32. [[CrossRef](#)]
51. Bessette, P.H.; Aslund, F.; Beckwith, J.; Georgiou, G. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13703–13708. [[CrossRef](#)] [[PubMed](#)]
52. Hatahet, F.; Nguyen, V.D.; Salo, K.E.; Ruddock, L.W. Disruption of reducing pathways is not essential for efficient disulfide bond formation in the cytoplasm of *E. coli*. *Microb. Cell Fact.* **2010**, *9*, 67. [[CrossRef](#)] [[PubMed](#)]
53. Miroux, B.; Walker, J.E. Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **1996**, *260*, 289–298. [[CrossRef](#)] [[PubMed](#)]
54. Schlegel, S.; Lofblom, J.; Lee, C.; Hjelm, A.; Klepsch, M.; Strous, M.; Drew, D.; Slotboom, D.J.; de Gier, J.W. Optimizing membrane protein overexpression in the *Escherichia coli* strain Lemo21(DE3). *J. Mol. Biol.* **2012**, *423*, 648–659. [[CrossRef](#)]
55. Tegel, H.; Tourle, S.; Ottosson, J.; Persson, A. Increased levels of recombinant human proteins with the *Escherichia coli* strain Rosetta(DE3). *Protein Expr. Purif.* **2010**, *69*, 159–167. [[CrossRef](#)]
56. Wu, X.C.; Lee, W.; Tran, L.; Wong, S.L. Engineering a *Bacillus subtilis* expression-secretion system with a strain deficient in six extracellular proteases. *J. Bacteriol.* **1991**, *173*, 4952–4958. [[CrossRef](#)]
57. Thwaite, J.E.; Baillie, L.W.; Carter, N.M.; Stephenson, K.; Rees, M.; Harwood, C.R.; Emmerson, P.T. Optimization of the cell wall microenvironment allows increased production of recombinant *Bacillus anthracis* protective antigen from *B. subtilis*. *Appl. Environ. Microbiol.* **2002**, *68*, 227–234. [[CrossRef](#)]
58. Kuipers, O.P.; Beerthuyzen, M.M.; de Ruyter, P.G.; Luesink, E.J.; de Vos, W.M. Autoregulation of nisin biosynthesis in *Lactococcus lactis* by signal transduction. *J. Biol. Chem.* **1995**, *270*, 27299–27304. [[CrossRef](#)]
59. Sohlmann, P.; Soppa, J.; Oesterhelt, D.; Lohse, M.J. Expression of beta 2-adrenoceptors in halobacteria. *Naunyn Schmiedebergs Arch. Pharmacol.* **1997**, *355*, 150–160. [[CrossRef](#)]

60. Lobasso, S.; Vitale, R.; Lopalco, P.; Corcelli, A. Haloferax volcanii, as a Novel Tool for Producing Mammalian Olfactory Receptors Embedded in Archaeal Lipid Bilayer. *Life* **2015**, *5*, 770–782. [[CrossRef](#)]
61. Suzuki, H.; Yoshida, K.; Ohshima, T. Polysaccharide-degrading thermophiles generated by heterologous gene expression in *Geobacillus kaustophilus* HTA426. *Appl. Environ. Microbiol.* **2013**, *79*, 5151–5158. [[CrossRef](#)] [[PubMed](#)]
62. Wacker, M.; Feldman, M.F.; Callewaert, N.; Kowarik, M.; Clarke, B.R.; Pohl, N.L.; Hernandez, M.; Vines, E.D.; Valvano, M.A.; Whitfield, C.; et al. Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 7088–7093. [[CrossRef](#)] [[PubMed](#)]
63. Valderrama-Rincon, J.D.; Fisher, A.C.; Merritt, J.H.; Fan, Y.Y.; Reading, C.A.; Chhiba, K.; Heiss, C.; Azadi, P.; Aebi, M.; DeLisa, M.P. An engineered eukaryotic protein glycosylation pathway in *Escherichia coli*. *Nat. Chem. Biol.* **2012**, *8*, 434–436. [[CrossRef](#)] [[PubMed](#)]
64. Schaffer, C.; Messner, P. Emerging facets of prokaryotic glycosylation. *FEMS Microbiol. Rev.* **2017**, *41*, 49–91. [[CrossRef](#)] [[PubMed](#)]
65. Thurmond, J.M.; Hards, R.G.; Seipelt, C.T.; Leonard, A.E.; Hansson, L.; Stromqvist, M.; Bystrom, M.; Enquist, K.; Xu, B.C.; Kopchick, J.J.; et al. Expression and characterization of phosphorylated recombinant human beta-casein in *Escherichia coli*. *Protein Expr. Purif.* **1997**, *10*, 202–208. [[CrossRef](#)] [[PubMed](#)]
