CrossMark

MINI-REVIEW

Tools to cope with difficult-to-express proteins

Paolo Saccardo[1,2] · José Luís Corchero[1,2,3] · Neus Ferrer-Miralles[1,2,3]

Abstract The identification of DNA coding sequences contained in the genome of many organisms coupled to the use of high throughput approaches has fueled the field of recombinant protein production. Apart from basic research interests, the growing relevance of this field is highlighted by the global sales of the top ten biopharmaceuticals on the market, which exceeds the trillion USD in a steady increasing tendency. Therefore, the demand of biological compounds seems to have a long run on the market. One of the most popular expression systems is based on *Escherichia coli* cells which apart from being cost-effective counts with a large selection of resources. However, a significant percentage of the genes of interest are not efficiently expressed in this system, or the expressed proteins are accumulated within aggregates, degraded or lacking the desired biological activity, being finally discarded. In some instances, expressing the gene in a homologous expression system might alleviate those drawbacks but then the process usually increases in complexity and is not as cost-effective as the prokaryotic systems. An increasing toolbox is available to approach the production and purification of those difficult-to-express proteins, including different expression systems, promoters with different strengths, cultivation media and conditions, solubilization tags and chaperone coexpression, among others. However, in most cases, the process follows a non-integrative trial and error strategy with discrete success. This review is focused on the design of the whole process by using an integrative approach, taken into account the accumulated knowledge of the pivotal factors that affect any of the key processes, in an attempt to rationalize the efforts made in this appealing field.

Introduction

The expression of genes in heterologous expression systems has made available large amounts of recombinant proteins that are limited in their natural sources (Ferrer-Miralles et al. 2009). In addition to being used in functional, structural, and biochemical studies, recombinant proteins are being formulated to be administered to patients in an increasing number of medical conditions and even produced at an industrial scale (Lingg et al. 2012; Reichert and Paquette 2003). Recombinant DNA technology is not only able to produce naturally occurring proteins, but it can also provide with an almost infinite plethora of variants of such natural proteins by mutations refining interesting features (like pharmacokinetics, targeting, etc.) while maintaining their biological function (Huang et al. 2008; Kamionka 2011). One case that exemplifies the great potential of this technology is human insulin. Insulin was the first recombinant product launched to the market in 1982. Since then, some mutant insulin have appeared with relevant implications in its time span of action improving the quality of life of insulin-dependent patients (Ferrer-Miralles et al. 2009; Kamionka 2011; Toyoda et al. 2012). Moreover, recombinant DNA technology allows the production of de novo designed proteins, not present in nature but resulting from the fusion of different active peptides or

✉ Neus Ferrer-Miralles
neus.ferrer@uab.cat

1   Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

2   CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Bellaterra, 08193 Barcelona, Spain

3   Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Bellaterra, 08193 Barcelona, Spain

⌂ Springer

domains (each one from a different origin) in single-chain poly-peptides with a modular architecture. This offers intriguing possibilities for the development of multifunctional and smart drug vehicles at the nanoscale range (Vazquez et al. 2009).

One of the most critical decision points when facing gene expression in a heterologous system is the cost of the procedure. There is a huge gap between the derived costs from using prokaryotic or low eukaryotic expression systems from those using a complex eukaryotic expression system. The difference is related not only to the cost of the culture medium used in the process but also to the process itself that in the case of complex eukaryotic expression systems might include sophisticated infrastructures to maintain cell growth. At the end, most production procedures are designed to be performed in a prokaryotic system and, among them, the most used is based on *Escherichia coli* strains. This is easily illustrated when consulting the origin of the recombinant proteins in databases such as the Protein Data Base (PDB): even though more than 50 % of the genes have a mammalian origin, only 3 % are expressed in mammalian expression systems, being *E. coli* used instead in more than 85 % of the cases. This data indicates the great potential of *E. coli* as a robust and reliable expression system. However, *E. coli* cells are not always able to render functional proteins (especially for complex eukaryotic proteins including post-translational modifications, or hydrophobic patches as in the case of membrane proteins with transmembrane domains) or even production is impeded (Freigassner et al. 2009). In those cases, production of the recombinant protein can be addressed using other prokaryotic expression systems such as psychrophilic bacteria (Unzueta et al. 2015). It is quite difficult to give exact data about the percentage of proteins that have failed to be expressed in prokaryotic hosts, since those results are not usually published and probably only the tip of the iceberg is available to the scientific community. This review is intended to give key clues to approach the expression of heterologous genes specifically for difficult-to-express (DTE) recombinant proteins in prokaryotic expression systems.

Where to start: analysis of properties of recombinant proteins with bioinformatic tools

In the design of the gene expression procedure, information about biological and physicochemical details from online bioinformatics resources is obtained just with the primary sequence of the protein (UNIPROT, EXPASY, see Fig. 1). Apart from those general parameters, a second line of important features are obtained when running programs to predict protein secondary structure and the presence of protein domains at Expasy (PROSITE, JPRED, CFSSP). Then, the tertiary structure can be either predicted (SWISS MODEL) or obtained if already deposited at the RSCB Protein Data Bank

(PDB). At that point, the specific posttranslational modifications as glycosylations or disulfide bridges are revealed and the protein engineering design can be approached. In addition, computational modeling can be applied to increase protein stability by specific point mutations, as it has been demonstrated modifying solvent-exposed domains of complex proteins by using modeling programs. (Kim et al. 2016).

