Delft University of Technology

Genetically encoded phospholipid production for autonomous synthetic cell proliferation

Blanken, D.M.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Genetically encoded phospholipid production for autonomous synthetic cell proliferation

Genetically encoded phospholipid production for autonomous synthetic cell proliferation

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, Prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on Thursday 14 January 2021 at 12:30 o'clock

by

Duco Martijn BLANKEN

Master of Science in Applied Physics, Delft University of Technology, the Netherlands
born in Harderwijk, the Netherlands

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. Dr. A.M. Dogterom	Delft University of Technology, promotor
Dr. C.J.A. Danelon	Delft University of Technology, promotor

Independent members:

Prof. Dr. G.H. Koenderink	Delft University of Technology
Prof. Dr. B. Poolman	Rijksuniversiteit Groningen
Dr. P. Stano	Università del Salento, Italy
Dr. K. Ruiz-Mirazo	University of the Basque Country, Spain
Prof. Dr. Ir. S.J. Tans	Delft University of Technology, reserve member

Pleasure to me is wonder—the unexplored, the unexpected, the thing that is hidden and the changeless thing that lurks behind superficial mutability. To trace the remote in the immediate; the eternal in the ephemeral; the past in the present; the infinite in the finite; these are to me the springs of delight and beauty.
H.P. Lovecraft

There is nothing like looking, if you want to find something. You certainly usually find something, if you look, but it is not always quite the something you were after.
J.R.R. Tolkien, The Hobbit, or There and Back Again

Contents

Chapter 1: Introduction

All life consists of compartments called cells, is governed by information transfer following the central dogma of molecular biology, and is able to proliferate in an autonomous manner. Reconstitution of these features will result in the construction of a synthetic living cell. Building a synthetic cell, following the bottom-up approach, will provide a 'blueprint' of life, and is therefore an area of considerable interest in biology. Cell-free gene expression from DNA inside phospholipid vesicles called liposomes is the most suitable framework for synthetic cell assembly. Growth of the compartment is one of the important aspects of cell proliferation. Within the established framework, its reconstitution requires the gene-based expression of enzymes involved in the synthesis of phospholipids inside liposomes. The goal of this thesis is to realize synthetic cell compartment growth in this fashion. Based on literature, various requirements and necessary improvements on gene expression in liposomes and in vitro reconstitution of lipid synthesis are identified. These will guide the research in the remainder of this thesis, which is outlined here.

1.1 THE SYNTHETIC CELL

1.1.1 Life is complex and diverse

What is life? Throughout the ages, philosophers and scientists alike have asked this question[1–4]. However, even the efforts of so many great minds have not resulted in an unambiguous definition of a living system[5,6]. This is in no small part the consequence of life's staggering complexity. Even the simplest creatures we know, such as *Mycoplasma genitalium*, are governed by an intricate web of physicochemical reactions and processes[7]. This complexity confers the ability to maintain oneself, to reproduce oneself, and to evolve to increase fitness in a particular environment. Indeed, these three attributes are often employed to formulate a 'working definition' of life[4,8,9]. The variety of things that adhere to this definition is immense, from single-celled prokaryotes to huge elephants and enormous sequoia trees. Viruses, not able to perform self-maintenance or -reproduction, are not alive according to this definition, and most others[10,11].

A small set of features is shared amongst all living species currently known to biology. First of all, all life is built up from fundamental units called cells[12]. Some species consist of a single cell, others of billions. Every cell is a self-contained entity, clearly separated from its environment and other cells. This **compartmentalization** defines the cell by distinguishing it from the extracellular space[12]. Secondly, every living system is directed by some form of **information transfer**. In current biology, all information transfer takes places according to the central dogma of molecular biology[13]: DNA, which stores information in the form of genes, is transcribed into messenger RNA (mRNA), which is then translated into proteins[14] (**Fig. 1.1a**). Proteins are the 'workhorses' of life, as they provide catalysis and organisation[12,14]. Finally, all living systems are, in one way or another, capable of **proliferation**[12]. For single-celled organisms, this is represented by processes such as DNA replication, cell growth, and compartment division.

1.1.2 The synthetic cell

Of course, there is more to life than just these three features. However, it is hypothesized that a compartment encapsulating machinery to execute a genetic program encoding for proliferation by applying information transfer would represent the most simple living system possible: the minimal cell[8,15]. By building a minimal, or synthetic, cell, biology is studied according to a constructive paradigm[16,17]: learning by building. Constructing a synthetic cell is currently a topic of considerable scientific interest. This is exemplified by large consortia working on synthetic cells, such as BaSyC in the Netherlands and MaxSynBio in Germany[18]. Building a synthetic cell holds the promise to yield a fully understood living system and establish a detailed 'blueprint' of life. The availability of such a blueprint will give profound insight in the nature and evolution of life and elucidate the interplay between various modules of the cell. Furthermore, it will clarify what are the minimal requirements for a system to be alive. More practically, a synthetic cell could be used as a simple scaffold to investigate biological modules *in vivo* or to design custom species for biotechnological purposes[19].

Two general design strategies for the building of a synthetic cell can be distinguished. The first is the so-called top-down approach. Here, a living cell is used as starting point, and it is minimized by sequentially removing pieces of the DNA program encoded in the genome. Using this approach, the Craig Venter Institute has created a living organism, JCVI-3, with less genes than any naturally existing species known[20] (**Fig. 1.1b**). However, of these 473 essential genes, 149 are of unknown function, clearly demonstrating that the creation of a fully characterized synthetic cell via this approach is still beyond the horizon. Furthermore, the architecture of a cell minimized via the top-down approach will

always be dependent on the starting organism. Therefore, a cell minimized according to the top-down approach might reach a local complexity minimum, but not a global complexity minimum.

The alternative path towards a synthetic cell is called the bottom-up approach. Here, biomolecules, that are by itself not alive, are combined in such a way to create a living system[18,21]. These biomolecules can be of various biological origins, including viruses, as well as being (semi-) synthetic[22]. Currently, bottom-up minimal cell research has focused mostly on the reconstitution of individual biological modules, such as gene expression[16,23–28], DNA replication[29,30], inter-cell communication[31–34], cell growth[35–41], or cell division[22,42–45]. Combining these disparate modules, often built on incompatible synthetic cell scaffolds, in a functioning whole is a major challenge faced by the bottom-up approach[21,46,47]. Systematic integration of fully characterized modules, combined with extensive study of their interplay in synthetic cell intermediates, will provide a well-defined blueprint of the final synthetic cell. Thereby, black-box systems with many unknown parts, as resulting from the top-down approach, will be circumvented. Moreover, the wide range of starting points will make reaching a global complexity minimum more attainable than with the top-down approach.

1.1.3 Autonomy

As indicated by the requirements of *self*-maintenance and *self*-reproducibility, autonomy (from ancient Greek αυτονομος, self-governing), that is independence of outside control, is a key element of any living system. This autonomy principle has been explored by Varela et al. by defining a living system as a so-called autopoietic system[48]: *"a unity by a network of productions of components which (i) participate recursively in the same network of productions of components which produced these components, and (ii) realize the network of productions as a unity in the space in which the components exist."* So, in other words, an autopoietic system is a system capable of self-maintenance. They continue to state that "*autonomy is the distinctive phenomenology resulting from an autopoietic system: the realization of the autopoietic organization is the product of its operation".* Also, other attempts to logically define life, such as Von Neumann's theory on cellular automata[49] and Gánti's chemoton theory[50], (implicitly) state autonomy as a defining feature of living systems. Therefore, one should strive to build an autonomous synthetic cell, which is able to sustain and proliferate without direct input from or control by the experimentalist.

Defining autonomy as a necessary attribute for life has several practical consequences for the design of a synthetic cell. First of all, internal control is a key requirement for an autonomous synthetic cell. In all biological systems, internal control is provided by the execution of genetic programs encoded on DNA according to the central dogma. Therefore, a synthetic cell can be imbued with autonomy by using, as a key design principle, the execution of genetic programs, encoded either on DNA or a molecule with equivalent information storage properties. All attempts at building a minimal cell that circumvent the use of genetics will suffer from a lack of autonomy, and thereby never be alive. For example, division of liposomes with a mechanical splitter[51] (**Fig. 1.1c**) is not controlled by the liposome itself, but externally imposed by the experimenter

a

b

c
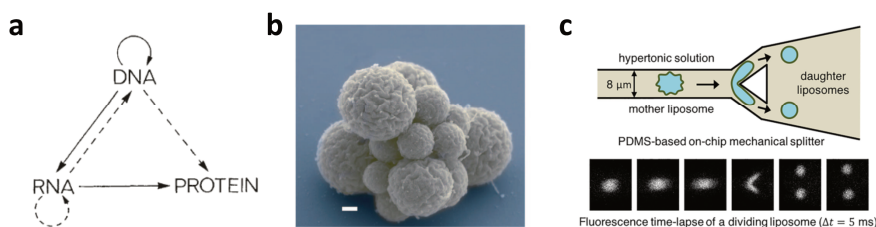
Figure 1.1. Autonomy is a central characteristic of life. (a) The central dogma of molecular biology, as schematically represented by Francis Crick in 1970[13]. The solid arrows represent the main modes of information transfer in biology. Information transfer via the central dogma imbues life with autonomy. **(b)** Scanning electron microscopy image of a small cluster of JCVI-3 cells, the most minimal autonomous organism[20]. Scale bar is 200 nm. **(c)** Dividing liposomes with a microfluidic flow splitter is an example of a reconstitution of a cellular process (division in this case) in a non-autonomous manner. Taken from Deshpande et al.[51]

switching on the flow. Therefore, the division process is not autonomous, and can never recapitulate the *self*-replication requirement of life.

Moreover, an autonomous minimal cell requires compartmentalization, to define the self and separate it from the external world. Indeed, to speak about internal control, it is necessary to separate the internal and external. A compartment furthermore provides an interface for interactions with the environment, such as the selective uptake of nutrients. Autonomy does not prohibit interactions with the environment but requires these interactions to be controlled by the autonomous living system.

1.1.4 Our synthetic cell design

Based on the aforementioned considerations, it is possible to sketch a design for an autonomous synthetic cell (**Fig. 1.2**). The design proposed here will motivate all decisions made in the subsequent parts of this thesis and serves as a guiding principle for all work in the Danelon lab.

Genetic programs in biology are ubiquitously encoded on DNA. Using RNA as a genetic template can simplify expression and replication schemes[52,53], but the instability of RNA makes it an unreliable carrier of information through large numbers of generations. In contrast, DNA is stable, and a plethora of methods to edit, process, and analyze DNA are readily available. Synthetic nucleic acid analogues, or XNA's[54], can also store genetic information, but lack the highly evolved gene expression and replication machinery of DNA, as well as standardized methods.

The use of DNA allows the direct application of gene expression machinery from current biology to generate proteins in a process called cell-free gene expression or *in vitro* transcription-translation (IVTT). Cell extracts are a simple and cost-effective system for cell-free gene expression[55–59], but are ill-defined, and suffer from contaminations and lack of reproducibility[60–63]. Cell extract formulations of increasing sophistication have been developed to counter these issues[24]. An alternative to cell extracts is the Protein synthesis Using Recombinant Elements (PURE) system[60,61]. The PURE system is a reconstitution of essential transcription and translation machinery, purified from *Escherichia coli*. All components originate from *E. coli*, except the T7 phage RNA polymerase and one enzyme in the energy regeneration pathway. Being based on purified components, the PURE system is well-defined and minimal, and therefore particularly suitable for synthetic cell research[16].
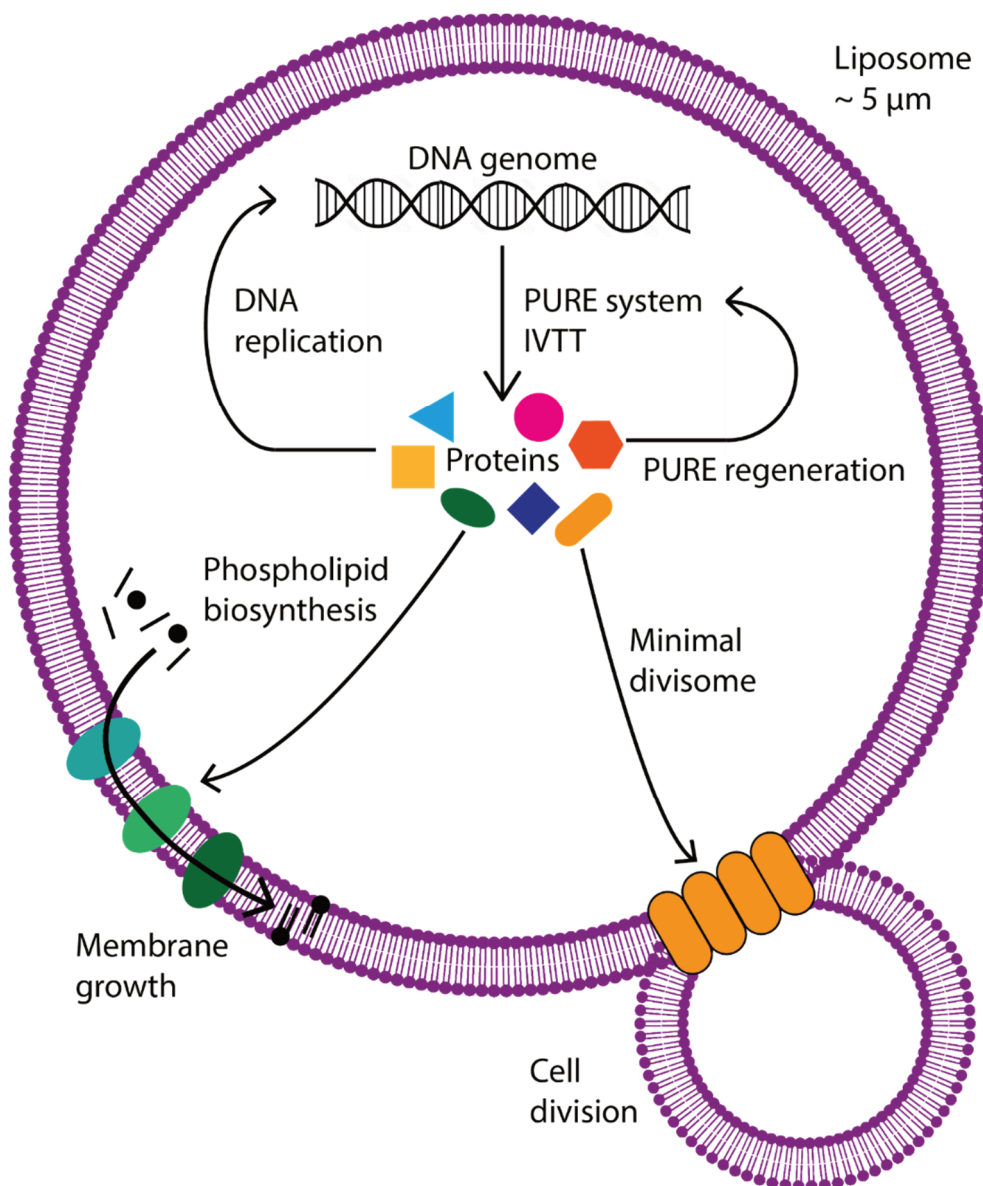
Figure 1.2: Synthetic cell design propounded by the Danelon lab and applied in the research presented in this thesis. Phospholipid vesicles with a diameter around 5 µm will be used as synthetic cell compartment. *In vitro* transcription translation (IVTT) of a DNA genome by the PURE system leads to the formation of proteins. These proteins are involved in regeneration of the PURE system, DNA replication, compartment growth by phospholipid synthesis, and cell division by expression of a minimal divisome. Regulatory elements, membrane pores, and deforming protein filaments are not shown, but are likely needed to successfully execute the core functionalities displayed in the figure.

Compartmentalization in nature is provided by cell membranes, bilayers of lipids (ether lipids in archaea, phospholipids in bacteria and eukaryotes) with embedded proteins[12]. When exposed to an aqueous solution, lipids self-assemble into bilayers due to their amphiphilic nature. A closed sphere is

1

the most energetically favorable state of a lipid bilayer in aqueous solution[64]. Such a sphere is called a liposome and closely resembles the cell membrane[15,28]. Liposomes are therefore particularly suited as minimal cell compartment[15,65]. Similar structures can be generated from amphiphilic polymers[66], or peptides[39], but these membranes might interfere with modules that are evolved to function within a lipid compartment *in vivo*. Vesicles formed of fatty acids are more permeable than liposomes and of high interest in origins-of-life research[67], but their low stability limits their use in synthetic cell research. Hybrid fatty acid-phospholipid vesicles can potentially be designed to have the stability and permeability required for synthetic cell purposes[68]. Water-in-oil droplets provide a simple compartment[27,69–71], but can only exist in oil-based environments, excluding interaction with the generally aqueous[72,73] world of biology. Moreover, Ostwald ripening severely compromises the lifetime of droplets as independent cells[74]. Coacervates have received considerable interest in synthetic- and protocell research[69,75], but they lack a defined interface with the environment, compromising their representation as an autonomous unit. Microfluidic chambers[76,77] are unsuitable as synthetic cells, since they cannot replicate, autonomously or otherwise.

Concluding, a liposome compartment, encapsulating DNA as information carrier, and performing cell-free gene expression with the PURE system, might represent the best-suited candidate to serve as a general scaffold for synthetic cell research. It is therefore applied as the framework for all research described in this thesis, and the Danelon lab as a whole. Reconstitution of proliferation is built on this framework. In order to proliferate, all constituents of the cell need to be replicated. To maintain autonomy, these processes need to be guided by the cell itself; in this framework, by the genes contained in the DNA. It is therefore necessary to have DNA programs for the (re)generation of cell-free gene expression machinery, for compartment growth by phospholipid synthesis, and the replication of the DNA itself. These processes together constitute the growth of the cell; moreover, a process needs to be in place to divide the mother cell into two or more daughter cells. This could be DNA-encoded, by expression of divisome proteins[44,45,78], but division might also spontaneously occur as a consequence of growth. For example, in nature L-form bacteria generate offspring by enhanced lipid synthesis, without the need for dedicated division systems[79].

1.1.5 Synthetic cell research without gene expression

Many research performed under the synthetic cell umbrella reconstitutes modules by using purified proteins instead of cell-free gene expression[22,43,80,81]. The utility of these projects in the greater synthetic cell field could be questioned. First of all, the environment in which the module is reconstituted might be incompatible with gene expression, for example due to buffer composition, nature of the compartment, or applied temperature. Moreover, the proteins used might not be expressible by a cell-free gene expression system, especially eukaryotic proteins needing post-translational modifications. Converting the purified protein-based approach to a gene expression-based approach might therefore be very difficult, if not impossible.

However, research based on purified proteins presents major advantages too. The method is particularly suited to study the effect of protein stoichiometry[82], a parameter that is often ill-controlled in gene-based experiments[45]. Moreover, *in vitro* reconstitution of a system with purified proteins facilitates detailed study into system behavior in a very minimal environment. The 'physics' of a module, comprising parameters like typical time and length scales, and applied forces, can be studied excellently by *in vitro* module reconstitution with purified proteins[83]. These are however answers to questions that are very different than the questions asked by synthetic cell research. Instead of investigating how all modules in a cell work together, which is the grand goal of synthetic cell research, *in vitro* reconstitution-type research attempts to answer how they work in isolation. I

would therefore argue for a clearer distinction between *in vitro* reconstitution and synthetic cell research.

Some synthetic cell research forgoes the use of proteins altogether. Compartment growth has been realized by inclusion of externally supplied fatty acids[84], or fusion with other liposomes[85,86]. Such growth mechanisms are simple, and therefore preferred by parts of the synthetic cell community[18]. However, since this manner of growth cannot be genetically encoded, it cannot proceed autonomously, and I therefore consider it an unviable approach to grow synthetic cells. In contrast, the synthesis of lipid analogues by native chemical ligation[87] has potential to be executed autonomously by coupling to the activity of a cell-free expressed enzyme[41].

1.2 TOWARDS AN AUTONOMOUSLY GROWING SYNTHETIC CELL MEMBRANE

One of the most important features of a synthetic cell is the ability to grow its liposome compartment. Since liposomes are made up of phospholipids, this requires the reconstitution of a phospholipid synthesis pathway. For the growth to be autonomous, the phospholipid synthesis pathway needs to be encoded on DNA and expressed inside liposomes. Moreover, continuous proliferation cycles require that synthesized phospholipids and the phospholipids that make up the initial synthetic cell membrane are identical. Therefore, the choice of lipid composition for the scaffold liposome and the choice of pathway to reconstitute are intertwined. The choice for lipid composition is also important to ensure liposome stability and proper interaction between reconstituted membrane-associated proteins and the membrane. Here, we will discuss pioneering studies on cell-free gene expression in liposomes, review phospholipid synthesis and membrane composition *in vivo*, and discuss efforts to reconstitute phospholipid synthesis *in vitro*.

1.2.1 Cell-free gene expression inside liposomes

Cell-free gene expression inside liposomes is a central part of synthetic cell research and is under investigation for almost two decades. Already in 2003, Nomura et al.[28] realized the expression of red shifted green fluorescent protein (rsGFP) by encapsulating an *E. coli*-based cell-free gene expression reaction inside liposomes formed by natural swelling (**Fig. 1.3a**). In 2004, Noireaux and Libchaber[57] provided the first kinetics of gene expression in liposomes, prepared by double emulsion transfer. Moreover, expression of the alpha-hemolysin (αHL) pore protein activated the membrane and facilitated *in vesiculo* expression of an αHL-GFP fusion for up to 100 h (**Fig. 1.3b**). A similar kinetics analysis, albeit which much shorter expression times, was performed in 2009 by Saito et al.[88], with expression of GFP in the PURE system. A major step forward was the study of a population (several hundreds) of liposomes, instead of one liposome at the time, resulting in the observation of significant liposome-to-liposome variability with respect to expressed GFP fluorescence (**Fig. 1.3c**). This observation has since been supported by theory[89] and is extensively studied experimentally[17]. Besides the expression of fluorescent proteins, in Liu et al.[90] the outflow of small fluorescent molecules via the expressed connexin-43 pore protein is harnessed as a sensitive measure of cell-free gene expression (**Fig. 1.3d**).

In the Danelon lab, a method based on lipid-coated glass beads has been developed to generate liposomes encapsulating the PURE system machinery. This method represents a modification of the traditional liposome preparation method, where a glass surface, such as the bottom of a round-bottom flask, is coated with lipids, which spontaneously swell to form liposomes upon hydration with

1

aqueous solution. Using ~300 µm glass beads as a coating surface greatly improves the ratio of lipid film surface to swelling solution, facilitating the use of small volumes of PURE system to be encapsulated inside liposomes. In contrast to emulsion transfer methods such as applied by Noireaux and Libchaber[57], no surfactants or oil phases, that could compromise biocompatibility, are used. Moreover, liposomes prepared by the lipid-coated beads method are smaller than when prepared with other techniques, bringing their volume in the range of bacterial cells. Using this technique, GFP expression inside liposomes was realized[16], gene expression stochasticity was investigated with a dual-reported assay[25] (**Fig. 1.3e**) and Spinach RNA aptamer technology was applied to enable concurrent measurement of mRNA and fluorescent protein concentration inside liposomes[26,91] (**Fig. 1.3f**).

All research described above studies the expression of fluorescent proteins, motivated by the expected direct correlation between measured fluorescence and protein concentration. However, direct quantification of expressed protein concentration in a population of liposomes has not been provided. Moreover, in most studies, only a low number of liposomes is studied. Because of the highly heterogenous nature of gene expression in liposomes, this will result in overlooking potentially relevant phenotypes with low incidence. Main reasons for this include the fact that most liposome preparation methods are optimized towards liposome quality more than quantity, and a lack of image analysis methods to perform high-content microscopy analysis.

1.2.2 Membrane composition and phospholipid synthesis *in vivo*

Most investigations of cell-free gene expression inside liposomes use vesicles with a simple lipid composition. Especially membranes consisting of zwitterionic phosphatidylcholine (PC) are a popular choice. However, reconstituted machinery is often derived from *E. coli*, and *E. coli* membranes do not contain any PC. Instead, their membranes consist of ~75% zwitterionic phosphatidylethanolamine (PE), and ~20% anionic lipids phosphatidylglycerol (PG) and cardiolipin (CL)[92,93]. Careful reconstitution of *E. coli*-based machinery requires recapitulation of the lipidic environment the machinery has evolved in, especially the net negative charge introduced by PG and CL.

In the minimized organism JCVI-3[20], the main phospholipid constituents of the membrane are PG, CL, and phosphatidic acid (PA)[94]. Since PA is an intermediate in the synthesis of PG and CL, this represents a very minimal cell membrane. Equally minimal is the enzymatic pathway to generate these lipids. The minimized genome of JCVI-3 encodes for six phospholipid synthesis proteins[20]. PlsX and PlsY catalyze the first step of phospholipid synthesis, the acylation of the 1 position of glycerol-3-phosphate (G3P), to form lysophosphatidic acid (LPA). The 2 position is subsequently acylated by PlsC to form PA. Acyl-CoA or acyl-acyl carrier protein (ACP) serves as the acyl chain donor in these reactions[95]. Then, CdsA activates the molecule by incorporating cytidine diphosphate (CDP) at the 3 position of the glycerol, yielding CDP-diacylglycerol (CDP-DAG)[96]. Next, PgsA catalyzes the replacement of cytidine monophosphate (CMP) by G3P[97]. PgpA then dephosphorylates the molecule to form PG[98,99]. Cardiolipin synthases (Cls) then bind two PG molecules together to form CL[100].
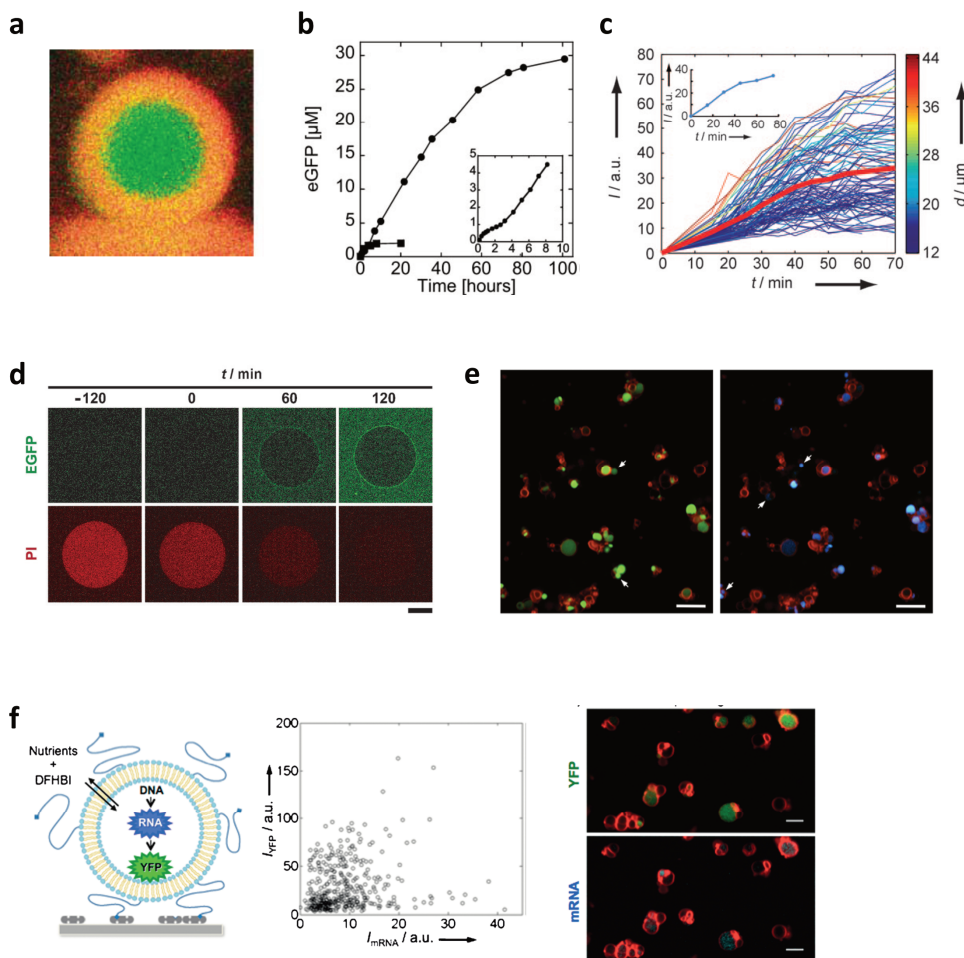
Figure 1.3: Cell-free gene expression inside liposomes. (a) A liposome, stained by hydrophobic fluorochrome, 4-Di-10-ASP (red), with internally expressed rsGFP, from Nomura et al.[28] **(b)** Concentration of expressed αHL-GFP (circles) and GFP (squares), over time, in a single liposome, from Noireaux & Libchaber[57]. **(c)** Fluorescence intensity of expressed GFP inside a population of liposomes, from Saito et al.[88] Thick red line indicates the population average, the inset represents GFP expression in a corresponding bulk reaction. **(d)** A liposome containing the fluorescent dye propidium iodide (PI, red) over time. DNA encoding for the fluorescent fusion pore protein Cx43-EGFP (green) is present outside of the liposome. At $t = 0$, the temperature is increased to 37 °C, and PURE system gene expression of the pore protein allows outflow of PI. Scale bar is 20 μm. From Liu et al.[90] **(e)** Liposomes prepared with the lipid-coated beads method (red), containing PURE system and DNA encoding for GFP (green, left panel) and CFP (blue, right panel). Liposomes indicated by an arrow express both DNAs. Scale bar is 10 μm. From Nourian & Danelon[25]. **(f)** In-liposome gene expression of DNA encoding for YFP and RNA with Spinach aptamer. Gene expression nutrients and 3,5-difluoro-4-hydroxybenzylidene (DFHBI), which binds to the Spinach aptamer, resulting in fluorescence, are supplied from the outside. The scatter plot shows mRNA (Spinach) fluorescence versus YFP fluorescence for individual liposomes and shows no clear correlation. The micrographs show a population of liposomes (red) with YFP (green, top) and mRNA (blue, bottom) fluorescence. Scale bar is 5 μm. From Van Nies et al.[26].

1
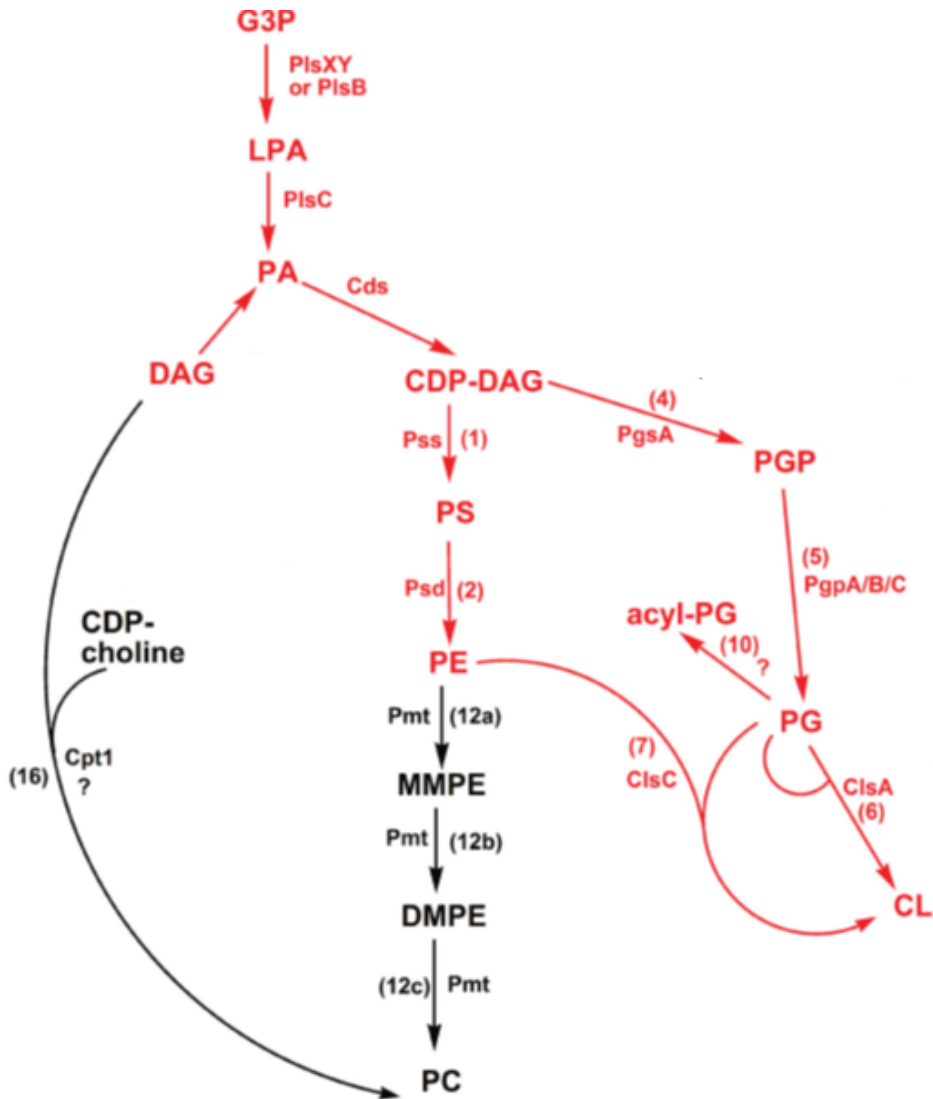
Figure 1.4: Bacterial phospholipid synthesis pathways. Highlighted in red is the Kennedy pathway, the canonical lipid synthesis pathway present in *E. coli*. In black are included the synthesis of PC from PE by Pmt(A) via a triple methylation reaction, and the PC salvage pathway employed by Schmidli et al.[104], whom, however, converted PA to DAG with phosphatidate phosphatase, whereas here, the opposite reaction is displayed, which is natively occurring in *E. coli*. Abbreviations not defined in the main text: MMPE: monomethyl phosphatidylethanolamine. DMPE: dimethyl phosphatidylethanolamine. CDP-choline: Cytidine diphosphate choline. Cpt1: cytidinediphosphocholine phosphocholinetransferase. Adapted from Sohlenkamp & Geiger[93].

In *E. coli*, the phospholipid synthesis pathway is slightly more advanced than the minimized pathway found in JCVI-3. An alternative pathway branch is present, where CDP-DAG is converted into phosphatidylserine (PS) by replacing CMP with the amino acid L-serine. This reaction is catalyzed by the enzyme PssA, which also regulates the ratio between anionic and zwitterionic lipids[101]. Psd then decarboxylates PS to form phosphatidylethanolamine (PE)[102]. Moreover, in *E. coli*, the role of PlsX/Y is performed by PlsB[95]. The complete, eight-enzyme, branched pathway leading to PE, PG, and CL (**Fig. 1.4**), is referred to as the Kennedy pathway, and is considered the canonical lipid synthesis pathway in bacteria[93,103].

1.2.3 Cell-free phospholipid synthesis

In vitro reconstitution of phospholipid synthesis with the aim of growing liposomes has been studied for almost three decades. In 1991, Luisi and coworkers managed to synthesize PC, reconstituting the PC salvage pathway (**Fig. 1.4**) by embedding four purified enzymes into the outer bilayer of PC liposomes[104]. In 1996, the same group managed to perform the first acyl transfer step inside liposomes by micro-injection of purified PlsB[38]. The synthesis of LPA resulted in a change in morphology of the liposomes[38].