66. Christensen, D.G.; Meyer, J.G.; Baumgartner, J.T.; D'Souza, A.K.; Nelson, W.C.; Payne, S.H.; Kuhn, M.L.; Schilling, B.; Wolfe, A.J. Identification of Novel Protein Lysine Acetyltransferases in *Escherichia coli*. *mBio* **2018**, *9*, e01905-18. [[CrossRef](#)]
67. Amrein, K.E.; Takacs, B.; Stieger, M.; Molnos, J.; Flint, N.A.; Burn, P. Purification and characterization of recombinant human p50csk protein-tyrosine kinase from an *Escherichia coli* expression system overproducing the bacterial chaperones GroES and GroEL. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 1048–1052. [[CrossRef](#)]
68. Nishihara, K.; Kanemori, M.; Yanagi, H.; Yura, T. Overexpression of trigger factor prevents aggregation of recombinant proteins in *Escherichia coli*. *Appl. Environ. Microbiol.* **2000**, *66*, 884–889. [[CrossRef](#)]
69. Prinz, W.A.; Aslund, F.; Holmgren, A.; Beckwith, J. The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. *J. Biol. Chem.* **1997**, *272*, 15661–15667. [[CrossRef](#)]
70. Bosnjak, I.; Bojovic, V.; Segvic-Bubic, T.; Bielen, A. Occurrence of protein disulfide bonds in different domains of life: A comparison of proteins from the Protein Data Bank. *Protein Eng. Des. Sel.* **2014**, *27*, 65–72. [[CrossRef](#)]
71. Freudl, R. Signal peptides for recombinant protein secretion in bacterial expression systems. *Microb. Cell Fact.* **2018**, *17*, 52. [[CrossRef](#)] [[PubMed](#)]
72. Stewart, E.J.; Aslund, F.; Beckwith, J. Disulfide bond formation in the *Escherichia coli* cytoplasm: An in vivo role reversal for the thioredoxins. *EMBO J.* **1998**, *17*, 5543–5550. [[CrossRef](#)] [[PubMed](#)]
73. Faulkner, M.J.; Veeravalli, K.; Gon, S.; Georgiou, G.; Beckwith, J. Functional plasticity of a peroxidase allows evolution of diverse disulfide-reducing pathways. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6735–6740. [[CrossRef](#)] [[PubMed](#)]
74. Gaciarz, A.; Khatri, N.K.; Velez-Suberbie, M.L.; Saaranen, M.J.; Uchida, Y.; Keshavarz-Moore, E.; Ruddock, L.W. Efficient soluble expression of disulfide bonded proteins in the cytoplasm of *Escherichia coli* in fed-batch fermentations on chemically defined minimal media. *Microb. Cell Fact.* **2017**, *16*, 108. [[CrossRef](#)]
75. Schlegel, S.; Genevaux, P.; de Gier, J.W. De-convoluting the Genetic Adaptations of *E. coli* C41(DE3) in Real Time Reveals How Alleviating Protein Production Stress Improves Yields. *Cell Rep.* **2015**, *10*, 1758–1766. [[CrossRef](#)]
76. Giacalone, M.J.; Gentile, A.M.; Lovitt, B.T.; Berkley, N.L.; Gunderson, C.W.; Surber, M.W. Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. *Biotechniques* **2006**, *40*, 355–364. [[CrossRef](#)]
77. Schallmeyer, M.; Singh, A.; Ward, O.P. Developments in the use of *Bacillus* species for industrial production. *Can. J. Microbiol.* **2004**, *50*, 1–17. [[CrossRef](#)]
78. Westers, L.; Westers, H.; Quax, W.J. *Bacillus subtilis* as cell factory for pharmaceutical proteins: A biotechnological approach to optimize the host organism. *Biochim. Biophys. Acta* **2004**, *1694*, 299–310. [[CrossRef](#)]
79. Dong, H.; Zhang, D. Current development in genetic engineering strategies of *Bacillus* species. *Microb. Cell Fact.* **2014**, *13*, 63. [[CrossRef](#)]

80. Boratynski, J.; Szermer-Olearnik, B. Endotoxin Removal from *Escherichia coli* Bacterial Lysate Using a Biphasic Liquid System. *Methods Mol. Biol.* **2017**, *1600*, 107–112.