Revisiting the recombinant gene

When a DTE protein is suspected or detected as an aggregation-prone protein, one of the most used strategies to overcome such problem is to modify the recombinant gene adding a solubilization tag, with the aim of obtaining the protein in its soluble form (Sun et al. 2011) (Fig. 1). However, the transfer of the insoluble protein to the soluble cell fraction does not guarantee the improvement in its conformational quality (Unzueta et al. 2015). In fact, it has been described that solubility is not a coincident parameter with conformational quality (Martinez-Alonso et al. 2008). Therefore, this strategy might not render the desired product at least when it comes to quality. An alternative to that strategy would be the refolding from IBs or its resolubilization under mild conditions (Singh and Panda 2005; Vallejo and Rinas 2004). Unfortunately both strategies use to give low efficiency in protein recovery.

One of the major problems observed in recombinant protein production is the difficulty of cell factories to manage the rare codons. As a consequence of the low availability of certain transfer RNA (tRNA) molecules, their associated codons are slowly translated. In recombinant protein production, the expression rates, yield, and final product quality can then be affected by the codon pattern of the gene of interest and by the codon bias of the expression system (Quax et al. 2015). If the synthesis is dramatically slowed, ribosomes could release prematurely the uncompleted polypeptide (Shah and Gilchrist 2010). In prokaryotes, the newly synthesized proteins start to fold in the ribosomal tunnel, protected by the ribosomal structure environment (O'Brien et al. 2014). This cotranslational folding is influenced by the speed of polypeptide chain elongation and rare codons play a key role in the protein synthesis speed (Varenne et al. 1984). Codon usage then is another important parameter that should be taken into account for recombinant protein production.

Frequent and rare triplets are mixed within the same expression sequences and the analysis of elongation speed gives a codon landscape graph, which is typical for each gene. Because of this, the analysis and the optimization of the codon usage of the gene of interest and adjustment of the tRNA's abundancy in cell factories are key steps to follow in order to overcome troubleshooting during the production process. In some cases, this strategy represents a significant increase in the production yield (Gustafsson et al. 2012) but in some
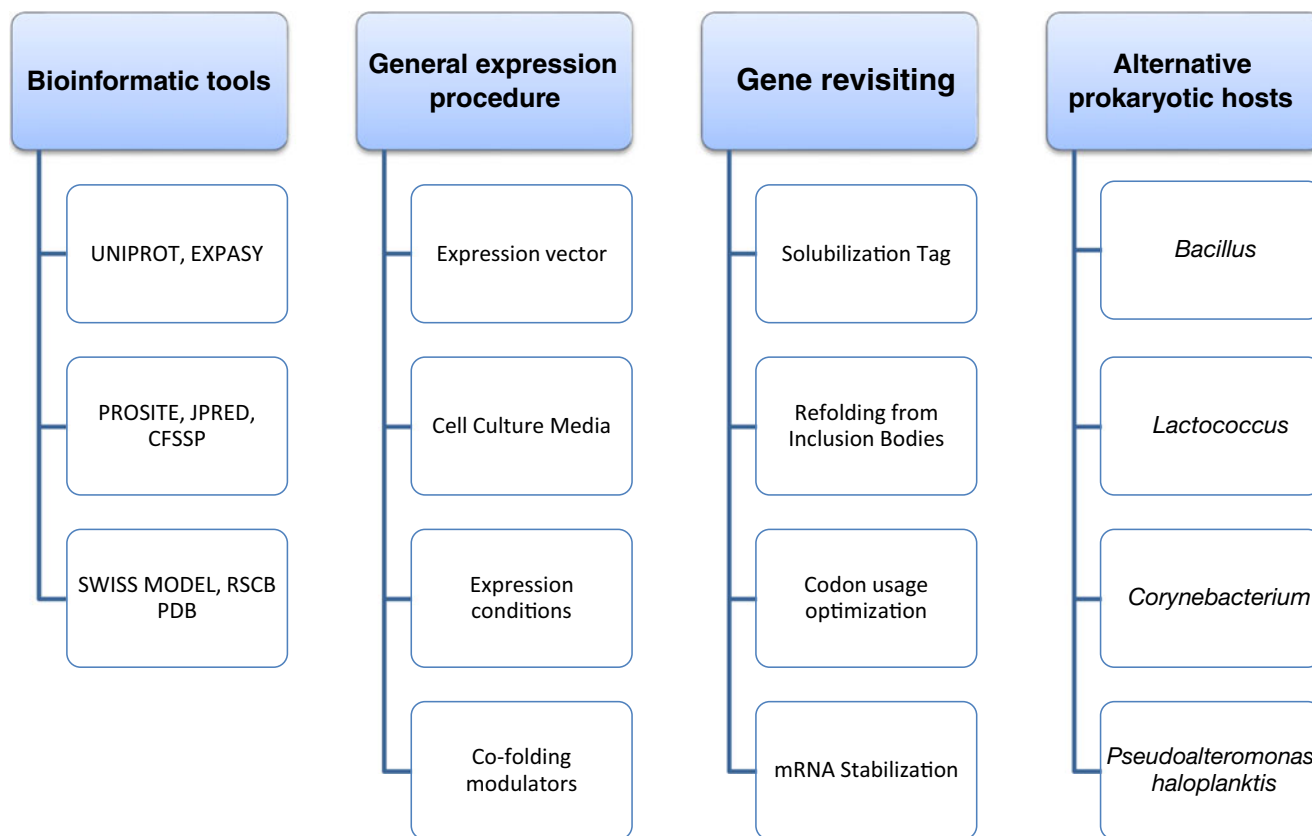
Bioinformatic tools	General expression procedure	Gene revisiting	Alternative prokaryotic hosts
UNIPROT, EXPASY	Expression vector	Solubilization Tag	*Bacillus*
PROSITE, JPRED, CFSSP	Cell Culture Media	Refolding from Inclusion Bodies	*Lactococcus*
SWISS MODEL, RSCB PDB	Expression conditions	Codon usage optimization	*Corynebacterium*
	Co-folding modulators	mRNA Stabilization	*Pseudoalteromonas haloplanktis*

Fig. 1 Alternative strategies to cope with difficult-to-express proteins. Recommended bioinformatics analysis as well as changes in the expression procedure and recombinant gene and alternative prokaryotic hosts are shown

instances this effect was not achieved (Gustafsson et al. 2012; Maertens et al. 2010).

