Optimising protein synthesis in cell-free systems, a review

Angelo C Batista, Paul Soudier, Manish Kushwaha, Jean-loup Faulon

Received: 4 September 2020 | Revised: 3 December 2020 | Accepted: 9 December 2020

DOI: 10.1049/enb2.12004

Engineering Biology

The Institution of Engineering and Technology WILEY

REVIEW

Optimising protein synthesis in cell-free systems, a review

Angelo C. Batista[1] | **Paul Soudier**[1] | **Manish Kushwaha**[1] | **Jean-Loup Faulon**[1,2]

[1]Université Paris-Saclay, INRAE, AgroParisTech, Micalis Institute, Jouy-en-Josas, France

[2]SYNBIOCHEM Center, School of Chemistry, Manchester Institute of Biotechnology, The University of Manchester, Manchester, UK

Correspondence

Jean-Loup Faulon, Université Paris-Saclay, INRAE, AgroParisTech, Micalis Institute, 78350, Jouy-en-Josas, France.
Email: jean-loup.faulon@inrae.fr

Funding information

Agence Nationale de la Recherche, Grant/Award Number: ANR-17-CE07-0046

Abstract

Over the last decades, cell-free systems have been extensively used for in vitro protein expression. A vast range of protocols and cellular sources varying from prokaryotes and eukaryotes are now available for cell-free technology. However, exploiting the maximum capacity of cell free systems is not achieved by using traditional protocols. Here, what are the strategies and choices one can apply to optimise cell-free protein synthesis have been reviewed. These strategies provide robust and informative improvements regarding transcription, translation and protein folding which can later be used for the establishment of individual best cell-free reactions per lysate batch.

1 | INTRODUCTION

Cell-free protein synthesis systems are cellular abstractions derived from lysed cells that allow continued performance of many important molecular functions in a growth-independent and relatively fast in vitro protein synthesis. They reduce considerably the time needed for prototyping recombinant protein expression over traditional in vivo procedures. Furthermore, cell-free reactions are usually run in very small volumes, reducing lab space and very suitable for liquid-handling automation. Over the past 60 years, cell-free systems have been applied to several applications such as bio-manufacturing of chemicals or biosimilars [1], biosensing [2,3], biological computing [4], deciphering genetic and protein functions [5] and paving the way for complete synthetic cells [6]. During this time, the technology has increasingly become more simple, affordable and robust. Many cell-free systems can be readily purchased, or even prepared in the lab. Furthermore, the cost of lab-made cell-free reactions is becoming lower, in addition to the ease of preparation, leading the way for optimisation of reactions against commercially available ones [7].

Cell-free systems have been made from a range of cells, both prokaryotic and eukaryotic. The choice of organisms for cell-free extracts have spanned many model organisms such as *Escherichia coli*, *Bacillus subtilis*, insects (*Spodoptera frugiperda*), wheat germ and mammalian-derived extracts such as rabbit reticulocyte lysates, or humans CHO and HeLa cell lines. Each cellular source provides unique characteristics for protein expression, for example production or post-translational modifications in vitro [8]. Depending on the cell type and equipment availability, several methods can be used for lysate preparation including physical (like sonication, centrifugation, flash freezing, bead beating, french-press etc.) and enzymatic (lysozyme-based) techniques [9].

Despite lab-scale methods being valuable for testing bio-manufacturing strategies, cell lysis preparation exhibits considerable batch-to-batch variation and requires repeated buffer optimisation [10]. In spite of commercial kits, it is still possible to achieve efficient and robust lab-made cell-free systems. Lavickova and Maerkl managed to develop an updated version of the PURE system, which was cheaper in comparison to the commercial PURE system (New England Biolabs) or their lab-made methods [11]. Importantly, the one-pot PURE system showed 14-fold improvement in cost of protein yield over the commercialised ones.

Other protocols frequently applied for metabolic engineering do not perform coupled expression (transcription and translation) in vitro. Commonly called Cell-Free Protein Synthesis driven Metabolic Engineering (CFPS-ME), the protocol induces expression of enzymes by the cell prior to the lysis step. Subsequently, the enzyme-enriched lysate can be used either directly or after purification in an in vitro reaction [12]. Thus, enzymatic reactions can still catalyse reactions over a combined mix of lysates according to an estimated protein content in the final cell-free reaction.

Despite lower price for lab-based cell-free lysates methods, these protocols can be laborious for a less equipped laboratory. Furthermore, most cell-free systems lack reproducibility of lysate preparation when using diverse methods and total gene expression may not achieve the maximum protein production yield in every batch. In order to achieve high yield despite poor reproducibility, many research groups have been using T7 systems for boosting gene expression in vitro; tuning the translation initiation sequences or the direct use of mRNA for straight translation in the reaction [13]. In addition, recent studies have shown adjusting several components in buffer reaction could lead to a 34-fold increase in cell-free systems protein yield [14]. In this review, we discuss the challenges and highlight recent developments in the state-of-the-art in protein production optimisation in lab-made cell-free systems. We focus on five main aspects one can tackle to improve protein yield in cell-free systems. These aspects concern constraints such as choice of chassis, transcription-translation, post-translation modifications; and rationalising the concentration of reaction elements for sustaining the in vitro protein synthesis.

2 | DIVERSIFYING CELL-FREE EXTRACTS FOR IMPROVED PROTEIN PRODUCTION

From its early development, cell free protein synthesis systems have not been limited to *Escherichia coli*-based extract. Recent works are now also focussing on the development of cell-free systems for synthetic biology applications for these other organisms including fast lysate preparation, implementation of unique regulatory elements or prototyping of pathways. *Vibrio natriegens* is a non-conventional fast-growing strain with potential for speed-up lysate preparation by 1–2 days and even a considerable fast protein synthesis [15,16]. Additionally, one can choose a strain for its capability of unique regulatory elements present in the cell extract. *Cupriavidus metallidurans* CH34 was used for creating a cell-free biosensor to detect heavy metals (e.g. Hg^{+}, Pb^{2+}, Cu^{2+}, Zn^{2+} and Ni^{2+}) since this strain carries naturally transcription factors for a broad range of these molecules [17].

Generally, in most of non-model systems have not been able to reach the titres observed in *E. coli*-based systems but their interest lies mostly in their ability to generate specific proteins such as hard to fold ones, proteins with specific post-translational modifications (PTMs), difficult to synthesise proteins and membrane proteins. However, a recent work demonstrated that a non-model organism, *Clostridium autoethanoge*, showed high yield protein synthesis in cell-free reactions [18]. Furthermore, the *Clostridium*-based cell-free can be considered an interesting strain for metabolic engineering, to be further discussed in Section 5. Together those extracts constitute a bank of diversity that can be used for optimising the synthesis of functional proteins.

