Cell-Free Protein Synthesis: Chassis toward the Minimal Cell

Ke Yue 1, Yiyong Zhu 2 and Lei Kai 1,3*

1 The Key Laboratory of Biotechnology for Medicinal Plants of Jiangsu Province, School of Life Sciences, Jiangsu Normal University, Shanghai Road 101, Xuzhou 221116, China; kyue@jsnu.edu.cn
2 Jiangsu Provincial Key Lab for Organic Solid Waste Utilization, National Engineering Research Center for Organic-based Fertilizers, Jiangsu Collaborative Innovation Center for Solid Organic Waste Resource Utilization, Nanjing Agricultural University, Nanjing 210095, China; yiyong1973@njau.edu.cn
3 Department of Cellular and Molecular Biophysics, Max Planck Institute of Biochemistry, D-82152 Martinsried, Germany
* Correspondence: kai@biochem.mpg.de or lkai@jsnu.edu.cn; Tel.: +49-(0)89-8578-2319

Received: 19 February 2019; Accepted: 4 April 2019; Published: 5 April 2019

Abstract: The quest for a minimal cell not only sheds light on the fundamental principles of life but also brings great advances in related applied fields such as general biotechnology. Minimal cell projects came from the study of a plausible route to the origin of life. Later on, research extended and also referred to the construction of artificial cells, or even more broadly, as in vitro synthetic biology. The cell-free protein synthesis (CFPS) techniques harness the central cellular activity of transcription/translation in an open environment, providing the framework for multiple cellular processes assembling. Therefore, CFPS systems have become the first choice in the construction of the minimal cell. In this review, we focus on the recent advances in the quantitative analysis of CFPS and on its advantage for addressing the bottom-up assembly of a minimal cell and illustrate the importance of systemic chassis behavior, such as stochasticity under a compartmentalized micro-environment.

Keywords: cell-free protein synthesis; chassis; gene circuits; micro-compartments; stochasticity

1. Introduction

Fascinated by the emergence of life from non-living matter through billions of years of evolution, scientists began to comprehend and reconstruct how this occurred. Advances in multi-disciplinary research fields, crossing physics, chemistry, and biology, allowed us to explore the essence of life via experimental approaches. Although, until today, the definition of the full characteristics of life and what are essentials for a minimal cell are debated, though a number of features have been agreed upon, i.e., regarding compartmentalization, metabolism (energy and mass exchange), self-organization, growth and division, adaptability and mobility, and information processing [1–4]. Those processes have become target functions to be reconstructed using in vitro modular systems and have integrated stepwise toward the minimal cellular mimicry. From this point of view, aiming at understanding the essence of life, minimal cell projects seek to reconstitute cellular processes controllably and predictably via a minimum set of compounds [1,5]. When designing and building such a minimal biological system, one can proceed either ‘top-down’ or ‘bottom-up’ [6]. In a top-down approach, a bacterial target genome is continuously reduced to a minimal gene set in vivo; a bottom-up approach relies on the production and purification of functional molecules, which are then combined in vitro with the goal of assembling a minimal cell [7].
Since its first application in deciphering genetic codes [8], cell-free protein synthesis (CFPS) has emerged as an important recombinant protein production method [9–16]. CFPS, being a framework for understanding, harnessing, and expanding biological systems in vitro, has also been used as an important toolbox in other fields of synthetic biology [2,3,17,18]. CFPS, as indicated by its name, refers to the expression of recombinant proteins without living cells. Either cell extracts or individual purified enzymes are used as the machinery for protein transcription/translation [11]. The development of minimal cell projects via a bottom-up approach came along with the quest for pre-biotically plausible routes to the origin of life [19], experimentally repeating the transition from pure chemical compounds to living systems [20]. Hence, the CFPS system as a fully reconstituted system naturally became useful for such projects.

However, we are still far from the ultimate goal of obtaining a minimum cellular system. As mentioned above, full agreement on the essential properties for a minimal cell has not been reached. Although the critical characters of life are still debated, such features after extensive discussion could lead to the design and assembly of minimal cells [21,22]. Moreover, the stepwise construction of a minimal cell could provide new insights into the essence of living systems. Finally, revealing the fundamental principles of a living system will accelerate related applications, certainly in this sense beneficial for biotechnology in general.

As shown in Figure 1, in this review, we set our focus on the characterization of a CFPS system, in particular about the regulation of protein expression via genetic circuits and designed micro-compartments, as well as the principles to reconstitute biological patterns, moving toward a self-organizing system. Additionally, systematic stochasticity, molecular crowding effects, and their important roles in a stepwise assembly of the multi-functional minimal system are briefly discussed. Finally, we point out the trend toward the quantitative analysis of CFPS systems, which will be beneficial to the integration and hierarchical assembly of a minimal cellular system in vitro.