As it is widely accepted, optimizing codon usage of the gene of interest in the chosen heterologous expression systems is a way to optimize its translation rate by exploiting the natural ratios in tRNAs content in the expression system (Fig. 1). Among the 64 combinations of triplets, the 20 standard amino acids are coded by 61 codons; meanwhile, three are translation stop signals. This redundancy means that more than one triplet can be used as "synonym" for the same amino acid. Depending on the organism, different biased frequency of synonym triplet is observed. Moreover, the tRNA species have variation in the expression ratio; consequently, synonymous codons are not equivalently represented in tRNA population (Ikemura 1985).

Codon optimization is then not referred as changing all synonymous triplets to the one that is prevalently represented in tRNA battery. If this happens, the codon landscape would be altered, compromising the rhythm of protein synthesis. This speed alteration could led to an incorrect folding of the recombinant protein, and the consequently degradation of the misfolded protein or accumulation in the insoluble fraction (Quax et al. 2015). In fact, it has been described that the elongation of nascent messenger RNA (mRNA) is a non-uniform process and speed seems to be linked to the need to modulate the time required to allow the correct folding of domains in multidomain proteins. On the one hand, optimizing the codon usage would mean to increase the speed of the elongation through all the sequence, but on the other hand promoting then the chances to obtain partially folded species that would be either degraded or sent to the inclusion bodies. Novel algorithms to harmonize the translation pattern of the gene with the one of the expression host are proposed (Bartholomaus et al. 2016; Hess et al. 2015).

Designing the recombinant gene expression procedure

In most of the cases, the recombinant genes are expressed in *E. coli* under standard conditions, described elsewhere (Graslund et al. 2008; Rosano and Ceccarelli 2014). In general, once the primary sequence is decided, several parameters are set at the initial steps for the expression experiments as for example the host strain, the promoter and media, among others (Gopal and Kumar 2013). However, in the case of a DTE protein, those basic procedures are not enough to render either the required amount or quality of the recombinant protein or the protein is not even produced. It is then when a series of trial and error proofs is performed. Many factors seem to be

involved in this phenomenon. The saturation of the folding machinery (Martinez-Alonso et al. 2010), the RNA stability (Ahn et al. 2008), and the codon usage are among the most relevant (Rosano and Ceccarelli 2014). Therefore, when trying to approach the production of heterologous proteins in prokaryotic expression systems, a specific gene expression design has to be developed.

As many factors might be beneath the low production or the aggregation of the recombinant protein, a multifactorial approach has to be studied. The parameters to be taken into account include the change of the expression vector mainly to get a weaker promoter in case of toxic proteins (Lebendiker and Danieli 2014), the expression strain able to fulfill some specific post-translational modifications or able to supplement tRNA for rare codons (Tegel et al. 2010), the growth media and induction conditions (Marini et al. 2014), and the coexpression of folding modulators (Martinez-Alonso et al. 2010) (Fig. 1). In this sense, high-throughput expression screening and purification protocols can be established to include several variables at a time, increasing the chances to obtain the recombinant protein in the soluble cell fraction. (Nozach et al. 2013; Saez et al. 2014; Saez and Vincentelli 2014).

Emerging microbial systems as cell factories for "difficult-to-express" proteins

Taking into consideration the abovementioned obstacles when facing the production of DTE proteins, there is a clear consensus in the urgent need of new cell factories able to overcome these issues. Apart from sticking to current systems pursuing the development of improved strains by adding new properties to old systems (what in any case is an interesting and useful approach) (Makino et al. 2011), the wide diversity of the microbial forms of life and metabolism offers an almost unlimited source to explore new microorganisms. Their specific properties as unconventional cell factories offer intriguing opportunities to deal with the increasing demand of more complex DTE proteins. In this line, the prokaryotic universe is under continuous exploration in the search of new protein cell factories (Fig. 1). Among those systems, some of the most promising under current study are described in the following sections.

Bacillus

Bacillus species (a non-pathogenic, toxin-free Gram-positive bacterium) have a recognized history of safe use in food industry. The use of *Bacillus subtilis* has many advantages, such as the "generally recognized as safe" (GRAS) status and the easy and inexpensive culture methods, that can result in high cell densities. These characteristics and significant progress in

genetic manipulation pushes *Bacillus* as a promising production platform (Schallmey et al. 2004; Wenzel et al. 2011), and therefore, a variety of expression systems have been developed for the efficient expression of heterologous genes in *B. subtilis* (Schumann 2007).

Among recent advances in *B. subtilis* tools development, new expression systems have been established. For example, the T7 expression system, a highly efficient system widely used in *E. coli* (Chen et al. 2010), or the IPTG-inducible strong promoter $P_{grac}100$ (Phan et al. 2012) have been introduced in *B. subtilis.* Since IPTG-inducible promoters are not suitable for industrial-scale production (IPTG is expensive and toxic), maltose-inducible (Ming et al. 2010) and mannitol-inducible (Heravi et al. 2011) expression systems, based on cheaper and safer inducers, have been also developed for *Bacillus*. *B. subtilis* is also known for its capacity to secrete large amounts of endogenous enzymes, like proteases and lipases, many of them showing industrially relevant applications. However, when dealing with heterologous proteins, secretion usually shows a substantially lower efficiency because of the lack of specific secretion pathways, insufficient protein-conducting channels, or secretion stress response. To overcome these limitations, recent research has focused on engineering secretion pathways, like the Sec-SRP pathway (Diao et al. 2012; Kakeshita et al. 2010) or the twin-arginine translocation (Tat) pathway (Zhu et al. 2008), in order to avoid rate-limiting factors and enhance recombinant protein secretion. Many other tools for genetic modification of *Bacillus* species have been developed in recent years (Dong and Zhang 2014).