Despite these improvements in bacteria-base cell-free systems, some kinds of proteins cannot be produced in them. Thus, eukaryotic are the preferred choice for some features once these systems could contain several endogenous microsomes. Microsomes are lipidic bodies originating from endoplasmic reticulum, Golgi apparatus and plasmatic membrane. They enable better yield for the production of membrane proteins (that can be translocated into microsomal structure, reducing their inherent toxicity for the system and allowing a better accumulation and extraction of them), the addition of certain PTMs and the folding of some specific proteins [19].

Other eukaryotic specific features like the presence of chaperones [20] or of a specific PTM machinery [21] have been engineered to some extent in *E. coli* with various results but as of today the vast majority of eukaryotic proteins are still synthesised using non-coli systems. A last application where the need of organism specific extract is required is in the production of proteins coming from special biological organisms. For example, the high GC-content genes present in *Pseudomonas putida* have been mostly expressed in CFS derived from this specific organism due to the incompatibility of this gene to a classical expression machinery [22]. An overview of some of the commonly used cell free systems and their respective advantages/disadvantages is shown below (see Table 1).

The large-scale deployment of most of these systems having been limited by their productivity lower than the one observed in *E. coli*, the improvement of these protein synthesis reactions is a key challenge for the generalisation of non-model cell-free systems for bioproduction. The improvement of protein titre in them has been developed along the following two axes: First by tackling specific issues for each organism by identifying and correcting each process that could inhibit the protein production. This approach led to the identification of cellular compounds hindering the protein production in certain organisms [23], revealed the inefficiency of some lysis methods for others [24] and pointed out issues specifically related to transcription or translation in certain organisms [23]. Improvement solving these issues included host organism engineering for in vitro protein production maximisation [25], development of new specifics extract preparation protocol protecting the expression machinery and avoiding having inhibitors in the final extract [24] and development of expression vectors adapted for cell-free production of proteins in a specific host-based extract.

The second axis involves the development of standard methods and cross species tools enabling both a continuous improvement of all the non-model cell free systems. Thus, allowing fast optimisation of new organisms-based extracts and unlocking the transferability of protein synthesis systems from one species' extract to another [26,27]. Despite these advances, there is still a big potential of improvement for non-model cell-free systems as it is suggested by the subsequent number of recent publications obtaining multiple fold of protein productions improvement [28–30] and also by the number of improvement methods demonstrated in *E. coli* but never tested on other platforms [10,14].

The last approach for improving protein synthesis using the potential of various species extract lies in the mixing of various extracts inside the same reaction tube creating a new cell-free system with original features [31] (see Figure 1a). Such

TABLE 1 Overview of cellular sources and their applications in cell-free systems

Organism	Protein expression yield (μg/ml)	Advantages of this system	Disadvantages	Examples of applications in protein production	Existing level of optimisation[a]
E. coli	2300 μg/ml (batch) [32]	Low cost High yield Easy to prepare Adaptability Most documented system	Prokaryote (limited potential for PTM, folding, no microsomes)	Cytokines [33], urokinase [34], aglycosylated antibodies [35]	+++++
S. cerevisiae	8 μg/ml (batch) 17 μg/ml (continuous) [36]	High chassis knowledge for bioproduction Simple culture for extract production	Low protein yield No mammalian-like PTM	Virus-like particles [28]	++
P. pastoris	50 μg/ml [29]	High chassis knowledge for bioproduction	Low protein yield Recently developed system	Human serum albumin [37]	+
Spodoptera frugiperda	285 μg/ml [38]	High microsomes level helping PTM, membrane protein production	Relatively low protein yield High cultivation cost	Human epidermal growth factor receptor (EGFR) [38]	++++
CHO cells	980 μg/ml (continuous) [39]	High acceptance ER-derived microsomes	Relatively low protein yield High cultivation cost	Monoclonal antibodies [40]	+++
Human cell line	50 μg/ml (continuous) [41]	Good natural codon usage ER-derived microsomes	Laborious and expensive cultivation	mTOR [41]	+++
Wheat germ (*Triticum aestivum*)	20000 μg/ml [42]	Better folding than *E. coli* More PTM than coli Suitable for disulphide bound proteins	Requires exogenous microsome addition Laborious and expensive lysate preparation Limited PTM	Human replication protein A complex [26]	+++++

[a]Level of optimisation of the system according to our judgement on the literature.