Figure 1. Cell-free protein synthesis (CFPS) within various compartments and its regulation via gene circuits. The CFPS system hosts the core transcription and translation processes, providing the chassis/framework for different cellular mimicry modules/systems. A number of regulatory elements were introduced and validated to manipulate the protein synthesis within CFPS systems. Both RNA and protein-based gene circuits were built to regulate target protein expression on the transcription level via tuning corresponding mRNA concentration. With such design principles, large genetic networks were successfully realized, i.e., 3- and 5-node ring oscillators. On the translational level, RNA thermometers were employed and were able to control translation initiation via tuning the availability of the ribosomal binding sites (RBS). Different materials were applied for creating the...
physical boundary to encapsulate the CFPS reactions, including coacervates, water in oil droplets, and lipid vesicles. System stochasticity starts to influence the output of gene expression when CFPS reactions were encapsulated. RBS: ribosomal binding site; Anti-RBS: anti-ribosomal binding site; sfGFP: superfolder green fluorescent protein; tetR: gene sequence coding Tet Repressor proteins; cl: gene sequence coding cl protein that binds OR1 and OR2 sites within P1 promoter; lacI: gene coding lac repressor; PwxO1, PwxO2, and POR1-OR2-Pr promoter sequences that can be regulated via corresponding repressor proteins.

2. Regulation of the CFPS System

2.1. Genetic Circuits

Cells develop a set of regulatory tools to sense and process stimuli (information) from the external environment and internal physiological states [23]. In response to constant environmental changes, cellular activities are tuned through a set of regulatory elements, controlling various gene expressions. Such a regulatory system is encoded within genetic networks, interconnected webs of regulatory molecules, synchronizing gene expression in defined patterns, namely ‘gene circuits’ [24]. Similar to electrical circuits, gene circuits are analogies abstracted from well-characterized genes and gene products that respond to a stimuli signal [25,26]. Since the pioneering work from Elowitz, Leibler, and Gardner et al., the single cell system has been conceived as the framework which was composed of standard interaction circuits capable of receiving input signals, executing a serial logical computation, and producing output signals [27]. Therefore, most known gene circuits using such a cell system were discovered via introducing genetic or phrenological perturbation of the model system via a top-down approach [28]. The discovery of such gene circuits did not necessarily give a clear answer on the design and selection principles for a particular functional gene circuit unit [29]. The initial goal of such a synthetic approach was to create autonomous genetic circuits, functioning independently from endogenous cellular circuitry, and finally replacing the endogenous circuitry completely [29]. In addition, continuous efforts in developing computational tools greatly accelerated the characterization and design of genetic regulators [23], which resulted in a number of well-characterized regulatory elements and design principles. However, generic limitations of in vivo modular systems greatly hamper the designing of new circuits, so a limited set of molecules can be successfully implemented—far less than those contained in the simplest organisms [23,24]. The chassis behavior of a cell system requires high compatibility with the existing regulatory elements, often resulting in an unpredictable output; on the other hand, the implementation of new regulatory gene circuits into cells often requires a long procedure until the output signal can be characterized. In addressing the above challenges, a complementary in vitro approach was employed and developed, offering a more flexible chassis as a simplified cell mimicry environment for characterization of the output of designed gene circuits [30–34]. Different in vitro systems were applied for testing designed gene circuits, including nucleic acid systems, hybrid systems and transcription, and translation systems (we direct our readers to a detailed review [35]). Such a complementary in vitro approach, namely the bottom-up approach in synthetic biology, followed by the design–build–test workflow, helps to reveal the fundamental regulatory mechanism and is devoid of the influence of cellular chassis behavior. Next, we focus on the gene circuits investigated in the in vitro transcription and translation system, also referred to as a CFPS or Cell-free TX-TL (Transcription-Translation) system.