81. Simonen, M.; Palva, I. Protein secretion in *Bacillus* species. *Microbiol. Rev.* **1993**, *57*, 109–137. [[CrossRef](#)] [[PubMed](#)]
82. Tjalsma, H.; Antelmann, H.; Jongbloed, J.D.; Braun, P.G.; Darmon, E.; Dorenbos, R.; Dubois, J.Y.; Westers, H.; Zanen, G.; Quax, W.J.; et al. Proteomics of protein secretion by *Bacillus subtilis*: Separating the secrets of the secretome. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 207–233. [[CrossRef](#)] [[PubMed](#)]
83. Diao, L.; Dong, Q.; Xu, Z.; Yang, S.; Zhou, J.; Freudl, R. Functional implementation of the posttranslational SecB-SecA protein-targeting pathway in *Bacillus subtilis*. *Appl. Environ. Microbiol.* **2012**, *78*, 651–659. [[CrossRef](#)] [[PubMed](#)]
84. Kakeshita, H.; Kageyama, Y.; Ara, K.; Ozaki, K.; Nakamura, K. Enhanced extracellular production of heterologous proteins in *Bacillus subtilis* by deleting the C-terminal region of the SecA secretory machinery. *Mol. Biotechnol.* **2010**, *46*, 250–257. [[CrossRef](#)] [[PubMed](#)]
85. Chen, J.; Zhao, L.; Fu, G.; Zhou, W.; Sun, Y.; Zheng, P.; Sun, J.; Zhang, D. A novel strategy for protein production using non-classical secretion pathway in *Bacillus subtilis*. *Microb. Cell Fact.* **2016**, *15*, 69. [[CrossRef](#)]
86. Zhao, L.; Chen, J.; Sun, J.; Zhang, D. Multimer recognition and secretion by the non-classical secretion pathway in *Bacillus subtilis*. *Sci. Rep.* **2017**, *7*, 44023. [[CrossRef](#)]
87. He, X.S.; Bruckner, R.; Doi, R.H. The protease genes of *Bacillus subtilis*. *Res. Microbiol.* **1991**, *142*, 797–803. [[CrossRef](#)]
88. Westers, L.; Dijkstra, D.S.; Westers, H.; van Dijl, J.M.; Quax, W.J. Secretion of functional human interleukin-3 from *Bacillus subtilis*. *J. Biotechnol.* **2006**, *123*, 211–224. [[CrossRef](#)]
89. Wong, S.L.; Ye, R.; Nathoo, S. Engineering and production of streptokinase in a *Bacillus subtilis* expression-secretion system. *Appl. Environ. Microbiol.* **1994**, *60*, 517–523. [[CrossRef](#)]
90. Murashima, K.; Chen, C.L.; Kosugi, A.; Tamaru, Y.; Doi, R.H.; Wong, S.L. Heterologous production of *Clostridium cellulovorans* engB, using protease-deficient *Bacillus subtilis*, and preparation of active recombinant cellulosomes. *J. Bacteriol.* **2002**, *184*, 76–81. [[CrossRef](#)]
91. Ji, S.; Li, W.; Baloch, A.R.; Wang, M.; Li, H.; Cao, B.; Zhang, H. Efficient biosynthesis of a Cecropin A-melittin mutant in *Bacillus subtilis* WB700. *Sci. Rep.* **2017**, *7*, 40587. [[CrossRef](#)] [[PubMed](#)]
92. Zweers, J.C.; Barak, I.; Becher, D.; Driessen, A.J.; Hecker, M.; Kontinen, V.P.; Saller, M.J.; Vavrova, L.; van Dijl, J.M. Towards the development of *Bacillus subtilis* as a cell factory for membrane proteins and protein complexes. *Microb. Cell Fact.* **2008**, *7*, 10. [[CrossRef](#)] [[PubMed](#)]
93. Steidler, L.; Hans, W.; Schotte, L.; Neiryneck, S.; Obermeier, F.; Falk, W.; Fiers, W.; Remaut, E. Treatment of murine colitis by *Lactococcus lactis* secreting interleukin-10. *Science* **2000**, *289*, 1352–1355. [[CrossRef](#)] [[PubMed](#)]
94. Mierau, I.; Kleerebezem, M. 10 years of the nisin-controlled gene expression system (NICE) in *Lactococcus lactis*. *Appl. Microbiol. Biotechnol.* **2005**, *68*, 705–717. [[CrossRef](#)]
95. Boumaiza, M.; Colarusso, A.; Parrilli, E.; Garcia-Fruitos, E.; Casillo, A.; Aris, A.; Corsaro, M.M.; Picone, D.; Leone, S.; Tutino, M.L. Getting value from the waste: Recombinant production of a sweet protein by *Lactococcus lactis* grown on cheese whey. *Microb. Cell Fact.* **2018**, *17*, 126. [[CrossRef](#)]