Kuruma et al.[36] reconstituted PlsB and PlsC inside liposomes by cell-free gene expression with the PURE system. Although both enzymes were active individually, one-pot synthesis of PA from G3P and acyl-CoA was not realized. Scott et al.[37] resolved this problem, creating liposomes that synthesize PA internally based on a genetic program (**Fig. 1.5a**). Furthermore, liquid chromatography-coupled mass spectrometry was applied to quantitatively assess the synthesis of LPA and PA. Moreover, headgroup modification to form PE and PG was reconstituted by cell-free gene expression of the associated genes (*pssA, psd, pgsA, pgpA*) on the outside of phospholipid vesicles **(Fig. 1.5b)**. PE and PG synthesis were also reconstituted by Exterkate et al.[81], employing purified proteins embedded into liposome membranes from the outside. Moreover, they expanded the phospholipid synthesis pathway upstream by incorporation of FadD. This allowed for the use of high concentrations of oleic acid as acyl chain precursor, resulting in liposome growth visible with transmission electron microscopy **(Fig. 1.5c)**. However, only few example liposomes were shown, and no statistical analysis was performed, complicating interpretation.

The next step towards an autonomously growing synthetic cell membrane would be the reconstitution, inside liposomes, of the full pathways to generate the phospholipids that make up the liposome membrane. Several approaches to create a liposome that grows by synthesizing its own membrane constituents can be envisioned, balancing the trade-off between membrane stability and pathway complexity. Firstly, one can attempt to generate PA liposomes containing cell-free gene expression machinery to express PlsB and PlsC, as shown in Scott et al[37]. Alternatively, liposomes could be made of PE and PG, recapitulating inside the liposome lumen the headgroup modifying pathways expressed externally by Scott et al.. Expression of a cardiolipin synthase[100] would allow liposomes with membranes of PE, PG, and CL, which is the native composition of the *E. coli* membrane and likely interacts best with *E. coli* membrane-associated proteins. However, with all these lipid species, generating stable, micrometer-sized vesicles will be very challenging. To enhance liposome stability, a fraction of PC could be included. Synthesis of PC would require the expression of the enzyme PmtA[105], not found in *E. coli*.

Figure 1.5: Cell-free reconstitution of the Kennedy pathway in liposomes. (a) PURE*frex* synthesis inside phospholipid vesicles of PlsB and PlsC. Expression outside the liposomes is prevented by proteinase K inactivation of the PURE system. Oleoyl-CoA (o-CoA) and G3P are supplied from the outside of the liposome. Bar plots show concentration of synthesized LPA and PA after incubation. The yield of synthesized PA is not significantly lower when the reaction is confined to the liposome lumen, compared to when it is synthesized inside and outside. **(b)** Cell-free reconstitution of PE (top) and PG (bottom) headgroup modifying pathways. The bar plot indicates synthesis of PE and PG in the appropriate conditions. Panel (a) and (b) are adapted from Scott et al.[37] **(c)** Transmission electron micrographs of liposomes (1), liposomes with phospholipid synthesis enzymes including FadD and 0.3 mM oleic acid (2), and liposomes with phospholipid synthesis enzymes including FadD and 1.5 mM oleic acid (3). The two big liposomes in panel 3 are interpreted as having grown as a result of conversion of oleic acid into phospholipids. Taken from Exterkate et al.[81].

1.3 THESIS OUTLINE

The goal of this thesis is to reconstitute autonomous membrane growth of a synthetic cell. The core scaffold of an autonomous synthetic cell will be a cell-sized phospholipid vesicle encapsulating the PURE system as cell-free gene expression machinery and a DNA program. In **Chapter 2**, we present an improvement of the liposome preparation protocol based on lipid-coated beads, that can robustly generate tens of thousands of gene-expressing liposomes. The liposome membrane consists of PC, PE, PG, and CL. This system is quantitatively investigated by expressing simple DNA programs encoding for a fluorescent reporter protein and/or a pore-forming protein. High-content imaging of large populations of liposomes reveals rare favorable phenotypes, exemplified by high protein yield or long expression times.

The experimental framework developed in chapter 2 will be applied to reconstitute phospholipid synthesis inside liposomes. In **Chapter 3**, we demonstrate the functional expression, inside liposomes, of seven enzymes from the Kennedy phospholipid synthesis pathway from a single piece of DNA, resulting in the synthesis of PE and PG. Genetic and enzymatic regulation of the relative amounts of these membrane phospholipids represents a big step towards homeostatic membrane growth. Furthermore, we apply fluorescence microscopy methods to analyze the synthesis of lipids and their membrane incorporation on the single vesicle level.

Combined with PE and PG, PC makes up 98% of our liposome membrane. Therefore, conversion of PE into PC by the enzyme PmtA from *R. sphaeroides*, as described in **Chapter 4,** is a major milestone in the realization of autonomous membrane growth. Moreover, we found that cell-free expressed PmtA can remodel liposomes on the micrometer scale, resulting in an average size decrease and the formation of so-called 'septumed' liposomes.

In **Chapter 5**, we discuss various attempts to increase the yield of phospholipid production, in order to achieve liposomes that display significant growth. Expanding the phospholipid synthesis pathway upstream with the enzyme FadD enabled us to use oleic acid as a more soluble precursor but did not improve the final lipid yield. Using acyl-ACP as a more soluble precursor was unsuccessful, due to low concentration and purity of the substrate. Synthesizing lipids using a click-chemistry-based approach was fast and had a high yield, but greatly compromised cell-free gene expression activity.

Chapter 6 concerns the use of topologically active lipids to deform and divide gene-expressing liposomes subject to temperature cycles. Phase transition of lipid membranes resulted in liposome elongation, mediated by excess surface area, but division was not observed.

In **Chapter 7**, we will present a perspective on module integration, the process of combining separate synthetic cell parts into a functioning whole. Some of the challenges of module integration are highlighted in a test-case: the integration of phospholipid synthesis and φ29 DNA replication. Two approaches towards module integration are discussed: the traditional approach and the evolutionary approach. A hybrid between these two approaches is proposed as a potentially attractive compromise.

1

REFERENCES

1. Berryman, S. Democritus. *The Stanford Encyclopedia of Philosophy* (2016).

2. Kenny, A. J. P. & Amadio, A. H. Aristotle. *Encyclopædia Britannica* (2019).

3. Schrodinger, E. *What is Life?* (1944).

4. Oparin, A. I. *The origin of life on the earth*. (1957).

5. Schwille, P. Jump-starting life? Fundamental aspects of synthetic biology. *J. Cell Biol.* **210,** 687–690 (2015).

6. Szostak, J. W. Attempts to define life do not help to understand the origin of life. *J. Biomol. Struct. Dyn*. **29,** 599–600 (2012).

7. Fraser, C. M. et al. The Minimal Gene Complement of Mycoplasma genitalium. *Science*. **270,** 397–404 (1995).

8. Luisi, P. L., Ferri, F. & Stano, P. Approaches to semi-synthetic minimal cells: A review. *Naturwissenschaften* **93**, 1–13 (2006).

9. Nourian, Z., Scott, A. & Danelon, C. Toward the assembly of a minimal divisome. *Syst. Synth. Biol.* **8,** 237–247 (2014)

10. Moreira, D. & López-García, P. Ten reasons to exclude viruses from the tree of life. *Nat. Rev. Microbiol.* **7,** 306–311 (2009).

11. Luisi, P. L. & Stano, P. *The minimal cell.* (2013).

12. Alberts et al. *Essential cell biology (3rd edition).* Garland Science (2010).

13. Crick, F. Central Dogma of Molecular Biology. *Nature* **227,** 561–563 (1970).

14. Crick, F. H. On protein synthesis. *Symp. Soc. Exp. Biol.* **12,** 138–163 (1958).

15. Szostak, J. W., Bartel, D. P. & Luisi, P. L. Synthesizing life. *Nature.* **409,** 387–390 (2001).

16. Nourian, Z., Roelofsen, W. & Danelon, C. Triggered gene expression in fed-vesicle microreactors with a multifunctional membrane. *Angew. Chemie - Int. Ed.* **51,** 3114–3118 (2012).

17. Altamura E, Carrara P, D'Angelo F, Mavelli F and Stano P. Extrinsic stochastic factors (solute partition) in gene expression inside lipid vesicles and lipid-stabilized water-in-oil droplets: a review *Synthetic Biology* **3,** ysy011 (2018).

18. Schwille, P. et al. MaxSynBio: Avenues Towards Creating Cells from the Bottom Up. *Angew. Chemie Int. Ed.* **57,** 2-13 (2018).

19. Deamer, D. & Pohorille, A. Artificial cells: prospects for biotechnology. *TRENDS Biotechnol.* **20,** 123–128 (2002).

20. Hutchison, C. A. et al. Design and synthesis of a minimal bacterial genome. *Science*. **351,** aad6253 (2016).

21. Forster, A. C. & Church, G. M. Towards synthesis of a minimal cell. *Mol. Syst. Biol.* **2,** 45 (2006).

22. Osawa, M. & Erickson, H. P. Liposome division by a simple bacterial division machinery. *Proc. Natl. Acad. Sci. U. S. A*. **110,** 11000–4 (2013).

23. Blanken, D., van Nies, P. & Danelon, C. Quantitative imaging of gene-expressing liposomes reveals rare favorable phenotypes. *Phys. Biol*. **16,** 045002 (2019).

24. Garamella, J., Marshall, R., Rustad, M. & Noireaux, V. The all E. coli TX-TL toolbox 2.0: a platform for cell-free synthetic biology. *ACS Synth. Biol*. **5,** 344–55 (2016).

25. Nourian, Z. & Danelon, C. Linking genotype and phenotype in protein synthesizing liposomes with external supply of resources. *ACS Synth. Biol*. **2,** 186–193 (2013).

26. van Nies, P. et al. Unbiased tracking of the progression of mRNA and protein synthesis in bulk and in liposome-confined reactions. *ChemBioChem* **14,** 1963–1966 (2013).

27. Hansen, M. M. K. et al. Macromolecular crowding creates heterogeneous environments of gene expression in picolitre droplets. *Nat. Nanotechnol.* **11,** 1–8 (2015).

28. Nomura, S. I. M. et al. Gene Expression within Cell-Sized Lipid Vesicles. *ChemBioChem* **4,** 1172–1175 (2003).

29. Van Nies, P. et al. Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. *Nat. Commun*. **9,** (2018).

30. Libicher, K., Hornberger, R., Heymann, M. & Mutschler, H. In vitro self-replication and multicistronic expression of large synthetic genomes. *Nat. Commun*. **11,** 904 (2020).

31. Adamala, K. P., Martin-Alarcon, D. A., Guthrie-Honea, K. R. & Boyden, E. S. Engineering genetic circuit interactions within and between synthetic minimal cells. *Nat. Chem*. **9,** 1–9 (2016).

32. Niederholtmeyer, H., Chaggan, C. & Devaraj, N. K. Communication and quorum sensing in non-living mimics of eukaryotic cells. Nat. Commun. **9,** 5207 (2018)

33. Dupin, A. & Simmel, F. C. Signalling and differentiation in emulsion-based multi-compartmentalized in vitro gene circuits. *Nat. Chem.* **11,** 32-39 (2018).

34. Booth, M. J., Schild, V. R., Graham, A. D., Olof, S. N. & Bayley, H. Light-activated communication in synthetic tissues. *Sci. Adv*. **2,** 1–11 (2016).

35. Schmidli, P. K., Schurtenberger, P. & Luisi, P. L. Liposome-mediated enzymatic synthesis of phosphatidylcholine as an approach to self-replicating liposomes. *J. Am. Chem. Soc.* **113,** 8127–8130 (1991).

36. Kuruma, Y., Stano, P., Ueda, T. & Luisi, P. L. A synthetic biology approach to the construction of membrane proteins in semi-synthetic minimal cells. *Biochim. Biophys. Acta - Biomembr.* **1788,** 567–574 (2009).

37. Scott, A. et al. Cell-Free phospholipid biosynthesis by gene-encoded enzymes reconstituted in liposomes. *PLoS One* **11,** e0163058 (2016).

38. Wick, R. & Luisi, P. L. Enzyme-containing liposomes can endogenously produce membrane-constituting lipids. *Chem. Biol.* **3,** 277–285 (1996).

39. Vogele, K. et al. Towards synthetic cells using peptide-based reaction compartments. *Nat. Commun*. **9,** 3862 (2018)

1

40. Hardy, M. D. et al. Self-reproducing catalyst drives repeated phospholipid synthesis and membrane growth. *Proc. Natl. Acad. Sci.* **112,** 8187–8192 (2015).

41. Bhattacharya, A., Brea, R. J., Niederholtmeyer, H. & Devaraj, N. K. A minimal biochemical route towards de novo formation of synthetic phospholipid membranes. *Nat. Commun.* **10,** 300 (2019)

42. Caspi, Y. & Dekker, C. Divided we stand: splitting synthetic cells for their proliferation. *Syst. Synth. Biol.* **8,** 249–269 (2014).

43. Litschel, T., Ramm, B., Maas, R., Heymann, M. & Schwille, P. Beating vesicles: encapsulated protein oscillations cause dynamic membrane deformations. *Angew. Chemie - Int. Ed.* **57,** 16286–16290 (2018)

44. Furusato, T. et al. De novo synthesis of basal bacterial cell division proteins FtsZ, FtsA, and ZipA inside giant vesicles. *ACS Synth. Biol*. **7,** 953–961 (2018).

45. Godino, E. et al. De novo synthesized Min proteins drive oscillatory liposome deformation and regulate FtsA-FtsZ cytoskeletal patterns cytoskeletal patterns. *Nat. Commun.* **10,** 4969 (2019) .

46. Jia, H., Heymann, M., Bernhard, F., Schwille, P. & Kai, L. Cell-free protein synthesis in micro compartments: building a minimal cell from biobricks. *N. Biotechnol.* **39,** 199–205 (2017).

47. Göpfrich, K., Platzman, I. & Spatz, J. P. Mastering complexity: towards bottom-up construction of multifunctional eukaryotic synthetic cells. *Trends Biotechnol.* **36,** 938–951 (2018).

48. Varela, F. G., Maturana, H. R. & Uribe, R. Autopoiesis: The organization of living systems, its characterization and a model. *BioSystems* **5,** 187–196 (1974).

49. Von Neumann, J. The general and logical theory of automata. *Syst. Res. Behav. Sci.* 289–326 (2017).

50. Gánti, T. Organization of chemical reactions into dividing and metabolizing units: The chemotons. *BioSystems* **7,** 15–21 (1975).

51. Deshpande, S., Spoelstra, W. K., Van Doorn, M., Kerssemakers, J. & Dekker, C. Mechanical division of cell-sized liposomes. *ACS Nano* **12,** 2560–2568 (2018).

52. Ichihashi, N. et al. Darwinian evolution in a translation-coupled RNA replication system within a cell-like compartment. *Nat. Commun.* **4,** 2494 (2013).

53. Sumami, I., Ichihashi, N., Nishikawa, T., Kazuta, Y. & Yomo, T. Effect of liposome size on internal RNA replication coupled with replicase translation. *ChemBioChem* **17,** 1–9 (2016).

54. Pinheiro, V. B. & Holliger, P. The XNA world: Progress towards replication and evolution of synthetic genetic polymers. *Curr. Opin. Chem. Biol.* **16,** 245–252 (2012).

55. Mark S. Bretscher. Direct translation of a circular messenger DNA. *Nature* **220,** 1088–1091 (1968).

56. Nevin, D. E. & Pratt, J. A coupled in vitro transcription-translation system for the exclusive synthesis of polypeptides expressed from the T7 promoter. *FEBS Lett.* **291,** 259–263 (1991).

57. Noireaux, V. & Libchaber, A. A vesicle bioreactor as a step toward an artificial cell assembly. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 17669–74 (2004).

58. Krinsky, N. et al. A simple and rapid method for preparing a cell-free bacterial lysate for protein synthesis. *PLoS One* **11,** 1–13 (2016).

59. Kwon, Y. C. & Jewett, M. C. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci. Rep*. **5,** 1–8 (2015).

60. Shimizu, Y. et al. Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* **19,** 751–5 (2001).

61. Shimizu, Y., Kanamori, T. & Ueda, T. Protein synthesis by pure translation systems. *Methods* **36,** 299–304 (2005).

62. Noireaux, V., Maeda, Y. T. & Libchaber, A. Development of an artificial cell, from self-organization to computation and self-reproduction. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 3473–3480 (2011).

63. Doerr, A. et al. Modelling cell-free RNA and protein synthesis with minimal systems. *Phys. Biol.* **16,** 025001 (2019).

64. Phillips, R., Kondev, J., Theriot, J., Garcia, H. G. & Orme, N. *Physical Biology of the Cell.* (2012).

65. Blanken, D., Nies, P. Van & Danelon, C. Quantitative imaging of gene-expressing liposomes reveals rare favorable phenotypes. *Phys. Biol.* **16,** 045002 (2019).

66. Martino, C. et al. Protein expression, aggregation, and triggered release from polymersomes as artificial cell-like structures. *Angew. Chemie - Int. Ed*. **51,** 6416–6420 (2012).

67. Chen, I. A., Roberts, R. W. & Szostak, J. W. The emergence of competition between model protocells. *Science*. **305,** 1474–1476 (2004).

68. Jin, L., Kamat, N. P., Jena, S. & Szostak, J. W. Fatty acid/phospholipid blended membranes: A potential intermediate state in protocellular evolution. *Small* **14,** 1–9 (2018).

69. Qiao, Y., Li, M., Booth, R. & Mann, S. Predatory behaviour in synthetic protocell communities*. Nat. Chem.* **9,** 110–119 (2016).

70. Sakamoto, R., Noireaux, V. & Maeda, Y. T. Anomalous scaling of gene expression in confined cell-free reactions. *Sci. Rep.* **8,** 1–8 (2018).

71. Fanalista, F. et al. Shape and Size Control of Artificial Cells for Bottom-Up Biology. *ACS Nano* **13,** 5439–5450 (2019).

72. Spoelstra, W. K., Deshpande, S. & Dekker, C. Tailoring the appearance: what will synthetic cells look like? *Curr. Opin. Biotechnol*. **51,** 47–56 (2018).

73. Krieger, M. S., Sinai, S., Ferrari, R. & Nowak, M. A. Turbulent coherent structures and early life below the Kolmogorov scale. *Nat. Commun.* **11,** 2192 (2019)

74. Zwicker, D., Seyboldt, R., Weber, C. A., Hyman, A. A. & Jülicher, F. Growth and division of active droplets: A model for protocells. *Nat. Phys.* **13,** 408-414 (2017).

75. Deshpande, S. et al. Spatiotemporal control of coacervate formation within liposomes. *Nat. Commun.* **10,** 1800 (2019).

76. Karig, D. K., Jung, S. Y., Srijanto, B., Collier, C. P. & Simpson, M. L. Probing cell-free gene expression noise in femtoliter volumes. *ACS Synth. Biol.* **2,** 497–505 (2013).

1

1

77. Deshpande, S. & Dekker, C. Synthetic life on a chip. *Emerg. Top. Life Sci.* **3,** 559–566 (2019).

78. Godino, E. et al. Cell-free biogenesis of bacterial division proto-rings that can constrict liposomes. *bioRxiv* 2020.03.29.009639 (2020)

79. Mercier, R., Kawai, Y. & Errington, J. Excess membrane synthesis drives a primitive mode of cell proliferation. *Cell* **152,** 997–1007 (2013).

80. Schwille, P. Division in synthetic cells. *Emerg. Top. Life Sci.* (2019)

81. Exterkate, M., Caforio, A., Stuart, M. C. A. & Driessen, A. J. M. Growing membranes in vitro by continuous phospholipid biosynthesis from free fatty acids. *ACS Synth. Biol.* **7,** 153–165 (2017)

82. Loose, M., Fischer-friedrich, E., Ries, J., Kruse, K. & Schwille, P. Spatial regulators for bacterial cell. *Science*. **320,** 789–792 (2008).

83. Zieske, K. & Schwille, P. Reconstitution of pole-to-pole oscillations of min proteins in microengineered polydimethylsiloxane compartments. *Angew. Chemie - Int. Ed.* **52,** 459–462 (2013).

84. Chen, I. A. & Szostak, J. W. A kinetic study of the growth of fatty acid vesicles. *Biophys. J.* **87,** 988–98 (2004).

85. Terasawa, H., Nishimura, K., Suzuki, H., Matsuura, T. & Yomo, T. Coupling of the fusion and budding of giant phospholipid vesicles containing macromolecules. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 5942–5947 (2012).

86. Deshpande, S., Wunnava, S., Hueting, D. & Dekker, C. Membrane tension-mediated growth of liposomes. *Small* 15, 1902898 (2019).

87. Brea, R. J., Cole, C. M. & Devaraj, N. K. In situ vesicle formation by native chemical ligation. *Angew. Chemie - Int. Ed.* **53,** 14102–14105 (2014).

88. Saito, H. et al. Time-resolved tracking of a minimum gene expression system reconstituted in giant liposomes. *ChemBioChem* **10,** 1640–1643 (2009).

89. Frazier, J. M., Chushak, Y. & Foy, B. Stochastic simulation and analysis of biomolecular reaction networks. *BMC Syst. Biol.* **3,** 1–21 (2009).

90. Liu, Y. J., Hansen, G. P. R., Venancio-Marques, A. & Baigl, D. Cell-free preparation of functional and triggerable giant proteoliposomes. *ChemBioChem* **14,** 2243–2247 (2013).

91. Van Nies, P., Canton, A. S., Nourian, Z. & Danelon, C. Monitoring mRNA and protein levels in bulk and in model vesicle-based artificial cells. *Methods Enzymol* **550,** 186-214 (2015).

92. Raetz, C. R. H. & Dowhan, W. Biosynthesis and function of phospholipids in Escherichia coli. *J. Biol. Chem.* 265, 1235–1238 (1990).

93. Sohlenkamp, C. & Geiger, O. Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiol. Rev.* fuv008 (2015)

94. Breuer, M. et al. Essential metabolism for a minimal cell. *Elife* **8,** 1–77 (2019).

95. Yao, J. & Rock, C. O. Phosphatidic acid synthesis in bacteria. *Biochim. Biophys. Acta* **1831,** 495–502 (2013).

96. Sparrow, C. P. & Raetz, C. R. H. Purification and properties of the membrane-bound CDP-diglyceride synthetase from Escherichia coli. *J. Biol. Chem.* **260,** 12084–12091 (1985).

97. Hirabayashi, T., Larson, T. J. & Dowhan, W. Membrane-associated phosphatidylglycerophosphate synthetase from Escherichia coli: Purification by substrate affinity chromatography on cytidine 5' diphospho-1,2-diacyl-sn-glycerol sepharose. *Biochemistry* **15,** 5205–5211 (1976).

98. Icho, T. & Raetz, C. R. H. Multiple genes for membrane-bound phosphatases in Escherichia coli and their action on phospholipid precursors. *J. Bacteriol.* **153,** 722–730 (1983).

99. Lu, Y.-H. H., Guan, Z., Zhao, J. & Raetz, C. R. Three phosphatidylglycerol-phosphate phosphatases in the inner membrane of Escherichia coli. *J. Biol. Chem*. **286,** 5506–5518 (2011).

100. Hiraoka, S., Nukui, K., Uetake, N., Ohta, A. & Shibuya, I. Amplification and substantial purification of cardiolipin synthase of Escherichia coli. *J. Biochem.* **110,** 443–449 (1991).

101. Saha, S. K., Nishijima, S., Matsuzaki, H., Shibuya, I. & Matsumoto, K. A regulatory mechanism for the balanced synthesis of membrane phospholipid species in Escherichia coli. *Biosci. Biotechnol. Biochem.* **60,** 111–6 (1996).

102. Li, Q. X. & Dowhan, W. Structural characterization of Escherichia coli phosphatidylserine decarboxylase. *J. Biol. Chem.* **263,** 11516–22 (1988).

103. Cronan, J. E. Bacterial membrane lipids: Where do we stand? *Annu. Rev. Microbiol.* **57,** 203–224 (2003).

104. Schmidli, P. K., Schurtenberger, P. & Luisi, P. L. Liposome-mediated enzymatic synthesis of phosphatidylcholine as an approach to self-replicating liposomes. *J. Am. Chem. Soc.* **113,** 8127–8130 (1991).

105. Aktas, M. & Narberhaus, F. In vitro characterization of the enzyme properties of the phospholipid n-methyltransferase PmtA from Agrobacterium tumefaciens. *J. Bacteriol.* **191,** 2033–2041 (2009).

Chapter 2: Quantitative imaging of gene-expressing liposomes reveals rare favorable phenotypes

The biosynthesis of proteins from genomic DNA is a universal process in every living organism. Building a synthetic cell using separate biological parts hence implies to reconstitute a minimal gene expression apparatus and to compartmentalize it in a cell-mimicking environment. Previous studies have demonstrated that the PURE (Protein synthesis Using Recombinant Elements) system could be functionally encapsulated inside lipid vesicles. However, quantitative insights on functional consequences of spatial confinement of PURE system reactions remain scarce, which has hampered the full exploitation of gene-expressing liposomes as the fundamental unit to build an artificial cell. We report on direct imaging of tens of thousands of gene-expressing liposomes per sample allowing us to assess sub-population features in a statistically relevant manner. Both the vesicle size (diameter <10 μm) and lipid composition (mixture of phospholipids with zwitterionic and negatively charged headgroups, including cardiolipin) are compatible with the properties of bacterial cells. Therefore, our liposomes provide a suitable chassis to host the Escherichia coli-derived PURE translation machinery and other bacterial processes in future developments. The potential of high-content imaging to identify rare phenotypes is demonstrated by the fact that a subset of the liposome population exhibits a remarkably high yield of synthesized protein or a prolonged expression lifespan that surpasses the performance of ensemble liposome-averaged and bulk reactions. Among the three commercial PURE systems tested, PUREfrex2.0 offers the most favorable phenotypes displaying both high yield and long protein synthesis lifespan. Moreover, probing membrane permeability reveals a large heterogeneity amongst liposomes. In situ expression and membrane embedding of the pore-forming connexin leads to a characteristic permeability time profile, while increasing the fraction of permeable liposomes in the population. We see diversity in gene expression dynamics and membrane permeability as an opportunity to complement a rational design approach aiming at further implementing biological functions in liposome-based synthetic cells.

2.1 INTRODUCTION

2

Compartmentalization of cell-free protein synthesis has emerged as a scaffold to build a whole (semi)synthetic cell following a bottom-up approach[1-6]. Encoding functions onto genomic DNA[6] or RNA[7,8] provides a unique route to self-replication of the core machinery as alternative approaches relying exclusively on purified proteins and RNAs are unavoidably destined to death by dilution upon division events. This gene-directed strategy comes with a great system's complexity since the entire gene expression network itself comprises about hundred different components in its minimal form[9,10]. An example of minimal gene expression platform is the PURE (Protein synthesis Using Recombinant Elements) system, an *E. coli*-based *in vitro* transcription-translation (IVTT) system[9,10]. Its utilization in the context of building a minimal synthetic cell is particularly suited as it bypasses the endogenous complexity of cell lysates, thereby channeling resources (energy-rich phosphate donor molecule, nucleoside triphosphates and amino acids) for gene expression-related processes only. PURE system reaction kinetics have been monitored under a variety of initial conditions using different fluorescent reporters of mRNA and synthesized protein (see our accompanying paper[11] and references therein).

Various types of cell-like compartments have been proposed to enclose the biosynthesis apparatus, such as phospholipid vesicles (called liposomes)[1-6,12], water-in-oil emulsion droplets[7,13], peptide[14] and polymersome vesicles[15]. In particular, spatial confinement of gene expression systems inside liposomes provides a chassis that closely emulates the bacterial cytoplasmic environment. The mechanical properties of the phospholipid bilayers can be exploited to mimic cellular processes, e.g. membrane deformation and subsequent compartment division[16]. Moreover, the lipid membrane can serve as a native-like platform to integrate proteins and regulate the ionic and molecular exchanges with the environment[3,17-20]. Finally, recent progress on *in vitro* phospholipid synthesis by reconstituted bacterial enzymes represents an encouraging step towards liposome growth through internal production of its membrane constituents[5,21,22]. Hence, we postulate that liposome-based compartments are uniquely suited to recapitulate biological functions in synthetic cell research.

We identify three important requirements to develop a robust in-liposome gene expression platform, which will serve as a scaffold to further implement cellular functionalities. (i) One needs a predictive model describing the kinetics of mRNA and protein synthesis from single and multiple genes, as well as the key factors that govern performance (see our accompanying paper[11] and references therein). (ii) Liposome preparation methods need to be compatible with a large variety of lipids (e.g. scaffolding, signaling, membrane remodeling active lipids) and they should not alter the biological integrity of the entrapped compounds or the mechanical properties of the membrane. Moreover, the methods should generate cell-sized (1 µm < diameter <10 µm) unilamellar vesicles at a yield enabling high-throughput quantitative analysis of gene expression dynamics at the single liposome level. (iii) The new system's properties associated with liposome compartmentalization should be apprehended and harnessed. These include compositional heterogeneity or stochastic effects arising from low-copy number of some reactants[23-27], possible interaction with the lipidic surface[13], macromolecular crowding[28-30], effects of confinement on reaction rates[31], and membrane permeability governing molecular exchange with the external environment[3].

In this chapter, we established a robust protocol for preparing PURE system-containing liposomes. The method was applied to all three commercially available PURE system variants, i.e. PURExpress, PURE*frex* and PURE*frex*2.0 (see our accompanying paper[11] for details on their composition and bulk reaction properties). Gene expressing liposomes with sizes and lipid composition relevant to bacterial cell physiology have been produced at high yield and immobilized for long-term (> 12 h) time-lapse

fluorescence imaging of the yellow fluorescent protein (YFP) reporter. A disparity in translation kinetics and end-point protein levels between liposomes within a population has previously been observed[4,12,23,26,27,32] and theoretically predicted[33] (for a recent review, see ref. 34). The high-content liposome analysis performed here unravels new quantitative insights on stochastic enhancement and on production differences between the three types of PURE system.

Previous attempts to regulate the molecular exchange between the liposome interior and its environment (requirement (iii)) utilized membrane-spanning protein channels or transporters, such as α-hemolysin[3], MScL[18], connexin[17,19] and EmrE[20], exploited transient bilayer defects[4,23] or mild detergent at sub-solubilizing concentration[35]. Stable permeability with pore forming proteins could potentially alleviate the problem of energy and nutrient shortage inherent to sealed liposomes, with the promise of sustained gene expression[3]. We found an intrinsic heterogeneity of membrane permeability, demonstrated by the fact that liposomes coexisting in the same population take up, or not, a fluorescent dye with slightly larger size than relevant IVTT nutrients. We then expressed inside liposomes the connexin Cx43, a protein channel with a molecular mass cutoff of about 1.5 kDa[36], which is suitable for selective translocation of low-molecular weight substances. We showed that internal production of Cx43 does not lower expression of a second gene coding for YFP. Interestingly, functional reconstitution of Cx43 leads to distinct permeability kinetics with a sharp increase after about 60 min expression. However, no prominent increase of the YFP yield or translation lifespan could be observed, suggesting that the concentration of permeable nutrients available in the external solution is not limiting PURE*frex* reactions in liposomes.

2.2 RESULTS

2.2.1 High-throughput quantitative imaging of gene expression in individual liposomes smaller than ten micrometer reveals stochastic enhancement compared to bulk reactions

We first challenged requirement (ii) (see Introduction) by preparing bacterial cell-sized liposomes containing a mixed lipid composition comprising PC, PE, PG and cardiolipin phospholipids. Micrometer glass beads were used as a support for the lipid film to increase the total surface area and improve the yield of liposome production by natural swelling. The addition of rhamnose in the interspace between the stacked bilayers, combined with the presence of PEG-conjugated lipids, facilitate the osmotically driven flow of the aqueous solution into the lipid film, resulting in a high yield of unilamellar liposomes even in physiological buffers with a high ionic strength. Liposomes encapsulating PURExpress, PURE*frex* or PURE*frex*2.0, along with a linear DNA construct coding for the fluorescent reporter protein YFP, were produced. Gene expression was triggered by raising the temperature, and both the liposome membrane and synthesized YFP were imaged by laser scanning confocal microscopy (LSCM).

Tens to hundreds of liposomes can be visualized in single fields of view, resulting in >10,000 vesicles per chamber even after 16 h incubation at 37 °C (**Fig. 2.1a**). Liposomes of various morphologies and sizes are observed. While many vesicles are unilamellar, a significant fraction of multivesicular and multilamellar liposomes are also generated. Most liposomes have a diameter below 10 μm, satisfying the size criteria that we formulated above. Notably, liposomes encapsulating PURE*frex* or PURE*frex*2.0 are on average larger than those containing PURExpress (mean/median radii are 2.3/2.0 μm, 2.6/2.1 μm and 1.8/1.5 μm, respectively) (**Fig. 2.1c**).

We then examined the activity of the PURE systems by imaging YFP fluorescence after overnight expression. We applied an image analysis script to automatically identify the liposome lumina and to

quantify the amount of synthesized active YFP after converting the average lumen intensity into protein concentration (**Supplementary Fig. 2.2**). In all three PURE systems, about 30% of liposomes exhibit gene expression (determined as the signal above an intensity threshold, see Methods). Histograms of YFP concentration reveal a high liposome-to-liposome heterogeneity with all three PURE system variants (**Fig. 2.1d**). The mean/median concentration values are 0.5/0.3 µM, 3.5/1.7 µM and 2.6/1.1 µM for PURE*frex*, PURE*frex*2.0 and PURExpress, respectively. Interestingly, the distributions display a long tail with a fraction of liposomes outperforming the corresponding reaction in bulk (PURE*frex*: 5% with YFP > 1 µM, PURE*frex*2.0: 1.9% with YFP > 11 µM, PURExpress: 2% with YFP > 8.5 µM). Heat maps of YFP concentration plotted against liposome radius indicate no obvious relationship in PURE*frex* and PURE*frex*2.0 within the range of sizes analyzed (**Fig. 2.1e**). In PURExpress, liposomes with a diameter ≤3 µm seem to show higher expression levels but the trend is not prominent (**Supplementary Fig. 2.3**).