All these advances have helped to realize the potential of *Bacillus* species as production hosts, making these microorganisms competitive with traditional cell factories like *E. coli*.

Lactococcus

Lactococcus lactis, a Gram-positive bacterium from the lactic acid bacteria (LAB) group and also with GRAS status, is becoming during the last decades another attractive candidate as a microbial cell factory for heterologous protein production. The huge potential of these GRAS microorganisms and the continuous development of an increasing genomic and proteomic toolbox are expected to convert *Lactococcus* into a promising cell factory for a wide range of applications (de Vos 2011; Teusink et al. 2011).

Within this available genetic toolbox, the nisin-inducible (NICE) expression system (de Ruyter et al. 1996; Kuipers et al. 1995; Teusink et al. 2011), from the *nis* operon present in some *L. lactis* strains, has been extensively used to produce recombinant proteins in *L. lactis* (Kunji et al. 2003; Le et al. 2005; Mierau and Kleerebezem 2005; Teusink et al. 2011). The NICE system offers numerous advantages like easy use, tightly controlled, and efficiently induced expression leading

to high protein yields (Mierau and Kleerebezem 2005; Teusink et al. 2011), and large-scale production process (Mierau et al. 2005). Other gene expression systems, like P170 (Jorgensen et al. 2014) or zinc systems (Llull and Poquet 2004), have been successfully used in *L. lactis*. Interestingly, *L. lactis* is also currently used as a life vector for drug, DNA, and other molecule delivery to mucosal surfaces (Braat et al. 2006; Pontes et al. 2011).

Studies on membrane proteins, a typical example of DTE proteins, have benefited from the introduction of *L. lactis* as a cell factory. Heterologous overproduction of membrane proteins, particularly eukaryotic ones, remains challenging, and even when a high-yield expression is achieved, membrane proteins are usually found in the form of inclusion bodies (Monne et al. 2005). In this line, several membrane proteins, including transporting proteins, receptors, and a quinone oxidoreductase, have been successfully overexpressed in lactococcal hosts (Kunji et al. 2003; Marreddy et al. 2011; Wieczorek and Martin 2010). Thus, this expression system can be used for the functional characterization of eukaryotic membrane proteins.

Moreover, a recent study showed that recombinant protein quality in *L. lactis* is deeply influenced by growth conditions and temperature (Cano-Garrido et al. 2014). By controlling these parameters, conformational quality of both soluble and insoluble protein fractions can be modulated, offering a high versatility of this system for the production of endotoxin-free soluble proteins of biomedical interest. These findings not only confirm *L. lactis* as an excellent producer of recombinant proteins but also reveal that there is still plenty of room for significant improvement by the exploitation of external protein quality modulators.

Corynebacterium

Corynebacterium glutamicum is another Gram-positive, non-pathogenic bacterium, traditionally used for the industrial production of L-amino acids, nucleic acids, antibiotics, and other biochemicals (Kimura 2003). *C. glutamicum* is also a promising cell factory for the production of recombinant proteins due to several features: (1) it has a single cellular membrane that allows direct secretion of target proteins into the culture medium; (2) it secretes small amounts of endogenous proteins into the culture medium, simplifying downstream processes; (3) extracellular hydrolytic enzyme activity is not detectable, which gives strong potential for expressing protease-sensitive proteins; and (4) it is an endotoxin-free, GRAS microorganism. Moreover, fermentation conditions for this host are well-known and established for large-scale production of various biomolecules, including recombinant proteins.

However, compared with *E. coli*, *C. glutamicum* showed some disadvantages, like lower transformation efficiency and only a few available expression vectors (Yim et al. 2014). To address these issues, a considerable effort has been put into genetic modification of *C. glutamicum*. Therefore, several techniques and vector components for genetic manipulation of *C. glutamicum* have been developed during the last years, including strong promoters for precise regulation of gene expression, various types of plasmid vectors, secretion systems and methods of genetically modifying the host strain genome (Nesvera and Patek 2011).

Thanks to all these advances, a variety of recombinant proteins have been successfully expressed in *C. glutamicum*, such as ovine gamma interferon (Billman-Jacobe et al. 1994), the human epidermal growth factor (Date et al. 2006), green fluorescent protein (Meissner et al. 2007), α-amylase (Tateno et al. 2007), or antibody fragments (Yim et al. 2014). In addition, a *C. glutamicum*-based protein expression system, with intellectual property rights, named CORYNEX has been developed by Ajinomoto Co., Inc. This system can secrete recombinant proteins into the medium, simplifying purification processes (http://www.ajinomoto.co.jp/corynex/en/characteristic/).

Pseudoalteromonas haloplanktis

Protein aggregation, a process mainly driven by sterospecific interactions between solvent-exposed hydrophobic patches (Carrio et al. 2005; Speed et al. 1996), has been traditionally seen as one of the main bottlenecks in recombinant protein production. Since interactions leading to aggregation are weakened when temperature decreases, production of recombinant proteins in psychrophilic bacteria (microorganisms cultured at 4 °C or below) represents a promising model to improve protein quality and/or solubility. In this context, a few cold-adapted bacterial species are under intense exploration as cell factories, with *Pseudoalteromonas haloplanktis* TAC125 as a representative example. *P. haloplanktis* TAC125 is a Gram-negative bacterium isolated from an Antarctic coastal seawater sample (Birolo et al. 2000), being able to grow in the range of 0–30 °C (Tutino et al. 2001) and even at lower temperatures, making it one of the faster growing psychrophiles so far characterized, and an attractive host as cell factory.