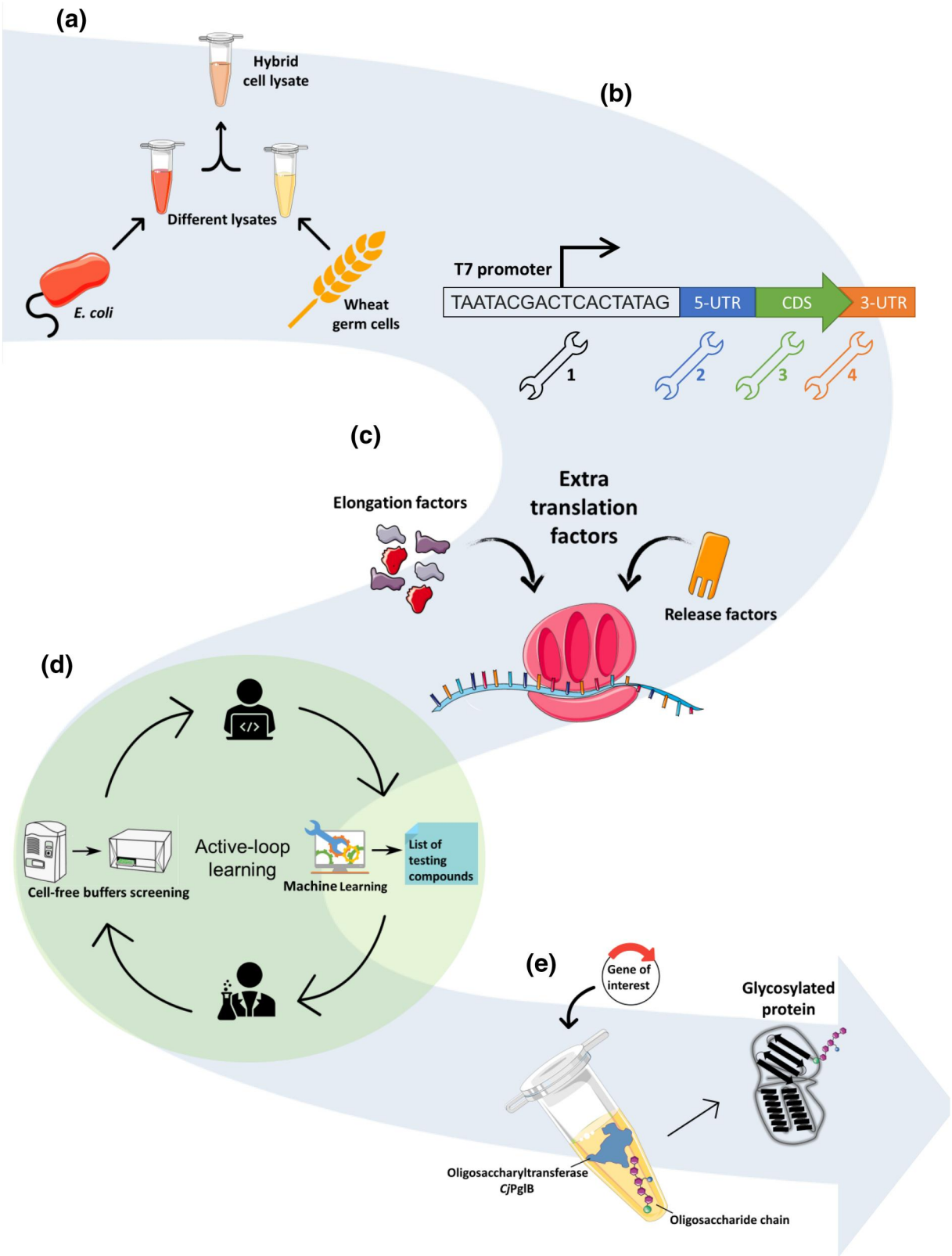

FIGURE 1 Schematic representation of optimisation approaches for improving cell-free protein expression. (a) The choice of the chassis for lysate preparation. Cell lysates from different sources (i.e. *E. coli* and wheat germ) can be combined for different characteristics. (b) For coupled (transcription and translation) cell-free expression, engineering the DNA added can improve both quantity and quality of protein synthesis. The strategy can be applied for a library of synthetic promoters, 5-Untranslated Regions (UTR) and Ribosome Binding Sites (RBS), at the level of Coding sequence (CDS) and even 3-UTR or poly-A tails. (c) Translation machinery can be supplemented or adjusted to improve protein synthesis. (d) Reaction buffers can be optimised with the use of Machine Learning and automation approaches to infer what composition needs better adjustment per lysate preparation. (e) Once established a cell-free protein synthesis protocol, reactions can be optimised for better protein function such as in vitro glycosylation by glycotransferases enriched lysates (Jaroentomeechai et al., 2018)

hybrid cell-free systems have been generated by mixing various commonly used extracts together or adding classically used extract to low adoption cell-free platforms. The idea supporting these systems lies on each individual cell extract might present specific potentials but also specific drawbacks, mixing two different extracts in a rational way could help compensate individual limitations and combine to some extent the advantages of each system.

A substantial amount of work has focussed on developing a hybrid system approaching the high productivity of an *E. coli*-based system while retaining properties present in eukaryotic cell free systems such as a good protein folding ability. This was demonstrated in a protein microarray synthesis work involving a combination of *E. coli* and wheat germ extract that allowed the production of large amounts of correctly folded proteins in a simple setup (see Figure 1a) [43].

Other works comprising hybrid cell-free rely on not mixing complete cell lysates but on just combining different intracellular elements from different organisms. Panthu et al. developed a system based on Rabbit reticulocyte Lysate (RRL) supernatant complemented with ribosomal fraction isolated from other backgrounds such as HeLa cells [44]. They demonstrated the potential of this hybrid mix to overcome traditional limitations of RRL in terms of translational characteristics. More recently, transcriptional potential of such hybrid combinations of lysate was also investigated for a diversity of mixes. Yim et al used their newly developed DNA Regulatory element Analysis by cell-Free Transcription and Sequencing (DRAFTS) method in order to characterise the transcriptional activity of a promoter bank in cell free reactions mixing *E. coli*, *B. subtilis* and *C. glutamicum* in various proportions [31]. Besides the transcriptional aspect, Yim et al., also interestingly found that most of their hybrid lysate exhibited protein synthesis level at average of their single-species counterparts showing the complementarity of various expression machinery and a limited inter-species inhibition. Taken together all those elements strongly support further development of hybrid cell-free systems for multi objective optimisation of protein synthesis.

3 | ENGINEERING EXPRESSION MACHINERY FOR CELL-FREE SYSTEMS

While most genetic functions are preserved from in vivo to in vitro, *E. coli* cell-free reactions have been greatly performed via T7 systems. Therefore, the construction of expression vectors can be done in monocistronic and/or in tandem under a T7 promoter [45] and even for expressing complex multimeric proteins in one operon [46]. Nevertheless, the positioning of genes in operons under promoter T7 does not affect the level of transcripts, but the translation context where the gene is positioned in the operon affects gene expression of individual genes whereas first genes will be firstly expressed [27].

Despite the preference, due to its transcriptional strength being eighth times higher than endogenous polymerase, RNAPT7 lacks a modularity, affecting the cell-free information processes [47], especially for gene networks. The issue could be attributed to secondary structures on mRNA while its fast transcription impacting mostly on protein translation and function, about 50% ± 20% of proteins are indeed functional [48]. Tuning gene expression under T7 systems can still be achieved by mutated versions of RNAPT7 (reduced activity) [48] or at the level of synthetic promoters T7 (see Figure 1b) [27]. Regulated promoters T7 can also be engineered by tuning the kinetics of operators to transcriptional factors. McManus et al. showed that tuning the position of TetO upon the PT7 could interfere with the transition of T7 RNAP from initiation to elongation [49]. Positioning tetO within the first 13 base pairs downstream PT7 can strongly down regulate transcription in contrast to positioning immediately upstream the promoter.

Remarkable improvements were made on several methods for lysates preparation, which were optimised for a high gene expression under *E. coli* endogenous promoters, therefore, conceivably dismissing the need of T7 systems [47,50,51]. Additionally, a synthetic set of the constitutive σ70 promoter was shown to have an equivalent strength during in vitro or in vivo expression, supporting its usage for expression constructs [52]. The use of different sigma-regulated promoters could lead to a passive transcriptional repression by competition of RNAP core, therefore, downregulating transcription levels [47]. Albeit, sigma factors, mostly σ70 and σ28 are robust enough for being integrated in logic gates such as oscillators [30].

Despite the modularity of promoters for coupled protein expression in cell-free systems, translation represents a main bottleneck for adjustment. Many constraints must be considered for translation efficiency starting from codon context. The presence of rare codons can impact the translation of a given protein [53], alternatively, their optimised version despite improving protein yield could lead to a decrease in protein function due to misfolding [54]. Another important concern is regarding the choice of RBS sequences. In general, RBS strength is well correlated to its strength in in vivo systems and cell-free systems can be used to predict in vivo burden in expression of genetic constructs [55].