To obtain further quantitative insights about the number of YFP-expressing liposomes and sample heterogeneity, we performed flow cytometry measurements[37]. Using PUREfrex2.0, we found that the amount of endpoint expressing liposomes per microliter of swelling solution was ~75,000, corresponding to a concentration of $7.5 \cdot 10^7$ active vesicles ml^{-1} (**Fig. 2.1b**). Thus, the number of expressing liposomes per imaging chamber is ~150,000. In Eq. 8, we derived the expected relationship between the YFP (volume marker) and Texas Red (membrane marker) fluorescence intensity values assuming simple geometrical and dye partitioning constrains. Accordingly, a regression line with slope 1.5 (representing the volume-to-surface scaling factor of a sphere) was calculated and displayed on the fluorescence scatter plot shown in **Fig. 2.1b**). The spread orthogonal to this line corresponds to the extent of heterogeneity, or uncorrelated variation, with respect to theoretical predictions. For instance, data points with $\Delta\log(I\_YFP)/\Delta\log(I\_TexasRed) < 1.5$ may correspond to multilamellar or elongated vesicles. The results show strong uncorrelated variation of the YFP-Texas Red paired signals, consistent with the observation that liposomes of similar sizes display YFP concentrations that can span up to two orders of magnitude (**Fig. 2.1d,e**).
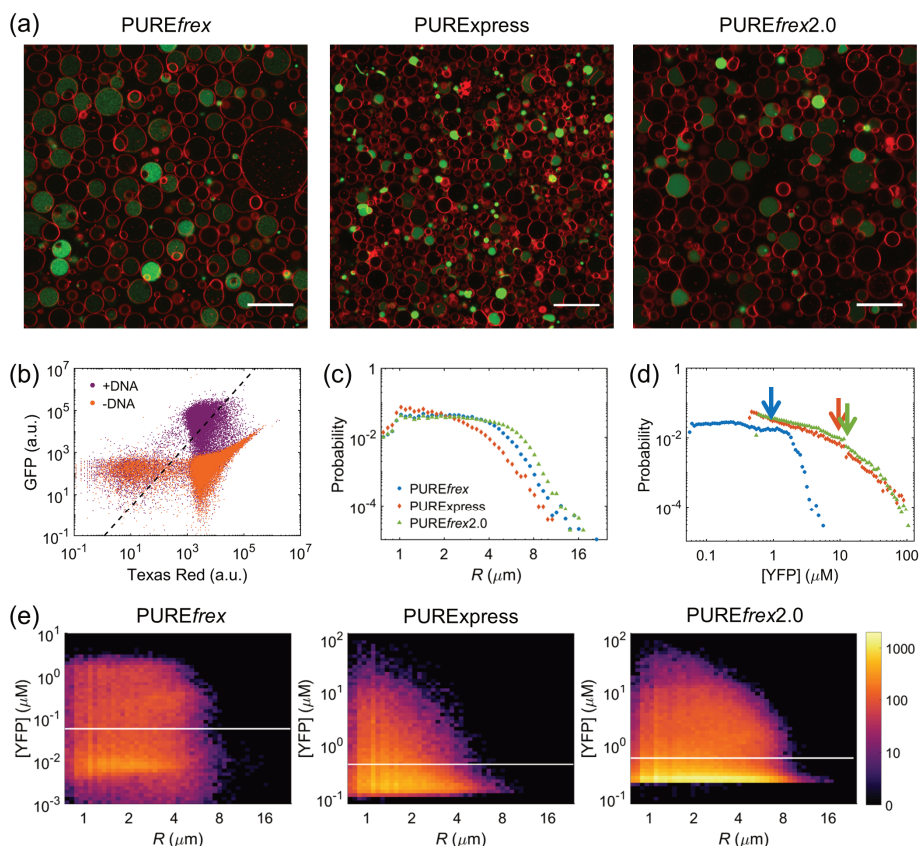
Figure 2.1: Confocal microscopy reveals stochastic enhancement of YFP expression in cell-sized vesicles. (a) Confocal micrographs of representative populations of liposomes (membrane in red) expressing YFP (green) in the three commercially available PURE systems. A higher laser power was used for PUREfrex, to compensate for the lower YFP signal. Scale bars are 20 μm. **(b)** Scatter plot of flow cytometry measurements of Texas Red and YFP fluorescence for liposomes containing PUREfrex 2.0 with (purple) and without (orange) *yfp* DNA. The appended dashed line with slope 1.5 represents the theoretical relationship between YFP and Texas Red signals in compliance with the assumptions used to derive Eq. 8. YFP-expressing liposomes correspond to the distinct purple cloud that appears in the +DNA sample. **(c)** Probability histogram of liposome radii (R) for PUREfrex (N = 96,924), PURExpress (N = 102,333) and PUREfrex2.0 (N = 209,870), each from six independent microscopy experiments. **(d)** Probability histogram of YFP concentrations for the liposome populations displayed in (c). Arrows indicate the corresponding bulk YFP yield[11]. The low-concentration limit for PURExpress and PUREfrex2.0 is shifted to higher concentration values compared to PUREfrex as a consequence of the lower laser power applied. **(e)** Heat maps of YFP concentration versus liposome radius for the three liposome populations shown in (d). White horizontal lines indicate the YFP concentration threshold above which liposomes are considered expressing, i.e.: 56.2 nM for PUREfrex, 437 nM for PURExpress and 573 nM for PUREfrex2.0. The color scale indicates the number of liposomes. Note that graphs are displayed with logarithmic scales in (c-e).

2.2.2 Single-vesicle gene expression kinetics reveal dynamic heterogeneity

2

Next, we asked whether the large heterogeneity in the end-point protein synthesis yield is related to differences in the apparent translation rate or expression lifespan[11]. We performed time-lapse LSCM and tracked the fluorescence intensity of YFP inside individual liposomes (**Fig. 2.2a**). Fluorescence time traces exhibit similar qualitative behavior as bulk reaction kinetics: an initial time lag, a linear regime with constant slope representing the apparent translation rate and a saturation phase with plateauing (**Fig. 2.2b**). Using phenomenological fitting, the apparent plateau time, translation rate and final yield of all individual kinetics were determined, their correlations analyzed, and their values compared to bulk measurements (**Fig. 2.2c**). For all three parameters, a large variability is observed within each PURE system condition. Expression times vary from ~45 min up to 4 h with both PURE*frex* and PURExpress, in analogy with bulk observations. In PURE*frex*2.0 prolonged translation times up to 20 h were measured, which is significantly longer than in bulk reactions (~4 h). Expression duration and YFP yield are found to be strongly correlated in all PURE systems (ρ = 0.60, 0.77 and 0.73 for PURE*frex*, PURE*frex*2.0 and PURExpress, respectively), meaning that longer-lived expression reactions lead on average to a higher yield of YFP. The apparent translation rate is also subject to large variability, with values ranging from ~4 nM min^{-1} to > 30 nM min^{-1}. Higher values are observed in PURExpress, but they remain on average lower than in bulk reactions. Apparent translation rate and YFP yield are slightly correlated (ρ = 0.40) for PURE*frex*, but not for PURExpress and PURE*frex*2.0. For the latter two PURE system variants, the apparent translation rate and the time to reach plateau are negatively correlated (ρ = − 0.62 and − 0.49, respectively), indicating that reactions with faster protein production cease earlier on average. No marked size dependence can be observed (**Fig. 2.2d**).

These results clearly underline the dynamical diversity of liposome-confined gene expression as a manifestation of the unique environment conferred by every single vesicle. Quantitative inspection of the kinetics parameters reveals that the factors governing protein synthesis are different for the three types of PURE systems. Yet the causes limiting gene expression might differ between in-liposome and bulk reactions, in particular with PURE*frex*2.0 and PURExpress (**Fig. 2.2c**).
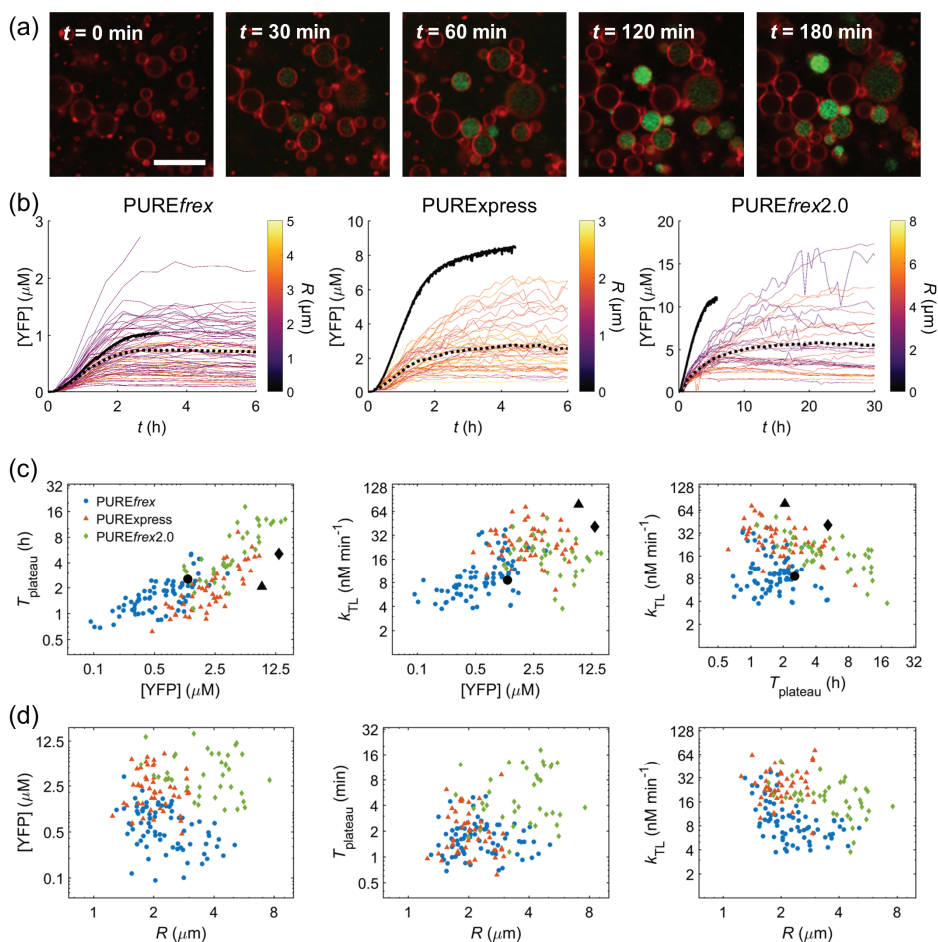
Figure 2.2: YFP concentration kinetics reveal heterogeneity in both expression lifespan and translation rate. (a) Time-lapse microscopy images of PURE*frex*-containing liposomes producing YFP at 37 °C. Scale bar is 10 μm and scale is identical in all frames. Some vesicles sedimented onto the glass surface in the time course of imaging. **(b)** Time traces of YFP concentration for PURE*frex* ($N = 74$), PURExpress ($N = 49$), and PURE*frex*2.0 ($N = 38$), each from two independent experiments. kinetic curves are color-coded with respect to the liposome radius R as indicated by the color bar. The black solid lines show the corresponding bulk gene expression kinetics. The dotted lines represent the average of the displayed liposome kinetics. **(c)** Scatter plots of final YFP concentration, expression lifespan $T_{plateau}$, and apparent translation rate k_{TL}, for the three PURE systems. Similar black symbols show the corresponding values from bulk expression kinetics. Parameters are determined by fitting a phenomenological model (see Methods). **(d)** Scatter plots of fitted parameters from (c), versus the radius R of the liposome. Color coding is the same as in (c). Note that logarithmic plots are shown in (c) and (d) to better visualize the spread of parameter values.

2

2.2.3 Differences in liposome permeability contribute to functional heterogeneity

Compositional diversity with regard to the copy number of DNA molecules and PURE system constituents are certainly factors contributing to the heterogeneity of gene expression in femtoliter liposomes. Another possible source of such a functional variability is the permeability state of the liposome membrane. We tested this hypothesis using PURE*frex*, as an extension of the detailed investigation of bulk reactions with this kit[11]. We first carried out a permeability assay using the Cy5 dye. It has a higher molecular mass than amino acids and nucleoside triphosphates but lower than any macromolecules in the PURE system. Therefore, it is a relevant fluorescent indicator of the permeability of low-molecular weight nutrients. The dye was supplied in the external solution, such that a fluorescence signal detected in the liposome lumen would reflect permeability across the lipid bilayer. No DNA was included in these experiments. We found that ~20 ± 6% (mean ± standard deviation, *n* = 2) of the liposomes do not contain the Cy5 dye and are thus impermeable (**Fig. 2.3a**). The same experiment was conducted with the *yfp* gene added (**Fig. 2.3b,c**). Combining the Cy5 signal (membrane permeability indicator) with the YFP fluorescence (gene expression reporter) and plotting the paired intensity values for every liposome on a heat map with four quadrants (Q1 impermeable/expressing, Q2 permeable/expressing, Q3 impermeable/non-expressing and Q4 permeable/non-expressing) reveal that ~40% of the expressing liposomes are impermeable (11% of the 29%), whereas only 16% of the non-expressing liposomes are impermeable (11% of 71%) (**Fig. 2.3c**). Said differently, 51% of the impermeable liposomes are expressing (11% of the 28%), whereas only 23% of the permeable liposomes are expressing (17% of the 77%). Within the expressing liposome category, no clear difference in the final YFP yield between the permeable and impermeable vesicles can be observed (**Fig. 2.3c**). Moreover, we did not find evidence for a correlation between liposome radius and permeability state.
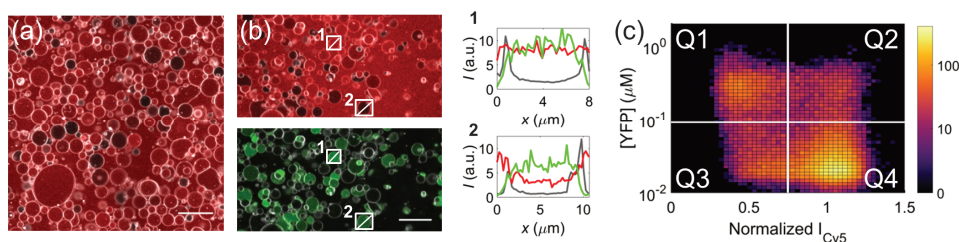
Figure 2.3: A significant fraction of YFP-expressing PURE*frex* liposomes shows spontaneous permeability to Cy5. **(a)** Fluorescence micrograph of a population of liposomes (membrane in white) with no encapsulated DNA. The Cy5 molecule (red) was added after liposome formation. Scale bar is 20 µm. **(b)** Microscopy images of liposomes (white) expressing YFP (green, bottom panel), with Cy5 (red, top panel) added after liposome formation. Scale bar is 20 µm. Line profiles of fluorescence intensity for the two liposomes highlighted: in grey the membrane signal, in red Cy5 and in green YFP. a.u., arbitrary unit. **(c)** Heat map of background-normalized Cy5 signal versus YFP concentration generated from 43,438 liposomes from one experiment. Color coding represents the number of liposomes. The white horizontal line corresponds to a YFP concentration threshold of 90 nM. The vertical white line corresponds to a normalized Cy5 intensity value of 0.75, which is defined as the threshold above which liposomes are considered filled with Cy5 and classified as permeable. Heat maps of four independent repeats are shown in **Supplementary Fig. 2.4**.

2.2.4 Expression of the pore-forming protein connexin-43 increases the fraction of permeable liposomes

We cannot exclude the possibility that the liposomes classified as non-expressing (~70% of the population) do actually enclose an active gene expression machinery but the concentration of synthesized YFP is too low for accurate detection (<103 nM). We reasoned that the functional expression and membrane incorporation of a pore-forming protein enabling the influx of the Cy5 dye (~790 Da) would provide a more sensitive readout. We chose the connexin-43 (Cx43) to stably permeabilize the liposome membrane[17]. We first validated the production of full-length Cx43 in a single-gene reaction or together with a second linear DNA coding for the YFP (**Fig. 2.4a**). Compartmentalized expression of the *Cx43* gene significantly reduces the number of Cy5-impermeable liposomes from ~22% (**Fig. 2.3a**) to less than 5% (**Fig. 2.4b**). This result indicates that a large fraction (~80%) of liposomes that are impermeable in the absence of pore-forming proteins can express the active Cx43, which elicits the uptake or efflux of permeable compounds along their concentration gradient. This percentage is significantly larger than the one calculated based on the correlation between the YFP and Cy5 signals, for which only 51% of the impermeable vesicles are found to express (**Fig. 2.3c**). This finding suggests the higher sensitivity of the Cx43-mediated uptake of a membrane-impermeable dye to probe gene expression activity in liposomes. Alternatively, the observation that a higher fraction of impermeable liposomes contains an active PURE system might be specific to Cx43 production and induced permeability, whereas individual *yfp* gene expression would be less effective in closed vesicles.

2.2.5 Functional expression of connexin-43 leads to membrane permeabilization without compromising YFP yield

Cx43-based gain-of-permeability was applied in combination with co-expression of the *yfp* gene to probe the influence of resource sharing and passive molecular diffusion across a protein channel on the IVTT performance. First, the bulk expression kinetics of active YFP were monitored with or without the Cx43-coding DNA template. Co-expression of the *Cx43* gene reduces the YFP final yield by ~40% (**Fig. 2.4c**), supporting the idea that some factors limiting productivity are common to the expression of the two genes (e.g. the concentration of translation initiation or elongation factors) and do not exclusively rely on the specific nucleotide or amino acid sequence. In contrast, the production period is not markedly affected (**Fig. 2.2b** and **Fig. 2.4h**). We next performed co-expression experiments with the *Cx43* and *yfp* genes inside liposomes. Given the bulk DNA concentrations (7.4 nM and 7.7 nM for the YFP- and Cx43-coding constructs, respectively) and assuming a Poisson distribution of DNA copies between liposomes, one expects on average ~150 DNA molecules of each gene inside a spherical liposome of 4 μm in diameter (volume 33.5 fL). This means that the probability for a liposome of that size to not contain at least one copy of each gene is negligible, even if one considers that only 10% of the DNA pool is transcriptionally active[23]. For smaller liposomes with a diameter of 2 μm, this probability increases to ~0.3 assuming 10% of active genes. Therefore, in the majority of liposomes, competition for nutrients and machinery is expected in order to produce both the YFP and Cx43 proteins.
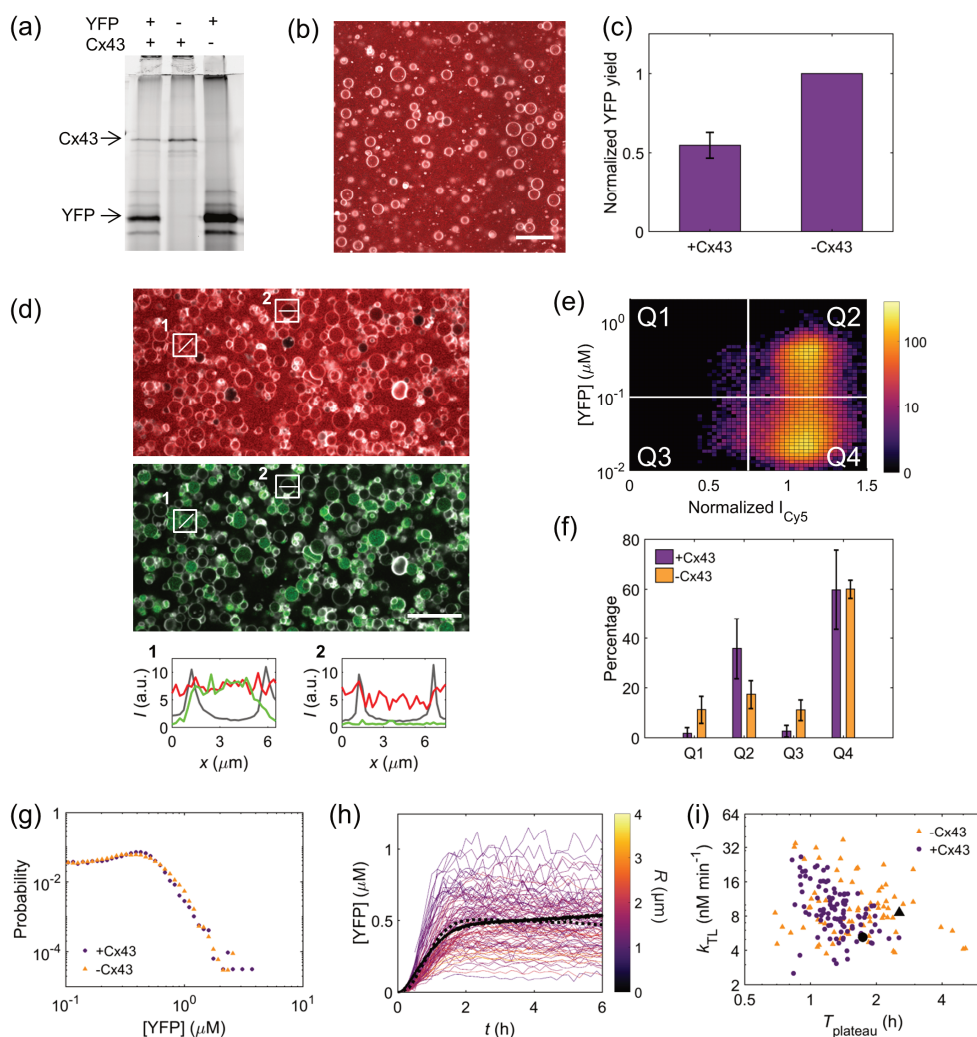
Figure 2.4: Co-expression of connexin-43 and YFP increases membrane permeability without compromising YFP yield. All experiments were performed with PUREfrex. (a) SDS-PAGE GreenLys gel of individually expressed or co-expressed YFP and Cx43 proteins. (b) LSCM image of liposomes (white) encapsulating PUREfrex and Cx43 DNA. The Cy5 dye (red) was added after liposome formation. Scale bar is 20 μM. (c) YFP expression yield obtained at the plateau of bulk reactions in the presence of Cx43 DNA, normalized to the end-point yield obtained in a minus Cx43 DNA bulk reaction conducted in parallel. Data are the average of two experiments, with error bar representing standard deviation. (d) LSCM images of liposomes (membrane in white) encapsulating the yfp and Cx43 genes. Both the Cy5 intensity (red, top panel) and YFP (green, bottom panel) signal were acquired simultaneously. Line profiles correspond to liposomes highlighted in the micrographs. (e) Heat map of normalized Cy5 signal versus YFP concentration for 39,046 liposomes from one experiment. Color coding represents the number of liposomes. The white horizontal line corresponds to a YFP concentration of 103 nM, above which liposomes are considered expressing. The vertical white line corresponds to a normalized Cy5 intensity value of 0.75, the threshold above which liposomes are considered filled with Cy5. Heat maps of two independent repeats are shown in **Supplementary Fig. 2.4**. (f) Mean fraction of liposomes in each quadrant, with roman numerals corresponding to the quadrants as indicated in (e), both in the presence (three independent repeats) and absence (five independent repeats, **Fig. 2.3c** and **Supplementary Fig. 2.4**) of Cx43. Error bar represents standard deviation. (g) Probability

histogram of YFP concentration in the presence (N = 84,260) and absence (N = 112,989) of Cx43. Logarithmic scales are displayed. (h) Time traces of YFP concentration for 100 liposomes containing YFP and Cx43 DNA. Color coding represents liposome radius R. Black line indicates bulk kinetics and the dotted line is the average of the displayed liposome kinetics. (i) Scatter plot of expression lifespan $T_{plateau}$ versus the apparent translation rate k_{TL}, as determined by fitting the time traces in (h) to a phenomenological model (purple symbols). The appended orange symbols correspond to PUREfrex data shown in **Fig. 2.2**. Black symbols are data from bulk experiments (c). Logarithmic scales are displayed.

Co-expressing *Cx43* and *yfp* genes within liposomes leads to a notable increase of the fraction of permeable liposomes (~72% without and ~95% with Cx43) (**Fig. 2.3c**, **Fig. 2.4b,d-f**), similarly as observed with the sole expression of the *Cx43* gene (**Fig. 2.3a** and **Fig. 2.4b**). This finding suggests that sharing the IVTT resources for YFP production does not compromise the PURE system's ability to generate membrane-permeabilizing Cx43 pores. On the other hand, co-expression of the Cx43 protein results only in a slight increase of YFP-producing liposomes (~30% without and ~37% with Cx43) (**Fig. 2.3c** and **Fig. 2.4e**), and the histograms of YFP concentration are similar with and without Cx43 co-synthesis (**Fig. 2.4g**). Together these results suggest that, in contrast to bulk reactions, sharing of resources does not limit expression of YFP in liposomes. Alternatively, other mechanisms could compensate for a possible loss of production due to resource sharing. For example, Cx43-mediated molecular permeability might ensure continuous supply of nutrients that would otherwise be depleted, while avoiding lethal accumulation of reaction byproducts in sealed liposomes.

To test whether the translation rate or/and the protein synthesis lifespan could be increased upon selective membrane permeability, we measured the kinetics of YFP expression with both *yfp* and *Cx43* genes concurrently expressed (**Fig. 2.4h**). No differences were observed compared to individual YFP expression experiments (**Fig. 2.4i**), ruling out the hypothesis that Cx43 functional insertion into the membrane could lead to prolonged gene expression in PURE*frex*.

2.2.6 Time course analysis of the fluorescent permeability indicator reveals distinct mechanisms causing liposome membrane permeability

To gain some insights on the time evolution of membrane permeability, we tracked the Cy5 fluorescence signal inside individual liposomes expressing both *yfp* and *Cx43* genes or solely *yfp*. In particular, we wondered whether the type of permeability induced by Cx43 differs from that occurring in the absence of a specialized protein channel. Different typical time profiles were monitored, which can each be associated with a certain permeability state of the membrane (**Fig. 2.5a**): (Phenotype I) Liposomes do not entrap Cy5 at the start of the kinetics and remain empty, indicating that the membrane is impermeable throughout the measurement period. (Phenotype II) Liposomes exhibit Cy5 signal that photobleaches at the same rate as the background (Cy5 in the outside solution), suggesting continuous membrane permeability. (Phenotype III) Liposomes initially contain Cy5 that photobleaches faster than the background signal, suggesting that the dye cannot freely exchange across the membrane. This category is more populated in liposomes that do not express YFP, hence likely not Cx43 (**Fig. 2.5c**). (Phenotype IV) Liposomes experience a sudden increase of Cy5 intensity at $t = t^*$ (**Fig. 2.5b**), after which bleaching occurs similarly as the background signal. This phenotype is most predominant in YFP-expressing liposomes with the encapsulated *Cx43* gene (**Fig. 2.5c**). Therefore, we attribute this behavior to the internal biosynthesis and functional assembly of stable Cx43 pores. Furthermore, t^* values are distributed around 60 min (**Fig. 2.5b**), a reasonable value for the expression and membrane insertion of the connexin pore. (Phenotype V) Liposomes display a sudden increase of Cy5 intensity at $t = t^*$, after which bleaching occurs at a faster rate than the background. Such transient permeability events occur randomly throughout the acquisition period

2

(**Fig. 2.5b**) and may arise from unstable membrane defects. (Phenotype VI) This category encompasses Cy5 fluorescence time traces that cannot be classified in one of the previous phenotypes, for example due to multiple filling events or to a filling event occurring at $t > 300$ min, for which the distinction between phenotypes IV and V cannot be assessed.
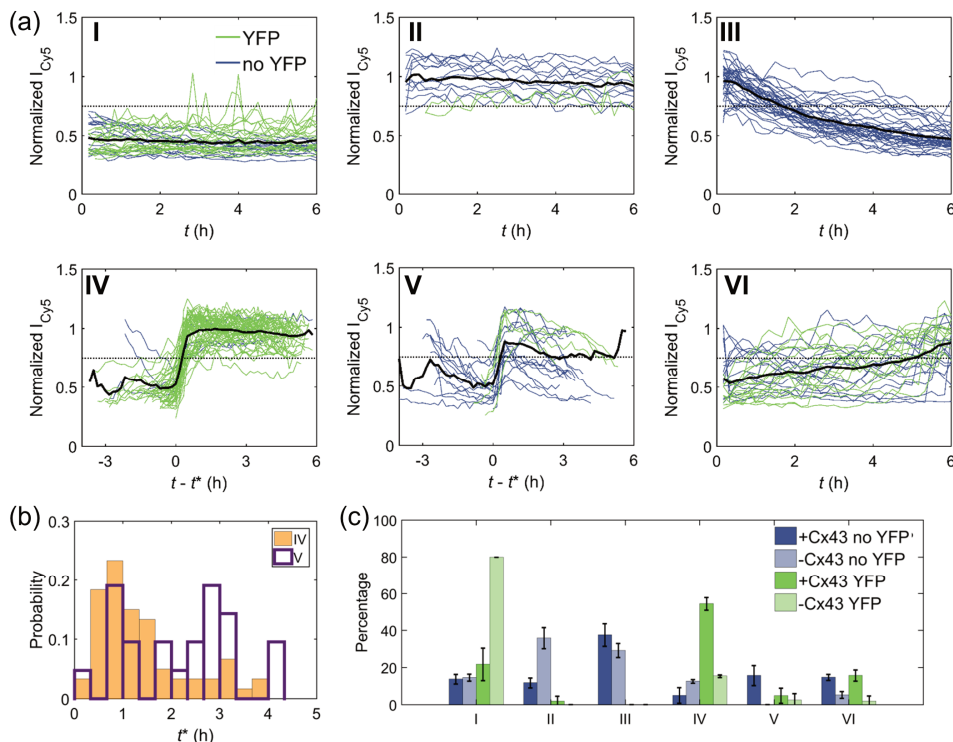
Figure 2.5: Expression of Cx43 inside liposomes leads to a specific permeability time profile. (a) Time traces of normalized Cy5 signal in liposomes expressing (green, $N = 100$) or not expressing (blue, $N = 100$) YFP, from two independent experiments. In both conditions, the *Cx43* gene was co-encapsulated. Black solid lines indicate average time traces. Dotted horizontal lines depict the Cy5 permeability threshold. Phenotypes, as indicated by roman numerals, are described in the main text. For **IV** and **V**, a time shift t^* is applied to align curves based on the occurrence of the filling event. **(b)** Probability histogram of t^* values, for phenotypes **IV** ($N = 60$ liposomes) and **V** ($N = 21$). **(c)** Occurrence of phenotypes, as percentage of the total number of time traces, for experiments with *Cx43* DNA (shown in a) and without (**Supplementary Fig. 2.5**). Error bars represent standard deviation between experiments. In total 400 traces were analyzed, 100 for each of the four conditions.

2.3 DISCUSSION

A milestone towards the realization of a synthetic minimal cell is to achieve robust gene expression inside liposomes, recapitulating cytoplasmic and membrane processes found in bacteria like E. coli. In the present study, we demonstrate the high-yielding production of cell-sized ≤10 μm) liposomes expressing genes by an encapsulated transcription-translation machinery. Protein-synthesizing phospholipid vesicles are stable for at least 16 h and can be immobilized at high density on a glass surface to facilitate long-time, high-throughput fluorescence imaging at the single liposome level. Most publications on gene expression inside liposomes display images with only a few vesicles, with severe limitations on the ability to derive quantitative information with statistical relevance, in particular when it comes to identify rare phenotypes. The high-content liposome analysis described here enabled quantification of essential properties, such as the expression yield, apparent kinetic parameters, membrane permeability states, and revealed large vesicle-to-vesicle heterogeneity. Using flow cytometry as a complementary analytical technique[37], we estimated the concentration of gene-expressing liposomes ($7.5 \cdot 10^7$ vesicles ml^{-1} with PURE*frex*2.0) and supported the uncorrelated variation in the amounts of expressed YFP and membrane dye. Regarding the protein synthesis yield and the apparent translation lifespan as measures of the IVTT system's performance, we recommend utilizing PUREfrex2.0 for liposome-compartmentalized reactions. Furthermore, triggered membrane permeability was achieved by expressing the non-*E. coli* protein Cx43 that acts as a general diffusion pore.

Various liposome preparation methods have been described in the literature (recently reviewed in ref. 38): lipid film hydration[39] and its many variants ((lipid-coated powder[40] or glass-bead[4] carriers, electroformation[41], hydrogel-based[42-44]), emulsion droplet-phase transfer[45-47], and microfluidic-based approaches[19,48-53]. In particular, the water-in-oil emulsion transfer and microfluidic-assisted methods have recently gained popularity in the liposome community, largely because of their ability to generate monodisperse and unilamellar vesicles. However, a vast majority of them produce giant vesicles with a diameter >10 μm. While such big liposomes can easily be manipulated and imaged by standard optical microscopes, they might not provide a physiological environment when it comes to recapitulate cytoplasmic and membrane processes (e.g. the E. coli-based translation machinery of the PURE system) occurring in bacteria with a size of a few micrometer only. In fact, a number of geometrical cues are known to drive biological functions in living cells[54,55]: the volume and the actual copy number of biochemical reactants, surface-to-volume ratio which regulates the molecular diffusion across the membrane as well as protein and RNA interactions with the lipid surface, confinement effects that influence biochemical reactions as a function of the order of the reaction, macromolecular crowding that is more prominent in small liposomes[28], and membrane curvature that affects lipid sorting and membrane protein functions[56].

Moreover, microfluidics and water-in-oil droplet phase transfer methods to form liposomes employ mineral oil or organic solvent that may remain in the form of residual traces in the hydrophobic core of the bilayer[48,57,58]. This may lead to altered membrane characteristics, such as lateral diffusion and flip-flop of lipids, microdomain formation, mechanical properties, permeability and biological activity of membrane-embedded proteins[38,47]. Actually, in the vast majority of the microfluidics-based vesicle formation studies, the possible presence of residual amounts of oil or solvent (not visible pockets) is not examined, despite the proven analytical efficacy of Raman spectroscopy and microscopy techniques[19,57]. In contrast, our glass bead-assisted natural lipid film swelling approach ensures unaltered membrane properties, which is essential to equip liposomes with more elaborate membrane-dependent functionalities. As a trade-off, however, a significant fraction of the liposomes

2

generated with the present method are multilamellar or multivesicular, and the vesicle size distribution is polydisperse (**Fig. 2.1**), which complicates fully automated image analysis and discards lipids in nonfunctional liposomes. This caveat is mitigated if one considers the large number of gene-expressing liposomes and the opportunity offered by such a heterogeneity. Indeed, in a single sample, a large combinatorial space of molecular inputs and geometrical parameters can be explored, increasing the likelihood to find liposomes exhibiting desired features, e.g. expression kinetics or vesicle morphology.

Our work shows that confinement of gene expression in femtoliter liposomes leads to higher heterogeneity of the apparent kinetic parameter values compared to bulk reactions (**Fig. 2.2** and accompanying article[11]). It is unclear whether the inhibitory mechanisms unraveled in bulk measurements similarly affect protein synthesis in liposome-confined reactions. The compositional heterogeneity and low-copy number effects arising from the stochastic entrapment of PURE system constituents make every liposome unique. Hence, the parameters that control the dynamics of the network and govern its performance might differ from one liposome to the other. Note that the absolute number of macromolecules can on average be > 100 (as calculated above for DNA molecules), yet the effective amount might be significantly lower if one considers partial inactivation[23], association with multiple partners into distinct complexes (e.g. translation factors, including tRNAs), or deviation from a Poisson encapsulation profile[28]. Intriguingly, a subset of liposomes express remarkably high protein concentration (**Fig. 2.1d**) or have a long translation lifespan (**Fig. 2.2**) surpassing the performance of any of the bulk reactions that we monitored so far[11]. These exceptional properties certainly result from peculiar initial conditions and membrane permeability state. Further studies should investigate the molecular determinants that confer such properties. For instance, individual PURE system constituents could be fluorescently labelled, and their encapsulation profile correlated with gene expression properties. The presence of distinct liposome sub-populations with permeable or impermeable membrane to small charged solutes has already been reported, albeit with a different lipid composition and liposome formation protocol[59]. Probing membrane tension with a fluorescent lipid reporter[60] could unveil the mechanical and structural properties associated to the permeability of gene expressing liposomes. Moreover, we propose that this heterogeneity should be exploited – instead of being seen as a drawback – e.g. by mapping in-liposome activity (amount of synthesized protein or expression kinetic parameters) as a function of the internal concentration of key components in one-pot reactions. This approach will aid the implementation of biological subsystems whose functioning relies on a narrow range of initial conditions that would hardly be tractable in batch mode screening reactions.

Extending this work to achieve sustainable RNA and protein synthesis through continuous liposome feeding from externally supplied nutrients (i.e. creatine phosphate, nucleoside triphosphates, amino acids) will be an important milestone towards the creation of a semisynthetic minimal cell. Long-lived gene-expressing liposomes, as reported in one study with an E. coli cell lysate entrapped inside liposomes with membrane-embedded α-hemolysin pores[3], was not realized in PUREfrex under the tested conditions (**Fig. 2.4h,i**). Next to the different type of membrane channels used in the two studies, it is likely that PUREfrex reactions in the tracked liposomes are limited by other factors than shortage of low-molecular weight substrates or accumulation of inhibiting byproducts. This hypothesis is consistent with results obtained for PUREfrex bulk reactions in a dialysis chamber[11]. Further developments of liposomes actively interfacing with their environment may take advantage of in situ expressed Cx43 to permit the entrance of membrane-impermeable molecules involved in transcriptional regulation, signaling, sensing or communication between liposomes.