96. Mierau, I.; Olieman, K.; Mond, J.; Smid, E.J. Optimization of the *Lactococcus lactis* nisin-controlled gene expression system NICE for industrial applications. *Microb. Cell Fact.* **2005**, *4*, 16. [[CrossRef](#)]
97. Steen, A.; Wiederhold, E.; Gandhi, T.; Breitling, R.; Slotboom, D.J. Physiological adaptation of the bacterium *Lactococcus lactis* in response to the production of human CFTR. *Mol. Cell Proteom.* **2011**, *10*, M000052MCP200. [[CrossRef](#)]
98. Singh, S.K.; Tiendrebeogo, R.W.; Chourasia, B.K.; Kana, I.H.; Singh, S.; Theisen, M. *Lactococcus lactis* provides an efficient platform for production of disulfide-rich recombinant proteins from *Plasmodium falciparum*. *Microb. Cell Fact.* **2018**, *17*, 55. [[CrossRef](#)]
99. Cano-Garrido, O.; Rueda, F.L.; Sanchez-Garcia, L.; Ruiz-Avila, L.; Bossler, R.; Villaverde, A.; Garcia-Fruitos, E. Expanding the recombinant protein quality in *Lactococcus lactis*. *Microb. Cell Fact.* **2014**, *13*, 167. [[CrossRef](#)]
100. Geertsma, E.R.; Poolman, B. High-throughput cloning and expression in recalcitrant bacteria. *Nat. Methods* **2007**, *4*, 705–707. [[CrossRef](#)]

101. Ventosa, A.; Nieto, J.J.; Oren, A. Biology of moderately halophilic aerobic bacteria. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 504–544. [[CrossRef](#)] [[PubMed](#)]
102. Arakawa, T.; Timasheff, S.N. The stabilization of proteins by osmolytes. *Biophys. J.* **1985**, *47*, 411–414. [[CrossRef](#)]
103. Tokunaga, H.; Arakawa, T.; Tokunaga, M. Novel soluble expression technologies derived from unique properties of halophilic proteins. *Appl. Microbiol. Biotechnol.* **2010**, *88*, 1223–1231. [[CrossRef](#)] [[PubMed](#)]
104. Nagayoshi, C.; Ishibashi, M.; Tokunaga, M. Purification and characterization of human brain serine racemase expressed in moderately halophilic bacteria. *Protein Pept. Lett.* **2009**, *16*, 201–206. [[CrossRef](#)] [[PubMed](#)]
105. Zeldes, B.M.; Keller, M.W.; Loder, A.J.; Straub, C.T.; Adams, M.W.; Kelly, R.M. Extremely thermophilic microorganisms as metabolic engineering platforms for production of fuels and industrial chemicals. *Front. Microbiol.* **2015**, *6*, 1209. [[CrossRef](#)] [[PubMed](#)]
106. Cava, F.; Hidalgo, A.; Berenguer, J. *Thermus thermophilus* as biological model. *Extremophiles* **2009**, *13*, 213–231. [[CrossRef](#)]
107. Velez, A.M.; Horta, A.C.; da Silva, A.J.; Iemma, M.R.; Giordano Rde, L.; Zangirolami, T.C. Enhanced production of recombinant thermo-stable lipase in *Escherichia coli* at high induction temperature. *Protein Expr. Purif.* **2013**, *90*, 96–103. [[CrossRef](#)]
108. Sorensen, H.P.; Sperling-Petersen, H.U.; Mortensen, K.K. Production of recombinant thermostable proteins expressed in *Escherichia coli*: Completion of protein synthesis is the bottleneck. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2003**, *786*, 207–214. [[CrossRef](#)]
109. Albers, S.V.; Jonuscheit, M.; Dinkelaker, S.; Urich, T.; Kletzin, A.; Tampe, R.; Driessen, A.J.; Schleper, C. Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Appl. Environ. Microbiol.* **2006**, *72*, 102–111. [[CrossRef](#)]
110. Drejer, E.B.; Hakvag, S.; Irla, M.; Brautaset, T. Genetic Tools and Techniques for Recombinant Expression in Thermophilic Bacillaceae. *Microorganisms*. **2018**, *6*, 42. [[CrossRef](#)]
111. Khalil, A.S.; Collins, J.J. Synthetic biology: Applications come of age. *Nat. Rev. Genet.* **2010**, *11*, 367–379. [[CrossRef](#)] [[PubMed](#)]
112. Vickers, C.E.; Blank, L.M.; Kromer, J.O. Grand challenge commentary: Chassis cells for industrial biochemical production. *Nat. Chem. Biol.* **2010**, *6*, 875–877. [[CrossRef](#)] [[PubMed](#)]