Although RBS strength impacts protein synthesis directly, fluctuation in translation levels can be observed due to other factors such as molecular crowding of translation machinery [56], mRNA secondary structure [32] and GC content in coding sequence (see Figure 1b) [57]. On the other hand, a supplementation of translation related factors enhances the protein yield of a high GC-content target in eight times from the initial system, for example in *Streptomyces* sp. cell-free systems [58].

In addition, a big concern for eukaryotic cell-free platforms regards UTR sequences (see Figure 1b). Furthermore, efficient translation is generally achieved when UTR elements are present in the mRNA. For example it is known for some eukaryotic lysates that capping is inefficient in coupled expression systems and results in a retard in translation initiation [59], hence, in vitro capping of mRNA is expensive [8]. Methods for adding UTR elements to the coding sequences such as 5′ UTR, 3′ UTR, enhancer sequences and poly-A tails

can be easily constructed by overlap extension PCR and have been shown to increase protein production in yeast and wheat germ platforms [60,61]. Interestingly, the addition of 3′ UTR tails leads to up 26-fold increase in protein yield even in *E. coli*-based platforms such as PURE systems [57].

Furthermore, alternative capping-independent translation could be performed by mediating initiation by a viral IRES. In fact, IRES sequences improve considerably complex proteins in systems derived from CHO or insect cells [38,62]. Moreover, the development of Species-Independent Translational Sequences (SITS), a synthetic structure of 5′-UTR enriched in poly-A, is a landmark in engineering UTR sequences since it works merely as a universal 5′-UTR tag [63,64]. Despite its initial design for *Leishmania tarentolae*, SITS can be used for other platforms such as yeast, wheat germ, insect cells, rabbit reticulocytes and even *E. coli*.

4 | OPTIMISATION OF REACTION COMPONENTS FOR CELL-FREE PROTEIN SYNTHESIS

One of the main issues concerning cellular lysates preparation, if not the most significant, regards batch-to-batch variation and its implications on buffer composition of cell-free reactions [65]. Concerning the optimisation of this buffer, two different strategies have been used to increase the efficiency of the TX-TL reactions: combinatorial approaches in order to determine what concentration of each commonly used substrate is optimal and replacement approaches in order to find molecules other than the classically used ones that optimise cell-free reactions. We will discuss this aspect here but cover the combinatorial optimisation approach in the next section about computational approaches.

A diverse range of protocols is available for the most different types of cells and available equipment. To some extent, protocols vary according to lysis procedure but more importantly by the energy components provided for the cost of transcription and translation, buffering the reaction, and co-factors for the genetic expression. A recurrent issue regarding the supporting energy source for ATP regeneration concerns the accumulation of inorganic phosphate in the reaction, therefore, leading to a decrease in magnesium ion concentration and halting of protein synthesis [66]. Additionally, an excess of NTPs inhibits *E. coli* cell-free reactions whereas supplementation of Mg^{2+} is also required to recover translation function [67]. For resolving such problems, the energy supplies must be balanced and the alternative secondary energy sources for ATP regeneration have been shown as a solution for some cases.

Despite its lower cost the use of glucose in excess could lead to an accumulation of organic acids, ethanol and consequently lower pH. Kim et al. used glycogen and soluble starch to slow down the glucose release for the reaction in *E. coli* systems [68]. This retardation could prolong the protein synthesis over 12 h (against the regular 8–10 h of stated protein expression) and consequently an increase in protein yield and also its solubility and activity. On the other hand, some eukaryotic platforms are limited on the usage of phosphated secondary energy substrates [32]. Anderson et al. were able to recreate a native glycolytic pathway for ATP regeneration by the addition of cyclic AMP (cAMP) and exogenous phosphate [32]. The strategy led to a 16% increase in relative protein yield. However, the best energy source for cell-free systems to date is 3-phosphoglycerate (3-PGA) capable of reaching 2.3 mg/ml in *E. coli* systems [69]. The key improvement relies on recycling of inorganic phosphate by maltose or maltodextrin catabolism through addition of two enzymes. Later, the system was improved by addition of hexametaphosphate, thus, allowing the recycling of phosphates endogenously and still leading to a high protein yield [70].

Balancing translation using additional elements is another valuable approach for improving protein yield in cell-free systems. After supplementation with a series of translation factors involved in elongation, release and/or recycling of ribosomes, chaperones, tRNA and BSA, it was possible to enhance up to fivefold the protein yield in PURE systems [29]. Later, authors demonstrated that a balance between elongation of the polypeptide chain and recycling of ribosomes are crucial for obtaining a considerable amount of functional proteins [37]. With the additional translation factors, the authors prevented ribosome stalling and a 2-fold increase in functional proteins from the already modified recipe.

Moreover, lab-made *E. coli* lysates can be improved by systematically supplementing a combination of transcription and/or translation factors (see Figure 1c). Zhang et al. have tested the external addition of over 31 macromolecular elements for expression of GFP. From those 31 transcription and translation factors 11 showed a consistent approximately twofold increase in protein synthesis [71]. Such manual titration methods, despite being laborious, provide strong information about customising cell-free composition and, therefore, help to recover variability among batches. A recent work relied on preparing *E. coli* from a consortium of transformed cells with plasmids to overexpress translation factors. A consortium of seven strains overexpressing each of 11 initiation, elongation, and termination factors (IETs) and one aminoacyl-tRNA transferases (AAT) was capable of a fivefold higher expression level than classical cell-free systems [72]. This work skips the need for isolating and purifying such elements and provides a robust and easy-to-do protocol for boosting protein expression.

In order to optimise the efficiency of protein synthesis systems, other properties of the reaction media have also been investigated through the replacement of components. Molecular crowding has been early described as a key element influencing bio-chemical processes in general including cell-free systems. Although historically this problem was solved through the addition of polyethylene glycol, Ge et al. showed that it may not be the best option [73]. Indeed, optimal concentrations of Filicoll-70 showed improved results mostly for transcriptional optimisation but also for full protein synthesis process compared with an absence of crowding agent but also to the traditionally used PEG 8000. However, since a high concentration of these same elements were shown to have a deleterious effect on protein production, an extract specific optimisation must be performed.