2.4 MATERIALS AND METHODS

2.4.1 DNA constructs

A linear DNA template encoding for the eYFP (enhanced yellow fluorescent protein) was used to enable direct fluorescence readout of the level of synthesized protein[32,61]. Transcriptional elements include the T7 promoter and terminator, and the eYFP sequence was codon-optimized for expression in the *E. coli*-based PURE system[61]. The encoded Spinach RNA aptamer was not exploited for quantitative analysis of mRNA production (**Supplementary Fig. 2.2e**). Linear DNA constructs were regularly prepared by polymerase chain reaction using a circular plasmid as template, as also described in our accompanying article[11]. The linear Cx43 construct was prepared from a plasmid coding for the Cx43-eGFP fusion protein[17]. Forward (5'-GCGAAATTAATACGACTCACTATAGGGAGACC-3') and reverse (5'-CAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCTAGTCAGATTTCCAGATCATCAGGACGCG-3') primers were used to remove the eGFP sequence.

2.4.2 Liposome preparation

To prepare liposomes containing a PURE system reaction mixture, we optimized a method previously described by our lab[4,5]. The experimental workflow consists of four main steps: (i) preparation of lipid-coated beads, (ii) swelling of the lipid film by gentle rehydration with the PURE system solution, (iii) freeze-thaw cycles to break multilamellar structures and increase encapsulation efficiency, (iv) immobilization of liposomes on a coverslip, followed by triggering and monitoring gene expression by time-lapse fluorescence microscopy.

The lipid mixture deposited on the glass beads (212-300-µm acid washed glass beads, Sigma Aldrich) consists of 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC), 1,2-dioleoyl-*sn*-glycero-3-phosphatidylethanolamine (DOPE), 1,2-dioleoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) (DOPG), 1',3'-bis[1,2-dioleoyl-*sn*-glycero-3-phospho]-*sn*-glycerol (cardiolipin), supplemented with 1,2-distearoyl-*sn*-glycero-3-phosphoethanolamine-N-[biotinyl(polyethylene glycol)-2000] (DSPE-PEG-biotin) and Texas Red 1,2-dihexadecanoyl-*sn*-glycero-3-phosphoethanolamine, triethylammonium salt (TR-DHPE). All lipids were from Avanti Polar Lipids, except TR-DHPE that was purchased from Thermo Fisher Scientific. A lipid matrix with oleoyl (18:1) acyl chains and sub-zero phase transition temperature was chosen as it enables natural lipid film swelling on ice, which prevents gene expression from starting. A mixture of 2.54 mg DOPC (50 mol %), 1.68 mg DOPE (36 mol %), 0.58 mg DOPG (12 mol %), 0.20 mg cardiolipin (2 mol %), 0.05 mg DSPE-PEG(2000)-biotin (1 mass %) and 0.025 mg TR-DHPE (0.5 mass %), all dissolved in chloroform, was prepared in a round-bottom flask. The lipid-containing chloroform solution was mixed with 100 mM rhamnose dissolved in methanol in a 2.5 to 1 volume ratio, corresponding to ten rhamnose molecules per lipid. Then, 1.5 g of glass beads was poured to the lipid-sugar mixture and the solution was subjected to 2 h of rotary evaporation at 200 mbar and room temperature, followed by overnight desiccation, resulting in beads coated with stacked lipid bilayers interspersed with rhamnose. The addition of the hydrophilic sugar rhamnose to the lipid film, combined with the presence of PEG-modified lipids, promotes separation of adjacent bilayers and formation of unilamellar vesicles[5,62].

PURE*frex* and PURE*frex*2.0 are from GeneFrontier Corporation (Chiba, Japan). PURExpress is from New England Biolabs. PURE*frex*, PURE*frex*2.0 and PURExpress reaction mixtures were prepared in 0.5-mL Eppendorf tubes according to the guidelines provided by the manufacturers. The *yfp* DNA template

2

was supplemented at a concentration of 5 ng/μL (final concentration 7.4 nM). When indicated, 6.25 ng/μL of the *Cx43* construct was added. All reactions were supplemented with SUPERase RNAse inhibitor (0.75 U/μL final, Invitrogen). An amount of lipid-coated beads corresponding to 1.25 mg/μL was added to the 10-μL PURE system solution. Liposomes were formed by spontaneous swelling of the lipid film for 2 h on ice, during which the samples were regularly subjected to manual tumbling. Four freeze-thaw cycles were finally applied by dipping samples in liquid nitrogen in order to increase liposome unilamellarity and encapsulation efficiency.

2.4.3 Surface immobilization of liposomes

Reaction chambers were custom-made by sandblasting two holes of 5 mm in diameter in a 1-mm thick microscopy slide (Thermo Scientific). A 150-μm thick coverslip (Thermo Scientific) was glued on the microscopy slide with Norland Optical Adhesive 61 (Norland) to cover the holes on one side forming the bottom of the chamber. The sample reservoir was functionalized with 1:1 molar ratio of BSA and BSA-biotin (Thermo Fisher Scientific), and then with Neutravidin (Sigma Aldrich), to tether the biotinylated liposomes[4]. About 2 μL of the liposome solution was carefully pipetted (with a cut tip to avoid damaging the liposomes) in the imaging chamber and diluted with 5 μL of feeding solution consisting of 4:7 vol/vol of PURE*frex* Sol I, PURE*frex*2.0 Sol I or PURExpress A and Milli-Q supplemented with 83 mg/L Proteinase K. For permeability assays, 2 μL of Milli-Q was replaced by an equal volume of the dye Cy5 (stock 20 μM), giving a final concentration of 3.33 μM. The second reservoir was filled with water to saturate the air thereby limiting sample evaporation during incubation and imaging. Chambers were closed using a 0.5-mm thick double-sided adhesive silicone sheet (Life Technologies) and a cover slide, creating an interconnecting air volume, which enables oxygen supply to the reaction.

2.4.4 In-liposome gene expression and image acquisition

Image acquisition was performed using a Nikon A1R Laser scanning confocal microscope (LSCM) equipped with an SR Apo TIRF 100 × oil immersion objective. Samples were mounted in a stage preheated to 37 °C and imaged using the laser lines 514 nm (YFP), 561 nm (Texas Red) and 640 nm (Cy5), with appropriate emission filters. The distance between the sample and the objective was adjusted manually in order to equatorially dissect as many liposomes as possible. For kinetics measurements, the liposome sample was mounted in the microscope directly after sealing the chamber. An automated image acquisition protocol was run for at least 6 h, taking an image every 10 min. For end-point measurements, the sample chambers were first incubated overnight at 37 °C in a plate incubator, before being transferred to the microscope. Several YFP acquisition settings were applied in order to make optimal use of the dynamic range of the detector. Each setting was calibrated individually (see below). Montages of six-by-six fields of view, stitched with 20% overlap, were acquired. For kinetics experiments, gene expression was triggered while liposomes were immobilizing on the surface to minimize the delay time before imaging (about 10 min).

2.4.5 High-throughput image analysis

High-throughput analysis of microscopy images was performed using a custom-made MATLAB (MathWorks) code based on the *flood fill* (imfill.m) algorithm supplied with the MATLAB Image Analysis toolbox. The main steps are illustrated in **Supplementary Fig. 2.1**. In short, after applying a sharpening filter on the Texas Red image, the *flood fill* algorithm increases the intensity value of dark pixels surrounded by lighter pixels, effectively 'filling up' the liposomes. After subtraction of the original Texas Red image and applying a binary threshold, the resulting image serves as a marker for the liposome lumina (**Supplementary Fig. 2.1**). To ensure that only liposomes are marked, and not debris or noise, the image is first eroded, and regions that do not satisfy the following criterion are dismissed:

$$0.5 < \frac{P^2}{4\pi A} < 2 \qquad\qquad (\text{Eq. } 1)$$

where *P* is the perimeter of the region-of-interest (ROI) and *A* the area. This criterion is established to only select circle-like shapes. The resulting marker is used to identify pixels corresponding to liposome lumina in the YFP and Cy5 channels. The average intensity value of each region is measured in both channels. In the case of YFP the intensity is converted to concentration (see below). In the case of Cy5, this intensity value is normalized by dividing by the Cy5 fluorescence of five randomly selected extra-liposomal ROI's per field of view. For analysis of time-lapse images, the lumen of individual liposomes was traced manually using Fiji[63], and the average fluorescence intensity in both YFP and Cy5 channels was recorded at every time point.

The radius of each liposome in micrometer is estimated as:

$$R = 0.25 \times \left(2 + \sqrt{\frac{3A}{2\pi}}\right) \qquad\qquad (\text{Eq. } 2)$$

where *A* is the area of the liposome cross-section in pixel. The area is multiplied by a factor of 3/2 to account for an underestimation by non-equatorial dissection of the liposomes (**Supplementary Text**). A value of two pixels is added in order to account for the discrepancy in size between marker and liposome lumen as a result of erosion. To convert pixel values into micrometers, the conversion factor of 0.25 μm/pixel specific to the imaging setup is applied.

2.4.6 Phenomenological fitting of gene expression kinetics

Measured YFP expression kinetics from single liposomes were fit with a phenomenological model to estimate salient dynamic parameters, such as the final yield, YFP production rate and translation lifespan (or time to plateau). The following sigmoid equation was used[64]:

$$y = k' + k\frac{t^n}{t^n + K^n} \qquad\qquad (\text{Eq. } 3)$$

where *t* is the time in minutes, *y* the YFP concentration at a given time point, and *k'*, *k*, *K* and *n* are fit parameters. Using this expression, the final yield corresponds to *k* and the plateau time – or expression lifespan – (not corrected for the maturation time[11,34]) is expressed as:

2

$$T_{plateau} = \frac{2K}{n} + K \qquad\qquad (Eq.\,4)$$

The apparent translation rate, which is defined as the steepness at time $t = K$, is:

$$rate = \frac{kn}{4K} \qquad\qquad (Eq.\,5)$$

Note that these observed kinetic parameters do not account for the synthesis of non-fluorescent products.

2.4.7 Calibration of YFP fluorescence intensity into protein concentration

Conversion formulas that relate microscopy fluorescence to YFP concentrations were obtained via a two-step process. First, a bulk PURExpress reaction with the *yfp* DNA construct was run overnight. A dilution series of the end-point PURE system reaction solution was performed by addition of buffer (20 mM HEPES, 180 mM potassium glutamate, 14 mM magnesium acetate, pH 7.6). For every dilution sample, YFP fluorescence was measured in a spectrophotometer (Cary Eclipse from Agilent, exc/em 515/528 nm), using a 20-μL quartz cuvette (Hellma) at 37 °C. Owing to a previously established conversion factor[11,61], the final concentration of synthesized YFP can be determined. The same samples were measured by LSCM under conditions (chamber, settings) identical to those employed for liposome experiments. A YFP fluorescence-to-concentration conversion factor could be derived for each laser power setting used. To correct for significant fluctuations of laser power over the period of this study, regular calibration was performed.

Fluorescence intensity thresholds were determined with Fiji[63] to classify liposomes as expressing (with YFP) or permeable (with Cy5). For YFP, the background fluorescence intensity values of sixty randomly selected ROI's outside liposomes from six independent experiments were measured at various laser power settings. The threshold value below which the YFP signal within liposomes was considered negligible was defined as the mean plus two standard deviations of the calculated background intensity value. This calibration was repeated several times and slightly different threshold values were obtained. Similarly, the Cy5 background intensity value was determined and used for normalization of the liposome intensity values. Liposomes with a normalized Cy5 signal below 0.75 were considered empty (impermeable).

2.4.8 Flow cytometry

Liposomes containing PURE*frex* 2.0 with or without 100 ng of DNA were prepared as described above. After swelling, 2 μL of liposomes was diluted in 5.5 μL of feeding solution (see above) and incubated overnight at 37 °C in a BioRad C1000 Touch Thermal Cycler. Next, 1 μL of liposome mixture was diluted in 299 μL of isotonic buffer (20 mM HEPES, 180 mM potassium glutamate, 14 mM magnesium acetate, pH 7.6). Of the diluted liposome solution, 200 μL was injected in a BD FACSCelesta flow cytometer (BD Biosciences). Fluorescence was acquired with the pre-set mCherry (for Texas red) and GFP (for YFP) channels. Data was analyzed using MATLAB and the online Cytobank application (https://community.cytobank.org/cytobank/experiments). The YFP vs Texas Red fluorescence intensity scatter plot was compared to the expected correlation between the signals of a lumen dye and a membrane dye assuming basic geometrical considerations. Let assume a population of spherical and unilamellar vesicles. Let also assume that both the YFP and the Texas Red dyes are randomly distributed inside the lumen (following a Poisson distribution) or in the membrane of liposomes, respectively. Using simple scaling geometrical factors, one obtains:

$$I_{\text{YFP}} \propto \frac{4}{3}\pi r^3 \qquad\qquad\qquad (\text{Eq. 6})$$

$$I_{\text{TexasRed}} \propto 4\pi r^2 \qquad\qquad\qquad (\text{Eq. 7})$$

where, r is the liposome radius, I_{YFP} and I_{TexasRed} are the fluorescence intensities in the YFP and Texas Red channels, respectively. Rearranging to eliminate the dependency with respect to r and linearizing on a logarithmic scale, one obtains:

$$log(I_{\text{YFP}}) = 1.5\, log(I_{\text{TexasRed}}) + \text{constant} \qquad\qquad (\text{Eq. 8})$$

This theoretical correlation was compared to actual fluorescence intensity data. Note that I_{YFP} represents the integrated intensity of YFP in a given liposome and cannot be compared to the mean YFP intensity inferred from confocal images.

2.4.9 Bulk gene expression kinetics

Bulk gene expression kinetics were measured with a spectrophotometer (Cary Eclipse from Agilent, exc/em 515/528 nm). PURE system reactions with 5 ng/µL (7.4 nM) of the *yfp* construct were assembled according to the manufacturers' instructions. When indicated, the *Cx43* DNA template was added at a final concentration of 6.25 ng/µL by substituting the equivalent volume of Milli-Q water. Samples were immediately transferred to 20-µL cuvettes (Hellma) and loaded into the spectrometer sample holder, preheated to 37 °C. Fluorescence intensity was converted into YFP concentration using a previously determined conversion factor[11].

2.4.10 Polyacrylamide gel electrophoresis

PURE*frex* reaction mixtures containing either or both *yfp* and *Cx43* DNA constructs (same concentrations as indicated for bulk kinetics and liposome experiments) were supplemented with 0.5 µL BODIPY-Lys-tRNA$_{\text{Lys}}$ (FluoroTect$^{\text{TM}}$ GreenLys, Promega), an *in vitro* translation labelling system. The sample was incubated for 3 h at 37 °C, denatured for 10 min at 90 °C and analysed on a 12% SDS-PAGE using a fluorescence gel imager (Typhoon, Amersham Biosciences).

REFERENCES

2

1. Luisi, P. L. & Stano, P. *The minimal cell.* (2013).

2. Nomura, S. M., Tsumoto K., Hamada, T., Akiyoshi K., Nakatani, Y. and Yoshikawa, K., Gene expression within cell-sized lipid vesicles *ChemBioChem.* **4,** 1172-75 (2003)

3. Noireaux, V., and Libchaber, A. A vesicle bioreactor as a step toward an artificial cell assembly *Proc Natl Acad Sci USA.* **101,** 17669-74 (2004)

4. Nourian, Z., Roelofsen, W., and Danelon, C. Triggered gene expression in fed-vesicle microreactors with a multifunctional membrane *Angew Chem Int Ed.* **51,** 3114-18 (2012)

5. Scott A, Noga M J, de Graaf P, Westerlaken I, Yildirim E and Danelon C. Cell-free phospholipid biosynthesis by gene-encoded enzymes reconstituted in liposomes *PLoS One*. **11,** e0163058 (2016)

6. Van Nies P, Westerlaken I, Blanken D, Mencía M, Salas M and Danelon C. Self-replication of DNA by its encoded proteins in liposome-based synthetic cells *Nat. Commun.* **9,** 1583 (2018)

7. Ichihashi N, Usui K, Kazuta Y, Sunami T, Matsuura T and Yomo T. Darwinian evolution in a translation-coupled RNA replication system within a cell-like compartment *Nat Commun.* **4,** 2494 (2013)

8. Sunami T, Ichihashi N, Nishikawa T, Kazuta Y and Yomo T. Effect of liposome size on internal RNA replication coupled with replicase translation *Chembiochem.* **17,** 1282-9 (2016)

9. Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K and Ueda T. Cell-free translation reconstituted with purified components *Nat Biotechnol* **19,** 751-5 (2001)

10. Shimizu Y, Kanamori T and Ueda T. Protein synthesis by pure translation systems *Methods.* **36,** 299-304 (2005).

11. Doerr A, de Reus E, van Nies P, van der Haar M, Wei K, Kattan J, Wahl A and Danelon C. Modelling cell-free RNA and protein synthesis with minimal systems *Phys. Biol.* **16,** 025001 (2019).

12. Saito H, Kato Y, Le Berre M, Yamada A, Inoue T, Yosikawa K and Baigl D. Time-resolved tracking of a minimum gene expression system reconstituted in giant liposomes *Chembiochem.* **10,** 1640-3 (2009).

13. Sakamoto R, Noireaux V and Maeda Y T. Anomalous scaling of gene expression in confined cell-free reactions *Sci Rep.* **8,** 7364 (2018).

14. Vogele K, Frank T, Gasser L, Goetzfried M A, Hackl M W, Sieber S A, Simmel F C and Pirzer T. Towards synthetic cells using peptide-based reaction compartments *Nat Commun.* 2018 **9,** 3862 (2018).

15. Martino C, Kim S-H, Horsfall L, Abbaspourrad A, Rosser S J, Cooper J and Weitz D A. Protein expression, aggregation, and triggered release from polymersomes as artificial cell-like structures *Angew. Chem. Int. Ed*. **51,** 6416-20 (2012).

16. Nourian Z, Scott A and Danelon C. Toward the assembly of a minimal divisome *Syst Synth Biol.* **8,** 237-47 (2014).

17. Liu Y J, Hansen G P, Venancio-Marques A and Baigl D. Cell-free preparation of functional and triggerable giant proteoliposomes *Chembiochem.* **14,** 2243-7 (2013).

18. Majumder S, Garamella J, Wang Y L, DeNies M, Noireaux V and Liu A P. Cell-sized mechanosensitive and biosensing compartment programmed with DNA *Chem Commun (Camb).* **53,** 7349-52 (2017).

19. Kamiya K, Kawano R, Osaki T, Akiyoshi K and Takeuchi S. Cell-sized asymmetric lipid vesicles facilitate the investigation of asymmetric membranes *Nat Chem.* **8,** 881-9 (2016).

20. Uyeda A, Nakayama S, Kato Y, Watanabe H and Matsuura T. Construction of an in vitro gene screening system of the E. coli EmrE transporter using liposome display. *Anal Chem.* **88,** 12028-35 (2016).

21. Kuruma Y, Stano P, Ueda T and Luisi P L. A synthetic biology approach to the construction of membrane proteins in semi-synthetic minimal cells. *Biochim Biophys Acta.* **1788,** 567-74 (2009)

22. Exterkate M, Caforio A, Stuart M C A and Driessen A J M. Growing membranes in vitro by continuous phospholipid biosynthesis from free fatty acids *ACS Synth Biol.* **7,** 153-65 (2018).

23. Nourian Z and Danelon C. Linking genotype and phenotype in protein synthesizing liposomes with external supply of resources *ACS Synth Biol.* **2,** 186-93 (2013).

24. Mavelli F, Marangoni R and Stano P. A simple protein synthesis model for the PURE system operation *Bull Math Biol.* **77,** 1185-212 (2015).

25. Lazzerini-Ospri L, Stano P, Luisi P L and Marangoni R. Characterization of the emergent properties of a synthetic quasi-cellular system *BMC Bioinformatics* **13,** Suppl 4:S9 (2012).

26. Caveney P M, Norred S E, Chin C W, Boreyko J B, Razooky B S, Retterer S T, Collier C P and Simpson M L. Resource sharing controls gene expression bursting *ACS Synth Biol.* **6,** 334-43 (2017).

27. Nishimura K, Tsuru S, Suzuki H and Yomo T. Stochasticity in gene expression in a cell-sized compartment. *ACS Synth Biol.* **4,** 566-76 (2015)

28. Pereira de Souza T, Steiniger F, Stano P, Fahr A and Luisi P L. Spontaneous crowding of ribosomes and proteins inside vesicles: a possible mechanism for the origin of cell metabolism *Chembiochem.* **12,** 2325-30 (2011).

29. Hansen M, Meijer L H, Spruijt E, Maas R J, Rosquelles M V, Groen J, Heus H A and Huck W T. Macromolecular crowding creates heterogeneous environments of gene expression in picolitre droplets *Nat Nanotechnol.* **11,** 191-7 (2016).

30. Norred S E, Caveney P M, Chauhan G, Collier L K, Collier C P, Abel S M and Simpson M L. Macromolecular crowding induces spatial correlations that control gene expression bursting patterns *ACS Synth Biol.* **7,** 1251-8 (2018).

2

31. Okano T, Matsuura T, Suzuki H and Yomo T. Cell-free protein synthesis in a microchamber revealed the presence of an optimum compartment volume for high-order reactions *ACS Synth Biol.* **3,** 347-52 (2014).

32. Van Nies P, Nourian Z, Kok M, van Wijk R, Moeskops J, Westerlaken I, Poolman J M, Eelkema R, van Esch J H, Kuruma Y, Ueda T and Danelon C. Unbiased tracking of the progression of mRNA and protein synthesis in bulk and in liposome-confined reactions *Chembiochem.* **14,** 1963-6 (2013).

33. Frazier J M, Chushak Y and Foy B. Stochastic simulation and analysis of biomolecular reaction networks *BMC Syst Biol.* **3,** 64 (2009).

34. Altamura E, Carrara P, D'Angelo F, Mavelli F and Stano P. Extrinsic stochastic factors (solute partition) in gene expression inside lipid vesicles and lipid-stabilized water-in-oil droplets: a review *Synthetic Biology* **3,** ysy011 (2018).

35. Dezi M, Di Cicco A, Bassereau P and Lévy D Detergent-mediated incorporation of transmembrane proteins in giant unilamellar vesicles with controlled physiological contents *Proc Natl Acad Sci U S A.* **110,** 7276-81 (2013).

36. Kumar N M and Gilula N B. The gap junction communication channel *Cell* **84,** 381-8 (1996).

37. Sunami T, Sato K, Matsuura T, Tsukada K, Urabe I and Yomo T. Femtoliter compartment in liposomes for in vitro selection of proteins *Anal Biochem.* **357**, 128-36 (2006).

38. Stein H, Spindler S, Bonakdar N, Wang C and Sandoghdar V. Production of isolated giant unilamellar vesicles under high salt concentrations *Front Physiol.* **8,** 63 (2017).

39. Reeves J P and Dowben R M. Formation and properties of thin-walled phospholipid vesicles *J Cell Physiol.* **73,** 49-60 (1969).

40. Payne N I, Timmins P, Ambrose C V, Ward M D and Ridgway F. Proliposomes: a novel solution to an old problem *J. Pharm. Sci.* **75,** 325-9 (1986).

41. Angelova M I and Dimitrov D S. Liposome electroformation *Faraday Discuss. Chem. Soc.* **81,** 303 (1986).

42. Horger K S, Estes D J, Capone R and Mayer M. Films of agarose enable rapid formation of giant liposomes in solutions of physiologic ionic strength *J Am Chem Soc.* **131,** 1810-9 (2009).

43. Weinberger A, Tsai F C, Koenderink G H, Schmidt T F, Itri R, Meier W, Schmatko T, Schröder A and Marques C. Gel-assisted formation of giant unilamellar vesicles *Biophys J.* **105,** 154-64 (2013).

44. López Mora N, Hansen J S, Gao Y, Ronald A, Kieltyka R, Malmstadt N and Kros A. Preparation of size tunable giant vesicles from cross-linked dextran(ethylene glycol) hydrogels *Chem Commun (Camb).* **50** 1953-5 (2014).

45. Pautot S, Frisken B J and Weitz D A. Engineering asymmetric vesicles *Proc Natl Acad Sci U S A.* **100** 10718-21 (2003).

46. Abkarian M, Loiseau E and Massiera G. Continuous droplet interface crossing encapsulation (cDICE) for high throughput monodisperse vesicle design *Soft Matter* **7,** 4610-4 (2011).

47. Blosser M C, Horst B G and Keller S L. cDICE method produces giant lipid vesicles under physiological conditions of charged lipids and ionic solutions *Soft Matter* **12,** 7364-71 (2016).

48. Funakoshi K, Suzuki H and Takeuchi S. Formation of giant lipid vesicle like compartments from a planar lipid membrane by a pulsed jet flow *J. Am. Chem. Soc.* **129,** 12608-9 (2007).

49. Shum H C, Lee D, Yoon I, Kodger T and Weitz D A. Double emulsion templated monodisperse phospholipid vesicles *Langmuir* **24,** 7651-3 (2008).

50. Matosevic S and Paegel B M. Stepwise synthesis of giant unilamellar vesicles on a microfluidic assembly line *J Am Chem Soc.* **133,** 2798-800 (2011)

51. Karamdad K, Law R W, Seddon J M, Brooks N J and Ces O. Preparation and mechanical characterisation of giant unilamellar vesicles by a microfluidic method *Lab Chip* **15,** 557-62 (2015).

52. Deshpande S, Caspi Y, Meijering A E and Dekker C. Octanol-assisted liposome assembly on chip *Nat Commun.* **7,** 10447 (2016).

53. Deng N N, Yelleswarapu M and Huck W T. Monodisperse uni- and multicompartment liposomes *J Am Chem Soc.* **138,** 7584-91 (2016).

54. Wingreen N S and Huang K C. Physics of intracellular organization in bacteria *Annu Rev Microbiol.* **69,** 361-79 (2015).

55. Govindarajan S, Nevo-Dinur K and Amster-Choder O. Compartmentalization and spatiotemporal organization of macromolecules in bacteria *FEMS Microbiol Rev.* **36,** 1005-22 (2012).

56. Iversen L, Mathiasen S, Larsen J B and Stamou D. Membrane curvature bends the laws of physics and chemistry *Nat Chem Biol.* **11,** 822-5 (2015).

57. Kirchner S R, Ohlinger A, Pfeiffer T, Urban A S, Stefani F D, Deak A, Lutich A A and Feldmann J. Membrane composition of jetted lipid vesicles: a Raman spectroscopy study *J Biophotonics* **5,** 40-6 (2012).

58. Van Swaay D and deMello A. Microfluidic methods for forming liposomes *Lab Chip* **13,** 752-67 (2013).

59. Nishimura K, Matsuura T, Sunami T, Fujii S, Nishimura K, Suzuki H and Yomo T. Identification of giant unilamellar vesicles with permeability to small charged molecules *RSC Advances* **4,** 35224-32 (2014).

60. Colom A, Derivery E, Soleimanpour S, Tomba C, Molin M D, Sakai N, González-Gaitán M, Matile S and Roux A. A fluorescent membrane tension probe *Nat Chem.* **10,** 1118-25 (2018).

61. Van Nies P, Canton A S, Nourian Z and Danelon C. Monitoring mRNA and protein levels in bulk and in model vesicle-based artificial cells *Methods Enzymol.* **550**, 187-214 (2015).

2

62. Tsumoto K, Matsuo H, Tomita M and Yoshimura T. Efficient formation of giant liposomes through the gentle hydration of phosphatidylcholine films doped with sugar *Colloids and Surfaces B: Biointerfaces* **68,** 98-105 (2009).

64. Schindelin J, Arganda-Carreras I, Frise E. et al. Fiji: an open-source platform for biological-image analysis *Nature methods* **9,** 676-82 (2012).

64. Ang J, Harris E, Hussey B, Kil R and McMillen D. Tuning response curves for synthetic biology *ACS Synth. Bio*l. **2**, 547-67 (2013).

SUPPLEMENTARY INFORMATION

Estimation of the size of liposomes

Estimating the radius of liposomes from confocal images requires to take into account non-equatorial liposome dissection and calculate the equatorial area. Assume a liposome represents a sphere with radius R. Then, any cross-section at height z from the equator, will, by the Pythagorean theory, have a radius r given by:

$$r = \sqrt{R^2 - z^2}. \tag{1}$$

The resulting area A is given by:

$$A = \pi r^2 = \pi(R^2 - z^2). \tag{2}$$

The average area $\bar{A}$ of an arbitrary cross-section then follows, by symmetry, by integrating over z from 0 to R:

$$\bar{A} = \frac{1}{R}\int_0^R \pi(R^2 - z^2)dz = \frac{2}{3}\pi R^2. \tag{3}$$

Thus, the average area of an arbitrary cross-section is 2/3 of the actual equatorial area. To compensate for this underestimation, the measured area will be multiplied by 3/2 in the calculation of the radius (see Methods in main text).

2
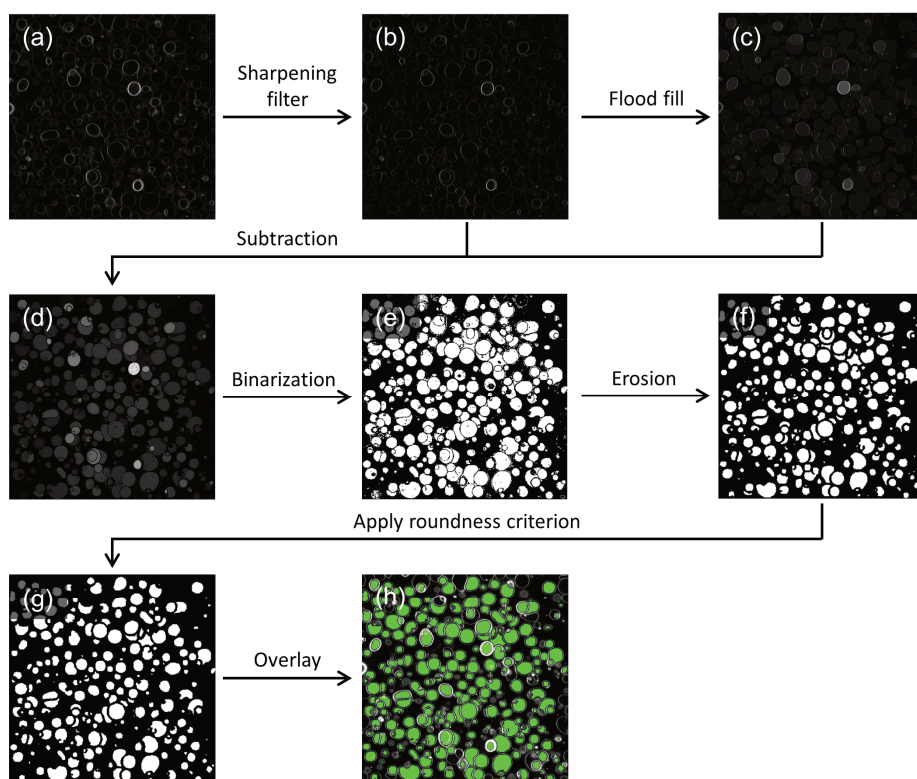
Supplementary Figure 2.1: Graphical representation of the image analysis algorithm developed to obtain liposome size and YFP signal from LSCM micrographs. **(a)** Texas red channel of the original micrograph. **(b)** A sharpening filter is employed. **(c)** Flood-fill algorithm, as encoded by the MATLAB function *floodfill.m*, is applied. **(d)** Sharpened image (b) is subtracted from filled image (c), resulting in an image with non-black areas corresponding to liposome lumina. **(e)** Binarization of (d). Binarization threshold was adjusted if necessary. **(f)** Erosion, using an erosion disk with a size of two pixels, as encoded by the MATLAB function *imerode.m*, is applied to remove 'debris'. **(g)** A roundness criterion (see Methods in main text) is employed to remove objects deviating to far from a circle, which are most likely not liposomes, but for example inter-vesicular space. Image (g) serves as marker to determine liposome size and to extract fluorescence intensity from the corresponding YFP channel. **(h)** Overlay of final marker (g) and original image (a), demonstrating successful detection of almost all liposomes.
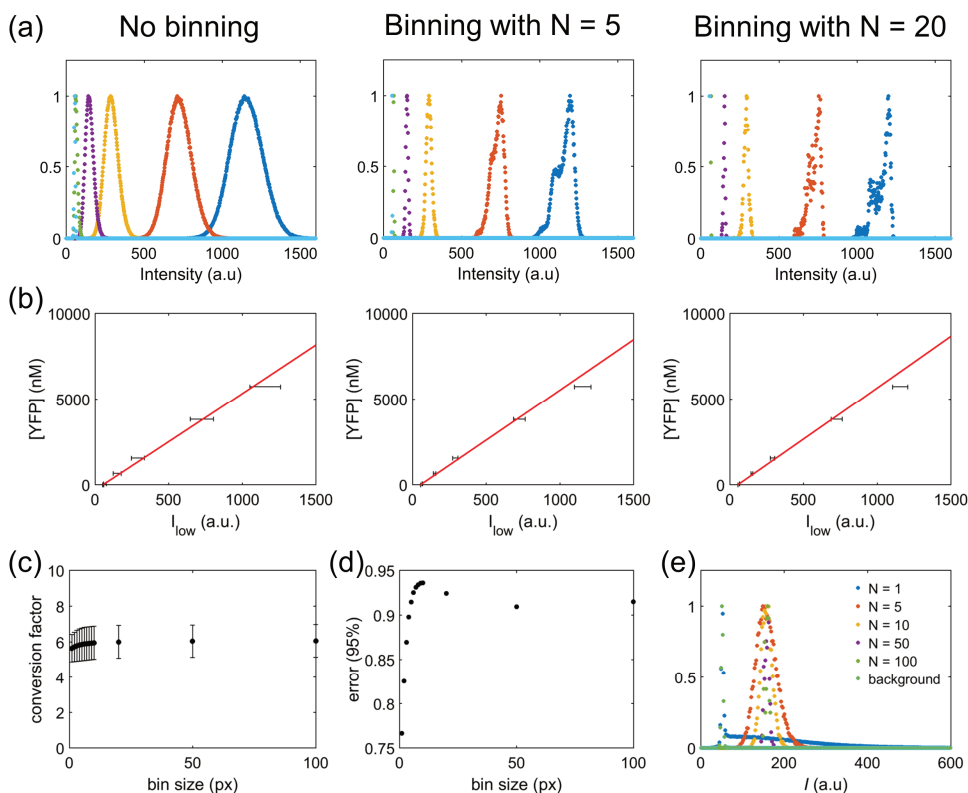
Supplementary Figure 2.2: Converting YFP intensity values as measured by LSCM into concentrations. (a) Histograms of fluorescence intensity of a six-by-six field of view in the YFP channel (regular settings) for various dilutions of a pre-run bulk PURExpress reaction sample. Binning has been investigated as a method to reduce pixilation effects, which mimics the way the average pixel intensity within liposome cross-section areas (or lumina) is calculated (Fig. S2). Histograms without binning, with binning into squares of 5x5 pixels and 20x20 pixels are shown. Binned histograms show narrower distributions, but reveal a shoulder that is a result of the stitching applied during acquisition. **(b)** Linear fit of fluorescence intensity versus YFP concentration for the three aforementioned binning conditions. YFP concentration was determined by spectrofluorometry as described in the main text. Error bars indicate standard deviation values. **(c)** Conversion factor (slope of fit) for various bin sizes. A slight increase is observed with increasing bin sizes. **(d)** Width of 95% confidence interval of slope of fit, for various bin sizes. Within error margin all conversion factor values are identical. We opted for using the unbinned conversion factor as it is the most accurate and its lowest value ensures that we do not overestimate YFP concentrations. **(e)** Histograms of the Spinach fluorescence intensity (exc/em wavelengths 457/482 nm) of a six-by-six field of view of a pre-run bulk PURE*frex* sample. Ribosomes were omitted and gene expression occurred overnight. Various binning conditions were tested. Without binning (*N* = 1), the histogram peak overlaps with that of the background signal, indicating very strong pixilation effects. Binning shifts the peak to the right and makes it more Gaussian; however, the intensity values are very low (5% of range), making it impossible to reliably quantify Spinach signal. Therefore, we have opted to not quantitatively analyses the mRNA concentration in this research.