2.2. Protein Based Gene Circuits

Since the first gene circuit in the CFPS system was established by Noireaux, Bar-Ziv and Libchaber in 2003 [30], a broad range of genetic circuits have been characterized and championed by different groups (see Table 1 for details). Early CFPS systems, especially the T7 polymerase-based system, employed phage polymerases as a strong and efficient transcriptional machinery to provide a sufficient amount of mRNA for translation [36]. However, due to the highly efficient phage polymerase-based transcription machinery, a limited number of regulatory elements could be used,
which hampered the design of large and complex gene circuits. In addressing such a challenge, the research group of Noireaux used the endogenous RNA polymerase from *E. coli* instead of T7 polymerase to support transcription in the CFPS system. Due to the well-studied transcriptional control elements in *E. coli*, a variety of control elements, i.e., sigma factor-based regulators, could be used in such a CFPS system. After extensive and comprehensive characterization in the CFPS system, the transcription repertoire of the CFPS system based on endogenous *E. coli* RNA polymerase was greatly extended, resulting in a large number of transcription-regulatory factors [37,38]. Based on the above-verified transcription factors in the CFPS system, modular circuit motifs, such as the logical AND gate, multiple stage cascades, negative feedback loops, positive feedback loops, RNA transcriptional cascades with a protein regulated incoherent feed-forward loop, and in vitro ring oscillators, were successfully implemented (see Table 1 for details). Such in vitro gene circuits based on CFPS allow systems to operate in a synthetic environment considerably more simply than do in vivo model systems, though the two systems are functionally similar. Furthermore, the rapid circuit design–build–test workflow allows one to probe fundamental aspects of gene circuit operation which are otherwise masked by the complex cellular environment in vivo [39].

### Table 1. Regulate motifs tested in the CFPS system and their regulatory functions.

<table>
<thead>
<tr>
<th>Regulate Elements</th>
<th>Regulation Description</th>
<th>Control Level</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> sigma factors</td>
<td>Transcriptional activation units</td>
<td>Transcription</td>
<td>[37]</td>
</tr>
<tr>
<td>SsrA-ClpXP</td>
<td>Positive degradation of reporter protein</td>
<td>Transcription</td>
<td>[31,40,41]</td>
</tr>
<tr>
<td>TetR, LacI, AraC, and lambda repressors Cl and Cro</td>
<td>Inducible transcriptional repression and activation; bistable switches; genetic oscillators</td>
<td>Transcription</td>
<td>[23,33,37]</td>
</tr>
<tr>
<td>Pr, Pr1, and Pr2</td>
<td>Provides constant transcription</td>
<td>Transcription</td>
<td>[39]</td>
</tr>
<tr>
<td>pT181 and its derivatives</td>
<td>RNA transcriptional attenuator; antisense RNA mediated transcriptional repressor; three level RNA transcription cascades</td>
<td>Transcription</td>
<td>[42]</td>
</tr>
<tr>
<td>small transcription activating RNAs</td>
<td>Small RNAs that activate the transcription of a specific gene regulated by a terminator (T181, AD1)</td>
<td>Transcription</td>
<td>[43]</td>
</tr>
<tr>
<td>BetI, PhiF and SrpR, TetR, LacI, and QacR</td>
<td>3,4 and 5-node oscillators</td>
<td>Translation</td>
<td>[34]</td>
</tr>
<tr>
<td>RNA thermometers as well as other rational designed sequences</td>
<td>control the secondary structure of mRNA leading to control of the ribosomal binding rate</td>
<td>Translation</td>
<td>[44]</td>
</tr>
</tbody>
</table>

### 2.3. RNA-Based Gene Circuits

Not only restricted to proteins, regulatory molecules can also consist of RNAs. Takahashi, et al. successfully established an RNA transcriptional cascade via RNA transcriptional attenuators (pT181 and its mutants) as the central regulator, performing as a transcriptional on/off switch [42]. Beyond the regulatory effect on the transcriptional level, RNAs can also regulate gene expression on a translation level. Classical regulation control mainly focused on mutating ribosomal binding sites so as to turn the translation rate via the binding kinetics of ribosomes [45]. Recently, noncoding RNA, such as riboswitch (sRNA and RNA thermometers) [44,46,47], and catalysis (ribozymes) can also act as regulatory elements in relation to the translation to tune the expression of the specific gene [48,49]. For instance, riboswitches, located within the 5′-UTR regions of mRNA, can regulate downstream gene expression in response to ligand binding directly to the mRNA [50,51].

### 3. Programming Spatiotemporal Patterns—Toward the Minimal Cell Division System

Biological systems are highly organized. Even in the simplest prokaryotic cell system, synchronized molecular rearrangement can be found. Besides the regulatory machinery encoded via
genetic circuits, biological systems also develop another strategy to organize protein expression on large space and time scales. As first described in the Turing/reaction-diffusion (RD) model, the simple interaction of two components with different diffusion coefficients lead to a spatiotemporal pattern formation under certain theoretical conditions [52]. Such pattern formation exists in different biological systems, from single cellular to animal embryo development processes [53]. One of the well-studied examples of self-organization and pattern formation both in vivo and in vitro was the bacteria MinCDE system [54]. Min proteins constantly oscillate from pole to pole (long axis in vitro) merely via the biochemical properties of MinD and MinE. Upon ATP binding and dimerization, MinD cooperatively binds the membrane via an amphipathic C-terminal membrane targeting sequence (mts) [55]. Both in vivo and in vitro experiments showed that the Min protein oscillation system was able to sense and react to morphological changes via dynamic Min protein patterns. Here, we would like to direct our readers to several comprehensive reviews on this topic [52,53,56]. Such a Min protein oscillation system is key to correctly positioning the contract ring—the ‘Z-ring’—at the mid-cell of \( E. coli \) [57–59]. Encapsulated Min proteins can even act as additional mechanical forces in giant unilamellar vesicles (GUVs), resulting in a rapid deformation of GUVs. This may provide simple autonomous division machinery for lipid vesicles systems [60].