113. Morimoto, T.; Kadoya, R.; Endo, K.; Tohata, M.; Sawada, K.; Liu, S.; Ozawa, T.; Kodama, T.; Kakeshita, H.; Kageyama, Y.; et al. Enhanced recombinant protein productivity by genome reduction in *Bacillus subtilis*. *DNA Res.* **2008**, *15*, 73–81. [[CrossRef](#)] [[PubMed](#)]
114. Bu, Q.T.; Yu, P.; Wang, J.; Li, Z.Y.; Chen, X.A.; Mao, X.M.; Li, Y.Q. Rational construction of genome-reduced and high-efficient industrial *Streptomyces* chassis based on multiple comparative genomic approaches. *Microb. Cell Fact.* **2019**, *18*, 16. [[CrossRef](#)] [[PubMed](#)]
115. Gibson, D.G.; Benders, G.A.; Andrews-Pfannkoch, C.; Denisova, E.A.; Baden-Tillson, H.; Zaveri, J.; Stockwell, T.B.; Brownley, A.; Thomas, D.W.; Algire, M.A.; et al. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **2008**, *319*, 1215–1220. [[CrossRef](#)] [[PubMed](#)]
116. Hutchison, C.A.; Chuang, R.Y.; Noskov, V.N.; Assad-Garcia, N.; Deerinck, T.J.; Ellisman, M.H.; Gill, J.; Kannan, K.; Karas, B.J.; Ma, L.; et al. Design and synthesis of a minimal bacterial genome. *Science* **2016**, *351*, 6253. [[CrossRef](#)]
117. Fredens, J.; Wang, K.; de la Torre, D.; Funke, L.F.H.; Robertson, W.E.; Christova, Y.; Chia, T.; Schmied, W.H.; Dunkelmann, D.L.; Beranek, V.; et al. Total synthesis of *Escherichia coli* with a recoded genome. *Nature* **2019**, *569*, 514–518. [[CrossRef](#)]
118. Wang, K.; Fredens, J.; Brunner, S.F.; Kim, S.H.; Chia, T.; Chin, J.W. Defining synonymous codon compression schemes by genome recoding. *Nature* **2016**, *539*, 59–64. [[CrossRef](#)]
119. Wang, K.; Sachdeva, A.; Cox, D.J.; Wilf, N.M.; Lang, K.; Wallace, S.; Mehl, R.A.; Chin, J.W. Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nat. Chem.* **2014**, *6*, 393–403. [[CrossRef](#)]
120. Datsenko, K.A.; Wanner, B.L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6640–6645. [[CrossRef](#)]
121. Abremski, K.; Hoess, R.; Sternberg, N. Studies on the properties of P1 site-specific recombination: Evidence for topologically unlinked products following recombination. *Cell* **1983**, *32*, 1301–1311. [[CrossRef](#)]

122. Broach, J.R.; Guarascio, V.R.; Jayaram, M. Recombination within the yeast plasmid 2 μ circle is site-specific. *Cell* **1982**, *29*, 227–234. [[CrossRef](#)]
123. Hashimoto, M.; Ichimura, T.; Mizoguchi, H.; Tanaka, K.; Fujimitsu, K.; Keyamura, K.; Ote, T.; Yamakawa, T.; Yamazaki, Y.; Mori, H.; et al. Cell size and nucleoid organization of engineered *Escherichia coli* cells with a reduced genome. *Mol. Microbiol.* **2005**, *55*, 137–149. [[CrossRef](#)] [[PubMed](#)]
124. Posfai, G.; Plunkett, G.; Feher, T.; Frisch, D.; Keil, G.M.; Umenhoffer, K.; Kolisnychenko, V.; Stahl, B.; Sharma, S.S.; de Arruda, M.; et al. Emergent properties of reduced-genome *Escherichia coli*. *Science* **2006**, *312*, 1044–1046. [[CrossRef](#)] [[PubMed](#)]
125. Mizoguchi, H.; Sawano, Y.; Kato, J.; Mori, H. Superpositioning of deletions promotes growth of *Escherichia coli* with a reduced genome. *DNA Res.* **2008**, *15*, 277–284. [[CrossRef](#)]
126. Yu, B.J.; Sung, B.H.; Koob, M.D.; Lee, C.H.; Lee, J.H.; Lee, W.S.; Kim, M.S.; Kim, S.C. Minimization of the *Escherichia coli* genome using a Tn5-targeted Cre/loxP excision system. *Nat. Biotechnol.* **2002**, *20*, 1018–1023. [[CrossRef](#)]
127. Park, M.K.; Lee, S.H.; Yang, K.S.; Jung, S.C.; Lee, J.H.; Kim, S.C. Enhancing recombinant protein production with an *Escherichia coli* host strain lacking insertion sequences. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 6701–6713. [[CrossRef](#)]
128. Komatsu, M.; Uchiyama, T.; Omura, S.; Cane, D.E.; Ikeda, H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 2646–2651. [[CrossRef](#)]