2
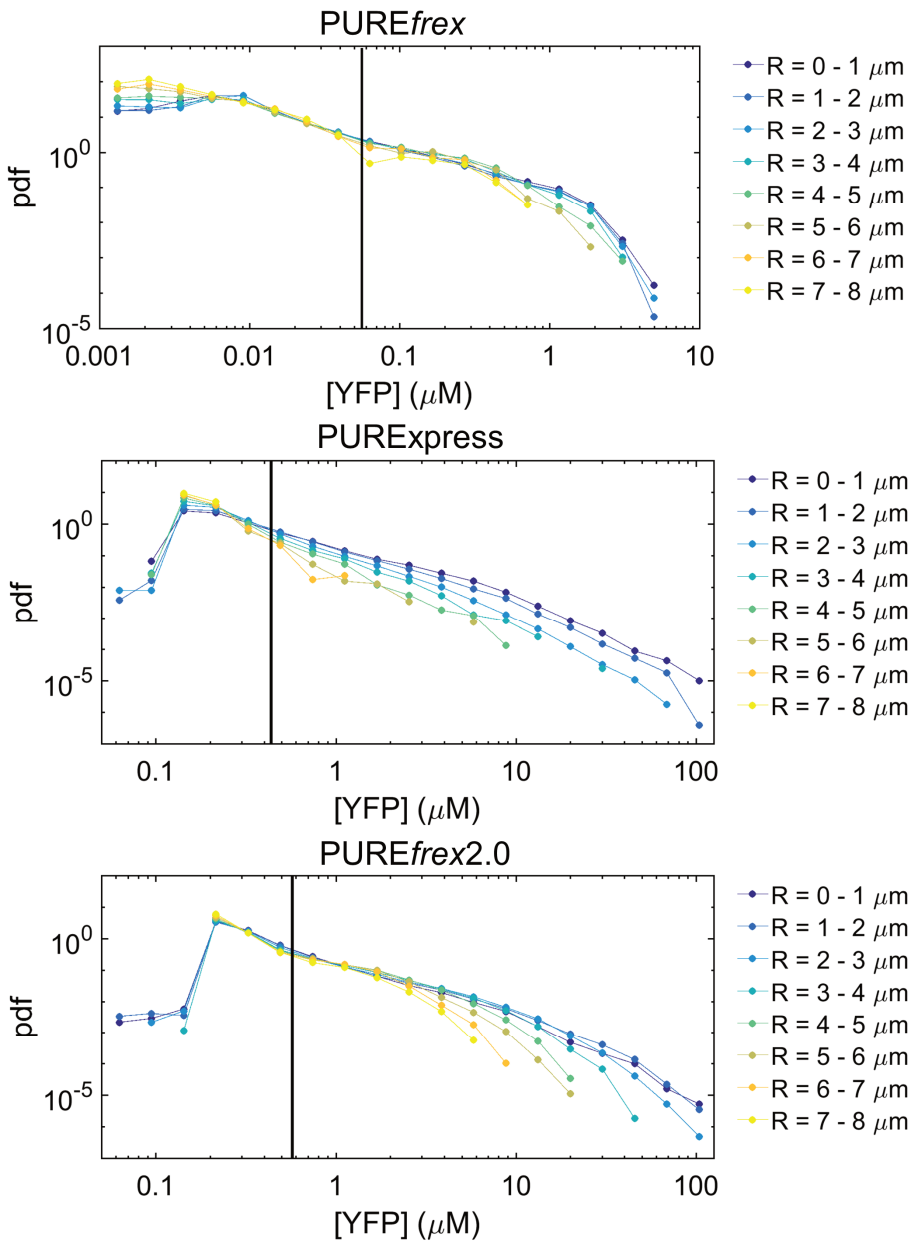
Supplementary Figure 2.3: Probability density functions of YFP concentrations for different liposome sizes. YFP concentration data shown in Fig. 1 of the main text was grouped according to the estimated liposome radius. For each size bin, a probability density histogram of the YFP concentration was plotted for all three PURE systems. Vertical black lines indicate the threshold concentration above which liposomes are considered expressing (see main text). In each type of PURE system, smaller liposomes are the ones with higher YFP concentrations, especially for PURExpress. However, a number of limiting factors – the rough estimate of the radius (Supplementary Text), pixilation effects and possible concentration overestimation which are more pronounced in smaller liposomes – prevent us from drawing strong conclusions about a possible effect of liposome size on gene expression performance.
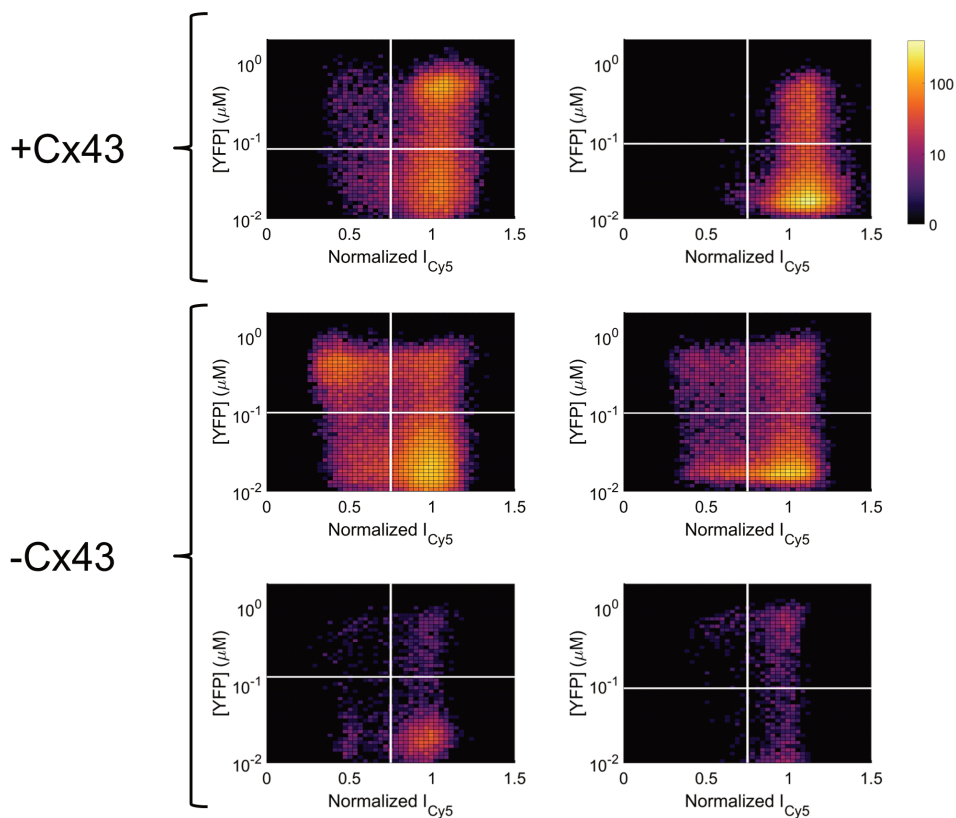
Supplementary Figure 2.4: Additional heat maps of YFP concentration versus normalized Cy5 signal. Heat maps are shown for two experiments in the presence of Cx43 and four in its absence, under similar conditions as in Fig. 3 and 4 in the main text. The color bar refers to the number of liposomes. Total number of liposomes analyzed per experiment (left-to-right, top-to-bottom): N = 22,946; 22,268; 40,271; 20,930; 4,114; and 4,236.
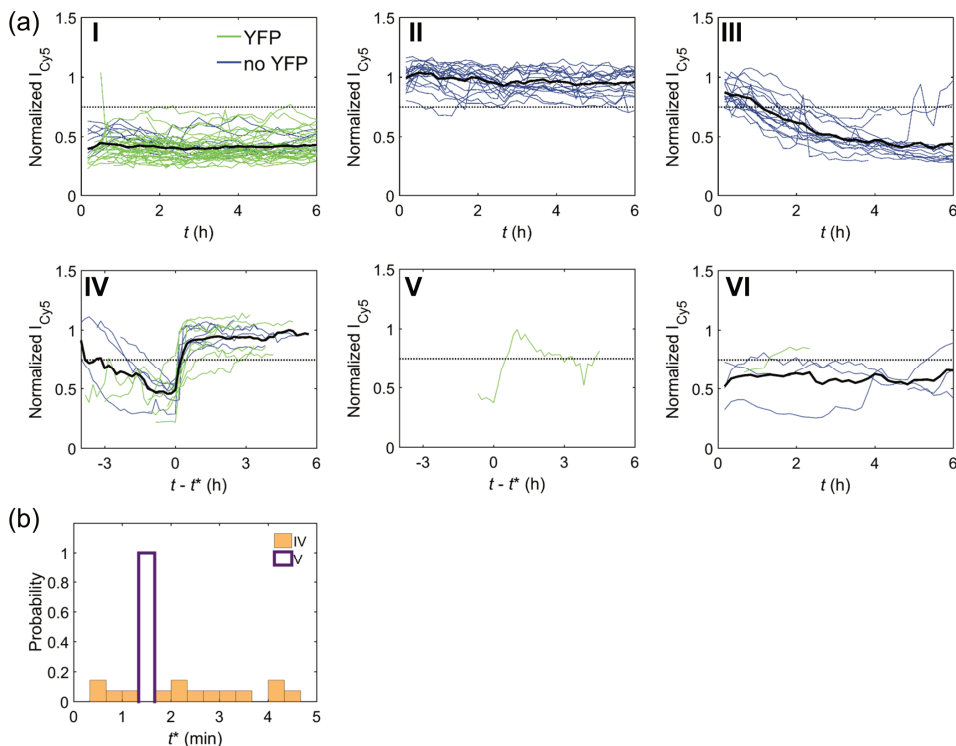
Supplementary Figure 2.5: Permeability phenotypes in the absence of Cx43 DNA. (a) Time traces of normalized Cy5 signal are sorted in six distinct phenotypes (see main text). Liposomes exhibiting YFP expression ($N = 100$, green lines) and no YFP expression ($N = 100$, blue lines) have been investigated. **(b)** Histogram of permeabilization time t^* as observed in phenotypes IV (14 liposomes) and V (one liposome). For phenotype IV, the peak at t^* ~60 min that was observed in the presence of Cx43 DNA (Fig. 5b), is clearly absent here.

Chapter 3: Genetically controlled membrane synthesis in liposomes

Lipid membranes, nucleic acids, proteins, and metabolism are essential for modern cellular life. Synthetic systems emulating the fundamental properties of living cells must therefore be built upon these functional elements. In this work, phospholipid-producing enzymes encoded in a synthetic minigenome are cell-free expressed within liposome compartments. The de novo *synthesized metabolic pathway converts precursors into a variety of lipids, including the constituents of the parental liposome. Balanced production of phosphatidylethanolamine and phosphatidylglycerol is realized, owing to transcriptional regulation of the activity of specific genes combined with a metabolic feedback mechanism. Fluorescence-based methods are developed to image the synthesis and membrane incorporation of phosphatidylserine at the single liposome level. Our results provide experimental evidence for DNA-programmed membrane synthesis in a minimal cell model. Strategies are discussed to alleviate current limitations toward more effective liposome growth and self-reproduction.*

3.1 INTRODUCTION

Life manifests itself as individual cellular entities. Biological cells are spatially delimited from their surrounding by a lipid membrane. While archaeal membranes are composed of ether lipids, other cell types use phospholipids as the most abundant membrane constituents. Most phospholipids self-assemble in aqueous solutions to form vesicles, called liposomes, under a wide range of experimental conditions. Spatial organization of biochemical processes within liposomes mimics the fundamental characteristics according to which natural cells are organized. Therefore, phospholipid vesicles provide a chassis for the construction of synthetic minimal cells representing comparatively simple model systems[1-5].

Also pertinent to a working definition of cellular life is the notion of self-maintenance, in line with the view of a basic cell as an autopoietic unit[6], whereby all the system's components are produced within its boundary. Substrates present in the external environment absorb to the membrane or diffuse across and are transformed into molecular building blocks by metabolic processes. Another aspect that is particularly relevant when describing the inner functioning of a biological cell is the coupling between the different subsystems[7], such as genetic information, protein synthesis, and metabolic synthesis of the membrane constituents. Herein, we apply this conceptual framework to the construction of a minimal cell that can produce its own membrane components. Cell-free protein and phospholipid synthesis, directed by a DNA program, is carried out inside a liposome, constituting a first integrative step on the way to the development of an autonomously growing and dividing artificial cell.

Various strategies have been described to grow liposomes. Membrane constituents directly supplied in the external medium in the form of monomers, micelles or small unilamellar vesicles can spontaneously adsorb or fuse to the liposome membrane, increasing its surface area[8-11]. Moreover, non-enzymatic mechanisms to produce membrane lipids from synthetic reactive precursors and catalysts are particularly effective, leading to substantial vesicle growth[12-15]. To establish a link between the lipid compartment and its internal content, liposome growth could be made conditional to encapsulated nucleic acids[12,16] or catalysts[17]. Such model systems are attractive for their molecular simplicity and may resemble primitive cells before the emergence of modern biology. Closer to processes occurring in contemporary cells, enzyme-catalyzed biosynthesis of phospholipids has been realized using purified proteins[18-21]. Further, the lipid-producing enzymes were encoded in DNA and expressed by in vitro protein synthesis inside liposomes, providing a genotype-to-phenotype linkage[22,23]. The *Escherichia coli* enzymes glycerol-3-phosphate (G3P) acyl transferase and lysophosphatidic acid (LPA) acyl transferase, respectively referred below as PlsB and PlsC from their gene names, were in situ expressed from two DNA templates[23]. The precursors G3P and fatty acyl coenzyme A (acyl-CoA) were sequentially converted into lysophosphatidic acid and phosphatidic acid (PA) lipids in a two-step enzymatic reaction. However, the output phospholipid PA was not part of the original membrane composition and the PA detection method was not compatible with single vesicle resolution[23]. Regeneration of the main constituents of the liposome membrane obligates the reconstitution of five additional headgroup-modifying enzymes, which together with PlsB and PlsC form the Kennedy metabolic pathway that produces phosphatidylethanolamine (PE) and phosphatidylglycerol (PG), the most abundant lipids in the E. coli membranes. Although the unregulated expression of the Kennedy pathway enzymes was enabled from the outside of liposomes[23], *in vesiculo* synthesis of membrane-forming lipids with controlled molecular ratios remains a challenge.

In the present work, we show that the synthesis of PE and PG lipids from simpler precursors can be genetically controlled inside PE- and PG-containing liposomes. Our results provide experimental evidence for DNA-encoded homeostatic growth of a liposome-based artificial cell. Because the metabolic pathway encompasses seven different enzymes, we first assemble all seven genes on a single plasmid. The PURE (Protein synthesis Using Recombinant Elements) system[24], here PUREfrex2.0, is used as a minimal cell-free protein synthesis platform that converts the DNA program into the whole enzymatic pathway. Phospholipid biosynthesis within liposomes is demonstrated by quantitative liquid chromatography-coupled mass spectrometry (LC-MS). Relative PE and PG content is tailored through transcriptional and metabolic regulation mechanisms. Moreover, we develop fluorescence-based probes to directly visualize membrane incorporation of synthesized phospholipids at the single vesicle level.

3.2 RESULTS

3.2.1 Design and construction of a minigenome for phospholipid biosynthesis

We aimed to reconstitute the Kennedy phospholipid synthesis pathway from E. coli starting from all seven enzyme-encoded genes (**Fig. 3.1a**). The membrane-bound protein PlsB uses acyl-CoA (or acyl carrier protein, ACP) as a donor to acylate the 1-position of G3P to form LPA[25]. The 2-position is subsequently acylated by the membrane protein PlsC to form diacyl PA, again using acyl-CoA as fatty acid donor, preferring unsaturated carbon chains[25]. Enzymes downstream in the pathway are involved in phospholipid headgroup modifications. The integral membrane protein CdsA catalyzes the activation of PA with cytosine triphosphate (CTP) to generate diacyl-sn-glycero-3-(cytidine diphosphate) (diacyl-CDP-DAG)[26] which serves as a precursor for two separate branches of the Kennedy pathway. One branch, which leads to the formation of PG as the final product, comprises the synthesis of phosphatidylglycerol phosphate (PGP) from G3P and CDP-DAG by the membrane-associated protein CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase (PgsA)[27], followed by a dephosphorylation step that is catalyzed by the phosphatidylglycerophosphatase A, B[28] or C (PgpA, B or C)[29]. The other branch generates PE as the end-product in a two-step reaction. First, phosphatidylserine (PS) production from CDP-DAG and L-serine is catalyzed by the CDP-diacylglycerol-serine O-phosphatidyltransferase (PssA). Then, PS is decarboxylated to form PE, a reaction that is catalyzed by the phosphatidylserine decarboxylase (Psd), a two-subunit protein resulting from the autocatalytic serinolysis of a single proenzyme[30].

All seven genes, namely *plsB, plsC, cdsA, pgsA, pgpA, pssA* and *psd* were concatenated into a single plasmid DNA as individual transcriptional cassettes, i.e. every gene is under control of its own promoter, ribosome binding site and transcription terminator (**Fig. 3.1a**). This design strategy ensures that all genes will be present at the same copy number upon plasmid encapsulation in liposomes, thus obviating the functional heterogeneity inherent to uneven partitioning of the separate DNA templates. Thirty base pair linker sequences were added to each gene and to a linearized pUC19 plasmid backbone by polymerase chain reaction (PCR) to enable a one-step Gibson-assembly of the final plasmid[31] (**Supplementary Fig. 3.1 and 3.2**). The successful assembly of the pGEMM7 minigenome was confirmed using Sanger sequencing and restriction digestion (**Supplementary Fig. 3.3**). The three genes of the common pathway *plsB, plsC* and *cdsA*, as well as the two genes *pssA* and *psd* of the PE synthesis branch are under control of a T7 promoter and are constitutively expressed in PURE*frex*2.0. The two genes *pgsA* and *pgpA*, encoding the enzymes for PG biosynthesis, are under control of an SP6 promoter and are encoded on the opposite strand to prevent read-through transcription by incomplete termination at the T7 terminator sites (**Supplementary Note 3.1,**

Supplementary Fig. 3.4). Orthogonality of the two promoter-RNA polymerase (RNAP) pairs in PURE*frex*2.0 was demonstrated using a fluorescent protein reporter (**Supplementary Fig. 3.5**).
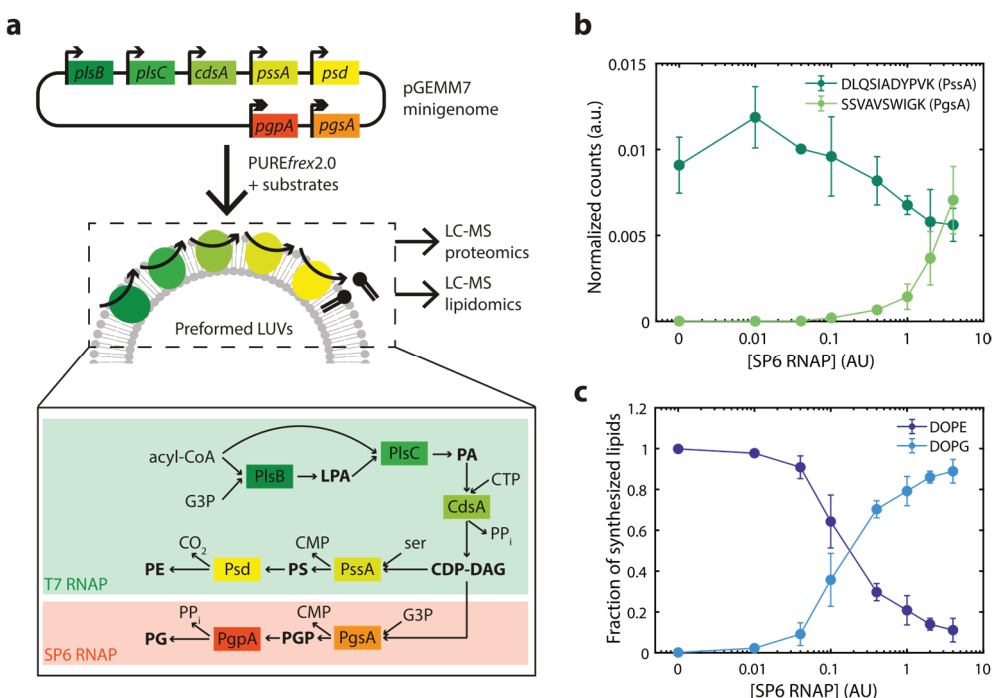
Figure 3.1: Genetically controlled production of PE and PG by de novo synthesized enzymes. (a) The pGEMM7 plasmid contains seven genes encoding *E. coli* lipid synthesis enzymes. Transcriptional regulation over the production of PE and PG lipids is provided by controlling the expression of specific enzymes with the orthogonal T7 (single arrow) and SP6 (double arrow) promoters. The main reaction products are in bold and the enzyme names are squared. The PgsA-PgpA branch of the pathway, which leads to PG synthesis, is activated upon addition of the SP6 RNAP. An acyl-CoA, the heavy isotope of G3P, CTP, and serine (Ser) are the input substrates. Expression of pGEMM7 with PURE*frex*2.0 occurred in the presence of preformed LUVs. Protein and lipid production were monitored by LC-MS. PPi, pyrophosphate. (**b**) LC-MS analysis of cell-free synthesized proteins. Normalized integrated peak intensity for representative peptides of PssA and PgsA, the first enzymes after the pathway branches out, for a range of SP6 RNAP concentrations (given in activity units as defined by the supplier, AU). Data are the mean ± SD of three independent experiments. (**c**) LC-MS analysis of de novo synthesized phospholipids. The fraction of synthesized DOPE and DOPG is plotted for a range of SP6 RNAP concentrations (in AU). Data are the mean ± SD of three independent experiments. Source Data are available for panels **b** and **c**.

3.2.2 Transcriptional regulation of PE and PG biosynthesis

Traditionally, cell-free translation products are characterized by one-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using isotopically or fluorescently labelled amino acids as a readout. While these methods are suitable to analyze single or a few gene expression products, they suffer from a poor resolution when multiple proteins are co-synthesized (**Supplementary Fig. 3.6**). Here, we applied a targeted LC-MS proteomics approach to detect the de novo synthesized enzymes and validate transcriptional activation of the PgsA-PgpA pathway by the SP6 RNAP.

Large unilamellar vesicles (LUVs) supplied in PURE*frex*2.0 reactions served as a scaffold for the expressed membrane-associated and integral membrane proteins. Several proteolytic peptides of the expressed proteins were identified (**Supplementary Table 3.4**, **Supplementary Fig. 3.7**) and the total ion current of their observed fragment ions was normalized to a peptide originated from elongation factor thermo unstable (EF-Tu), an abundant protein in PURE system. In-solution digestion of pre-ran PURE system reaction samples with trypsin failed to deliver detectable peptides for one of the seven proteins, namely PgpA (**Supplementary Note 3.2**). No detectable amount of PgsA was measured when the SP6 RNAP was omitted, indicating that unintended expression of the *pgsA* gene is negligible (**Fig. 3.1b**). Varying the concentration of SP6 RNAP between 0.01 U μL^{-1} and 4 U μL^{-1} is accompanied by a gradual increase in PgsA. Concurrently, the concentration of the PssA enzyme under T7 promoter control decreases upon increased SP6 RNAP concentration. These results show the power of targeted proteomics for relative quantification of cell-free protein synthesis. Moreover, they validate our design for tunable expression levels of different enzymes belonging to orthogonal transcriptional pathways.

Successful production of PE and PG lipids and its genetic modulation were confirmed by an LC-MS lipidomics analysis (**Fig. 3.1c**, **Supplementary Fig. 3.8**). To distinguish the newly produced lipids from those initially present in the liposome membrane, ^{13}C-labelled G3P was used as an isotopically heavy precursor. Oleoyl-CoA was used as the acyl donor. Absolute quantification was achieved by measuring DOPG and DOPE standards prior and posterior to data acquisition of PURE system samples. In agreement with proteomics data, synthesized DOPG was detected exclusively in the presence of SP6 RNAP (**Fig. 3.1c**). The only intermediate species that significantly accumulates is DOPA (**Supplementary Fig. 3.9**).

3.2.3 Metabolic regulation of PE and PG biosynthesis

PssA is unique among the proteins of the Kennedy pathway since it is found both associated with the membrane and in the cytosol[32,33]. PssA is thought to maintain the ratio between acidic (PG and cardiolipin, CL) and zwitterionic (PE) lipids in *E. coli* by being activated upon association with PG/CL-rich membranes, whereas the cytosolic form is latent (**Fig. 3.2a**)[34,35]. We sought to exploit this feedback mechanism to provide membrane content homeostasis without relying on genetic control. LUVs with different amounts of DOPE and DOPG were prepared, and synthesis of ^{13}C-labelled DOPE and DOPG was determined by LC-MS, both in the presence and absence of SP6 RNAP (**Fig. 3.2b,c**). A clear positive correlation between initial PG content and yield of synthesized PE was observed, both in the presence ($\rho = 0.91 \pm 0.07$, mean ± SD of three independent repeats) and absence ($\rho = 0.94 \pm$

3

0.04, mean ± SD of three independent repeats) of SP6 RNAP. Moreover, a negative correlation between initial PG content and yield of synthesized PG was observed (ρ = −0.95 ± 0.03, mean ± SD of three independent repeats). These results confirm the model of allosteric regulation of PssA activity by PG content, providing non-genetic homeostasis of mixed lipid composition to our system. Interestingly, PE synthesis was reduced at low PG content, independent of the expression of the PG-synthesizing pathway branch (**Fig. 3.2b**). This result indicates that the regulatory mechanism is not solely driven by competition between the two pathway branches, but it relies also on the association-dissociation of PssA to the membrane (**Fig. 3.2a**)[30-32]. We also found that the total amount of synthesized PE and PG is ~2-fold higher at a higher mol% of initial PG (~18 µM at 0 mol% PG vs. ~28 µM at 35 mol% PG in the experiment shown in **Fig. 3.2d**). This result is in line with previous observations that PlsB activity is promoted by PG[36,37].
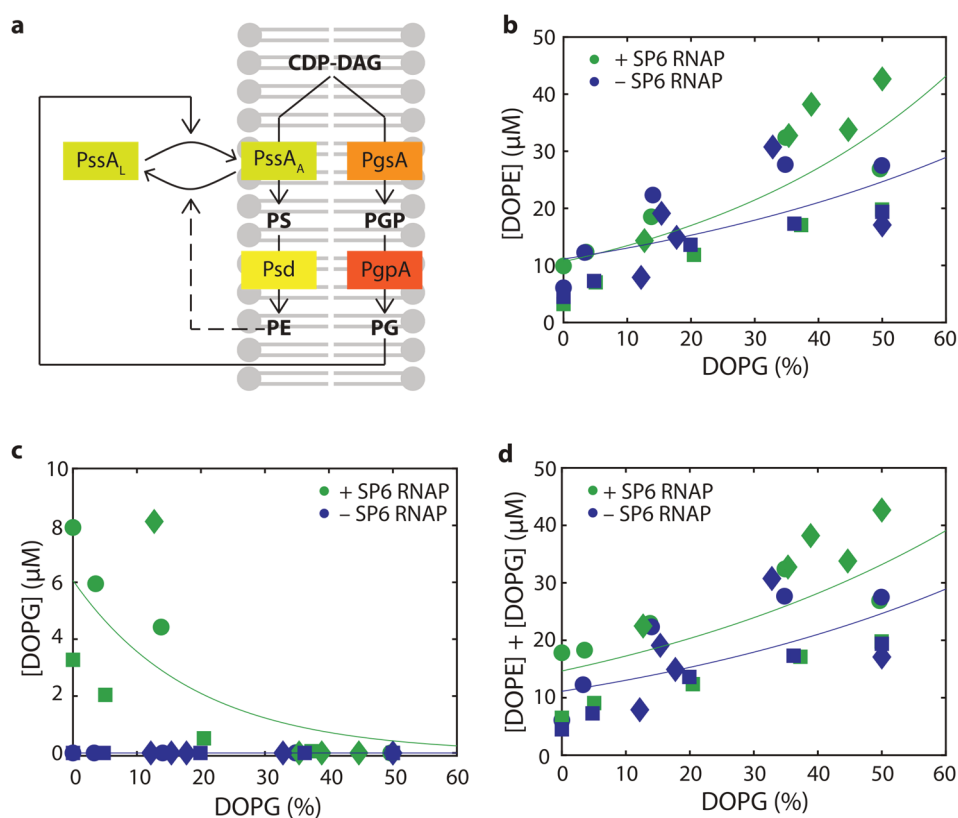
Figure 3.2: Metabolic feedback as a regulator for cell-free synthesis of PE and PG. (a) Schematic illustration of PssA activity regulation by membrane content. PssA exists in the membrane-bound, activated state (PssA$_A$) and in the cytosolic, inactivated or latent state (PssA$_L$). High amounts of PG favor PssA$_A$ by promoting membrane recruitment, thus increasing the yield of synthesized PE. Low PG content (i.e. high fraction of PE) shifts the equilibrium to PssA$_L$, channeling resources to the production of PG and reducing the fraction of synthesized PE. **b-d**, Concentrations of synthesized PE (**b**), PG (**c**) and PE+PG (**d**) for different initial compositions of SUVs in the presence (green symbols) or absence (blue symbols) of SP6 RNAP. The three different markers represent data from three independent experiments. The solid lines are appended to guide the eye. Membranes always contain 50 mol% DOPC and varying fractions of DOPG and DOPE. Source Data are available for panels **b-d**.

3.2.4 Compartmentalized biosynthesis of PE and PG in liposomes

Lipid synthesis localized inside individual liposomes is of paramount importance in the realization of autonomously growing artificial cells. The successful reconstitution of the seven gene-encoded enzymes for PE and PG synthesis in the presence of LUVs prompted us to confine the entire chain of reactions inside cell-sized liposomes that initially contain PE and PG lipids. PURE system, pGEMM7 minigenome and soluble phospholipid precursors were encapsulated inside large and giant liposomes with a membrane consisting of DOPC/DOPE/DOPG/CL/DHPE-Texas Red/DSPE-PEG-biotin (50 mol%/36 mol%/12 mol%/2 mol%/0.5 mass%/1 mass%). Acyl-chain precursors were supplied as a dried film and, when suspended in the aqueous solution, partitioned in the membrane of liposomes. Cell-free gene expression was restricted to the liposome lumen by adding either proteinase K or DNase I in the external medium. In-liposome gene expression was first validated using the yellow fluorescent protein (YFP) as a reporter (**Fig. 3.3a,b**). Quantitative mass spectrometry analysis of synthesized lipid products showed that it is possible to synthesize up to 20 µM of phospholipid end products, corresponding to an acyl-CoA conversion yield of 40%, when all reactions are confined to the liposome lumen (**Fig. 3.3d**). Both acyl-chain precursors palmitoyl-CoA (16:0) and oleoyl-CoA (18:1) could be used as substrates, resulting in the synthesis of dipalmitoyl and dioleoyl phospholipids, respectively (**Fig. 3.3d**, **Supplementary Fig. 3.10**). Because the newly synthesized DOPE and DOPG are also constituents of the parental liposomes, this result represents a milestone towards homeostatic membrane growth directed from genomic DNA. Control experiments without proteinase K (**Fig. 3.3d**, **Supplementary Fig. 3.10**) result in only slightly higher phospholipid yields, despite the much larger reaction volume of the extravesicular space. This could suggest a possible enhancement of gene expression and/or lipid synthesis by encapsulation inside liposomes[4,38].

PG was only observed when the SP6 RNAP was co-encapsulated (**Fig. 3.3d,e**), demonstrating that genetic regulation of phospholipid synthesis occurs inside liposomes. In accordance with LUV experiments, activation of PG synthesis does not substantially decrease the amount of synthesized PE. In all cases, the final yield of PG was about two times lower than that of PE, mirroring the initial PE/PG ratio of the vesicle membranes. This result suggests that the homeostatic mechanism mediated by PssA takes place when lipid synthesis is compartmentalized inside liposomes. Moreover, we found evidence for accumulation of the phospholipid intermediates LPA, PA, and CDP-DAG, but not of PGP and PS (**Supplementary Fig. 3.10**).

We then aimed to expand the repertoire of synthesized phospholipids by mixing the 16:0 and 18:1 acyl-CoA precursors in equimolar amounts. We found that 82.9 ± 0.4% (without SP6 RNAP) and 79 ± 11% (with SP6 RNAP) of the total synthesized phospholipid end products contained mixed-chain products (PO) (**Fig. 3.3d**, **Supplementary Fig. 3.10**), which is significantly higher than the expected 50% assuming random chain incorporation. The fraction of synthesized dioleoyl (8.8% ± 0.4% without SP6 RNAP, 11% ± 9% with SP6 RNAP) and dipalmitoyl (8.2% ± 0.3% without SP6 RNAP, 9% ± 5% with SP6 RNAP) species was consequently low but appreciable. Concluding, it has been possible to selectively produce up to six different lipid species (DOPE, DOPG, DPPE, DPPG, POPE, POPG) with a one-pot reaction coupling gene expression and phospholipid synthesis within cell-sized liposomes.