P. haloplanktis TAC125 versatility and uses as a cell factory have been expanded by the development of genetically engineered strains with improved features (Cusano et al. 2006; Parrilli et al. 2010). *P. haloplanktis* TAC125 is also the first Antarctic bacterium in which an efficient gene-expression technology was set up, by the proper assembly of psychrophilic molecular signals (Duilio et al. 2004; Tutino et al. 2001) into a modified *E. coli* cloning vector (Tutino et al. 2002). Several generations of cold-adapted expression vectors allow protein production either by constitutive (Duilio et al. 2004) or inducible regimes (Papa et al. 2007), and to address the product towards any cell compartment or to the extra-cellular medium (Parrilli et al. 2008).

Interestingly, insoluble aggregates of recombinant proteins have never been observed in *P. haloplanktis* TAC125, even at high expression levels, suggesting that its cellular physico-chemical conditions and/or folding processes are quite different from those observed in mesophilic bacteria (Piette et al. 2010). Beneficial effects in using this cold-adapted platform versus conventional *E. coli* have been seen during the production of antibody fragments (Dragosits et al. 2011; Giuliani et al. 2014), or in the production of DTE proteins such as the human nerve growth factor, h-NGF (Vigentini et al. 2006) or the alpha-glucosidase from *Saccharomyces cerevisiae*. When produced in *E. coli*, these two proteins failed to correctly fold and accumulated into insoluble aggregates, but their production in *P. haloplanktis* TAC125 results in fully soluble proteins.

Recently, a synthetic, optimized medium for *P. haloplanktis* TAC125 chemostat cultivation was introduced (Giuliani et al. 2014), and a fed-batch fermentation strategy, feasible for industrial purposes, has been established (Wilmes et al. 2010). These achievements will surely launch the industrial application of *P. haloplanktis* TAC125 as a new, non-conventional system for DTE protein production.

The choice of the final expression system is not an easy or straightforward decision, and it usually will need customized answers. There are no "magical recipes", and consequently, one cannot expect to find a universal production protocol, but rather should be aware of these different available options to improve the final results. In this context, a "function-driven" or "function-optimized" strategy is usually taken. For example, therapeutic proteins (requiring higher standards of quality, safety, etc.) are mainly produced in mammalian cells, but the appearance of GRAS prokaryotic systems can easily change this rule in the future. This has been nicely exemplified by the case of therapeutic antibodies, where the efforts in their production are moving from the obtention of generic "naked" antibodies to application-adapted immune-reagents that do not need further in vitro modification before their use (de Marco, 2015).

Apart from the abovementioned emerging prokaryotic expression systems, eukaryotic cells are becoming the preferred cell factories for proteins needing specific modifications, like glycosilations or disulfide bridges. These eukaryotic systems (beyond the scope of this review), together with the abovementioned emerging prokaryotic systems, are fueling a progressive diversification concerning available expression platforms.

Conclusions

E. coli is a robust microbial cell factory for the production of recombinant proteins. Even in the case of difficult-to-express proteins this expression system provides valuable tools as well as novel strategies to explore in a trial-and-error approach the improvement in protein production and protein solubility. In addition, development of other microbial cell factories is

making progress towards their use as effective protein production platforms. Since the number of approved therapeutic proteins is steadily growing, prokaryotic expression systems would continue to play a significant role as cost-effective, tunable, and easy to operate source of recombinant protein.

Compliance with ethical standards This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest The authors declare that they have no conflict of interest.

References

Ahn JH, Kang TJ, Kim DM (2008) Tuning the expression level of recombinant proteins by modulating mRNA stability in a cell-free protein synthesis system. Biotechnol Bioeng 101:422–427

Bartholomaus A, Del CC, Ignatova Z (2016) Mapping the non-standardized biases of ribosome profiling. Biol Chem 397:23–35

Billman-Jacobe H, Hodgson AL, Lightowlers M, Wood PR, Radford AJ (1994) Expression of ovine gamma interferon in *Escherichia coli* and *Corynebacterium glutamicum*. Appl Environ Microbiol 60: 1641–1645

Birolo L, Tutino ML, Fontanella B, Gerday C, Mainolfi K, Pascarella S, Sannia G, Vinci F, Marino G (2000) Aspartate aminotransferase from the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC 125. Cloning, expression, properties, and molecular modelling. Eur J Biochem 267:2790–2802

Braat H, Rottiers P, Hommes DW, Huyghebaert N, Remaut E, Remon JP, van Deventer SJ, Neirynck S, Peppelenbosch MP, Steidler L (2006) A phase I trial with transgenic bacteria expressing interleukin-10 in Crohn's disease. Clin Gastroenterol Hepatol 4:754–759

Cano-Garrido O, Rueda FL, Sanchez-Garcia L, Ruiz-Avila L, Bosser R, Villaverde A, Garcia-Fruitos E (2014) Expanding the recombinant protein quality in *Lactococcus lactis*. Microb Cell Factories 13:167

Carrio M, Gonzalez-Montalban N, Vera A, Villaverde A, Ventura S (2005) Amyloid-like properties of bacterial inclusion bodies. J Mol Biol 347:1025–1037

Chen PT, Shaw JF, Chao YP, David Ho TH, Yu SM (2010) Construction of chromosomally located T7 expression system for production of heterologous secreted proteins in *Bacillus subtilis*. J Agric Food Chem 58:5392–5399

Cusano AM, Parrilli E, Marino G, Tutino ML (2006) A novel genetic system for recombinant protein secretion in the Antarctic *Pseudoalteromonas haloplanktis* TAC125. Microb Cell Factories 5:40