5 | MACHINE-LEARNING APPROACHES TO INCREASE CELL-FREE SYSTEMS PRODUCTIVITY

Notwithstanding the efficacy of cell-free systems, assessing parameters for improving cell-free reactions could rely on a considerable number of variables and be impractical for manual experimentation. The amount of work required for exploring variables toward transcription or translation elements in cell-free systems can be reduced by using computational Design of Experiment (DoE) approaches. Cascheral et al. (2011), have used a learning loop oriented by models generated from DoE for optimising cell-free protein synthesis reactions and achieved within 0.014% of the combinatorial space calculated to be tested [74]. In their assessment the authors tested (i) a combination of genetic designs of few constructs by varying spacers (around RBS and before Terminator sequences), and (ii) other reaction components (amino acids, NAD, Phosphoenolpyruvate, glucose) to an incomplete commercial cell-free reaction buffer. By optimising one variable at a time with a learning algorithm, the authors managed to reach more than threefold increase in protein production in eight iterations of experimentation. This work illustrated for the first time the power of computational prediction for empowering protein synthesis in cell-free reactions.

Despite algorithm predictions that accomplished the overall increase in protein production, it is difficult to extrapolate the systems for future applications when authors used dependent variables such as DNA construction, buffer composition and time of addition of such extra components. A convenient method suitable for a protein production routine could be more reliable if only the buffer composition could be taken as standard for any eventual DNA constructions. Moreover, the same DoE was applied for improving the cell-free buffer composition to sustain a technique for ribosomes construction in S150 (ribosomes-free extract) lysates [75,76] and consequently increase protein yield. For achieving this goal, authors attempted to use a cheaper poly-sugar (maltodextrin) substrate for sustaining nucleotide regeneration along with 19 other molecules as inputs for the buffer. In the work they achieved a fourfold increase in protein yield over seven iterations and reduced cost of reaction by four times.

DoE can also be used to evaluate and improve the protein yield among two different cell lysates extracts preparation, for example S12 or S30 protocols. Pedersen et al. obtained a twofold to threefold improvement in protein yield for both protocols [77]. In addition, this approach does not only apply to *E. coli*-based cell-free systems whereas the computational approach was also used to improve cell-free systems from other organisms such as *Pichia pastoris* [37], Tobacco BY-2 [78] or insect cells [79]. Moreover, DoE was also implemented to not only improve cell-free buffer composition but also establish an automated lysate preparation in insect cells [79].

A novel approach uses active learning for optimising the reagents in cell-free reactions uniquely for protein yield by taking into consideration batch-to-batch variation even among users in the same lab (see Figure 1d) [14]. Differently from previous approaches, this method applies a randomisation of different concentrations of each of a list of 11 compounds in the buffer per iteration, which is, then, used by a machine-learning algorithm to predict the next iterations. Over seven iterations, a 34-fold increase in protein yield was obtained through the machine-learning predictions. Furthermore, the authors also developed a simplification of the method by using 102 combinations of most informative points (for the learning algorithm) out of combinatorial space higher than 411 possibilities. In addition, authors managed to develop a single-step loop with a recipe of 102-buffer composition, which is sufficiently informative to provide the right buffer composition for the lysate batch. This active-learning workflow could be extended to other kinds of cellular sources for cell-free systems and could be highly valuable for the future of buffer composition optimisation.

Computational approaches could also be used for improving enzymatic reactions in vitro. Karim et al. (2019), demonstrated DoE could also be applied for improving two established pathways of 3-hydroxybutyrate (3-HB) and n-butanol in a non-model cell-free based reaction of the aforementioned *Clostridium autoethanogenum* [80]. In this work, a workflow called iPROBE was developed and used for a cell-free metabolic engineering (CFME) reaction where the titration of enzymes for performing each pathway was oriented by neural network algorithms. Moreover, approximately 15 g/L of 3-HB was achieved in *C. autoethanogenum* cell-free reactions. While the production of butanol was not satisfying, the approach did reached a fourfold increase in production. More importantly, the authors show a correlation between the titre of enzymes in the cell-free reaction in *C. autoethanogenum* and expression levels in vivo regarding genetic expression (promoter strength). This approach can lead the way for an accurate prediction from in vitro to in vivo bioproduction in metabolic engineering, for example when demonstrated for the production of Limonene [81].

However, strategies involving machine-learning requires qualified personnel for the computational/engineering part. In addition, the lab should, preferably, obtain of automation facilities for distributing the different buffer compositions since a large amount of data is needed to be generated in order to cover the combinatorial space one will be probing. On the other hand, pipetting automation or machine-learning approaches become more and more well-spread in synthetic biology and, in the near future, such methodologies will become more accessible.

6 | ENHANCING ACTIVE PROTEINS AND THEIR APPLICATIONS

Achieving quality in protein products in an in vitro environment could often be arduous since protein synthesis is performed without compartmentalisation and/or lack of crucial intracellular elements for post-translation steps. *E. coli* lysates

are the most protein productive systems for cell-free, albeit they show particular issues with protein folding. Some strategies rely on balancing the redox potential with glutathione or inhibiting endogenous reductases with alkaline reagents such as iodoacetamide may influence on disulphide bonds formation [82]. The addition of chaperones, especially DnaK and GroEL, can lead to a considerable increase in solubility of more than 50% of approximately 800 aggregation-prone cytosolic *E. coli* proteins [83].

Some therapeutic proteins require correct glycosylation prior to their medical use, a complicated task that is mostly achieved in *in vivo* systems. Nevertheless, achieving glycosylation in cell-free protein synthesis, when possible, is usually preferable in eukaryotic-based extracts. CHO-based cell-free systems are capable of N-glycosylation in microsomal vesicles and have been used for production of antibodies [40]. Furthermore, erythropoietin (EPO), a complex glycosylated therapeutic hormone with considerable medical importance, could also be produced in insect-based cell-free systems via endoplasmic reticulum-derived microsomes [84].

However, a novel technique recently developed is now able to compensate for the difficulty glycosylation in *E. coli* systems but still benefit from its capacity of high protein production titres. Jaroentomeechai et al. were able to create a one-pot glycosylation system for *E. coli* extracts-based cell-free [85]. Authors used CLM24 cells [86] for the lysate preparation, a strain already engineered for glycosylation, therefore, overexpressing *Cj*PglB enzymes before the lysis step. By adding a plasmid (for coupled expression), the authors achieved 100% of glycosylation in vitro of the protein of interest.

Furthermore, the proper functionality of enzymes influences directly the efficiency for complex biochemical reactions involving harm conditions or biocatalysis. Kay and Jewett showed cell-free could use a relatively toxic biomass hydrolysate as a feedstock for producing 2,3-butanediol in cell-free [22]. Besides, the pathway could tolerate a range of polar solvents, which, on the other hand, tend to affect enzyme stability. Yet, Beer et al. showed ammonium bicarbonate could be used as a volatile buffer in order to facilitate downstream processes for producing 2.8 g/L/h α-ketoglutarate for a large-scale application [87]. Nevertheless, long-term reactions regularly require a NAD^+ supply for such oxidation reactions especially when reactions demand high temperatures [88]. For this aim, the use of selected thermophilic enzymes for NAD^+ regeneration could sustain high temperature reactions up to 15 h.