3
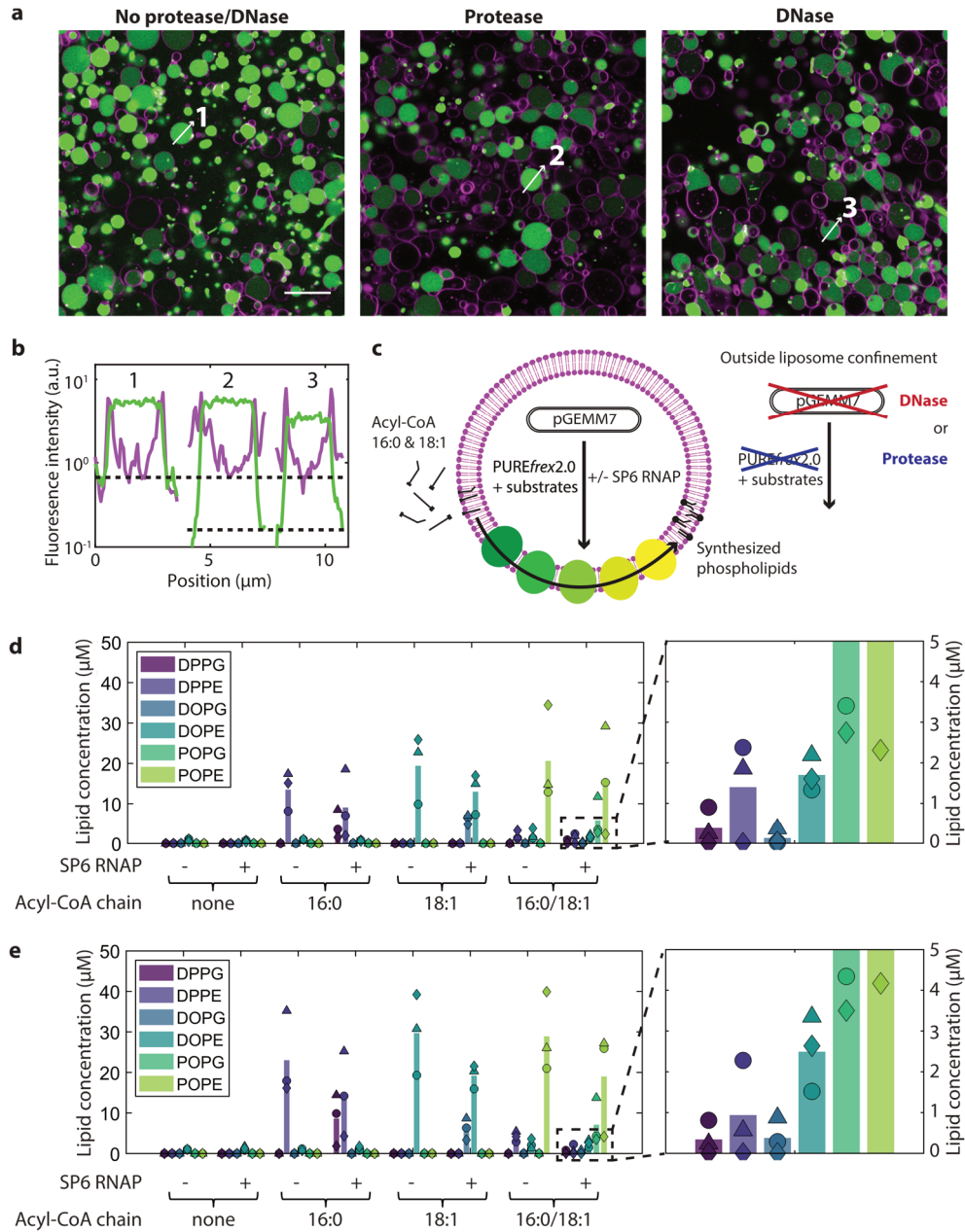
Figure 3.3: DNA-programmed phospholipid synthesis inside giant vesicles. (a) Fluorescence images of liposomes (membrane in magenta) encapsulating PURE*frex*2.0, 5 mM β-mercaptoethanol and 7 nM DNA encoding for the YFP (in green). Gene expression was confined inside liposomes by external addition of either proteinase K (middle) or DNase I (right), or was allowed to also occur outside liposomes (left). Liposomes were diluted (2 µL in 7.5 µL total) to reduce their surface density and aid visualization. Scale bar indicates 20 µm; all pictures have the same size and were acquired with identical imaging settings. **(b)** Fluorescence intensity line profiles for the liposomes indicated in **a**. Color coding is the same as in **a**. Dotted black lines indicate the background YFP level without (top) and with (bottom) protease/DNase. **(c)** Schematic representation of in-liposome gene expression coupled to

phospholipid synthesis. (**d,e)**, Concentration of synthesized phospholipids inside (d) or both inside and outside liposomes (e), as determined by LC-MS. Compartmentalization of gene expression was ensured by addition of proteinase K in the external environment. Different combinations of acyl-CoA precursors and SP6 RNAP were used. Concentrations of oleoyl-CoA and palmitoyl-CoA were 100 µM when added separately, and 50 µM each when added together. Symbols indicate measurements from three independent experiments and the bars represent mean values. A small amount of DOPE was measured in samples where no acyl-CoA was supplied. This represents the naturally occurring heavy-isotope fraction of the DOPE contained in the initial liposome membrane. The right panels are blow-up graphs of the indicated area in the left panels. Source Data are available for panels **d** and **e**.

3.2.5 Direct visualization of gene-encoded membrane synthesis in individual liposomes

In-liposome gene expression is subjected to high heterogeneity even when a single protein is produced from a high copy number of encapsulated DNA molecules[38]. While LC-MS methods provide sensitive detection of multiple lipid species in a liposome population, information about lipid composition at the single vesicle level is lost due to vesicle solubilization. To overcome this limitation and to quantify the fraction of phospholipid-producing liposomes as well as the degree of heterogeneity, we established two fluorescence-based imaging assays. Moreover, optical microscopy methods gave us the opportunity to confirm our assumption that synthesized lipids are incorporated into the liposome membrane.

The first approach was based on the use of the nitrobenzoxadiazole (NBD)-labelled palmitoyl-CoA as a fluorescent substrate for phospholipid synthesis (**Fig. 3.4a**). The integration of the NBD-labelled acyl chain into the different enzymatic products was analyzed by high-performance liquid chromatography (HPLC) (**Fig. 3.4b**). Peak assignment was realized by monitoring chromatograms of samples when only parts of the enzymatic pathway were expressed in the presence of LUVs. New peaks appearing after addition of a gene coding for an enzyme downstream the pathway were assumed to correspond to the final reaction product. In this way, signatures for the NBD-labelled PA, PS, and PE could unambiguously be identified (**Fig. 3.4b,c**). Furthermore, NBD-labelled PA and PE were detected when pGEMM7 was expressed inside cell-sized liposomes (**Fig. 3.4b**). These results demonstrate the versatility of our platform to synthesize novel lipid species.

Next, we performed fluorescence microscopy experiments to image the membrane localization of newly synthesized NBD-labelled phospholipid species from the interior of liposomes. We reasoned that two-acyl chain phospholipid products conjugated to NBD are more stably inserted in the bilayer than mono acyl species (NBD-palmitoyl-CoA and NBD-LPA) that have a faster exchange rate between the membrane and the bulk phase. Therefore, a more intense NBD signal at the liposome membrane is expected upon successful lipid production. A mixture of palmitoyl-CoA and NBD-palmitoyl-CoA (9:1 molar ratio) was used as acyl-chain precursors. This ratio was chosen to minimize the chance of incorporating two NBD-labelled chains in one phospholipid, which might result in fluorophore quenching, whilst yielding a sufficiently high fraction of NBD-labelled phospholipids for imaging. After pGEMM7 expression, the liposomes were diluted to reduce the membrane signal coming from NBD-palmitoyl-CoA and NBD-LPA. Background signal resulting from the transient interaction of NBD-palmitoyl-CoA with the vesicles was assayed in control samples where proteinase K was supplemented both inside and outside liposomes to totally inhibit gene expression (**Supplementary Fig. 3.11**). NBD-enriched liposomes, i.e. liposomes that successfully converted NBD-palmitoyl-CoA into two-acyl compounds were analyzed. Expression of pGEMM7 inside liposomes led to a higher NBD signal at the

membrane (**Supplementary Fig. 3.11**) and to a higher percentage of NBD-enriched liposomes than in the control sample (**Fig. 3.4d**) demonstrating phospholipid biosynthesis at the single vesicle level. In addition, the moderate increase (~50%) of the fraction of NBD-enriched liposomes when omitting the proteinase K (**Fig. 3.4d**) might be explained by an enhancement of enzymatic activity in liposome-confined reactions, as suggested above for lipid production at the population level (**Fig. 3.3d,e**).

Figure 3.4: Cell-free biosynthesis of NBD-labelled phospholipids. (**a**) Schematic illustration of the conversion of NBD-labelled acyl chain precursors into newly synthesized membrane-bound phospholipids. NBD-labelled acyl-CoA molecules undergo dynamic exchange between the membrane and the aqueous solution. Upon processing of the lipid precursors by the five-enzyme pathway, membrane-stable NBD-labelled PE is formed and an increase in NBD fluorescence at the liposome membrane is expected. Either PCR fragments of the individual genes or pGEMM7 were used as DNA templates and no SP6 RNAP was introduced. The precursor mix consisted of 10% (in mol) NBD-labelled palmitoyl-CoA and 90% palmitoyl-CoA. Cell-free gene expression was either performed outside LUVs or it was compartmentalized inside giant vesicles, as indicated. (**b**) HPLC chromatograms of NBD-conjugated species detected when single and multiple genes of the pathway are expressed in the presence of LUVs (top), and when pGEMM7 in expressed inside giant vesicles (bottom). The different gene combinations allowed us to assign peaks to specific lipid end products. Clear peaks were found for NBD-labelled DPPA (peak 1, *plsB* and *plsC*), NBD-labelled DPPS (peak 2, *plsB, plsC, cdsA,* and *pssA*), and NBD-labelled DPPE (peak 3, *plsB, plsC, cdsA, pssA,* and *psd,* labelled 'all DNA'). In the giant vesicle assay (bottom), proteinase K was either added to the inside of liposomes (no lipid synthesis), to the outside of liposomes (lipid synthesis confined to the liposome lumen), or it was omitted. NBD-labelled DPPA and NBD-labelled DPPE could be observed. (**c**) Chemical structures of NBD-labelled DPPA (1), NBD-labelled DPPS (2), and NBD-labelled DPPE (3). (**d**) Percentage of NBD-enriched liposomes, i.e. liposomes that successfully converted NBD-palmitoyl-CoA into two-acyl compounds, was calculated by analyzing line profiles of single liposomes imaged by fluorescence confocal microscopy (**Supplementary Fig. 3.11**). Liposomes were selected for unilamellarity based on the Texas Red channel. NBD-enrichment was defined when the mean NBD peak fluorescence for a given liposome is higher than the mean plus two standard deviations of the signal in the 'proteinase K in' negative control. Conditions correspond to the giant vesicle experiment shown in **b** (bottom). The samples were washed three times to remove non-reacted NBD-palmitoyl-CoA. Bars are mean values from three independent experiments. Symbols indicate data points from individual repeats. A total of 741, 613, and 505 line profiles were analyzed (from left to right). Source Data are available for panel **d**.

The second strategy to detect lipid synthesis and membrane incorporation relies on the C2-domain of lactadherin fused to eGFP (LactC2-eGFP) as a PS-specific fluorescent reporter[39] (**Fig. 3.5a**, **Supplementary Figs. 3.12** and **3.15**). At a concentration of 150 nM, LactC2-eGFP binds to PS-containing membranes, but not to membranes where PS was substituted by PG (**Supplementary Fig. 3.15**). PS is not an end-product of our reconstituted lipid synthesis pathway and is rapidly converted by Psd into PE (**Supplementary Figs. 3.9** and 3.**10**). To enable accumulation of PS, the plasmid DNA pGEMM7 was linearized using EcoRI that cuts at a unique restriction site located in the *psd* gene (**Supplementary Fig. 3.16**). The only end-product of the pathway encoded by the resulting construct (named pGEMM7Δ*psd*) is PS, when the SP6 RNAP is not added. Using pGEMM7Δ*psd* as a template for in-liposome gene expression led to significant accumulation of PS, as detected by LC-MS (**Supplementary Fig. 3.16**). Some residual PE synthesis was also measured, most likely as the result of incomplete restriction of the *psd* gene (**Supplementary Fig. 3.16**). When LactC2-eGFP was added to the feeding solution to probe PS production in individual liposomes, a clear recruitment to the membrane of some liposomes was observed (**Fig. 3.5b,c**), indicating PS-enrichment. No significant membrane binding of LactC2-eGFP was observed when omitting either oleoyl-CoA or the pGEMM7 template (**Fig. 3.5b,d**), corroborating the high PS specificity. Automated image analysis allowed us to extract the average rim intensity of eGFP in a large number of liposomes. A wide distribution of eGFP intensity values in PS-synthesizing liposomes was measured (**Fig. 3.5d**). The coefficient of variation is ~2-fold higher than in control samples with a predetermined fraction of PS (**Supplementary Fig. 3.17**). This result further supports the highly heterogeneous nature of liposome-encapsulated lipid synthesis. Moreover, we found that ~50% of the liposomes exhibited PS enrichment (**Fig. 3.5e**). Similar results were obtained when LactC2 was fused to mCherry in place of eGFP (**Supplementary Note 3.3**, **Supplementary Figs. 3.13, 3.14, 3.17** and **3.18**). We noticed that this approach is more robust and provides higher signal-to-background ratio than the use of an NBD-labelled acyl precursor. Moreover, no washing steps are necessary, making LactC2-eGFP a superior lipid probe to obtain kinetic information by real-time fluorescence imaging of individual liposomes. **Fig. 3.5f,g** shows a representative liposome imaged at six time points. Between 0.5 and 6 h, a clear increase in LactC2-eGFP signal at the membrane can be observed. Plotting fluorescence intensity over time for 47 liposomes from three independent experiments shows a sigmoidal profile representing synthesis and membrane incorporation of PS, with a plateau time of ~4.5 h ± 2.5 h and a rate of 9.2 a.u. ± 6.9 a.u. per minute (**Supplementary Note 3.4**, **Supplementary Fig. 3.19**). No clear dependency of the kinetic parameters with respect to the liposome size was observed for vesicles with an apparent diameter ranging between 4 and 12 µm (**Fig. 3.5g**, **Supplementary Fig. 3.19**). In addition, the amount of de novo synthesized lipids incorporated in the membrane was not sufficient for directly observing liposome growth under an optical microscope (**Fig. 3.5h**). When oleoyl-CoA was omitted, no increase of the LactC2-eGFP signal intensity was observed, confirming the specificity for synthesized PS. Further investigations will be necessary to elucidate the rate limiting step of the LactC2-eGFP signal increase and the cause of saturation. In particular, it would be insightful to examine if LactC2-eGFP recruitment saturates due to cessation of PS production.
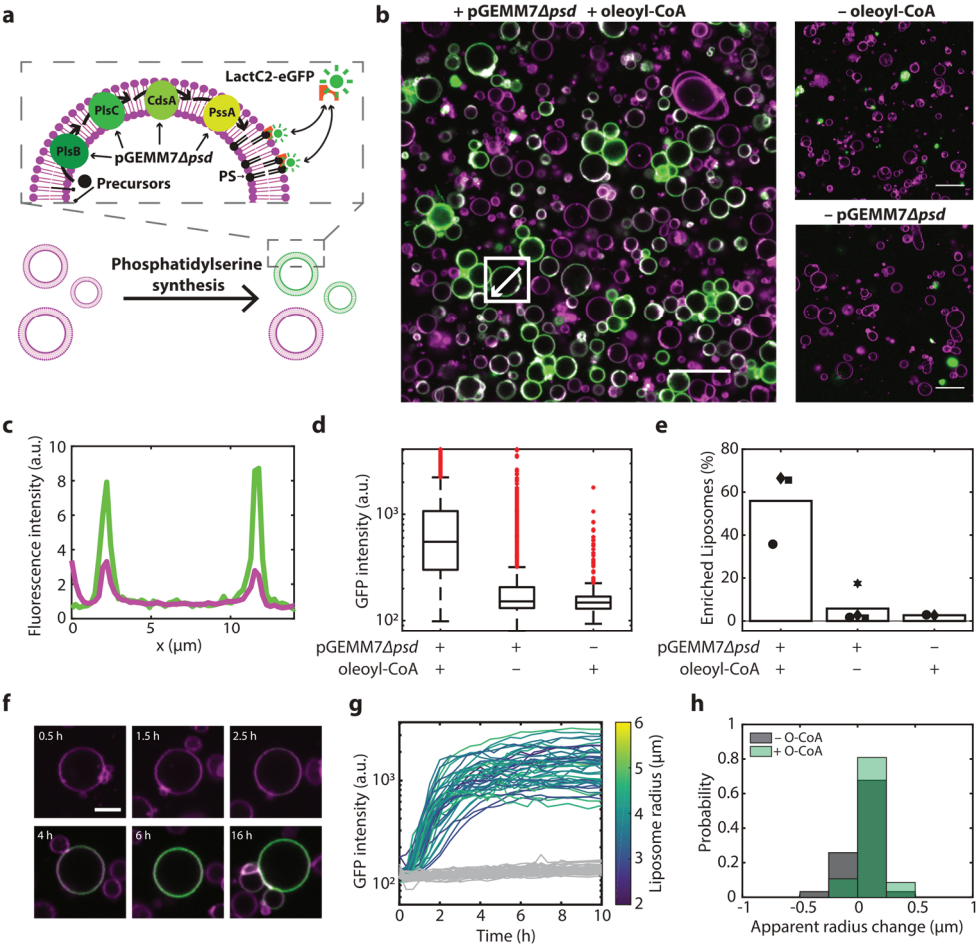
Figure 3.5: Single-vesicle imaging of internal PS production using LactC2-eGFP. (a) Schematic representation of gene expression-coupled PS biosynthesis inside liposomes and fluorescence imaging using the PS-specific LactC2-eGFP probe. The linear pGEMM7Δ*psd* template was expressed within giant vesicles to produce the PlsB, plsC, CdsA and PssA enzymes which catalyze formation of PS from acyl-CoA and G3P. Membrane-exposed PS recruits the fluorescent reporter LactC2-eGFP, resulting in accumulated GFP signal in PS-enriched liposomes. **(b)** Fluorescence confocal images of liposomes (membrane dye in magenta) producing DOPS from oleoyl-CoA, as illustrated in a. The externally added LactC2-eGFP binds to PS-containing liposomes and stains the membrane in green. In a series of negative control experiments, oleoyl-CoA was omitted, or the pGEMM7Δ*psd* DNA was replaced by a DNA coding for an unrelated protein, namely the terminal protein of the φ29 phage[40]. Bright spots of clustered LactC2-eGFP molecules that do not co-localize with liposomes are sometimes visible. The LactC2-mCherry variant showed less propensity to form clusters than the eGFP fusion protein (**Supplementary Fig. 3.18**) and similar quantitative results were obtained with the two reporters (**Supplementary Figs. 3.17**, **3.18**). (**c**) Line profiles of LactC2-eGFP intensity (green) and Texas Red membrane dye intensity (magenta) of the liposome highlighted in b. **(d)** Box-plot representation of the single-vesicle average LactC2-eGFP intensity values for the indicated samples. Data were pooled from four independent experiments for each condition, corresponding to 4048, 3642 and 569 liposomes analyzed (from left to right). Membrane-localized GFP fluorescent intensity is significantly higher when both DNA and oleoyl-CoA were present compared to negative controls, with *p* < 0.0001 (two-sample Welch's t-test). **(e)** Percentage of PS-enriched liposomes for the three types of samples analyzed in d. Values represent the mean ± inter-sample standard deviation from four biological repeats. Liposomes were defined as enriched in PS if their average LactC2-eGFP intensity is higher than the mean plus two standard deviations of the intensity distribution obtained with liposomes containing 0% PS (see **Supplementary Fig. 3.15**). The amount of PS-enriched liposomes

is significantly higher when both DNA and oleoyl-CoA were present compared to the two negative controls, with p < 0.04 (two-sample Welch's t-test). **(f)** Time-lapse images of a liposome exhibiting increasing LactC2-eGFP signal over time. Scale bar is 5 μm; picture size is identical for all images. **(g)** Single-liposome kinetics of LactC2-eGFP binding for 47 PS-synthesizing liposomes, color-coded with respect to the liposome radius (aggregated data from three independent repeats), and for 28 liposomes in a control sample that does not contain oleoyl-CoA (grey curves). **(h)** Probability distributions of the apparent radius change for the liposomes analyzed in panel **(g)**. The apparent radius change was determined by calculating the difference between the apparent radius at 0 h and 16 h. The apparent radius was calculated based on the perimeter of the liposome selection marker under the assumption of a perfect circle. A bin width of 250 nm was chosen. Identical imaging settings were used for all acquired data. Source Data are available for panels **d**, **e**, **g**, **h**.

3.3 DISCUSSION

We demonstrated here that an entire bacterial phospholipid synthesis pathway can be reconstituted inside liposomes by expressing seven membrane-associated enzymes from their genes concatenated on a DNA minigenome. Because the internally synthesized PE and PG lipids are also constituents of the liposome membrane, our synthetic cell platform satisfies the key requirements for self-maintenance. Moreover, higher-level regulation of membrane composition was provided through genetic control and metabolic feedback mechanisms, two processes that have so far been considered to be exclusive attributes of living organisms. The average PE-to-PG ratio could be maintained within the liposome population during phospholipid production, which is important to achieve homeostatic membrane growth.

What is the distribution of internally produced lipids in the bilayer? Phosphatidylserine is likely synthesized on the inner leaflet of the liposome membrane[41]. Nevertheless, synthesized PS is detected on the outer leaflet, where it is exposed to the LactC2-eGFP probe. Flipping of phospholipids is not energetically favorable and requires assistance of specialized enzymes in vivo. However, the artificial bilayer of our liposomes is not as rigid as the bacterial cell membrane and is more prone to transient defects[38,42]. Therefore, membrane dynamic processes, such as lipid flip-flop and translocation of small molecules, may be less impaired in liposomes, facilitating partitioning of PS (and possibly of other synthesized lipids too) in the outer leaflet. Although the possibility that LactC2-GFP permeates across the membrane and binds PS exposed to the lumen cannot be excluded, this process is severely hindered by the bulky fusion protein.

To realize a full proliferation cycle, all membrane constituents should be co-synthesized. Here, we showed internal production of DOPE and DOPG, two out of the four membrane building blocks. Synthesis of the two other lipids, namely PC and CL, would require the reconstitution of only two additional proteins, PmtA and ClsA, respectively[43,44].

Besides having clear implications to creating a synthetic minimal cell, we envisage that our engineered liposomes could serve as a versatile platform for tailored biosynthesis of natural and artificial phospholipids of industrial or biotechnological value, such as lipids with asymmetric acyl chain compositions and acyl-labelled phospholipid analogues (**Supplementary Table 3.7**).

Single vesicle imaging revealed that a significant fraction of the liposomes does not display lipid enrichment (**Fig. 3.4d**, **Fig. 3.5b**,**e**). Moreover, a large heterogeneity in the level of synthesized

3

phospholipids was observed among liposomes (**Fig. 3.5d**, **Supplementary Fig. 3.17**). Such a compositional and functional heterogeneity within a population of PURE system-containing liposomes has been reported in other studies[38,42,45,46] and is inherent to gene expression in cell-sized compartments. In the present experiments, other sources of heterogeneity in lipid enrichment may also contribute, such as a variability in the adsorption of acyl-CoA among liposomes upon resuspension of the precursor film. Investigating the mechanisms leading to phenotypic differences will be important to further optimize the chain of reactions from genes to output lipids.

Synthesis of phospholipids from an internal machinery and their incorporation in the lipid bilayer inevitably results in liposome growth. However, no visible membrane or volume expansion could unambiguously be measured by optical microscopy. It is clear that the amount of synthesized lipids per liposome should be increased to achieve doubling of the membrane surface area, which is necessary for sustainable proliferation. We envisage two complementary strategies to overcome this limitation, one acting at the gene expression level, the other at the lipid biosynthesis level. First, one could strive to improve the PURE system performance for producing larger amounts of the encoded enzymes in liposomes. Given the limited knowledge about the biochemical steps governing PURE system reactions[47], it remains nevertheless challenging to find generic solutions for improved DNA sequence design and composition of the PURE system. Alternatively, employing the native *E. coli* RNA polymerase along with sigma factors would expand the capabilities to regulate transcription of individual genes and fine-tune the level of individual proteins[48,49]. This could in turn ameliorate the production rate and yield of the output lipids.

Another factor that might limit the final amount of synthesized phospholipids is the initial concentration of acyl-CoA, absolute and relative with respect to the concentration of liposomes. Adding more than 100 µM acyl-CoA is hardly feasible due to its poor solubility in the presence of high concentration of Mg^{2+} contained in the PURE system and to its detergent effect on liposome membranes. One solution would be to provide a continuous supply of low-concentration acyl-CoA. Alternatively, more soluble precursors, such as acyl-ACP, fatty acids and malonyl-CoA could be used. Expanding the pathway upstream by introducing the FadD enzyme would enable to substitute acyl-CoA with a fatty acid and CoA[21]. Finally, the highly soluble malonyl-CoA could be used as a substrate provided the eukaryotic FASII mega-protein can be synthesized in a functional state in the PURE system[50].

Alternatively, chemical synthesis of non-natural phospholipids has emerged as an interesting strategy because of its high yield and quick conversion[12-14]. Chemical synthesis could potentially be coupled to one or more enzymatic reactions, resulting in a hybrid system equipped with genetic control facilitating rapid lipid synthesis[51]. A radically different approach would consist to use alternative membrane constituents, such as amphiphilic peptides, that would be expressed from the inside of the vesicle[52].

DNA-programmed lipid synthesis could be exploited as a rudimentary mechanism to trigger division of liposomes. Budding events reminiscent of the proliferation mode of L-form bacteria[53] could be stimulated through an excess membrane synthesis, potentially aided by gentle shear forces. Furthermore, internal synthesis of membrane remodeling phospholipids like DLPE, produced when starting from the short-chain 12:0 acyl-CoA, in combination with temperature cycling[54], might assist vesicle deformation and division.

It has not escaped our attention that liposome-confined DNA-based phospholipid production, combined with the fluorescently tagged LactC2 as a selection marker, is amenable to directed evolution experiments, owing to the linkage between genotype and phenotype. Activity of single or multiple enzymes in the pathway, or substrate selectivity, could be improved by generating a library of mutagenized genes and selecting for PS-enriched liposomes by fluorescence activated cell sorting[55]. This strategy may become decisive when combining membrane growth with other functional modules, such as DNA replication[40] and liposome division[56,57].

3.4 METHODS

All buffers and solutions were made using Milli-Q grade water with 18.2 MΩ resistivity (Millipore, USA). Chemicals were purchased from Sigma-Aldrich unless otherwise indicated.

3.4.1 Design and assembly of the pGEMM7 plasmid

The plasmid pGEMM7 was assembled from seven PCR fragments containing independent transcriptional cassettes and the plasmid backbone of pUC19 (New England Biolabs, USA) (**Supplementary Table 1**). The genes were used in a previous study with each gene inserted in a separate DNA construct[23]. Individual genes were amplified by PCR using primers containing linker sequences to determine the order and orientation of each cassette in the final plasmid. Linker sequences of 30 bp were designed by a random DNA generator such that they had no or little homologies to the *E. coli* genome (R20DNA designer, https://www.syntegron.org/R20/R20/R20.html, Imperial College London) to minimize unwanted recombination events. The vector backbone was amplified using primers 829 and 830 giving a 1,932 bp product bearing either linker site 1 or 13 on the ends. The transcriptional cassette of *plsB* was amplified using primers 628 and 629 introducing linker site 1 upstream of the gene and linker site 2 downstream. All other remaining transcriptional cassettes were made in the same way adding linker sequences both upstream and downstream of the cassette to enable each cassette to be recombined with the next one by Gibson assembly[31]. Primer 819 also adds an SP6 promoter to the *pgsA* gene as well as a linker sequence. The second SP6 promoter sequence was added to the *pgpA* gene in a previous step using primer 817. The homologous site that was added using primer 817 was deleted in a subsequent PCR using primer 851. **Supplementary Fig. 3.1** shows a schematic drawing of the two-step process to incorporate the homologous linker sites by PCR and then using the individual fragments to assemble pGEMM7. The primers, their targets and the homologous site they are bearing are listed in **Supplementary Table 3.2**.

3.4.2 Cloning of the pGEMM7 plasmid

All fragments for Gibson assembly were amplified using Phusion High-Fidelity DNA polymerase (New England Biolabs, USA) with the recommended standard reaction conditions from the supplier. Elongation times and primer annealing temperatures were varied according to primer length between 55 °C and 65 °C. Primers and remnants of the PCR reaction were removed using the Wizard PCR cleanup kit (Promega, USA). The concentration of the purified DNA was determined using an ND-2000 Nanodrop spectrophotometer. Purified PCR products were mixed following the pipetting scheme in **Supplementary Table 3.3** plus 15 µL of prepared Gibson assembly mix containing 100 mM Tris-HCl,

50 mM MgCl$_2$, 0.2 mM each dNTP, 10 mM dithiothreitol (DTT), 5% w/v PEG-8000, 1 mM NAD, 5.33 U mL^{-1} T5 Exonuclease, 33.3 U mL^{-1} Phusion polymerase and 5.33 U mL^{-1} Taq-ligase in a final volume of 20 µL. The Gibson assembly mixture was incubated at 50 °C for 1 h and 5 µL were subsequently used for transformation of 50 µL One Shot™ TOP10 Chemically Competent *E. coli* cells (ThermoFisher Scientific, USA).

Transformed cells were recovered in 1 mL LB medium for 1 h and transferred on LB-Agar plates containing 50 µg mL^{-1} ampicillin. After overnight incubation at 37 °C, ten colonies were selected for colony PCR using primers 91 and primers 397 which bind in the T7 terminator region and the RBS, respectively. Four of the tested colonies gave the expected pattern (**Supplementary Fig. 3.3a**) and were subsequently grown overnight in LB medium. Their plasmid DNA was isolated using a PureYield miniprep kit (Promega, USA) and was further analyzed with restriction digestion using the enzymes EcoRI-HF, SacI and DraI (New England Biolabs, USA). **Supplementary Fig. 3.3b** shows that all four colonies gave the expected pattern consisting of digestion products of 4,300 bp, 2,836 bp, 1,863 bp, 1,395 bp, 692 bp, and 19 bp (indicated by black stars, only the 19-bp product was not visible), plus some side products attributed to incomplete DNA digestion. The correct DNA sequence was finally confirmed with Sanger sequencing (Macrogen, South-Korea).

3.4.3 Cloning of *eGFP-lactC2* and plasmid purification

The original plasmid containing the *egfp-lactC2* gene was described in ref. 39 and was kindly provided by the lab of Dorus Gadella (University of Amsterdam, Netherlands). To enable expression and isolation from *E. coli*, regular PCR reactions were performed to amplify both the plasmid backbone of a pET11a vector and the *egfp-lactC2* gene construct. Primers 471 (forward) and 850 (reverse) were used for the amplification of the pET-11a backbone. Primers 848 (forward) and 849 (reverse) were used for the amplification of *egfp-lactC2*. The reaction was performed with 10 ng of template DNA, 1 U of Phusion High-Fidelity DNA Polymerase (New England Biolabs) in HF buffer and supplemented with 0.2 mM of dNTPs, and 0.2 µM of both forward and reverse primers in a final volume of 50 µL. An initial heating step at 95 °C for 5 min was applied to allow denaturation of DNA. The PCR reaction consisted of 34 cycles of 30 s steps for melting DNA at 95 °C, followed by the hybridization of the primers for 30 s at 55 °C and the elongation by the DNA polymerase at 72 °C for 30 s per kb template. After the 34 cycles, the temperature was kept at 72 °C for 5 min. Both PCR products were purified using the Wizard PCR cleanup kit (Promega, USA).

The size of the PCR products was verified on a TAE agarose gel (1% w/v) using SYBR safe staining (Thermo Fisher). The BenchTop 1-kb DNA Ladder from Promega was used. The fragments corresponding to the adequate sequence lengths of 1.3 kb and 5.6 kb were excised from the gel and purified using the Promega Wizard SV Gel and PCR Clean-Up System kit. DNA concentration of the eluate was determined by measuring the absorbance at 260 nm with a Nanodrop 2000c.

The pET-11a backbone and *egfp-lactC2* gene fragments were assembled using Gibson assembly [31]. 100 ng of backbone and an equimolar amount of the *egfp-lactC2* PCR fragment were mixed in a solution containing 100 mM Tris-HCl, 50 mM MgCl$_2$, 0.2 mM each dNTP, 10 mM DTT, 5% w/v PEG-8000, 1 mM nicotinamide adenine dinucleotide, 5.33 U mL^{-1} T5 Exonuclease, 33.3 U mL^{-1} Phusion polymerase and 5.33 U mL^{-1} Taq-ligase in a final volume of 20 µL. The assembly reaction was incubated at 50 °C for 60 min. Then, 20 U µL^{-1} of DpnI restriction enzyme (New England Biolabs, USA)

were added to digest possible methylated DNA left and the mixture was incubated for an additional 15 min at 37 °C.

Five microliters of the assembly mixture were transformed into 50 μL of One Shot™ TOP10 chemically competent *E. coli* cells using heat shock. The cells were heat shocked in a water bath at 42 °C for 45 s and then transferred back to ice for 2 min, to reduce cell damage. After incubation in 1 mL of LB medium (1:20 dilution) for 20 min at 37 °C, 50 μL of the cell suspension were spread in LB plates supplemented with 50 μg mL^{-1} ampicillin. The remaining sample was pelleted, re-suspended in 50 μL of LB medium and plated. All plates were incubated overnight at 37 °C.

Six colonies were picked to perform colony PCR and a replica plate was made. A PCR reaction was performed with 0.5 U of GoTaq DNA Polymerase in GoTaq Buffer (both from Promega) supplemented with primers and dNTPs to a final volume of 20 μL. Adequate forward and reverse primers (25 and 310, respectively) were chosen to amplify the gene region and part of the backbone sequence upstream and downstream of the gene (**Supplementary Table 3.2**). DNA was purified using the Promega Wizard® SV Gel and PCR Clean-Up System and analyzed on gel. Colonies leading to a band with the predicted length (6.9 kb) were grown in 5 mL LB medium overnight and plasmid DNA was isolated using the PureYield Plasmid Miniprep System (Promega). The plasmids were further tested by a restriction enzyme digestion analysis, in which 2.5 U of DraI and 2.5 U of StuI were mixed with 500 ng of DNA, in a final volume of 20 μL (both enzymes were from New England Biolabs). The mixture was then incubated at 37 °C for 1 h. Digested DNA was separated in TAE agarose gel (1%).

To infer the quality of the construct on the sequence level, DNA extracted from the six colonies was sequenced by Sanger sequencing (Macrogen). To 300 ng of plasmid DNA, 0.25 μM of adequate primers (288 and 25, **Supplementary Table 3.2**) were added, in a final volume of 10 μL. Plasmids with the correct sequence were selected.

3

3.4.4 Overexpression and purification of LactC2-eGFP and LactC2-mCherry

E. coli Rosetta ER2566 cells (New England Biolabs) and Rosetta 2 cells (Novagen) suited for protein overexpression were transformed with the plasmid for LactC2-eGFP by heat shock. The plasmid for LactC2-mCherry was transformed into Rosetta 2 cells and isolated in the same way as described below. A preculture of these strains was incubated overnight at 37 °C in LB medium supplemented with 50 µg L^{-1} ampicillin. Then, the cultures were diluted in the same medium in a ratio of 1:1000 and incubated at 37 °C with agitation (200 rpm) until an OD$_{600}$ of ~0.6 was reached. Protein production was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside. The cells were incubated at 30 °C for 3 h under agitation (200 rpm) and were pelleted by centrifugation at 13,000 rpm for 5 min. The pellet was resuspended in buffer A (150 mM NaCl, 20 mM imidazole, 20 mM Tris pH, 7.5) and the cells were disrupted by sonication using ten pulses of 10 s and 30 s of interval, with 30% amplitude. After centrifugation at 4 °C for 15 min and 13,000 rpm, the supernatant was cleared from debris.

Protein purification was done using Ni-NTA Spin Columns (Qiagen) following the supplier recommendations. The column was equilibrated and washed with buffer A and the protein was eluted with buffer B (150 mM NaCl, 500 mM imidazole, 20 mM Tris, pH 7.5). The elution buffer was exchanged for the storage buffer (10 mM HEPES-KOH, pH 7.5) using Zeba Spin Desalting Columns (ThermoFischer). This size exclusion chromatographic spin down columns retain small molecules (<1 kDa) and recover mostly large molecules (>7 kDa). Throughout all the steps of protein purification and buffer exchange, samples were harvested for subsequent analysis in polyacrylamide gels.

The 12% polyacrylamide resolving gel and the 4% stacking gel were prepared with final concentrations of 0.12% of sodium dodecylsulfate, 150 mM of Tris-HCl pH 8.8 for the resolving gel and 10 mM of Tris-HCl pH 6.8 for the stacking gel. Ammonium persulfate and tetramethylethylenediamine were added after to begin polymerization. The loading solution consisted of 15 µL of the protein sample mixed with 1 µL DTT and 15 µL Laemmli 2× Concentrate Loading Buffer (Sigma-Aldrich), and denatured at 95 °C for 10 min. The gel was run first at 100 V for 15 min and then at 180 V for approximately 45 min. Running buffer consisted of 250 mM Tris-HCl, 200 mM glycine, 1% w/v SDS, pH 8.3.