4. Toward Self-Organization in CFPS Systems

Moving beyond the Min protein oscillation pattern, how can one reconstitute a reaction-diffusion expression pattern in a CFPS system? What is the prerequisite for such a complex system? Considering the Turing model, a reaction-diffusion model for a chemical signal would be described by differential equation \( \frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2} + f(u) \) for local concentration \( u(x, t) \) in spatial coordinate \( x \) as a function of time \( t \). In this model, the rate of change, \( \frac{\partial u}{\partial t} \), is determined by a diffusion operator in space, \( D \frac{\partial^2 u}{\partial x^2} \), with a coefficient \( D \), and by a local nonlinear reaction function, \( f(u) \), which includes sources and degradation terms as well as molecular interactions and feedback regulation [56,61,62]. Thus, all the terms in this equation should be implemented in order to build a self-organization pattern in a CFPS system.

As a closed system, particularly within cell-size compartments, the CFPS system has suffered from the fast decay of protein production due to a loss of enzymatic activities, resource consumption, and product accumulation [62]. This has led to a chemical equilibrium, therefore limiting the complexity and size of the gene network [63]. Different approaches have been implemented in order to overcome this challenge: (1) the passive exchange of substrates via the incorporation of a pore-forming protein complex, i.e., \( \alpha \) hemolysin (\( \alpha \)HL) [30]; (2) the positive degradation of mRNAs and proteins using RNase and protease to improve the turnover of both mRNAs and proteins (however, only mRNA can be maintained in a steady state, which indicates that an extra mechanism might require the support of a steady translation rate) [40]; (3) periodic dilution of CFPS reactions via a fresh reaction mixture and a DNA template enabling continuous nutrient exchange, leading to steady transcriptional and translational reaction rates [63]; and (4) diffusive DNA compartments based on immobilization of DNA on the surface of circular micro-compartment connected via thin capillaries to a feeding channel of a CFPS reaction mixture [64]. Such a design would allow for a steady state expression via creating source-sink dynamics with a combination of synthesis and degradation [65,66].

Besides an effective turnover mechanism, maintaining a biochemical non-equilibrium is also essential for the construction of a dynamic system, which involves two fundamental principles: feedback and nonlinearity [34]. Feedback is at the center of the network design, which can be constructed with either two mutually inhibitory genes [37] or an autocatalytic gene and its inhibitor [65]. Non-linearity can be introduced by the cooperative binding of regulatory proteins [67], enzymatic degradation [68], and the network topology. In order to regulate gene expression in a heterogeneously distributed cellular environment, controllable diffusion communication should be established. Such sensing and communication designs were first verified through multi-compartment systems. For instance, two amphiphilic inducer molecules N-acyl-L-homoserine
lactones (AHLs) or isopropyl-β-D-thiogalactopyranoside (IPTG) were investigated as signaling molecules for the communication between water/oil droplets, artificial cells, and E. coli [69,70].

5. Compartmentalization

As one of the essential properties of life, physical boundaries distinguish living matter from non-living environments [71], which allow for the maintenance of non-equilibrium dynamics. Different materials can be used for the in vitro compartmentalization (IVC) process, starting from liquid–liquid phase separation to biomimetic lipid vesicles (examples shown in Figure 1). The first well-studied IVC process was achieved by simply mixing two immiscible fluids by either agitation or vortexing. For instance, an aqueous solution and oil could form self-assembled emulsion water/oil droplets [72,73], allowing for an aqueous environment for biochemical reactions. Later, biomimetic compartments such as lipid vesicles/liposomes [74–76], polymersomes [77,78], and proteinosomes [68,79] were developed for the IVC of various reactive components, ranging from reactions catalyzed by a single enzyme, to multi-step reactions driven by enzyme cascades [80,81], and finally to fully functional transcription and translation systems [30]. One major drawback of conventional methods based on processing bulk samples upon spontaneous self-assembly was the inhomogeneity [82], often resulting in poly-dispersed compartments and low encapsulation efficiency [83,84]. The development of micro-fabrication-based microfluidic devices provides a solution to address the above challenges, which could lead to homogeneous droplets and unilamellar vesicles with greatly improved encapsulation efficiency [85,86].