Date M, Itaya H, Matsui H, Kikuchi Y (2006) Secretion of human epidermal growth factor by *Corynebacterium glutamicum*. Lett Appl Microbiol 42:66–70

de Marco A (2015) Recombinant antibody production evolves into multiple options aimed at yielding reagents suitable for application-specific needs. Microb Cell Factories 14:125

de Ruyter PG, Kuipers OP, Beerthuyzen MM, van Alen-Boerrigter I, de Vos WM (1996) Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*. J Bacteriol 178:3434–3439

de Vos WM (2011) Systems solutions by lactic acid bacteria: from paradigms to practice. Microb Cell Fact 10(Suppl 1):S2

Diao L, Dong Q, Xu Z, Yang S, Zhou J, Freudl R (2012) Functional implementation of the posttranslational SecB-SecA protein-targeting pathway in *Bacillus subtilis*. Appl Environ Microbiol 78: 651–659

Dong H, Zhang D (2014) Current development in genetic engineering strategies of *Bacillus* species. Microb Cell Factories 13:63

Dragosits M, Frascotti G, Bernard-Granger L, Vazquez F, Giuliani M, Baumann K, Rodriguez-Carmona E, Tokkanen J, Parrilli E, Wiebe MG, Kunert R, Maurer M, Gasser B, Sauer M, Branduardi P, Pakula T, Saloheimo M, Penttila M, Ferrer P, Luisa TM, Villaverde A, Porro D, Mattanovich D (2011) Influence of growth temperature on the production of antibody fab fragments in different microbes: a host comparative analysis. Biotechnol Prog 27:38–46

Duilio A, Madonna S, Tutino ML, Pirozzi M, Sannia G, Marino G (2004) Promoters from a cold-adapted bacterium: definition of a consensus motif and molecular characterization of UP regulative elements. Extremophiles 8:125–132

Ferrer-Miralles N, Domingo-Espin J, Corchero JL, Vazquez E, Villaverde A (2009) Microbial factories for recombinant pharmaceuticals. Microb Cell Factories 8:17

Freigassner M, Pichler H, Glieder A (2009) Tuning microbial hosts for membrane protein production. Microb Cell Factories 8:69

Giuliani M, Parrilli E, Sannino F, Apuzzo GA, Marino G, Tutino ML (2014) Recombinant production of a single-chain antibody fragment in *Pseudoalteromonas haloplanktis* TAC125. Appl Microbiol Biotechnol 98:4887–4895

Gopal GJ, Kumar A (2013) Strategies for the production of recombinant protein in *Escherichia coli*. Protein J 32:419–425

Graslund S, Nordlund P, Weigelt J, Hallberg BM, Bray J, Gileadi O, Knapp S, Oppermann U, Arrowsmith C, Hui R, Ming J, dhe-Paganon S, Park HW, Savchenko A, Yee A, Edwards A, Vincentelli R, Cambillau C, Kim R, Kim SH, Rao Z, Shi Y, Terwilliger TC, Kim CY, Hung LW, Waldo GS, Peleg Y, Albeck S, Unger T, Dym O, Prilusky J, Sussman JL, Stevens RC, Lesley SA, Wilson IA, Joachimiak A, Collart F, Dementieva I, Donnelly MI, Eschenfeldt WH, Kim Y, Stols L, Wu R, Zhou M, Burley SK, Emtage JS, Sauder JM, Thompson D, Bain K, Luz J, Gheyi T, Zhang F, Atwell S, Almo SC, Bonanno JB, Fiser A, Swaminathan S, Studier FW, Chance MR, Sali A, Acton TB, Xiao R, Zhao L, Ma LC, Hunt JF, Tong L, Cunningham K, Inouye M, Anderson S, Janjua H, Shastry R, Ho CK, Wang D, Wang H, Jiang M, Montelione GT, Stuart DI, Owens RJ, Daenke S, Schutz A, Heinemann U, Yokoyama S, Bussow K, Gunsalus KC (2008) Protein production and purification. Nat Methods 5:135–146

Gustafsson C, Minshull J, Govindarajan S, Ness J, Villalobos A, Welch M (2012) Engineering genes for predictable protein expression. Protein Expr Purif 83:37–46

Heravi KM, Wenzel M, Altenbuchner J (2011) Regulation of mtl operon promoter of *Bacillus subtilis*: requirements of its use in expression vectors. Microb Cell Factories 10:83

Hess AK, Saffert P, Liebeton K, Ignatova Z (2015) Optimization of translation profiles enhances protein expression and solubility. PLoS One 10:e0127039

Huang YJ, Lundy PM, Lazaris A, Huang Y, Baldassarre H, Wang B, Turcotte C, Cote M, Bellemare A, Bilodeau AS, Brouillard S, Touati M, Herskovits P, Begin I, Neveu N, Brochu E, Pierson J, Hockley DK, Cerasoli DM, Lenz DE, Wilgus H, Karatzas CN, Langermann S (2008) Substantially improved pharmacokinetics of recombinant human butyrylcholinesterase by fusion to human serum albumin. BMC Biotechnol 8:50

Ikemura T (1985) Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol 2:13–34

Jorgensen CM, Vrang A, Madsen SM (2014) Recombinant protein expression in *Lactococcus lactis* using the P170 expression system. FEMS Microbiol Lett 351:170–178

Kakeshita H, Kageyama Y, Ara K, Ozaki K, Nakamura K (2010) Enhanced extracellular production of heterologous proteins in *Bacillus subtilis* by deleting the C-terminal region of the SecA secretory machinery. Mol Biotechnol 46:250–257

Kamionka M (2011) Engineering of therapeutic proteins production in *Escherichia coli*. Curr Pharm Biotechnol 12:268–274

Kim DN, Jacobs TM, Kuhlman B (2016) Boosting protein stability with the computational design of beta-sheet surfaces. Protein Sci 25:702–710