129. Gomez-Escribano, J.P.; Bibb, M.J. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microb. Biotechnol.* **2011**, *4*, 207–215. [[CrossRef](#)]
130. Zhou, M.; Jing, X.; Xie, P.; Chen, W.; Wang, T.; Xia, H.; Qin, Z. Sequential deletion of all the polyketide synthase and nonribosomal peptide synthetase biosynthetic gene clusters and a 900-kb subtelomeric sequence of the linear chromosome of *Streptomyces coelicolor*. *FEMS Microbiol. Lett.* **2012**, *333*, 169–179. [[CrossRef](#)]
131. Thanapipatsiri, A.; Claesen, J.; Gomez-Escribano, J.P.; Bibb, M.; Thamchaipenet, A. A *Streptomyces coelicolor* host for the heterologous expression of Type III polyketide synthase genes. *Microb. Cell Fact.* **2015**, *14*, 145. [[CrossRef](#)] [[PubMed](#)]
132. Myronovskyi, M.; Rosenkranzer, B.; Nadmid, S.; Pujic, P.; Normand, P.; Luzhetskyy, A. Generation of a cluster-free *Streptomyces albus* chassis strains for improved heterologous expression of secondary metabolite clusters. *Metab. Eng.* **2018**, *49*, 316–324. [[CrossRef](#)] [[PubMed](#)]
133. Liu, Q.; Xiao, L.P.; Zhou, Y.J.; Deng, K.H.; Tan, G.Y.; Han, Y.C.; Liu, X.H.; Deng, Z.X.; Liu, T.G. Development of *Streptomyces* sp. FR-008 as an emerging chassis. *Synth. Syst. Biotechnol.* **2016**, *1*, 207–214. [[CrossRef](#)] [[PubMed](#)]
134. Manabe, K.; Kageyama, Y.; Morimoto, T.; Shimizu, E.; Takahashi, H.; Kanaya, S.; Ara, K.; Ozaki, K.; Ogasawara, N. Improved production of secreted heterologous enzyme in *Bacillus subtilis* strain MGB874 via modification of glutamate metabolism and growth conditions. *Microb. Cell Fact.* **2013**, *12*, 18. [[CrossRef](#)] [[PubMed](#)]
135. Ara, K.; Ozaki, K.; Nakamura, K.; Yamane, K.; Sekiguchi, J.; Ogasawara, N. *Bacillus minimum* genome factory: Effective utilization of microbial genome information. *Biotechnol. Appl. Biochem.* **2007**, *46*, 169–178. [[PubMed](#)]
136. Westers, H.; Dorenbos, R.; van Dijl, J.M.; Kabel, J.; Flanagan, T.; Devine, K.M.; Jude, F.; Seror, S.J.; Beekman, A.C.; Darmon, E.; et al. Genome engineering reveals large dispensable regions in *Bacillus subtilis*. *Mol. Biol. Evol.* **2003**, *20*, 2076–2090. [[CrossRef](#)]
137. Reuss, D.R.; Altenbuchner, J.; Mader, U.; Rath, H.; Ischebeck, T.; Sappa, P.K.; Thurmer, A.; Guerin, C.; Nicolas, P.; Steil, L.; et al. Large-scale reduction of the *Bacillus subtilis* genome: Consequences for the transcriptional network, resource allocation, and metabolism. *Genome Res.* **2017**, *27*, 289–299. [[CrossRef](#)]
138. Koo, B.M.; Kritikos, G.; Farelli, J.D.; Todor, H.; Tong, K.; Kimsey, H.; Wapinski, I.; Galardini, M.; Cabal, A.; Peters, J.M.; et al. Construction and Analysis of Two Genome-Scale Deletion Libraries for *Bacillus subtilis*. *Cell Syst.* **2017**, *4*, 291–305. [[CrossRef](#)]
139. Zhu, D.; Fu, Y.; Liu, F.; Xu, H.; Saris, P.E.; Qiao, M. Enhanced heterologous protein productivity by genome reduction in *Lactococcus lactis* NZ9000. *Microb. Cell Fact.* **2017**, *16*, 1. [[CrossRef](#)]

140. Leprince, A.; de Lorenzo, V.; Voller, P.; van Passel, M.W.; Martins dos Santos, V.A. Random and cyclical deletion of large DNA segments in the genome of *Pseudomonas putida*. *Environ. Microbiol.* **2012**, *14*, 1444–1453. [[CrossRef](#)]
141. Lieder, S.; Nickel, P.I.; de Lorenzo, V.; Takors, R. Genome reduction boosts heterologous gene expression in *Pseudomonas putida*. *Microb. Cell Fact.* **2015**, *14*, 23. [[CrossRef](#)] [[PubMed](#)]