Concentration of the protein was measured with a Bradford assay. Bovine serum albumin was used as a standard spanning seven concentrations from 0.25 mg mL^{-1} to 2 mg mL^{-1}. Each sample was assayed in triplicate, including a Milli-Q sample, and the absorbance at a wavelength of 595 nm was measured by spectrophotometry.

3.4.5 Proteomics

A targeted proteomics approach was used following established in-house protocols. Samples of PURE*frex*2.0 (GeneFrontier, Japan) of 1 µL were taken and incubated at 55 °C for 20 min in 16.5 µL of 50 mM Tris-HCl, pH 7.6, 0.1% 2-octoglycoside, 12.5 mM DTT and 1 mM CaCl$_2$. Then, 32.6 mM final concentration of iodoacetamide was added and the solution was incubated for 30 min in the dark. Finally, 0.5 µg of trypsin was added, and the solution was incubated overnight at 37 °C. The following day, 2 µL of 10% trifluoroacetic acid was added, the sample was incubated at room temperature for 5 min, the solution was centrifuged at 16,200 r.c.f. for 30 min and the supernatant was transferred to an HPLC-vial for analysis.

Mass spectrometry analysis of tryptic peptides was conducted on a 6460 Triple Quad LC-MS system (Agilent Technologies, USA). From the samples prepared according to the protocol described above, 10 µL were injected into an ACQUITY UPLC® Peptide CSH™ C18 Column (Waters Corporation, USA). Peptides were separated in a gradient of buffer C (25 mM formic acid in Milli-Q) and buffer D (50 mM formic acid in acetonitrile) at a flow rate of 500 µL per minute at a column temperature of 40 °C. The column was equilibrated with 98:2 ratio of buffer C to D. After injection, over 20 min the ratio was changed to 75:25 buffer C to D after which, within 30 s, the ratio went to 20:80 buffer C to D and was held for another 30 s. Finally, the column was flushed for 5 min with 98:2 buffer C to D ratio. **Supplementary Table 3.3** shows the transitions of the MS/MS measurements that were observed in every experiment. EF-Tu is a constant component of the PURE system and served as a global internal standard for variations due to evaporation or sample handling. All data was represented as the peak integrated intensity of a given peptide normalized to that of the TTLTAAITTVLAK peptide of EF-Tu. All proteomics results were analyzed in Skyline-daily 4.1.1.18179 (MacCoss lab, University of Washington, USA).

Retention time was predicted after standard runs with the above-described method using the Pierce™ Peptide Retention Time Calibration Mixture (Catalog number 88320, Thermo Scientific, USA).

3.4.6 Precursor films

Palmitoyl-CoA, oleoyl-CoA, and NBD-palmitoyl-CoA were obtained from Avanti Polar lipids (USA) in powdered form. The powders were dissolved in chloroform:methanol:water (40:10:1 vol. fractions), aliquoted, dried, and stored under argon. Before use, the acyl-CoA's were resuspended and diluted in chloroform to a final concentration of 100 µM. Using Gilson Microman pipettes, the acyl-CoA solution was added to PCR tubes. Organic solvent was evaporated at ambient pressure and temperature for ~5 h, resulting in a dried precursor film. Acyl-CoA volumes were chosen such that the concentration of precursor after resuspension in the samples was 100 µM (50 µM for NBD experiments). For NBD experiments, films consisted of 10% NBD-palmitoyl-CoA and 90% palmitoyl-CoA and exposure to light was limited to prevent fluorophore bleaching.

3.4.7 LUV experiments

LUVs were prepared by extrusion of large multilamellar vesicles (LMVs). A 2 mg lipid mixture consisting of DOPC/DOPE/DOPG/CL/DSPE-PEG-biotin (50 mol%/36 mol%/12 mol%/2 mol%/1 mass%) dissolved in chloroform was prepared in a 2 mL glass vial, dried under gentle argon flow and subsequently desiccated for 1 h. The film was then resuspended in 250 µL buffer E (20 mM HEPES, 180 mM potassium glutamate, 14 mM magnesium acetate, pH 7.6) and vortexed to create LMVs. Four freeze-thaw cycles were applied, and samples were extruded with a 400 nm membrane using the Avanti mini-extruder, according to instructions provided by the manufacturer. LUVs were aliquoted, snap-frozen in liquid nitrogen and stored at −20 °C.

PURE*frex*2.0 reaction solutions were assembled according to the instruction provided by the manufacturer, and supplied with 0.75 U µL⁻¹ SUPERase (Invitrogen), 5 mM β-mercaptoethanol, 500 µM ^{13}C-labelled G3P, 1 mM CTP, 500 µM L-serine, 0.4 mg mL⁻¹ lipids from the LUV mixture, and 1 nM of pGEMM7 plasmid, unless stated otherwise. When indicated, 2 U µL⁻¹ SP6 RNAP was supplemented. The reaction mixture was then added to the dried precursor film and incubated overnight at 37 °C.

For the experiments shown in **Fig. 3.2**, LUVs with lipid compositions DOPC:DOPE (50:50 molar ratio) and DOPC:DOPG:CL (50:48:2 molar ratio) were prepared as described above, and were mixed in various ratios. Membrane fusion was promoted by applying four freeze-thaw cycles.

3.4.8 In-liposome gene expression assays

Giant vesicles were prepared according to the methods described in [38]. Briefly, 2 mg lipids consisting of DOPC/DOPE/DOPG/CL/DHPE-Texas Red/DSPE-PEG-biotin (50 mol%/36 mol%/12 mol%/2 mol%/0.5 mass%/1 mass%) were mixed with 25.4 µmol rhamnose in methanol and the mixture was added to 0.6 g of 212-300 µm glass beads (acid washed). Beads were rotary evaporated for 2 h at room temperature and 20 mbar and desiccated overnight to remove residual organic solvent. The lipid-coated beads were stored under argon at −20 °C up to one month.

PURE*frex*2.0 reaction solutions were assembled similarly to LUV experiments. Per 10 µL PURE*frex*2.0 reaction mixture, 5 mg (10 mg for LactC2-eGFP experiments) of lipid-coated beads were added. Lipid film swelling was performed for 2 h on ice with gentle tumbling every 30 min. Four freeze-thaw cycles were applied by dipping the sample in liquid nitrogen and thawing at room temperature. The supernatant (corresponding to about 50% of the total volume) was transferred to an Eppendorf tube using a cut pipette tip to avoid liposome breakage. To confine gene expression reactions to the inside of liposomes, 50 µg µL^{-1} proteinase K was added to the liposome sample, unless indicated otherwise. For experiments involving LactC2-eGFP, 2 µL of liposome-containing supernatant were diluted in 5.5 µL of a feeding solution consisting of PURE*frex*2.0 Solution I and Milli-Q (3:7), 150 nM of LactC2-eGFP and 0.07 U µL^{-1} RQ1 DNase (Promega). Liposomes were then transferred to the tube with deposited dried precursor films. Reactions were incubated overnight at 37 °C, or, in the case of time-lapse microscopy, liposomes were immediately immobilized for imaging (see below).

3.4.9 Sample preparation for LC-MS and HPLC

A solution consisting of methanol with 5 mM EDTA and 2 mM acetylacetone was prepared fresh for every experiment. Samples were diluted 10- (for HPLC) or 100-fold (for LC-MS) in the methanol solution, sonicated for 10 min, and centrifuged for 5 min at 16,000 g. The supernatant containing the lipid fraction was transferred to Agilent 2 mL glass mass spectrometry vials with a low-volume inset, flushed with argon and stored at −20 °C. Samples were analyzed within one week from preparation.

3.4 METHODS

3.4.10 Liquid chromatography and mass spectrometry analysis of lipids

Mass spectrometry measurements of phospholipid samples were performed using a 6460 Triple Quad LC-MS system equipped with a similar ACQUITY UPLC® Peptide CSH™ C18 Column as used in proteomics. However, different columns were used for each application, to prevent cross-contamination. Separation of lipids was performed using a gradient of mobile phase F (water with 0.05% ammonium hydroxide and 2 mM acetylacetone), and mobile phase G (80% 2-propanol, 20% acetonitrile, 0.05% ammonium hydroxide and 2 mM acetylacetone) at a flow rate of 300 µL min^{-1} and a column temperature of 60 °C. To equilibrate the column, a ratio of mobile phase F to mobile phase G of 70:30 was used. Upon injection, this ratio was gradually changed to 100% mobile phase G over the course of 8 min and then kept like that for 2 min. Subsequently, over the course of 1 min, the initial 70:30 ratio of mobile phase F and G was reset, which was then used for the last 4 min of the run. The built-in autosampler of the LC-MS system was used to inject 1 µL (quantitative analysis) or 5 µL (qualitative analysis of low-abundance compounds) of sample solution.

Transitions were established based on previous work[23], as well as scanning measurements of purified standards. The very regular fragmentation pattern (except for LPA and CDP-DAG fragmentation always occurs at the ester linkage between an acyl chain and the glycerol) could be used to determine transitions. Synthesized phospholipids were distinguished from phospholipids present at the start of the reaction as part of the liposome matrix by incorporation of ^{13}C-G3P, resulting in a 3 Da (or 6 Da for PG) mass shift.

Mass spectrometry data was analyzed using the Agilent Masshunter Quantitative analysis program, which automatically integrates peaks corresponding to the transitions set in the method. Integrated peak intensities were exported to MATLAB R2016b (MathWorks) for further analysis. For each transition in each sample, the average integrated counts of two injections was determined. For end products, integrated counts were converted to concentrations using linear calibration curves fitted to signals from a dilution series of standards ran before and after every mass spectrometry measurement series.

3.4.11 HPLC

High-pressure liquid chromatography (HPLC) was used to separate synthetized NBD-labelled lipid species. An Agilent Technologies 1260 Infinity HPLC system equipped with an HSS T3 2.5 µm column was used with mobile phase H (60% acetonitrile, 40% water, 0.0114% formic acid, 7 mM ammonium formate, and 2 mM acetylacetone) and mobile phase I (90% 2-propanol, 10% acetonitrile, 0.0378% formic acid, and 2 mM acetylacetone), as previously reported in Scott et al.[23]. The flow rate was 500 µL min^{-1} and the column temperature was 35 °C. Upon injection of 5 µL of sample, 100% mobile phase H was used, over the course of 1.5 min changing to a ratio of mobile phase H to mobile phase I of 35:65, which was then gradually changed in 8.5 min to 30:70, and then, in 2 min, to 5:95, which was retained for 1 min. Subsequently, in the final 2 min of the run, the initial gradient was restored. NBD fluorescence was detected with an excitation wavelength of 463 nm and an emission wavelength of 536 nm.

3

3.4.12 Microscopy

Liposomes were immobilized in custom-made glass imaging chambers pre-incubated for 10 min with BSA-biotin:BSA (1 mg mL^{-1}) and then with Neutravidin (1 mg mL^{-1}). When appropriate, free NBD-palmitoyl-CoA was removed by washing the sample three times with an equal volume of buffer E, followed by 30 min of incubation at 37 °C. Image acquisition was performed using a Nikon A1R Laser scanning confocal microscope using the following excitation/emission wavelengths: 457/525 nm (NBD), 488/509 nm (LactC2-eGFP), 514/540 nm (YFP), and 561/595 nm (Texas Red). The sample height was adjusted manually in order to equatorially dissect as many liposomes as possible. Within data sets, identical imaging settings were always used.

3.4.13 Image analysis

To determine NBD and LactC2-eGFP fluorescence intensity at the membrane, both manual and automated image analyses have been applied concurrently. For manual image analysis, Fiji [58] was used to obtain line profiles of Texas Red and NBD/eGFP intensity along cross-sections of liposomes selected for unilamellarity in the membrane dye channel. To prevent bias, the NBD/eGFP channel was not viewed during analysis. The two peaks in the NBD/eGFP line profiles were subsequently detected using a custom MATLAB R2016b script, and the average intensity of these peaks was calculated. Line profiles with less or more than two peaks were discarded from the analysis.

The automated image analysis script was written in MATLAB R2016b and was based on the image analysis procedure we previously developed[38]. In short, a *floodfill.m* algorithm was used to determine liposome lumina, based on the Texas Red membrane signal. To determine the NBD/eGFP intensity along the membrane, first the centroid and radius were determined for every detected liposome. Then, intensity profiles along a line from the centroid to 1.5 times the radius, along 63 different angles, were determined. For every line profile, the maximum intensity, corresponding to the membrane intersection, was recorded, and values were averaged to obtain the NBD/eGFP intensity of the membrane. Since this approach is quite sensitive to possible deviations from a spherical shape, a more stringent circularity criterion than previously reported was applied.

REFERENCES

1.   Pohorille, A., & Deamer, D. Artificial cells: prospects for biotechnology. *Trends Biotechnol* **20,** 123–128 (2002).

2.   Luisi, P. L. Toward the engineering of minimal living cells. *Anat. Rec.* **268**, 208–214 (2002)

3.   Luisi, P. L., Ferri, F., & Stano, P. Approaches to semi-synthetic minimal cells: a review. *Naturwissenschaften* **93,** 1–13 (2006).

4.   Nomura, S. et al. Gene expression within cell-sized lipid vesicles. *ChemBioChem* **4,** 1172–1175 (2003).

5.   Noireaux, V. & Libchaber, A. A vesicle bioreactor as a step toward an artificial cell assembly. *Proc. Natl. Acad. Sci. U.S.A* **101,** 17669-74 (2004).

6. Varela, F. G., Maturana, H. R., & Uribe, R. Autopoiesis: the organization of living systems, its characterization and a model. *BioSystems* **5,** 187–196 (1974).

7. Gánti, T. Organization of chemical reactions into dividing and metabolizing units: the chemotons. *BioSystems* **7,** 15–21 (1975).

8. Wick, R., Walde, P., & Luisi, P. L. Light microscopic investigations of the autocatalytic self-reproduction of giant vesicles. *J. Am. Chem. Soc.* **117,** 1435–1436 (1995).

9. Hanczyc, M. M., Fujikawa, S. M., & Szostak, J. W. Experimental models of primitive cellular compartments: encapsulation, growth, and division. *Science* **302,** 618–622 (2003).

10. Deshpande, S., Wunnava, S., Hueting, D. & Dekker, C. Membrane tension-mediated growth of liposomes. *Small* **15,** e1902898 (2019).

11. Tsuji, G., Fujii, S., Sunami, T. & Yomo, T. Sustainable proliferation of liposomes compatible with inner RNA replication. *Proc. Natl. Acad. Sci. U. S. A.* **113,** 590–595 (2016).

12. Kurihara, K., et al. Self-reproduction of supramolecular giant vesicles combined with the amplification of encapsulated DNA. *Nat. Chem.* **3,** 775–781 (2011).

13. Hardy, M. D., et al. Self-reproducing catalyst drives repeated phospholipid synthesis and membrane growth. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 8187–8192 (2015).

14. Brea, R. J., Cole, C. M. & Devaraj, N. K. In situ vesicle formation by native chemical ligation. *Angew. Chemie - Int. Ed.* **53,** 14102–14105 (2014).

15. Kurihara, K. et al. A recursive vesicle-based model protocell with a primitive model cell cycle. *Nat. Commun.* **6,** 8352 (2015).

16. Chen, I. A., Roberts, R. W. & Szostak, J. W. The emergence of competition between model protocells. *Science* **305,** 1474–1476 (2004).

17. Adamala, K. P., Engelhart, A. E. & Szostak, J. W. Collaboration between primitive cell membranes and soluble catalysts. Nat. Commun. **7,** 11041 (2016).

18. Deamer, D. W. & Gavino, V. Lysophosphatidylcholine acyltransferase: purification and applications in membrane studies. *Ann. N. Y. Acad. Sci*. **414,** 90–6 (1983).

19. Schmidli, P. K., Schurtenberger, P., & Luisi, P. L. Liposome-mediated enzymatic synthesis of phosphatidylcholine as an approach to self-replicating liposomes. *J. Am. Chem. Soc.* **113,** 8127–8130 (1991).

20. Wick, R. & Luisi, P. L. Enzyme-containing liposomes can endogenously produce membrane-constituting lipids. *Chem. Biol*. **3,** 277–285 (1996).

21. Exterkate, M., Caforio, A., Stuart, M. C. A., & Driessen, A. J. M. Growing membranes in vitro by continuous phospholipid biosynthesis from free fatty acids. *ACS Synth. Biol.* **7,** 153–165 (2018).

22. Kuruma, Y., Stano, P., Ueda, T., & Luisi, P. L. A synthetic biology approach to the construction of membrane proteins in semi-synthetic minimal cells. *Biochim. Biophys. Acta* **1788,** 567–574 (2009).

3

23. Scott, A., et al. Cell-free phospholipid biosynthesis by gene-encoded enzymes reconstituted in liposomes. *PLoS One* **11,** e0163058 (2016).

24. Shimizu, Y. et al. Cell-free translation reconstituted with purified components. *Nat. Biotechnol.* **19,** 751–755 (2001).

25. Yao, J., & Rock, C. O. Phosphatidic acid synthesis in bacteria. *Biochim. Biophys. Acta* **1831,** 495–502 (2013).

26. Sparrow, C. P., & Raetz, C. R. H. Purification and properties of the membrane-bound CDP-diglyceride synthetase from Escherichia coli. *J. Biol. Chem.* **260,** 12084–12091 (1985).

27. Hirabayashi, T., Larson, T. J., & Dowhan, W. Membrane-associated phosphatidylglycerophosphate synthetase from Escherichia coli: purification by substrate affinity chromatography on cytidine 5' diphospho-1,2-diacyl-sn-glycerol sepharose. *Biochemistry* **15,** 5205–5211 (1976).

28. Icho, T., & Raetz, C. R. H. Multiple genes for membrane-bound phosphatases in Escherichia coli and their action on phospholipid precursors. *J. Bacteriol.* **153,** 722–730 (1983).

29. Lu, Y.-H. H., Guan, Z., Zhao, J., & Raetz, C. R. Three phosphatidylglycerol-phosphate phosphatases in the inner membrane of Escherichia coli. *J. Biol. Chem.* **286,** 5506–5518 (2011).

30. Li, Q. X., & Dowhan, W. Structural characterization of Escherichia coli phosphatidylserine decarboxylase. *J. Biol. Chem.* **263,** 11516–22 (1988)

31. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–345 (2009).

32. Raetz, C. R. H., & Kennedy, E. P. The association of phosphatidylserine synthetase with ribosomes in extracts of Escherichia coli. *J. Biol. Chem* **247,** 2008–2014 (1972).

33. Louie, K., & Dowhan, W. Investigations on the association of phosphatidylserine synthase with the ribosomal component from Escherichia coli. *J. Biol. Chem.* **255,** 1124-1127 (1980).

34. Saha, S. K., Nishijima, S., Matsuzaki, H., Shibuya, I., & Matsumoto, K. A regulatory mechanism for the balanced synthesis of membrane phospholipid species in Escherichia coli. *Biosci. Biotechnol. Biochem.* **60,** 111–116 (1996).

25. Rilfors, L., et al. Reconstituted phosphatidylserine synthase from Escherichia coli is activated by anionic phospholipids and micelle-forming amphiphiles. *Biochim. Biophys. Acta* **1438,** 281–294 (1999).

36. Scheideler, M. A., & Bell, R. M. Phospholipid dependence of homogeneous, reconstituted sn-glycerol-3-phosphate acetyltransferase of Escherichia coli. *J. Biol. Chem.* **264,** 12455–12461 (1989).

37. Shibuya, I. Metabolic regulation and biological functions of phospholipids in Escherichia coli. *Prog. Lipid Res.* **31,** 245–299 (1992).

38. Blanken, D., Van Nies, P., & Danelon, C. Quantitative imaging of gene-expressing liposomes reveals rare favorable phenotypes. *Phys. Biol.* **16,** 045002 (2019).

39. Yeung, T., et al. Membrane phosphatidylserine regulates surface charge and protein localization. *Science* **319,** 210–213 (2008).

40. Van Nies, P., et al. Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. *Nat. Commun.* **9,** 1583 (2018).

41. Larson, T., & Dowhan, W. Ribosomal-associated phosphatidylserine synthetase from Escherichia coli: purification by substrate-specific elution from phosphocellulose using cytidine 5'-diphospho-1,2-diacyl-sn-glycerol. *Biochemistry* **15,** 5212-5218 (1976).

42. Nourian, Z., Roelofsen, W. & Danelon, C. Triggered gene expression in fed-vesicle microreactors with a multifunctional membrane. *Angew. Chemie - Int. Ed.* **51,** 3114–3118 (2012).

43. Aktas, M., & Narberhaus, F. In vitro characterization of the enzyme properties of the phospholipid N-methyltransferase PmtA from *Agrobacterium tumefaciens*. *J. Bacteriol.* **191,** 2033–2041 (2009).

44. Hiraoka, S., Nukui, K., Uetake, N., Ohta, A. & Shibuya, I. Amplification and substantial purification of cardiolipin synthase of Escherichia coli. *J. Biochem.* **110,** 443–449 (1991).

45. Nourian, Z. & Danelon, C. Linking genotype and phenotype in protein synthesizing liposomes with external supply of resources. *ACS Synth. Biol.* **2,** 186–193 (2013).

46. Saito, H. et al. Time-resolved tracking of a minimum gene expression system reconstituted in giant liposomes. *ChemBioChem* **10,** 1640–1643 (2009).

47. Doerr, A., et al. Modelling cell-free RNA and protein synthesis with minimal systems. *Phys. Biol.* **16,** 25001 (2019).

48. Shin, J. & Noireaux, V. An E. coli cell-free expression toolbox: application to synthetic gene circuits and artificial cells. *ACS Synth. Biol.* **1**, 29–41 (2011).

49. Maddalena, L.L., Niederholtmeyer, H., Turtola, M., Swank, Z.N., Belogurov, G.A. & Maerkl, S.J. GreA and GreB enhance expression of Escherichia coli RNA polymerase promoters in a reconstituted transcription-translation system. *ACS Synth Biol.* **5**, 929–935 (2016).

50. Murtas, G. Internal lipid synthesis and vesicle growth as a step toward self-reproduction of the minimal cell. *Syst. Synth. Biol.* **4,** 85–93 (2010).

51. Bhattacharya, A., Brea, R. J., Niederholtmeyer, H., & Devaraj, N. K. A minimal biochemical route towards de novo formation of synthetic phospholipid membranes. *Nat. Commun.* **10,** 300 (2019).

52. Vogele, K. et al. Towards synthetic cells using peptide-based reaction compartments. *Nat. Commun.* **9,** 3862 (2018).

53. Mercier, R., Kawai, Y., & Errington, J. Excess membrane synthesis drives a primitive mode of cell proliferation. *Cell* **152,** 997–1007 (2013).

54. Sakuma, Y., & Imai, M. Model system of self-reproducing vesicles. *Phys. Rev. Lett.* **107,** 1–5 (2011).

55. Fujii, S., Matsuura, T., Sunami, T., Kazuta, Y. & Yomo, T. In vitro evolution of α-hemolysin using a liposome display. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 16796–16801 (2013).

56. Godino, E., et al. De novo synthesized Min proteins drive oscillatory liposome deformation and regulate FtsA-FtsZ cytoskeletal patterns. *Nat. Commun.* **10,** 4969 (2019).

57. Furusato, T. *et al. De novo* synthesis of basal bacterial cell division proteins FtsZ, FtsA, and ZipA inside giant vesicles. *ACS Synth. Biol.* **7**, 953–961 (2018).

58. Schindelin, J., et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9,** 676–682 (2012).

3

SUPPLEMENTARY METHODS

Cloning of *yfp-spinach* constructs for gene regulation experiments

The *eYFP-LL-spinach* construct was synthesized by Eurogentech (Belgium) and supplied in a pUC57 vector backbone[1]. A lac operator site was introduced via PCR using primers 715 and 716, and *E. coli* TOP10 cells were transformed with the obtained PCR fragment. The T7 promoter in pUC57-T7p-LacO-meYFP-LL-spinach-T7t was substituted with the SP6 promoter using primers 719 and 720 (**Supplementary Table 3.2**) and the generated pUC57-SP6p-LacO-meYFP-LL-spinach-T7t was used for transformation of TOP10 *E. coli* cells. The identity of the plasmids was confirmed with restriction digestion and Sanger sequencing.

Promoter orthogonality assay

Reaction mixtures consisted of a custom made PURE*frex*2.0 devoid of the T7 RNA polymerase (Cosmo Bio, USA). Either 8.1 µg µL^{-1} of T7 RNA polymerase or 4 U µL^{-1} SP6 RNA polymerase (Promega, USA) were added. The T7 RNA polymerase was overexpressed from a gene cloned into pBAD33 and isolated following the protocol of He et al.[2]. 7 nM of plasmid DNA (either pUC57-T7p-LacO-meYFP-LL-spinach-T7t or pUC57-SP6p-LacO-meYFP-LL-spinach-T7t was added last to trigger transcription-translation. Reaction samples of 10 µL were loaded into Greiner 384-well glass bottom plates sealed with a transparent foil and measured in a Tecan Infinite 200 Pro plate reader (Tecan, Switzerland) with the temperature set to 37 °C. The excitation wavelength was 506 nm (9 nm bandwidth), the emission wavelength was 540 nm (20 nm bandwidth), and each measurement consisted of 25 flashes of 20 µs. Fluorescence of YFP was measured for 12 h with an interval time of 5 min. For each kinetic measurement, the data point of the highest fluorescence intensity was selected for further analysis.

SUPPLEMENTARY NOTES

Supplementary Note 3.1: Rational for gene orientation in pGEMM7

In a previous version of the plasmid (pGEMM6) both *pgsA* and *pgpA* genes were under control of an SP6 promoter but orientated in the same direction as the other genes on the plasmid (**Supplementary Fig. 3.4a**). We investigated whether PG production was conditional to the presence of SP6 RNAP. Plasmid pGEMM6 was expressed in PURE system containing LUVs and lipid synthesis was assessed by mass spectrometry. DOPG was produced also in the absence of SP6 RNAP (**Supplementary Fig. 3.4b**), showing the lack of transcriptional orthogonality. We suspected that read-through transcription might be the cause of unintended production of *pgsA* and *pgpA* transcripts, eventually leading to PG synthesis. To remedy this problem, we designed pGEMM7, where the two genes under SP6 promoter control were flipped in the opposite reading direction (**Supplementary Fig. 3.1**). As expected, expression of pGEMM7 resulted in no PgsA protein and no PG phospholipid production when the SP6 RNAP was omitted (main text **Fig. 3.1b,c**).

Supplementary Note 3.2: Absence of PgpA specific signal in MS data

We were not able to detect a suitable peptide for PgpA. Therefore, PgsA was used as a reporter of transcriptional activation of the PgsA-PgpA branch upon SP6 RNAP addition (main text **Fig. 1a,b**). As an alternative to PgpA, PgpC (Uniprot: P0AD42) can catalyze the same enzymatic reaction as PgpA. Four peptides of PgpC were detected with our protocol for trypsin digestion and LC-MS/MS as shown in **Supplementary Fig. 3.7**. They are listed in **Supplementary Table 3.5**.

Supplementary Note 3.3: Analysis of purified LactC2-mCherry by PAGE

The presence of two to three prominent bands of the LactC2-mCherry fusion protein observed in **Supplementary Fig. 3.12b** does not reflect a low purity. Instead, they correspond to different denaturation forms of LactC2-mCherry that migrate differently on gel. This observation has already been reported in the case of another mCherry fusion protein[3]. In their study, Mestrom et al. found that a stabilized form of the fusion protein was resistant to at least 2% SDS and ran at a lower apparent molecular weight by PAGE. This more native form of the fusion protein would correspond to one of the lower bands of LactC2-mCherry marked with a circle symbol in **Supplementary Fig. 3.12b**.

Further purification of LactC2-mCherry by gel filtration (Sephacryl S200 16/60 HiPrep mounted on an ÄKTA Pure system, both from GE Healthcare) failed to eliminate the extra bands visible by Coomassie staining (**Supplementary Fig. 3.13**, upper panel), indicating that they all correspond to the same protein.

Additionally, we analyzed these LactC2-mCherry samples using anion exchange chromatography (MonoQ 4.6/100 PE column on an ÄKTA Pure system, GE Healthcare). Isolated samples were treated with 2 M urea and 95 °C for different periods of time and were loaded on a 4-12% Bis-Tris SDS gel ran in MES buffer (**Supplementary Fig. 3.14**). The two prominent bands corresponding to fully and partially denatured states of the mCherry fusion protein can be observed. The middle one that is visible in the fluorescence channel disappeared after heat treatment, further indicating that it corresponds to a more folded, native form.

Concluding, the prominent bands observed in the lane of the purified LactC2-mCherry all correspond to the same protein and the level of purity is sufficient for utilization as a reporter probe for PS lipids (**Supplementary Figs. 3.17** and 3.**18**).

Supplementary Note 3.4: Kinetic analysis of LactC2-eGFP association to PS-producing liposomes

Phenomenological fitting was applied to extract kinetic parameters from the time traces reported in main text **Fig. 3.5g** (**Supplementary Fig. 3.19**). The following sigmoid equation was used:

$$y = k' + k \ \frac{t^n}{t^n + K^n} \qquad \text{(Eq. 1)}$$

Where t is the time in minutes, and y is the LactC2-eGFP fluorescence signal at the rim of a liposome at a given time point. k, k', K, and n are fitting parameters. This equation has previously been used to fit fluorescence measurements of cell-free gene expression of YFP in liposomes[4], and mass spectrometry measurements of cell-free Min protein expression[5]. The plateau time, i.e. the time until LactC2-eGFP fluorescence stops increasing, was defined as:

$$T_{plateau} = \frac{2K}{n} + K. \qquad\qquad \text{(Eq. 2)}$$

The rate, which is the steepness at time $t = K$, was defined as:

$$\text{rate} = \frac{kn}{4K}. \qquad\qquad \text{(Eq. 3)}$$

3

This apparent rate captures the rate of transcription and translation, the enzyme kinetics leading to the synthesis of PS, the membrane incorporation of PS, as well as the binding kinetics of LactC2-eGFP to PS, as illustrated in **Supplementary Fig. 3.19a**.

SUPPLEMENTARY TABLES

Supplementary Table 3.1: List of plasmids used in this study

Plasmid	Size (kb)	Encoded genes
pGEMM7	10.5	*plsB, plsC, cdsA, pssA, psd, pgsA, pgpA*
pGEMM6	10.5	*plsB, plsC, cdsA, pssA, psd, pgsA, pgpA*
pUC57-T7p-LacO-meYFP-LL-spinach-T7t	3.8	*meYFP*
pUC57-SP6p-LacO-meYFP-LL-spinach-T7t	3.8	*meYFP*
pUC57-T7p-fadD	4.5	*fadD*

Supplementary Table 3.2: List of primers used in this study

Primer	Sequence, 5' to 3'	Linker site	Target
25 ChD	GATGCTGTAGGCATAGGCTTGG		
91 ChD	AAAAAACCCCTCAAGACCCGTTTAGAGG		
288 ChD	CGATGCGTCCGGC		
310 ChD	GGATCTCGACGCTCTCCCTTATG		
397 ChD	CCTCTAGAAATAATTTTGTTTAACTTTAAGAAGG		
471 ChD	CAAAGCCCGAAAGGAAGCTGA		
507 ChD	ATGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGG		pUC57
535 ChD	TAGCATAACCCCTTGGGGCCTCTAAAC		pUC57
628 ChD	TGGGCCCGTAACAAAATCCTCCCAATAAGCTAATACGACTCACTATAGG	1 top	*plsB*
629 ChD	ACCAGGGCTGTTCAACCGACGCTCACGGGGCAAAAAACCCCTCAAGACC	2 bottom	*plsB*
630 ChD	CCCCGTGAGCGTCGGTTGAACAGCCCTGGTTAATACGACTCACTATAGG	2 top	*plsC*
631 ChD	AGCGATATATTCGGGCTTCTGGTCGGGCCGCAAAAAACCCCTCAAGACC	3 bottom	*plsC*
632 ChD	CGGCCCGACCAGAAGCCCGAATATATCGCTTAATACGACTCACTATAGG	3 top	*cdsA*
633 ChD	AGGTAACGCACCCCGGCCCAAGAGCCGTAACAAAAAACCCCTCAAGACC	4 bottom	*cdsA*
634 ChD	TTACGGCTCTTGGGCCGGGGTGCGTTACCTTAATACGACTCACTATAGG	4 top	*pssA*
635 ChD	AGGAATTAACGGACGGCCTCGATTTCTGCACAAAAAACCCCTCAAGACC	5 bottom	*pssA*
636 ChD	TGCAGAAATCGAGGCCGTCCGTTAATTCCTTAATACGACTCACTATAGG	5 top	*psd*
715 ChD	GGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTTAACTTT		
716 ChD	GAATTGTTATCCGCTCACAATTCCGGTCTCCCTATAGTGAGTCG		
719 ChD	ATTTAGGTGACACTATAGAAGAACCGGAATTGTGAGCGGA		
720 ChD	TCTTCTATAGTGTCACCTAAATATTTCGCATCTAGATGCATTCGC		
723 ChD	TCGATATTGGGGACTTCTCAAATCTCGTCACAAAAAACCCCTCAAGACC	SHR18 bottom	*psd*
848 ChD	TATACATATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGTGAGCAAGGGCGAGG		
849 ChD	AACTCAGCTTCCTTTCGGGCTTTGCTAACAGCCCAGCAGCTCC		

850 ChD	GATGATGGCTGCTGCCCATATGTATATCTCCTTCTTAAAGTTAAAC		
851 ChD	CCTGCAGGCCTTAGGCTCGAGCGGCCGCTGAGGACTAGTATTTAGGTGACACTATAGA	13 top	*pgpA*
852 ChD	TGACGAGATTTGAGAAGTCCCCAATATCGACAAAAAACCCCTCAAGACC	SHR18 top	*pgsA*
829 ChD	TGACGAGATTTGAGAAGTCCCCAATATCGAATTTAGGTGACACTATAGAAGAACCCCTCTAG AAATAATTTTGTTTAACTTTAAGAAGG	13 bottom	pUC19
830 ChD	GCTTATTGGGAGGATTTTGTTACGGGCCCAGAATCAGGGGATAACGCAGG	1 bottom	pUC19
818 ChD	GTGTCAGTCAATCACGGGCGGGTCCACTACCAAAAAACCCCTCAAGACCCGTTTAG	30 bottom	*pgpA*
819 ChD	GTAGTGGACCCGCCCGTGATTGACTGACACATTTAGGTGACACTATAGAAGAATCCCCTCTA GAAATAATTTTGTTTAAC	30 top	*pgsA*
873 ChD	GTTTAACTTTAAGAAGGAGATATACATATGAAGAAGGTTTGGCTTAACCGTTATCC		*fadD*
874 ChD	CCCGTTTAGAGGCCCCAAGGGGTTATGCTAGTCAGGCTTTATTGTCCACTTTGCCG		*fadD*

Supplementary Table 3.3: Pipetting scheme of Gibson assembly to obtain pGEMM7

Part	Size (kb)	Ratio length to backbone	Mass DNA added into reaction (ng)
pUC19	2	1	100
pgpA	0.8	0.4	40
pgsA	0.8	0.4	40
plsB	2.6	1.3	130
plsC	0.9	0.45	45
cdsA	1.2	0.6	60
pssA	1.4	0.7	70
psd	1	0.5	50

SUPPLEMENTARY TABLES

Supplementary Table 3.4: Full list of the tryptic peptides detected by LC-MS/MS in this study. The first column shows the Uniprot accession number as well as the gene name. The peptide sequence is shown in column 2 followed by the standard type used (internal for absolute quantification, global for normalization). The start and end positions of the peptide in the wild-type gene are indicated. In the last two columns, the predicted and measured retention time values are reported.