Using high temperature for cell-free reactions offers a considerable advantage for both breaking down of valuable biomass or metabolic engineering. Honda et al. demonstrated the use of thermophilic enzymes to degrade chitin and convert it into pyruvate with temperature rising 70°C (see Figure 1e) [89]. Despite the use of selected enzymes, protein engineering can still be used for generating more thermostable enzymes. Soh et al. engineered a Keto acid decarboxylase which was increased in stability at high temperature (60°C) and a 10.5-fold increased residual activity [90].

7 | CONCLUSION AND PERSPECTIVES

Cell-free technology is a growing tool for prototyping of biosynthesis of diverse biomolecules ranging from proteins to chemicals. On the other hand, cell-free systems have also potential for industrial or continuous reactors, specially, regarding the ease of downstream purification in a wide-open biological device. Nevertheless, scaling up cell-free protein synthesis is still challenging due to an increase cost of reagents for large-scale use, resources depletion and/or saturation of energy regeneration and accumulation of harmful components over time. With the availability of cell-free metabolic models [91] or computational methods for identifying metabolic modules [92] could provide helpful information to understand the metabolic bottlenecks for cell-free scaling-up. Finally, the application of machine-learning approaches could not only help to improve buffer composition but also a supply of metabolites [93] to boost a large-scale cell-free production via either compartmentalisation (such as synthetic cells) or cell-free continuous batches (bioreactors). Furthermore, hybrid cell-free systems or the use of external glycosylation apparatus in cell-free have potential for improvement in order to find a balance among both protein expression and protein functionality. Thus, the advent of a robust hybrid system is something commercial providers could be looking for. Finally, optimising human or mammalian protein expression for healthcare storage and fast synthesis is a great challenge for applying cell-free technology to point-of-care biosensing [3] and/or personalised medicine [94].

ACKNOWLEDGEMENTS

ACB and JLF acknowledge funding provided by the ANR SINAPUV grant number ANR-17-CE07-0046. JLF and PS acknowledge funding provided by the ANR SynBioDiag grant number ANR-18-CE33-0015. MK acknowledges funding support from INRAE's MICA department and Ile-de-France (IdF) region's DIM-RFSI.

ORCID

Angelo C. Batista https://orcid.org/0000-0003-0161-8462
Manish Kushwaha https://orcid.org/0000-0002-8850-3272

REFERENCES

1. Bundy, B.C., et al.: Cell-free biomanufacturing. Curr. Opin. Chem. Eng. 22, 177–183 (2018)
2. Koch, M., et al.: Custom-made transcriptional biosensors for metabolic engineering. Curr. Opin. Biotechnol. 59, 78–84 (2019)
3. Voyvodic, P.L., Bonnet, J.: Cell-free biosensors for biomedical applications. Curr. Opin. Biomed. Eng. 13, 9–15 (2020)
4. Pandi, A., et al.: Metabolic perceptrons for neural computing in biological systems. Nat. Commun. 10(1), 3880 (2019)
5. Silverman, A.D., Karim, A.S., Jewett, M.C.: Cell-free gene expression: an expanded repertoire of applications. Nat. Rev. Genet. 21(3), 151–170 (2020)
6. Caschera, F., Noireaux, V.: Integration of biological parts towards the synthesis of a minimal cell. Curr. Opin. Chem. Biol. 22, 85–91 (2014)
7. Levine, M.Z., et al.: Escherichia coli-based cell-free protein synthesis: protocols for a robust, flexible, and accessible platform technology. JoVE J. Vis. Exp. 144 (2019) e58882