6. Stochasticity

Cellular environments are often different from simplified in vitro environments. For instance, heterogeneity and stochasticity are commonly found in various prokaryotic and eukaryotic cells [87–90]. However, such properties, assuming a well-stirred bulk environment, were not considered to be significant, because most biochemical reactions are investigated under a diluted and homogenously distributed system in vitro. The research group of Luisi found that the process of macromolecules encapsulated within self-assembled lipid vesicles showed a power-law distribution rather than the expected Poisson distribution due to the solute self-condensing effect [91]. Similar solute self-concentration effects were also observed when encapsulating ribosomes and complete CFPS systems, consisting of more than 80 individual macromolecules (i.e., enzymes, ribosomes, and transfer RNAs) [92]. Besides the encapsulation process, the existence of systematic stochasticity is of great importance for the regulation of transcription and translation within a CFPS system, particularly when encapsulated within micro-compartments supposed to mimic cellular dimensions, which has been less studied. In a typical diluted in vitro environment, system stochasticity is often omitted due to the assumption that, under a well-stirred bulk environment, a low concentration of molecules and the random nature of their collisions can be eliminated, or at least greatly reduced [93]. However, the minimal CFPS system PURE hosts more than 900 reactions and involves the translation process [94], without considering the whole transcription process [95]. Therefore, this intrinsic randomness of the CFPS system cannot be completely ignored in micro-compartments, even under bulk conditions. Via applying the dual reporter model described by Elowitz and co-workers [96], the stochasticity of the CFPS system in terms of expression noise was investigated in microchambers, droplets, and vesicles, and such noise is a non-trivial factor when considering CFPS within compartments for assembly cellular mimicry systems [97,98]. The previous kinetic model of the CFPS system, established using the deterministic model under bulk environments, should be adapted by including a stochastic model when the CFPS system is applied to such micro-compartments.

7. Perspectives

The shift from qualitative to quantitative analysis of the CFPS system has greatly increased its applications, especially for the bottom-up assembly of minimal cells. Different experimental approaches have been taken to systematically investigate the kinetics of CFPS systems, which
resulted in a set of quantitative mathematical models that describe transcription and translation rates [99]. In addition, instead of using the overall fluorescence of a reporter protein, the performance of ribosomes in terms of the active fraction and the number of synthesizing cycles was determined, which could indicate the productivity of a particular CFPS system on a single ribosome level [100,101]. Such quantitative analysis of the CFPS system will contribute to more accurate prediction, together with the systematic development and establishment of the transcriptional and translational control toolbox, which allows for the construction of large and complex gene circuits. Furthermore, via the smart design and application of microfluidic devices, a steady state of the CFPS system can be achieved, which shows the success of assembling a self-organization CFPS system [63]. A number of modular cellular mimics were successfully achieved, i.e., an energy regeneration hybrid vesicle system [78], multi-enzyme cascades for CO₂ fixation (CETCH) [102], de novo synthesis of lipids [103], and recently even a self-replicating Φ29 virus DNA system [104]. Such advances offer versatile building blocks and allow for the attempt to synchronize multi-functional artificial systems under the common chassis of CFPS systems. Furthermore, the comprehensive studies of the chassis behavior of CFPS systems have allowed for the possibility to move beyond a single functional system and to synchronize the expression rhythm within. In spite of such success, this is just the beginning of the journey toward the ultimate goal of a minimum cell. The same is true for CFPS systems, being simplified cellular mimics, yet our understanding of such a fully reconstituted system is far from complete and systematic. There are still other physical properties that need to be investigated, i.e., systematic stochasticity and the molecular crowding effect on biochemical reactions in general. Thus, we do believe that further systematic investigation on the chassis behavior of CFPS systems will not only reveal the fundamental principles of higher-order cellular regulation but also accelerate the hierarchical assembly of a minimal cell.

Author Contributions: K.Y., Y.Z., and L.K. wrote the manuscript.

Funding: Lei Kai would like to thank for their support the Natural Science Foundation of Jiangsu Higher Education Institutions of China (Grant No. 17KJB180003) as well as the Natural Science Foundation of Jiangsu Normal University (Grant No. 17XLR037). This work is also part of the MaxSynBio consortium, which is jointly funded by the Federal Ministry of Education and Research of Germany and the Max Planck Society. Yiyong Zhu would like to thank the National Natural Science Foundation of China (31471937) for their support.

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

References


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).