142. Baumgart, M.; Unthan, S.; Ruckert, C.; Sivalingam, J.; Grunberger, A.; Kalinowski, J.; Bott, M.; Noack, S.; Frunzke, J. Construction of a prophage-free variant of *Corynebacterium glutamicum* ATCC 13032 for use as a platform strain for basic research and industrial biotechnology. *Appl. Environ. Microbiol.* **2013**, *79*, 6006–6015. [[CrossRef](#)] [[PubMed](#)]
143. Giga-Hama, Y.; Tohda, H.; Takegawa, K.; Kumagai, H. Schizosaccharomyces pombe minimum genome factory. *Biotechnol. Appl. Biochem.* **2007**, *46*, 147–155. [[PubMed](#)]
144. Kolisnychenko, V.; Plunkett, G.; Herring, C.D.; Feher, T.; Posfai, J.; Blattner, F.R.; Posfai, G. Engineering a reduced *Escherichia coli* genome. *Genome Res.* **2002**, *12*, 640–647. [[CrossRef](#)]
145. Aguilar Suarez, R.; Stulke, J.; van Dijl, J.M. Less Is More: Toward a Genome-Reduced Bacillus Cell Factory for Difficult Proteins. *ACS Synth. Biol.* **2019**, *8*, 99–108. [[CrossRef](#)]
146. Csorgo, B.; Feher, T.; Timar, E.; Blattner, F.R.; Posfai, G. Low-mutation-rate, reduced-genome *Escherichia coli*: An improved host for faithful maintenance of engineered genetic constructs. *Microb. Cell Fact.* **2012**, *11*, 11. [[CrossRef](#)]
147. Lee, J.H.; Sung, B.H.; Kim, M.S.; Blattner, F.R.; Yoon, B.H.; Kim, J.H.; Kim, S.C. Metabolic engineering of a reduced-genome strain of *Escherichia coli* for L-threonine production. *Microb. Cell Fact.* **2009**, *8*, 2. [[CrossRef](#)]
148. Choe, D.; Lee, J.H.; Yoo, M.; Hwang, S.; Sung, B.H.; Cho, S.; Palsson, B.; Kim, S.C.; Cho, B.K. Adaptive laboratory evolution of a genome-reduced *Escherichia coli*. *Nat. Commun.* **2019**, *10*, 935. [[CrossRef](#)]
149. Hirokawa, Y.; Kawano, H.; Tanaka-Masuda, K.; Nakamura, N.; Nakagawa, A.; Ito, M.; Mori, H.; Oshima, T.; Ogasawara, N. Genetic manipulations restored the growth fitness of reduced-genome *Escherichia coli*. *J. Biosci. Bioeng.* **2013**, *116*, 52–58. [[CrossRef](#)]
150. Mushegian, A.R.; Koonin, E.V. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10268–10273. [[CrossRef](#)]
151. Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K.A.; Tomita, M.; Wanner, B.L.; Mori, H. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol. Syst. Biol.* **2006**, *2*. [[CrossRef](#)] [[PubMed](#)]
152. Kobayashi, K.; Ehrlich, S.D.; Albertini, A.; Amati, G.; Andersen, K.K.; Arnaud, M.; Asai, K.; Ashikaga, S.; Aymerich, S.; Bessieres, P.; et al. Essential *Bacillus subtilis* genes. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 4678–4683. [[CrossRef](#)] [[PubMed](#)]
153. van Opijnen, T.; Bodi, K.L.; Camilli, A. Tn-seq: High-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat. Methods* **2009**, *6*, 767–772. [[CrossRef](#)] [[PubMed](#)]
154. Peters, J.M.; Colavin, A.; Shi, H.; Czarny, T.L.; Larson, M.H.; Wong, S.; Hawkins, J.S.; Lu, C.H.S.; Koo, B.M.; Marta, E.; et al. A Comprehensive, CRISPR-based Functional Analysis of Essential Genes in Bacteria. *Cell* **2016**, *165*, 1493–1506. [[CrossRef](#)] [[PubMed](#)]
155. Shlomi, T.; Eisenberg, Y.; Sharan, R.; Ruppin, E. A genome-scale computational study of the interplay between transcriptional regulation and metabolism. *Mol. Syst. Biol.* **2007**, *3*, 101. [[CrossRef](#)]
156. Rees, J.; Chalkley, O.; Landon, S.; Purcell, O.; Marucci, L.; Grierson, C. Designing genomes using whole-cell models. *bioRxiv* **2019**. [[CrossRef](#)]
157. Rugbjerg, P.; Myling-Petersen, N.; Porse, A.; Sarup-Lytzen, K.; Sommer, M.O.A. Diverse genetic error modes constrain large-scale bio-based production. *Nat. Commun.* **2018**, *9*, 787. [[CrossRef](#)]
158. Hoffmann, F.; Rinas, U. Stress induced by recombinant protein production in *Escherichia coli*. *Adv. Biochem. Eng. Biotechnol.* **2004**, *89*, 73–92.