8. Zemella, A., et al.: Cell-free protein synthesis: pros and cons of prokaryotic and eukaryotic systems. Chembiochem. 16(17), 2420–2431 (2015)
9. Shrestha, P., Holland, T.M., Bundy, B.C.: Streamlined extract preparation for Escherichia coli-based cell-free protein synthesis by sonication or bead vortex mixing. Biotechniques. 53(3), 163–174 (2012)
10. Dopp, J.L., Jo, Y.R., Reuel, N.F.: Methods to reduce variability in E. coli-based cell-free protein expression experiments. Synth. Syst. Biotechnol. 4(4), 204–211 (2019)
11. Lavickova, B., Maerkl, S.J.: A simple, robust, and low-cost method to produce the PURE cell-free system. ACS Synth. Biol. 8(2), 455–462 (2019)
12. Karim, A.S., Jewett, M.C.: A cell-free framework for rapid biosynthetic pathway prototyping and enzyme discovery. Metab. Eng. 36, 116–126 (2016)
13. Weise, L.I., et al.: Cell-free expression of RNA encoded genes using MS2 replicase. Nucleic Acids Res. 47(20), 10956–10967 (2019)
14. Borkowski, O., et al.: Large scale active-learning-guided exploration for in vitro protein production optimization. Nat. Commun. 11(1), 1–8 (2020)
15. Wiegand, D.J., et al.: Establishing a cell-free Vibrio natriegens expression system. ACS Synth. Biol. 7(10), 2475–2479 (2018)
16. Wiegand, D.J., et al.: Cell-free protein expression using the rapidly growing bacterium vibrio natriegens. JoVE J. Vis. Exp. 145 (2019) e59495
17. Millacura, F.A., et al.: TXO: transcription-only genetic circuits as a novel cell-free approach for synthetic biology. bioRxiv 2014: 826230 (2019)
18. Krüger, A., et al.: Development of a clostridia-based cell-free system for prototyping genetic parts and metabolic pathways. Metab. Eng. 62, 95–105 (2020)
19. Fenz, S.F., et al.: Cell-free synthesis of membrane proteins: tailored cell models out of microsomes. Biochim. Biophys. Acta BBA - Biomembr. 1838(5), 1382–1388 (2014)
20. Ryabova, L.A., et al.: Functional antibody production using cell-free translation: effects of protein disulfide isomerase and chaperones. Na. Biotechnol. 15(1), 79–84 (1997)
21. Sachse, R., et al.: Synthesis of membrane proteins in eukaryotic cell-free systems. Eng. Life Sci. 13(1), 39–48 (2013)
22. Kay, J.E., Jewett, M.C.: A cell-free system for production of 2,3-butanediol is robust to growth-toxic compounds. Metab. Eng. Commun. 10 (2020) e00114
23. Legault-Demare, L., Chambliss, G.H.: Natural messenger ribonucleic acid-directed cell-free protein-synthesising system of Bacillus subtilis. J. Bacteriol. 120(3), 1300–1307 (1974)
24. Li, J., Wang, H., Jewett, M.C.: Expanding the palette of Streptomyces-based cell-free protein synthesis systems with enhanced yields. Biochem. Eng. J. 130, 29–33 (2018)
25. Kelwick, R., et al.: Development of a Bacillus subtilis cell-free transcription-translation system for prototyping regulatory elements. Metab. Eng. 38, 370–381 (2016)
26. Beebe, E.T., et al.: Robotic large-scale application of wheat cell-free translation to structural studies including membrane proteins. New Biotechnol. 28(3), 239–249 (2011)
27. Chizzolini, F., et al.: Gene position more strongly influences cell-free protein expression from operons than T7 transcriptional promoter strength. ACS Synth. Biol. 3(6), 363–371 (2014)
28. Wang, X., et al.: An optimised yeast cell-free system: sufficient for translation of human papillomavirus 58 L1 mRNA and assembly of virus-like particles. J. Biosci. Bioeng. 106(1), 8–15 (2008)
29. Zhang, L., Liu, W.-Q., Li, J.: Establishing a eukaryotic Pichia pastoris cell-free protein synthesis system. Front. Bioeng. Biotechnol. 8, 536 (2020)
30. Yelleswarapu, M., et al.: Sigma factor-mediated tuning of bacterial cell-free synthetic genetic oscillators. ACS Synth. Biol. 7(12), 2879–2887 (2018)
31. Yim, S.S., et al.: Multiplex transcriptional characterisations across diverse bacterial species using cell-free systems. Mol. Syst. Biol. 15(8), e8875 (2019)
32. Anderson, M.J., et al.: Energising eukaryotic cell-free protein synthesis with glucose metabolism. FEBS Lett. 589(15), 1723–1727 (2015)
33. Zawada, J.F., et al.: Microscale to manufacturing scale-up of cell-free cytokine production-a new approach for shortening protein production development timelines. Biotechnol. Bioeng. 108(7), 1570–1578 (2011)
34. Kim, D.-M., Swartz, J.R.: Efficient production of a bioactive, multiple disulfide-bonded protein using modified extracts of Escherichia coli. Biotechnol. Bioeng. 85(2), 122–129 (2004)
35. Yin, G., et al.: Aglycosylated antibodies and antibody fragments produced in a scalable in vitro transcription-translation system. mAbs. 4(2), 217–225 (2012)
36. Schoborg, J.A., et al.: Substrate replenishment and byproduct removal improve yeast cell-free protein synthesis. Biotechnol. J. 9(5), 630–640 (2014)
37. Spice, A.J., et al.: Improving the reaction mix of a Pichia pastoris cell-free system using a design of experiments approach to minimise experimental effort. Synth. Syst. Biotechnol. 5(3), 137–144 (2020)
38. Quast, R.B., et al.: High-yield cell-free synthesis of human EGFR by IRES-mediated protein translation in a continuous exchange cell-free reaction format. Sci. Rep. 6(1), 30399 (2016)
39. Thoring, L., et al.: High-yield production of "difficult-to-express" proteins in a continuous exchange cell-free system based on CHO cell lysates. Sci. Rep. 7(1), 11710 (2017)
40. Stech, M., et al.: Cell-free synthesis of functional antibodies using a coupled in vitro transcription-translation system based on CHO cell lysates. Sci. Rep. 7(1), 12030 (2017)
41. Mikami, S., et al.: A human cell-derived in vitro coupled transcription/translation system optimised for production of recombinant proteins. Protein Expr. Purif. 62(2), 190–198 (2008)
42. Gregorio, N.E., Levine, M.Z., Oza, J.P.: A user's guide to cell-free protein synthesis. Methods Protoc. 2(1), 24 (2019)
43. Zárate, X., et al.: Development of high-yield autofluorescent protein microarrays using hybrid cell-free expression with combined Escherichia coli S30 and wheat germ extracts. Proteome Sci. 8, 32 (2010)
44. Panthu, B., et al.: In vitro translation in a hybrid cell free lysate with exogenous cellular ribosomes. Biochem. J. 467(3), 387–398 (2015)
45. Du, L., Gao, R., Forster, A.C.: Engineering multigene expression in vitro and in vivo with small terminators for T7 RNA polymerase. Biotechnol. Bioeng. 104(6), 1189–1196 (2009)
46. Katayama, Y., et al.: Cell-free synthesis of cytochrome c oxidase, a multicomponent membrane protein. J. Bioenerg. Biomembr. 42(3), 235–240 (2010)
47. Shin, J., Noireaux, V.: Efficient cell-free expression with the endogenous E. coli RNA polymerase and sigma factor 70. J. Biol. Eng. 4, 8 (2010)
48. Iskakova, M.B., et al.: Troubleshooting coupled in vitro transcription–translation system derived from Escherichia coli cells: synthesis of high-yield fully active proteins. Nucleic Acids Res. 34(19), e135 (2006)
49. McManus, J.B., et al.: A method for cost-effective and rapid characterisation of engineered T7-based transcription factors by cell-free protein synthesis reveals insights into the regulation of T7 RNA polymerase-driven expression. Arch. Biochem. Biophys. 674, 108045 (2019)
50. Sun, Z.Z., et al.: Protocols for implementing an Escherichia coli based TX-TL cell-free expression system for synthetic biology. JoVE J. Vis. Exp. 79 (2013), e50762
51. Silverman, A.D., et al.: Deconstructing cell-free extract preparation for in vitro activation of transcriptional genetic circuitry. ACS Synth. Biol. 8(2), 403–414 (2019)
52. Chappell, J., Jensen, K., Freemont, P.S.: Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. Nucleic Acids Res. 41(5), 3471–3481 (2013)
53. Yang, Q., et al.: eRF1 mediates codon usage effects on mRNA translation efficiency through premature termination at rare codons. Nucleic Acids Res. 47(17), 9243–9258 (2019)
54. Dondapati, S.K., et al.: Cell-free protein synthesis: a promising option for future drug development. BioDrugs. 34(3), 327–348 (2020)
55. Borkowski, O., et al.: Cell-free prediction of protein expression costs for growing cells. Nat. Commun. 9(1), 1457 (2018)
56. Norred, S.E., et al.: Macromolecular crowding induces spatial correlations that control gene expression bursting patterns. ACS Synth. Biol. 7(5), 1251–1258 (2018)