Kimura E (2003) Metabolic engineering of glutamate production. Adv Biochem Eng Biotechnol 79:37–57

Kuipers OP, Rollema HS, Siezen RJ, de Vos WM (1995) Lactococcal expression systems for protein engineering of nisin. Dev Biol Stand 85:605–613

Kunji ER, Slotboom DJ, Poolman B (2003) *Lactococcus lactis* as host for overproduction of functional membrane proteins. Biochim Biophys Acta 1610:97–108

Le LY, Azevedo V, Oliveira SC, Freitas DA, Miyoshi A, Bermudez-Humaran LG, Nouaille S, Ribeiro LA, Leclercq S, Gabriel JE, Guimaraes VD, Oliveira MN, Charlier C, Gautier M, Langella P (2005) Protein secretion in *Lactococcus lactis*: an efficient way to increase the overall heterologous protein production. Microb Cell Factories 4:2

Lebendiker M, Danieli T (2014) Production of prone-to-aggregate proteins. FEBS Lett 588:236–246

Lingg N, Zhang P, Song Z, Bardor M (2012) The sweet tooth of biopharmaceuticals: importance of recombinant protein glycosylation analysis. Biotechnol J 7:1462–1472

Llull D, Poquet I (2004) New expression system tightly controlled by zinc availability in *Lactococcus lactis*. Appl Environ Microbiol 70:5398–5406

Maertens B, Spriestersbach A, von Groll U, Roth U, Kubicek J, Gerrits M, Graf M, Liss M, Daubert D, Wagner R, Schafer F (2010) Gene optimization mechanisms: a multi-gene study reveals a high success rate of full-length human proteins expressed in *Escherichia coli*. Protein Sci 19:1312–1326

Makino T, Skretas G, Georgiou G (2011) Strain engineering for improved expression of recombinant proteins in bacteria. Microb Cell Factories 10:32

Marini G, Luchese MD, Argondizzo AP, de Goes AC, Galler R, Alves TL, Medeiros MA, Larentis AL (2014) Experimental design approach in recombinant protein expression: determining medium composition and induction conditions for expression of pneumolysin from *Streptococcus pneumoniae* in *Escherichia coli* and preliminary purification process. BMC Biotechnol 14:1

Marreddy RK, Pinto JP, Wolters JC, Geertsma ER, Fusetti F, Permentier HP, Kuipers OP, Kok J, Poolman B (2011) The response of

Lactococcus lactis to membrane protein production. PLoS One 6: e24060

Martinez-Alonso M, Garcia-Fruitos E, Ferrer-Miralles N, Rinas U, Villaverde A (2010) Side effects of chaperone gene co-expression in recombinant protein production. Microb Cell Factories 9:64

Martinez-Alonso M, Garcia-Fruitos E, Villaverde A (2008) Yield, solubility and conformational quality of soluble proteins are not simultaneously favored in recombinant *Escherichia coli*. Biotechnol Bioeng 101:1353–1358

Meissner D, Vollstedt A, van Dijl JM, Freudl R (2007) Comparative analysis of twin-arginine (tat)-dependent protein secretion of a heterologous model protein (GFP) in three different gram-positive bacteria. Appl Microbiol Biotechnol 76:633–642

Mierau I, Kleerebezem M (2005) 10 years of the nisin-controlled gene expression system (NICE) in *Lactococcus lactis*. Appl Microbiol Biotechnol 68:705–717

Mierau I, Olieman K, Mond J, Smid EJ (2005) Optimization of the *Lactococcus lactis* nisin-controlled gene expression system NICE for industrial applications. Microb Cell Factories 4:16

Ming YM, Wei ZW, Lin CY, Sheng GY (2010) Development of a *Bacillus subtilis* expression system using the improved Pglv promoter. Microb Cell Factories 9:55

Monne M, Chan KW, Slotboom DJ, Kunji ER (2005) Functional expression of eukaryotic membrane proteins in *Lactococcus lactis*. Protein Sci 14:3048–3056

Nesvera J, Patek M (2011) Tools for genetic manipulations in *Corynebacterium glutamicum* and their applications. Appl Microbiol Biotechnol 90:1641–1654

Nozach H, Fruchart-Gaillard C, Fenaille F, Beau F, Ramos OH, Douzi B, Saez NJ, Moutiez M, Servent D, Gondry M, Thai R, Cuniasse P, Vincentelli R, Dive V (2013) High throughput screening identifies disulfide isomerase DsbC as a very efficient partner for recombinant expression of small disulfide-rich proteins in *E. coli*. Microb Cell Factories 12:37

O'Brien EP, Ciryam P, Vendruscolo M, Dobson CM (2014) Understanding the influence of codon translation rates on cotranslational protein folding. Acc Chem Res 47:1536–1544

Papa R, Rippa V, Sannia G, Marino G, Duilio A (2007) An effective cold inducible expression system developed in *Pseudoalteromonas haloplanktis* TAC125. J Biotechnol 127:199–210

Parrilli E, De VD, Cirulli C, Tutino ML (2008) Development of an improved *Pseudoalteromonas haloplanktis* TAC125 strain for recombinant protein secretion at low temperature. Microb Cell Factories 7:2

Parrilli E, Giuliani M, Marino G, Tutino ML (2010) Influence of production process design on inclusion bodies protein: the case of an Antarctic flavohemoglobin. Microb Cell Factories 9:19

Phan TT, Nguyen HD, Schumann W (2012) Development of a strong intracellular expression system for *Bacillus subtilis* by optimizing promoter elements. J Biotechnol 157:167–172

Piette F, D'Amico S, Struvay C, Mazzucchelli G, Renaut J, Tutino ML, Danchin A, Leprince P, Feller G (2010) Proteomics of life at low temperatures: trigger factor is the primary chaperone in the Antarctic bacterium *Pseudoalteromonas haloplanktis* TAC125. Mol Microbiol 76:120–132