159. Borkowski, O.; Bricio, C.; Murgiano, M.; Rothschild-Mancinelli, B.; Stan, G.B.; Ellis, T. Cell-free prediction of protein expression costs for growing cells. *Nat. Commun.* **2018**, *9*, 1457. [[CrossRef](#)]
160. Thommen, M.; Holtkamp, W.; Rodnina, M.V. Co-translational protein folding: Progress and methods. *Curr. Opin. Struct. Biol.* **2017**, *42*, 83–89. [[CrossRef](#)]
161. Kudla, G.; Murray, A.W.; Tollervey, D.; Plotkin, J.B. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **2009**, *324*, 255–258. [[CrossRef](#)] [[PubMed](#)]

162. Lavner, Y.; Kotlar, D. Codon bias as a factor in regulating expression via translation rate in the human genome. *Gene* **2005**, *345*, 127–138. [[CrossRef](#)] [[PubMed](#)]
163. Collart, M.A.; Weiss, B. Ribosome pausing, a dangerous necessity for co-translational events. *Nucleic Acids Res.* **2019**. [[CrossRef](#)] [[PubMed](#)]
164. Harris, R.P.; Mattocks, J.; Green, P.S.; Moffatt, F.; Kilby, P.M. Determination and control of low-level amino acid misincorporation in human thioredoxin protein produced in a recombinant *Escherichia coli* production system. *Biotechnol. Bioeng.* **2012**, *109*, 1987–1995. [[CrossRef](#)] [[PubMed](#)]
165. Gustafsson, C.; Govindarajan, S.; Minshull, J. Codon bias and heterologous protein expression. *Trends Biotechnol.* **2004**, *22*, 346–353. [[CrossRef](#)]
166. Chung, B.K.; Lee, D.Y. Computational codon optimization of synthetic gene for protein expression. *BMC Syst. Biol.* **2012**, *6*, 134. [[CrossRef](#)]
167. Fu, W.; Lin, J.; Cen, P. 5-Aminolevulinate production with recombinant *Escherichia coli* using a rare codon optimizer host strain. *Appl. Microbiol. Biotechnol.* **2007**, *75*, 777–782. [[CrossRef](#)]
168. Ostrov, N.; Landon, M.; Guell, M.; Kuznetsov, G.; Teramoto, J.; Cervantes, N.; Zhou, M.; Singh, K.; Napolitano, M.G.; Moosburner, M.; et al. Design, synthesis, and testing toward a 57-codon genome. *Science* **2016**, *353*, 819–822. [[CrossRef](#)]
169. McManus, C.J.; May, G.E.; Spealman, P.; Shteyman, A. Ribosome profiling reveals post-transcriptional buffering of divergent gene expression in yeast. *Genome Res.* **2014**, *24*, 422–430. [[CrossRef](#)]
170. Jeong, Y.; Kim, J.N.; Kim, M.W.; Bucca, G.; Cho, S.; Yoon, Y.J.; Kim, B.G.; Roe, J.H.; Kim, S.C.; Smith, C.P.; et al. The dynamic transcriptional and translational landscape of the model antibiotic producer *Streptomyces coelicolor* A3(2). *Nat. Commun.* **2016**, *7*, 11605. [[CrossRef](#)]
171. Ebrahim, A.; Brunk, E.; Tan, J.; O'Brien, E.J.; Kim, D.; Szubin, R.; Lerman, J.A.; Lechner, A.; Sastry, A.; Bordbar, A.; et al. Multi-omic data integration enables discovery of hidden biological regularities. *Nat. Commun.* **2016**, *7*, 13091. [[CrossRef](#)] [[PubMed](#)]
172. Morgan, G.J.; Burkhardt, D.H.; Kelly, J.W.; Powers, E.T. Translation efficiency is maintained at elevated temperature in *Escherichia coli*. *J. Biol. Chem.* **2018**, *293*, 777–793. [[CrossRef](#)] [[PubMed](#)]
173. Niess, A.; Siemann-Herzberg, M.; Takors, R. Protein production in *Escherichia coli* is guided by the trade-off between intracellular substrate availability and energy cost. *Microb. Cell Fact.* **2019**, *18*, 8. [[CrossRef](#)] [[PubMed](#)]
174. Kustatscher, G.; Grabowski, P.; Rappsilber, J. Pervasive coexpression of spatially proximal genes is buffered at the protein level. *Mol. Syst. Biol.* **2017**, *13*, 937. [[CrossRef](#)] [[PubMed](#)]