57. Chizzolini, F., et al.: Cell-free translation is more variable than transcription. ACS Synth. Biol. 6(4), 638–647 (2017)
58. Xu, H., Liu, W.-Q., Li, J.: Translation related factors improve the productivity of a Streptomyces-based cell-free protein synthesis system. ACS Synth. Biol. 9, 1221–1224 (2020)
59. Alekhina, O.M., Vassilenko, K.S., Spirin, A.S.: Translation of non-capped mRNAs in a eukaryotic cell-free system: acceleration of initiation rate in the course of polysome formation. Nucleic Acids Res. 35(19), 6547–6559 (2007)
60. Gan, R., Jewett, M.C.: A combined cell-free transcription-translation system from Saccharomyces cerevisiae for rapid and robust protein synthesis. Biotechnol. J. 9(5), 641–651 (2014)
61. Nomoto, M., Tada, Y.: Cloning-free template DNA preparation for cell-free protein synthesis via two-step PCR using versatile primer designs with short 3′-UTR. Gene Cell. 23(1), 46–53 (2018)
62. Brödel, A.K., et al.: IRES-mediated translation of membrane proteins and glycoproteins in eukaryotic cell-free systems. PloS One. 8(12), (2013) e82234.
63. Mureev, S., et al.: Species-independent translational leaders facilitate cell-free expression. Nat. Biotechnol. 27(8), 747–752 (2009)
64. Swartz, J.R.: Universal cell-free protein synthesis. Nat. Biotechnol. 27(8), 731–732 (2009)
65. Miguez, A.M., McNerney, M.P., Styczynski, M.P.: Metabolic profiling of Escherichia coli -based cell-free expression systems for process optimization. Ind. Eng. Chem. Res. 58(50), 22472–22482 (2019)
66. Kim, D.M., Swartz, J.R.: Prolonging cell-free protein synthesis by selective reagent additions. Biotechnol. Prog. 16(3), 385–390 (2000)
67. Nagaraj, V.H., et al.: Translation inhibition and resource balance in the TX-TL cell-free gene expression system. Synth. Biol. 2(1), (2017)
68. Kim, H.-C., Kim, T.-W., Kim, D.-M.: Prolonged production of proteins in a cell-free protein synthesis system using polymeric carbohydrates as an energy source. Process Biochem. 46(6), 1366–1369 (2011)
69. Caschera, F., Noireaux, V.: Synthesis of 2.3 mg/ml of protein with an all Escherichia coli cell-free transcription–translation system. Biochimie. 99, 162–168 (2014)
70. Caschera, F., Noireaux, V.: A cost-effective polyphosphate-based metabolism fuels an all E. coli cell-free expression system. Metab. Eng. 27, 29–37 (2015)
71. Zhang, Y., et al.: Enhancing the efficiency of cell-free protein synthesis system by systematic titration of transcription and translation components. Biochem. Eng. J. 138, 47–53 (2018)
72. Contreras-Llano, L.E., et al.: Holistic engineering of cell-free systems through proteome-reprogramming synthetic circuits. Nat. Commun. 11(1), 3138 (2020)
73. Ge, X., Luo, D., Xu, J.: Cell-free protein expression under macromolecular crowding conditions. PloS One. 6(12), (2011) e28707
74. Caschera, F., et al.: Coping with complexity: machine learning optimization of cell-free protein synthesis. Biotechnol. Bioeng. 108(9), 2218–2228 (2011)
75. Jewett, M.C., et al.: In vitro integration of ribosomal RNA synthesis, ribosome assembly, and translation. Mol. Syst. Biol. 9, 678 (2013)
76. Caschera, F., et al.: High-throughput optimization cycle of a cell-free ribosome assembly and protein synthesis system. ACS Synth. Biol. 7(12), 2841–2853 (2018)
77. Pedersen, A., et al.: Rational improvement of cell-free protein synthesis. New Biotechnol. 28(3), 218–224 (2011)
78. Buntru, M., et al.: A versatile coupled cell-free transcription–translation system based on tobacco BY-2 cell lysates. Biotechnol. Bioeng. 112(5), 867–878 (2015)
79. Richter, C., et al.: High-throughput characterisation of an insect cell-free expression. Eng. Life Sci. 14(4), 409–417 (2014)
80. Karim, A.S., et al.: In vitro prototyping and rapid optimization of biosynthetic enzymes for cell design. Nat. Chem. Biol. 16(8), 912–919 (2020)
81. Dudley, Q.M., et al.: In vitro prototyping of limonene biosynthesis using cell-free protein synthesis. Metab. Eng. 61, 251–260 (2020)
82. Stech, M., Kubick, S.: Cell-free synthesis meets antibody production: a review. Antibodies. 4(1), 12–33 (2015)
83. Niwa, T., et al.: Global analysis of chaperone effects using a reconstituted cell-free translation system. Proc. Natl. Acad. Sci. 109(23), 8937–8942 (2012)
84. Zemella, A., et al.: Cell-free protein synthesis as a novel tool for directed glycoengineering of active erythropoietin. Sci. Rep. 8(1), 8514 (2018)
85. Jaroentomeechai, T., et al.: Single-pot glycoprotein biosynthesis using a cell-free transcription-translation system enriched with glycosylation machinery. Nat. Commun. 9(1), 1–11 (2018)
86. Feldman, M.F., et al.: Engineering N-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in Escherichia coli. Proc. Natl. Acad. Sci. 102(8), 3016–3021 (2005)
87. Beer, B., Pick, A., Sieber, V.: In vitro metabolic engineering for the production of α-ketoglutarate. Metab. Eng. 40, 5–13 (2017)
88. Honda, K., et al.: In vitro metabolic engineering for the salvage synthesis of NAD+. Metab. Eng. 35, 114–120 (2016)
89. Honda, K., et al.: In vitro bioconversion of chitin to pyruvate with thermophilic enzymes. J. Biosci. Bioeng. 124(3), 296–301 (2017)
90. Soh, L.M.J., et al.: Engineering a thermostable keto acid decarboxylase using directed evolution and computationally directed protein design. ACS Synth. Biol. 6(4), 610–618 (2017)
91. Horvath, N., et al.: Towards a genome scale sequence specific dynamic model of cell-free protein synthesis in Escherichia coli. Metab. Eng. Commun. 10 (2020) e00113
92. Kamp, A. von, Klamt, S.: MEMO: a method for computing metabolic modules for cell-free production systems. ACS Synth. Biol. 9(3), 556–566 (2020)
93. Miguez, A.M., et al.: Metabolic profiling of Escherichia coli-based cell-free expression systems for process optimization. Ind. Eng. Chem. Res. 58(50), 22472–22482 (2019)
94. Ogonah, O.W., Polizzi, K.M., Bracewell, D.G.: Cell free protein synthesis: a viable option for stratified medicines manufacturing? Curr. Opin. Chem. Eng. 18, 77–83 (2017)

How to cite this article: Batista AC, Soudier P, Kushwaha M, Faulon JL. Optimising protein synthesis in cell-free systems, a review. *Eng. Biol*. 2021;5:10–19. https://doi.org/10.1049/enb2.12004