Pontes DS, de Azevedo MS, Chatel JM, Langella P, Azevedo V, Miyoshi A (2011) *Lactococcus lactis* as a live vector: heterologous protein production and DNA delivery systems. Protein Expr Purif 79:165–175

Quax TE, Claassens NJ, Soll D, van der Oost J (2015) Codon bias as a means to fine-tune Gene expression. Mol Cell 59:149–161

Reichert JM, Paquette C (2003) Clinical development of therapeutic recombinant proteins. Biotechniques 35:176–175

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

Saez NJ, Nozach H, Blemont M, Vincentelli R (2014) High throughput quantitative expression screening and purification applied to

recombinant disulfide-rich venom proteins produced in *E. coli*. J Vis Exp 80:e51464

Saez NJ, Vincentelli R (2014) High-throughput expression screening and purification of recombinant proteins in *E. coli*. Methods Mol Biol 1091:33–53

Schallmey M, Singh A, Ward OP (2004) Developments in the use of *Bacillus* species for industrial production. Can J Microbiol 50:1–17

Schumann W (2007) Production of recombinant proteins in *Bacillus subtilis*. Adv Appl Microbiol 62:137–189

Shah P, Gilchrist MA (2010) Effect of correlated tRNA abundances on translation errors and evolution of codon usage bias. PLoS Genet 6: e1001128

Singh SM, Panda AK (2005) Solubilization and refolding of bacterial inclusion body proteins. J Biosci Bioeng 99:303–310

Speed MA, Wang DI, King J (1996) Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition. Nat Biotechnol 14:1283–1287

Sun P, Tropea JE, Waugh DS (2011) Enhancing the solubility of recombinant proteins in *Escherichia coli* by using hexahistidine-tagged maltose-binding protein as a fusion partner. Methods Mol Biol 705:259–274

Tateno T, Fukuda H, Kondo A (2007) Direct production of L-lysine from raw corn starch by *Corynebacterium glutamicum* secreting *Streptococcus bovis* alpha-amylase using cspB promoter and signal sequence. Appl Microbiol Biotechnol 77:533–541

Tegel H, Tourle S, Ottosson J, Persson A (2010) Increased levels of recombinant human proteins with the *Escherichia coli* strain Rosetta (DE3). Protein Expr Purif 69:159–167

Teusink B, Bachmann H, Molenaar D (2011) Systems biology of lactic acid bacteria: a critical review. Microb Cell Fact 10(Suppl 1):S11

Toyoda M, Kimura M, Yamamoto N, Miyauchi M, Umezono T, Suzuki D (2012) Insulin glargine improves glycemic control and quality of life in type 2 diabetic patients on hemodialysis. J Nephrol 25:989–995

Tutino ML, Duilio A, Parrilli R, Remaut E, Sannia G, Marino G (2001) A novel replication element from an Antarctic plasmid as a tool for the expression of proteins at low temperature. Extremophiles 5:257–264

Tutino ML, Parrilli E, Giaquinto L, Duilio A, Sannia G, Feller G, Marino G (2002) Secretion of alpha-amylase from *Pseudoalteromonas haloplanktis* TAB23: two different pathways in different hosts. J Bacteriol 184:5814–5817

Unzueta U, Vazquez F, Accardi G, Mendoza R, Toledo-Rubio V, Giuliani M, Sannino F, Parrilli E, Abasolo I, Schwartz S Jr, Tutino ML, Villaverde A, Corchero JL, Ferrer-Miralles N (2015) Strategies for the production of difficult-to-express full-length eukaryotic proteins using microbial cell factories: production of human alpha-galactosidase a. Appl Microbiol Biotechnol 99:5863–5874

Vallejo LF, Rinas U (2004) Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. Microb Cell Factories 3:11

Varenne S, Buc J, Lloubes R, Lazdunski C (1984) Translation is a non-uniform process: effect of tRNA availability on the rate of elongation of nascent polypeptide chains. J Mol Biol 180:549–576

Vazquez E, Ferrer-Miralles N, Mangues R, Corchero JL, Schwartz S Jr, Villaverde A (2009) Modular protein engineering in emerging cancer therapies. Curr Pharm Des 15:893–916

Vigentini I, Merico A, Tutino ML, Compagno C, Marino G (2006) Optimization of recombinant human nerve growth factor production in the psychrophilic *Pseudoalteromonas haloplanktis*. J Biotechnol 127:141–150

Wenzel M, Muller A, Siemann-Herzberg M, Altenbuchner J (2011) Self-inducible *Bacillus subtilis* expression system for reliable and inexpensive protein production by high-cell-density fermentation. Appl Environ Microbiol 77:6419–6425

Wieczorek AS, Martin VJ (2010) Engineering the cell surface display of cohesins for assembly of cellulosome-inspired

enzyme complexes on *Lactococcus lactis*. Microb Cell Factories 9:69

Wilmes B, Hartung A, Lalk M, Liebeke M, Schweder T, Neubauer P (2010) Fed-batch process for the psychrotolerant marine bacterium *Pseudoalteromonas haloplanktis*. Microb Cell Factories 9:72

Yim SS, An SJ, Choi JW, Ryu AJ, Jeong KJ (2014) High-level secretory production of recombinant single-chain variable fragment (scFv) in *Corynebacterium glutamicum*. Appl Microbiol Biotechnol 98:273–284

Zhu FM, Ji SY, Zhang WW, Li W, Cao BY, Yang MM (2008) Development and application of a novel signal peptide probe vector with PGA as reporter in *Bacillus subtilis* WB700: twenty-four tat pathway signal peptides from *Bacillus subtilis* were monitored. Mol Biotechnol 39:225–230