Protein (Uniprot No.)	Peptide	Standard type	Begin #AA	End #AA	Predicted retention time (min)	Average measured retention time (min)
PlsB (P0A7A7)	LLNLPLSILVK		10	20	17.87	20.31
PlsB (P0A7A7)	FSPSVSLR		166	173	7.77	8.09
PlsB (P0A7A7)	LAAVGPR		202	208	2.58	2.9
PlsB (P0A7A7)	DTIGDIIILPR	Internal standard	589	599	16.94	16.81
PlsC (P26647)	LAPLFGLK		44	51	13.66	13.48
PlsC (P26647)	GLLPFK		154	159	10.14	10.49
CdsA (P0ABG1)	DSGHLIPGHGGILDR		248	262	10.4	8.1
PssA (P23830)	GILNALYEAK		65	74	11.93	12.49
PssA (P23830)	DLQSIADYPVK	Internal standard	419	429	9.57	11
Psd (P0A8K1)	LSLQYILPK		6	14	13.64	14.89
Psd (P0A8K1)	LVIDLFVK		35	42	16.2	15.41
PgsA (P0ABF8)	EIIISALR		101	108	12.73	11.36
PgsA (P0ABF8)	SSVAVSWIGK	Internal standard	118	127	11.03	10.34
EF-Tu (P0CE47)	TTLTAAITTVLAK	Global standard	25	37	13.94	16.65
EF-Tu (P0CE47)	GITINTSHVEYDTPTR		59	74	9.11	9.56

Supplementary Table 3.5: Transitions of the MS/MS measurements for the proteomic analysis

Protein (Uniprot No.)	Peptide	Precursor ion *m/z*	MS1 Res	Product ion *m/z*	MS2 Res	Dwell time (ms)	Fragmentor (V)	Collision energy (V)	Ion
PlsB (P0A7A7)	LLNLPLSILVK	611.9103	Unit	769.5182	Unit	20	130	20	y7
PlsB (P0A7A7)	LLNLPLSILVK	611.9103	Unit	227.1754	Unit	20	130	20	b2
PlsB (P0A7A7)	LLNLPLSILVK	611.9103	Unit	341.2183	Unit	20	130	20	b3
PlsB (P0A7A7)	LLNLPLSILVK	611.9103	Unit	454.3024	Unit	20	130	20	b4
PlsB (P0A7A7)	FSPSVSLR	446.748	Unit	658.3883	Unit	20	130	14.8	y6
PlsB (P0A7A7)	FSPSVSLR	446.748	Unit	375.235	Unit	20	130	14.8	y3
PlsB (P0A7A7)	FSPSVSLR	446.748	Unit	288.203	Unit	20	130	14.8	y2
PlsB (P0A7A7)	FSPSVSLR	446.748	Unit	235.1077	Unit	20	130	14.8	b2
PlsB (P0A7A7)	LAAVGPR	342.2136	Unit	499.2987	Unit	20	130	11.6	y5
PlsB (P0A7A7)	LAAVGPR	342.2136	Unit	428.2616	Unit	20	130	11.6	y4
PlsB (P0A7A7)	LAAVGPR	342.2136	Unit	329.1932	Unit	20	130	11.6	y3
PlsB (P0A7A7)	LAAVGPR	342.2136	Unit	185.1285	Unit	20	130	11.6	b2
PlsB (P0A7A7)	DTIGDIIILPR	613.3612	Unit	896.5564	Unit	20	130	20	y8
PlsB (P0A7A7)	DTIGDIIILPR	613.3612	Unit	611.4239	Unit	20	130	20	y5
PlsB (P0A7A7)	DTIGDIIILPR	613.3612	Unit	498.3398	Unit	20	130	20	y4
PlsB (P0A7A7)	DTIGDIIILPR	613.3612	Unit	385.2558	Unit	20	130	20	y3
PlsB (P0A7A7)	DTIGDIIILPR	613.3612	Unit	272.1717	Unit	20	130	20	y2
PlsB (P0A7A7)	DTIGDIIILPR.heavy	618.3653	Unit	906.5646	Unit	20	130	20	y8

PlsB (P0A7A7)	DTIGDIIILPR.heavy	618.3653	Unit	621.4322	Unit	20	130	20	y5
PlsB (P0A7A7)	DTIGDIIILPR.heavy	618.3653	Unit	508.3481	Unit	20	130	20	y4
PlsB (P0A7A7)	DTIGDIIILPR.heavy	618.3653	Unit	395.264	Unit	20	130	20	y3
PlsB (P0A7A7)	DTIGDIIILPR.heavy	618.3653	Unit	282.18	Unit	20	130	20	y2
PlsC (P26647)	LAPLFGLK	429.776	Unit	674.4236	Unit	20	130	14.3	y6
PlsC (P26647)	LAPLFGLK	429.776	Unit	464.2867	Unit	20	130	14.3	y4
PlsC (P26647)	LAPLFGLK	429.776	Unit	337.7154	Unit	20	130	14.3	y6
PlsC (P26647)	LAPLFGLK	429.776	Unit	185.1285	Unit	20	130	14.3	b2
PlsC (P26647)	GLLPFK	337.7154	Unit	504.318	Unit	20	130	11.5	y4
PlsC (P26647)	GLLPFK	337.7154	Unit	391.234	Unit	20	130	11.5	y3
PlsC (P26647)	GLLPFK	337.7154	Unit	171.1128	Unit	20	130	11.5	b2
PlsC (P26647)	GLLPFK	337.7154	Unit	284.1969	Unit	20	130	11.5	b3
CdsA (P0ABG1)	DSGHLIPGHGGILDR	515.2707	Unit	630.3570	Unit	20	130	8.15	Y6
CdsA (P0ABG1)	DSGHLIPGHGGILDR	515.2707	Unit	397.1466	Unit	20	130	8.15	b4
CdsA (P0ABG1)	DSGHLIPGHGGILDR	515.2707	Unit	510.2306	Unit	20	130	8.15	b5
CdsA (P0ABG1)	DSGHLIPGHGGILDR	515.2707	Unit	623.3147	Unit	20	130	8.15	b6
PssA (P23830)	GILNALYEAK	546.3084	Unit	921.504	Unit	20	130	17.9	y8
PssA (P23830)	GILNALYEAK	546.3084	Unit	808.4199	Unit	20	130	17.9	y7
PssA (P23830)	GILNALYEAK	546.3084	Unit	623.3399	Unit	20	130	17.9	y5
PssA (P23830)	DLQSIADYPVK	624.8272	Unit	892.4775	Unit	20	130	20.4	y8

PssA (P23830)	DLQSIADYPVK	624.8272	Unit	692.3614	Unit	20	130	20.4	y6
PssA (P23830)	DLQSIADYPVK	624.8272	Unit	506.2973	Unit	20	130	20.4	y4
PssA (P23830)	DLQSIADYPVK	624.8272	Unit	357.1769	Unit	20	130	20.4	b3
PssA (P23830)	DLQSIADYPVK.heavy	628.8343	Unit	900.4917	Unit	20	130	20.4	y8
PssA (P23830)	DLQSIADYPVK.heavy	628.8343	Unit	700.3756	Unit	20	130	20.4	y6
PssA (P23830)	DLQSIADYPVK.heavy	628.8343	Unit	514.3115	Unit	20	130	20.4	y4
PssA (P23830)	DLQSIADYPVK.heavy	628.8343	Unit	357.1769	Unit	20	130	20.4	b3
Psd (P0A8K1)	LSLQYILPK	537.8315	Unit	633.397	Unit	20	130	17.7	y5
Psd (P0A8K1)	LSLQYILPK	537.8315	Unit	357.2496	Unit	20	130	17.7	y3
Psd (P0A8K1)	LSLQYILPK	537.8315	Unit	244.1656	Unit	20	130	17.7	y2
Psd (P0A8K1)	LSLQYILPK	537.8315	Unit	831.4975	Unit	20	130	17.7	b7
Psd (P0A8K1)	LVIDLFVK	473.8022	Unit	833.5131	Unit	20	130	15.7	y7
Psd (P0A8K1)	LVIDLFVK	473.8022	Unit	734.4447	Unit	20	130	15.7	y6
Psd (P0A8K1)	LVIDLFVK	473.8022	Unit	506.3337	Unit	20	130	15.7	y4
Psd (P0A8K1)	LVIDLFVK	473.8022	Unit	213.1598	Unit	20	130	15.7	b2
PgsA (P0ABF8)	EIIISALR	457.7871	Unit	559.3562	Unit	20	130	15.2	y5
PgsA (P0ABF8)	EIIISALR	457.7871	Unit	446.2722	Unit	20	130	15.2	y4
PgsA (P0ABF8)	EIIISALR	457.7871	Unit	359.2401	Unit	20	130	15.2	y3
PgsA (P0ABF8)	EIIISALR	457.7871	Unit	356.218	Unit	20	130	15.2	b3
PgsA (P0ABF8)	SSVAVSWIGK	517.2875	Unit	760.4352	Unit	20	130	17	y7

3

PgsA (P0ABF8)	SSVAVSWIGK	517.2875	Unit	689.3981	Unit	20	130	17	y6
PgsA (P0ABF8)	SSVAVSWIGK	517.2875	Unit	590.3297	Unit	20	130	17	y5
PgsA (P0ABF8)	SSVAVSWIGK	517.2875	Unit	345.1769	Unit	20	130	17	b4
PgsA (P0ABF8)	SSVAVSWIGK.heavy	521.2946	Unit	768.4494	Unit	20	130	17	y7
PgsA (P0ABF8)	SSVAVSWIGK.heavy	521.2946	Unit	697.4123	Unit	20	130	17	y6
PgsA (P0ABF8)	SSVAVSWIGK.heavy	521.2946	Unit	598.3439	Unit	20	130	17	y5
PgsA (P0ABF8)	SSVAVSWIGK.heavy	521.2946	Unit	345.1769	Unit	20	130	17	b4
EF-TU (P0CE47)	TTLTAAITTVLAK	652.3952	Unit	887.556	Unit	20	130	21.2	y9
EF-TU (P0CE47)	TTLTAAITTVLAK	652.3952	Unit	816.5189	Unit	20	130	21.2	y8
EF-TU (P0CE47)	TTLTAAITTVLAK	652.3952	Unit	745.4818	Unit	20	130	21.2	y7
EF-TU (P0CE47)	TTLTAAITTVLAK	652.3952	Unit	632.3978	Unit	20	130	21.2	y6
EF-TU (P0CE47)	GITINTSHVEYDTPTR	601.9672	Unit	752.3573	Unit	20	130	16.9	y6
EF-TU (P0CE47)	GITINTSHVEYDTPTR	601.9672	Unit	474.2671	Unit	20	130	16.9	y4
EF-TU (P0CE47)	GITINTSHVEYDTPTR	601.9672	Unit	710.3286	Unit	20	130	16.9	y12
EF-TU (P0CE47)	GITINTSHVEYDTPTR	601.9672	Unit	187.1133	Unit	20	130	16.9	y3
PgpC (P0AD42)	LQTLQADFVR	595.8300	Unit	949.5101	Unit	20	130	10.12	y8
PgpC (P0AD42)	LQTLQADFVR	595.8300	Unit	735.3784	Unit	20	130	10.12	y6
PgpC (P0AD42)	LQTLQADFVR	595.8300	Unit	242.1499	Unit	20	130	10.12	b2
PgpC (P0AD42)	LQTLQADFVR	595.8300	Unit	343.1975	Unit	20	130	10.12	b3
PgpC (P0AD42)	DNVTAFPLVQER	694.8620	Unit	959.5308	Unit	20	130	12.04	y8

3

PgpC (P0AD42)	DNVTAFPLVQER	694.8620	Unit	888.4937	Unit	20	130	12.04	y7
PgpC (P0AD42)	DNVTAFPLVQER	694.8620	Unit	741.4253	Unit	20	130	12.04	y6
PgpC (P0AD42)	DNVTAFPLVQER	694.8620	Unit	230.0771	Unit	20	130	12.04	b2
PgpC (P0AD42)	VNLIASQIQR	571.3380	Unit	631.3521	Unit	20	130	8.29	y5
PgpC (P0AD42)	VNLIASQIQR	571.3380	Unit	303.1775	Unit	20	130	8.29	y2
PgpC (P0AD42)	VNLIASQIQR	571.3380	Unit	214.1186	Unit	20	130	8.29	b2
PgpC (P0AD42)	VNLIASQIQR	571.3380	Unit	440.2867	Unit	20	130	8.29	b4
PgpC (P0AD42)	GYGGWVLTMR	570.2869	Unit	919.4818	Unit	20	130	13	y8
PgpC (P0AD42)	GYGGWVLTMR	570.2869	Unit	520.2911	Unit	20	130	13	y4
PgpC (P0AD42)	GYGGWVLTMR	570.2869	Unit	407.2071	Unit	20	130	13	y3
PgpC (P0AD42)	GYGGWVLTMR	570.2869	Unit	221.0920	Unit	20	130	13	b2

Supplementary Table 3.6: Mass spectrometry detection settings for phospholipid detection. Measurements in positive and negative modes were performed separately to circumvent polarity switching. Asterisk (*) indicates species incorporating ^{13}C-G3P, resulting in a 3 Da mass shift with respect to the regular species (6 Da for PG). Fragmentor voltage and collision energy were adapted from[6]. Novel species were assigned values of similar species. PGP products were not successfully detected. However, the transitions were scanned in all measurements and are therefore included here for completeness. Cell accelerator voltage was 4 V in all conditions.

3

Compound class	Compound name	Precursor ion m/z	Product ion m/z	Fragmentor (V)	Collision energy (eV)	Polarity
Matrix	DOPE	742.5	281.2	140	25	Negative
Matrix	DOPG	773.5	281.2	190	37	Negative
Matrix	Cardiolipin	727.5	281.2	140	29	Negative
Standards	DPPE	690.5	255.1	230	37	Negative
Standards	DPPG	721.5	255.1	220	45	Negative
Standards	POPE	716.5	281.2 255.1	140 230	25 37	Negative
Standards	POPG	750.5	281.2 255.1	190 220	37 45	Negative Negative
Standards	DOPS	786.5	281.2	130	40	Negative
16:0-product	16:0-LPA*	412.2	155.9	150	13	Negative
16:0-product	DPPA*	650.5	255.1	210	33	Negative
16:0-product	CDP-DPG*	955.5	712.4	135	55	Negative
16:0-product	DPPS*	737.5	255.1	180	41	Negative
16:0-product	DPPE*	693.5	255.1	230	37	Negative
16:0-product	DPPGP*	808	255.1	220	45	Negative
16:0-product	DPPG*	727.5	255.1	220	45	Negative
18:1-product	18:1-LPA*	438.3	155.9	160	17	Negative
18:1-product	DOPA*	702.5	281.2	190	37	Negative
18:1-product	CDP-DOG*	1007.5	764.5	135	55	Negative
18:1-product	DOPS*	789.5	281.2	130	40	Negative
18:1-product	DOPE*	745.5	281.2	140	25	Negative
18:1-product	DOPGP*	860	281.2	190	50	Negative
18:1-product	DOPG*	779.5	281.2	190	37	Negative
Mixed product	POPA*	676.5	281.2	190	37	Negative

			255.1	210	33	Negative
Mixed product	CDP-POG*	981.5	738.5	135	55	Negative
Mixed product	POPS*	763.5	281.2 255.1	130 180	40 41	Negative Negative
Mixed product	POPE*	719.5	281.2 255.1	140 230	25 37	Negative Negative
Mixed product	POPGP*	834	281.2 255.1	190 220	50 45	Negative Negative
Mixed product	POPG*	753.5	281.2 255.1	190 220	37 45	Negative Negative

Supplementary Table 3.7: Overview of synthesized lipid species. Symbols between square brackets correspond to residues as defined in the scheme below. Each species is a combination of two acyl chain residues (in numbers) and one head group residue (letter). Lipid species in bold have been quantitatively measured. Species in italic have not been unambiguously detected; however, their presence can be deduced from the successful measurements of subsequent reaction products.

Enzymes Precursors	PlsB	PlsC	CdsA	PssA	Psd	PgsA	PgpA
Palmitoyl-CoA	16:0-LPA <2,1,a>	DPPA <2,2,a>	*CDP-DPG* <2,2,b>	DPPS <2,2,c>	**DPPE** <2,2,d>	*DPPGP* <2,2,e>	**DPPG** <2,2,f>
Oleoyl-CoA	18:1-LPA <3,1,a>	DOPA <3,3,a>	*CDP-DOG* <3,3,b>	DOPS <3,3,c>	**DOPE** <3,3,d>	*DOPGP* <3,3,e>	**DOPG** <3,3,f>
Palmitoyl-CoA + Oleoyl-CoA	16:0-LPA <2,1,a> 18:1-LPA <3,1,a>	DPPA <2,2,a> DOPA <3,3,a> POPA <2,3,a>, <3,2,a>	*CDP-DPG* <2,2,b> *CDP-DOG* <3,3,b> *CDP-POG* <2,3,b>, <3,2,b>	DPPS <2,2,c> DOPS <3,3,c> POPS <2,3,c>, <3,2,c>	**DPPE** <2,2,d> **DOPE** <3,3,d> **POPE** <2,3,d>, <3,2,d>	*DPPGP* <2,2,e> *DOPGP* <3,3,e> *POPGP* <2,3,e>, <3,2,e>	**DPPG** <2,2,f> **DOPG** <3,3,f> **POPG** <2,3,f>, <3,2,f>
Palmitoyl-CoA + NBD-palmitoyl-CoA	16:0-LPA <2,1,a> *NBD-LPA* <4,1,a>	DPPA <2,2,a> NBD-PPA <2,4,a>, <4,2,a> *NBD-NDB-PA* <4,4,a>	*CDP-DPG* <2,2,b> *CDP-NBD-PG* <2,4,b>, <4,2,b> *CDP-NBD-NBD-G* <4,4,b>	DPPS <2,2,c> NBD-PPS <2,4,c>, <4,2,c> *NBD-NDB-PS* <4,4,c>	**DPPE** <2,2,d> NBD-PPE <2,4,d>, <4,2,d> *NBD-NDB-PA* <4,4,d>	-	-

Backbone

R₃

R₁, R₂
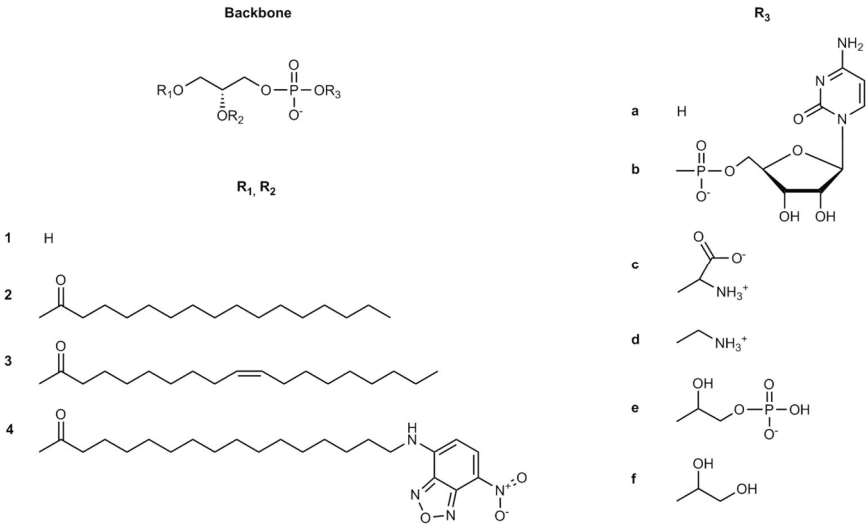
SUPPLEMENTARY FIGURES
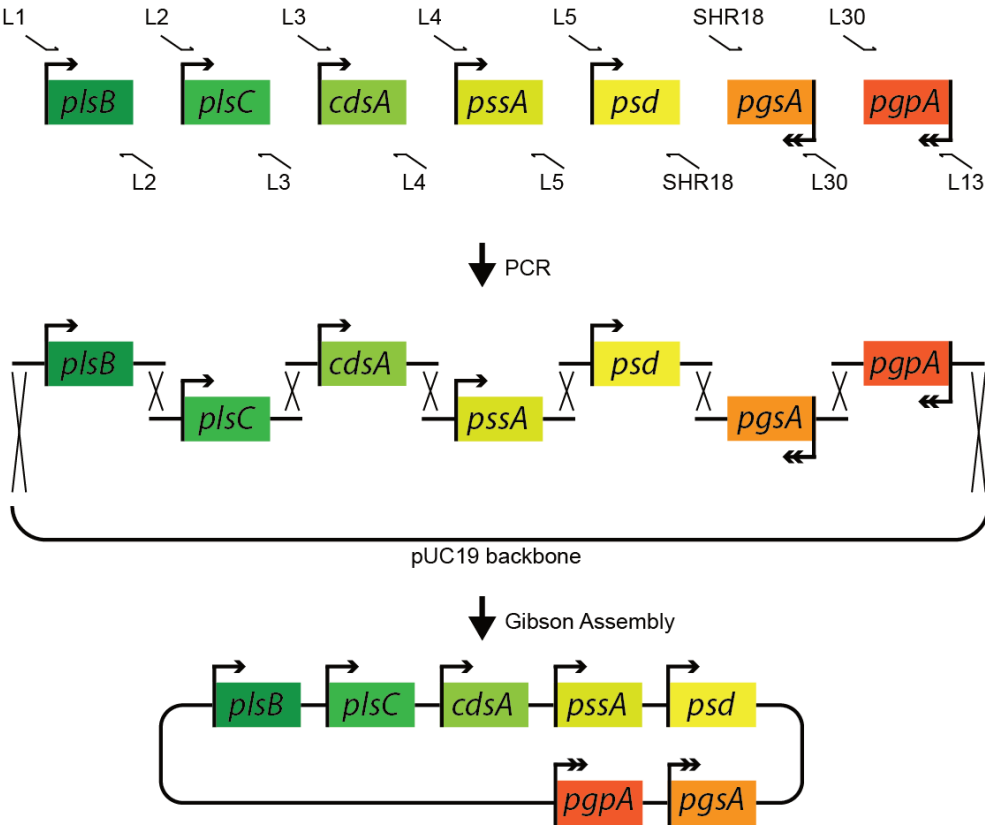
Supplementary Figure 3.1: Scheme for the construction of the pGEMM7 plasmid. Homologous linker sites are added to the transcriptional cassettes via PCR in the first step. After purification of the PCR fragments the entire plasmid is assembled by Gibson assembly in a single reaction.
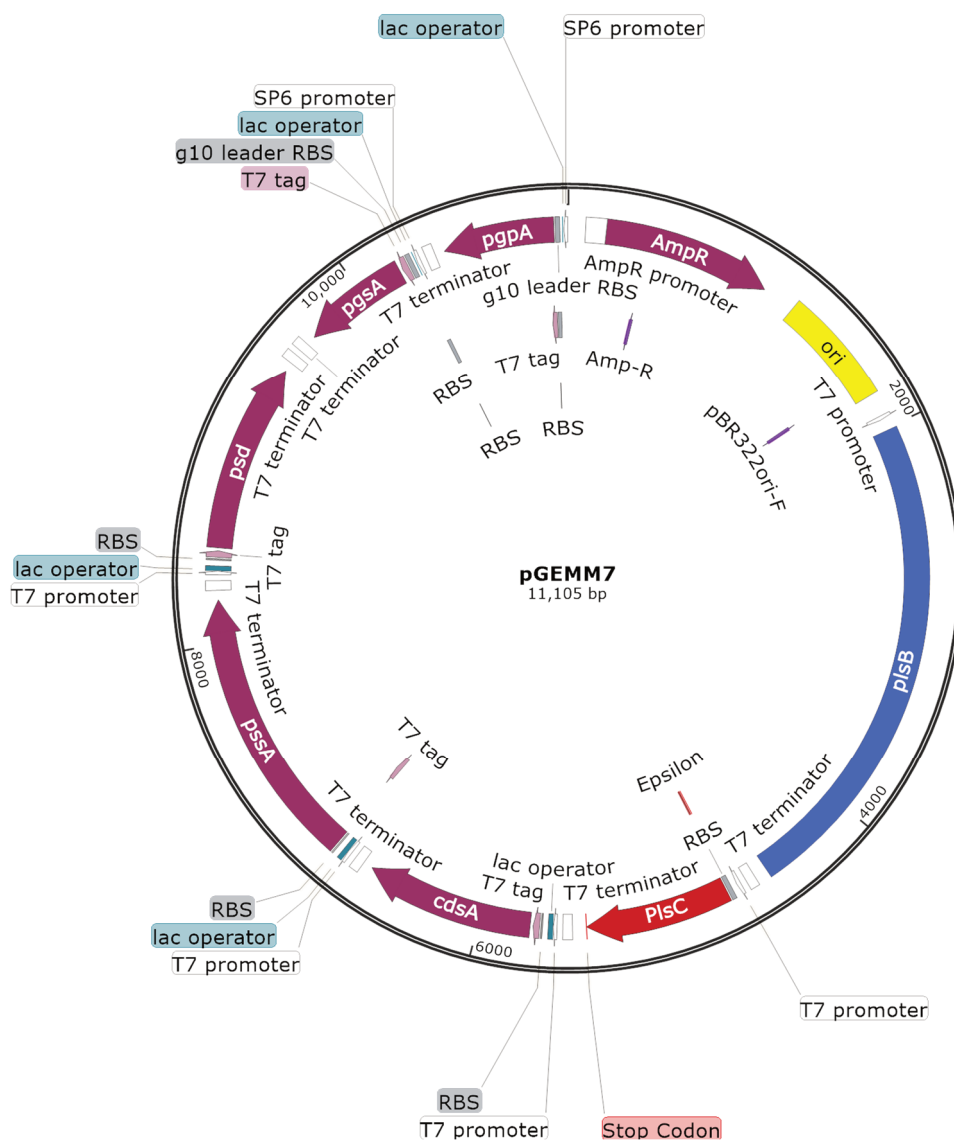
Supplementary Figure 3.2: Sequence map of the pGEMM7 construct. The plasmid map was generated with Snapgene.

a

b
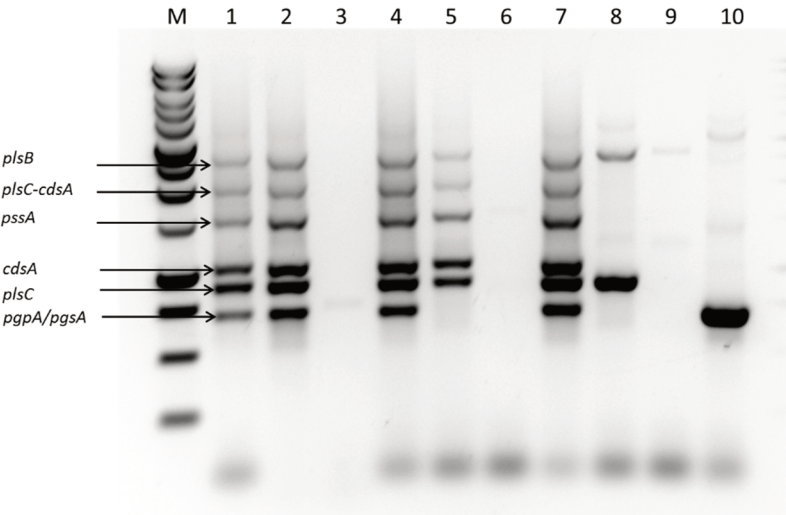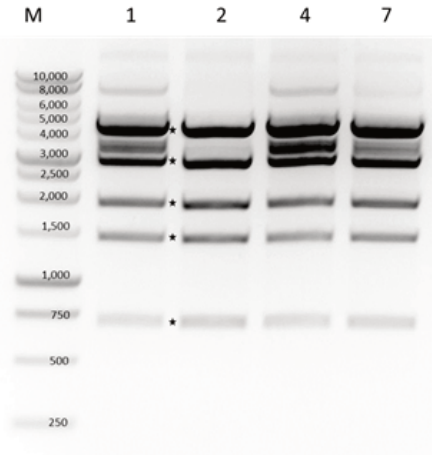
Supplementary Figure 3.3: Verification of pGEMM7 construct. (a) Agarose gel of a colony PCR using primers 91 and 397. Colonies 1, 2, 4 and 7 were chosen to be further analyzed due to the observed band pattern. **(b)** Agarose gel stained with SYBR safe DNA dye of a restriction digestion of pGEMM7. 1 µg of total DNA was loaded. 5 µL of Benchtop 1kb ladder (Promega) were loaded in the first lane. The stars indicate the location of the expected bands. The numbers above the lanes indicate the colony number that was tested.

a

pGEMM6
minigenome

b
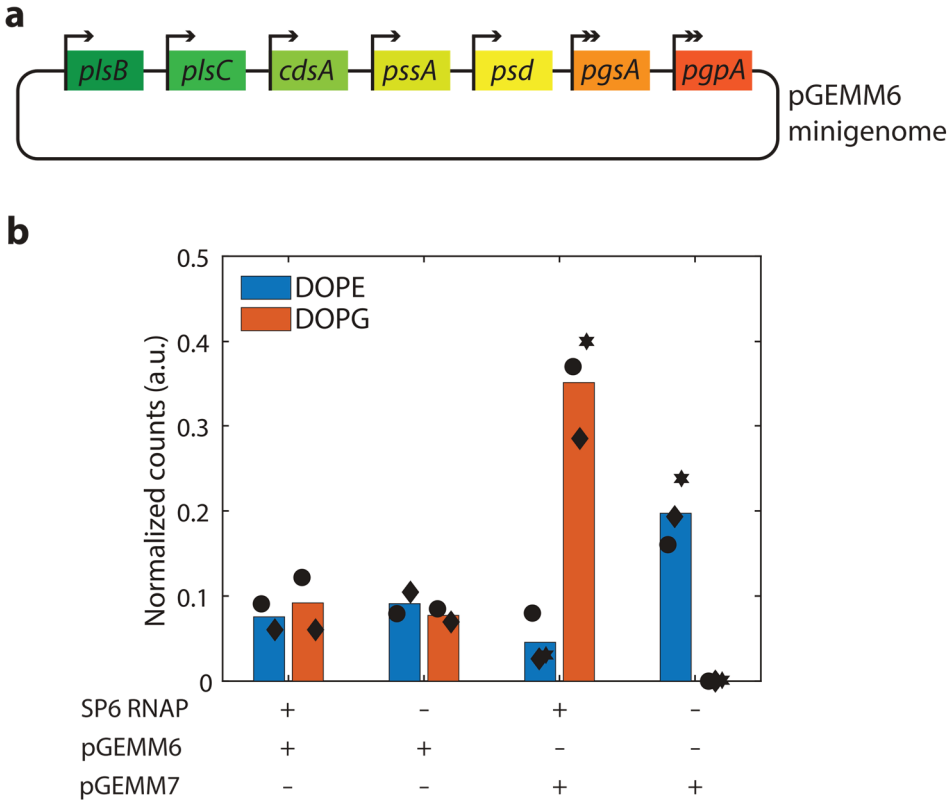
Supplementary Figure 3.4: Read-through causes non-specific PG synthesis with pGEMM6. (a) Schematic representation of the pGEMM6 plasmid where all five genes under T7 promoter control (single arrow) were assembled in the same orientation as the two genes *pgsA* and *pgpA* under SP6 promoter control (depicted by the double arrow). **(b)** Synthesis of phospholipids from pGEMM6 (data with pGEMM7 are appended for comparison), in the presence of LUVs and all necessary precursors, with and without SP6 RNAP. The integrated peak intensity of ^{13}C-labelled lipids has been normalized to the signal of ^{12}C-DOPE to account for differences in MS sensitivity. In the absence of SP6 RNAP, *pgsA* and *pgpA* should not be expressed, and therefore, no DOPG should be synthesized. However, an appreciable yield of DOPG was observed without SP6 RNAP in the case of pGEMM6. We hypothesized that this was caused by transcription termination read-through by the T7 RNAP, transcribing downstream PG synthesis genes even if those are under control of an orthogonal promoter. This finding prompted us to design the new construct pGEMM7, where the two sets of PE and PG pathway genes were cloned in opposite directions to ensure full orthogonality. Bars are average values from two independent repeats (three with pGEMM7), each represented by a different symbol. Source Data are available for panel **b**.
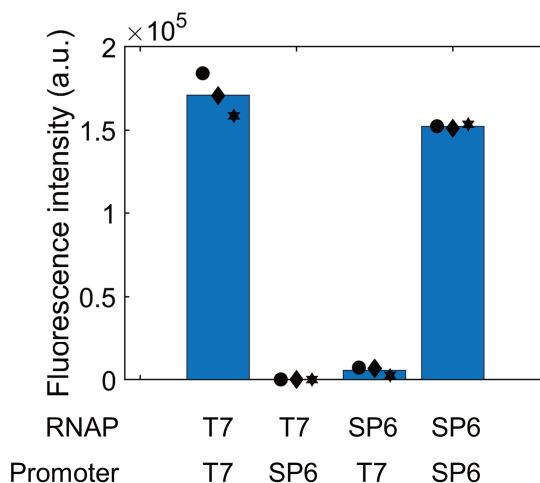
Supplementary Figure 3.5: Orthogonality diagram of T7 vs Sp6 promoters. The pUC57-T7p-LacO-meYFP-LL-spinach-T7t and pUC57-SP6p-LacO-meYFP-LL-spinach-T7t were expressed in PURE*frex*2.0 (ΔT7 RNAP) supplemented with either the T7 or SP6 RNAP. High end-point fluorescence signal from YFP was only observed with the canonical promoter/RNAP pairs, demonstrating strong orthogonality. Bars represent mean values from three repeats and the different symbols are individual data points. Source Data are available.
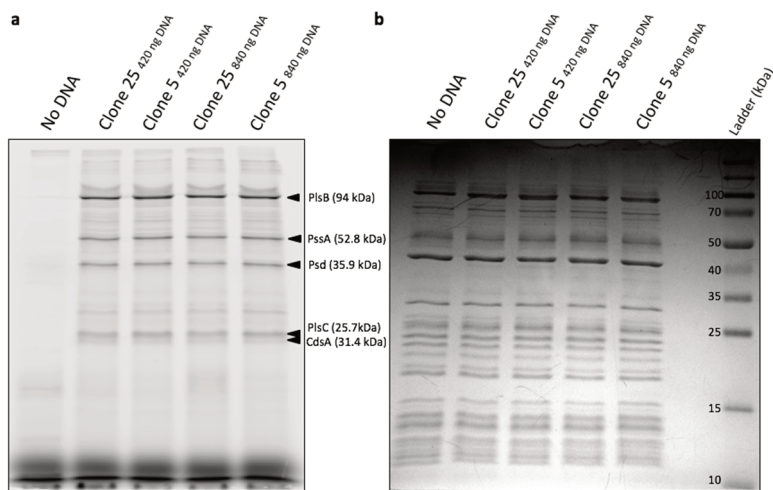

Supplementary Figure 3.6: Co-translational labelling of proteins with GreenLys. Early versions of pGEMM6 and pGEMM7 were screened for their ability to produce lipid-synthesizing enzymes. Clones 5 and 25 encode five genes that were successfully expressed in PUREfrex2.0 supplemented with 1 µL of GreenLys reagent (Promega). The translation products were analyzed by SDS-PAGE and Typhoon imaging. a, Fluorescence image with appended arrows pointing to the relevant protein bands. The protein names and theoretical molecular masses are indicated. Note that CdsA migrates significantly lower than predicted in the gel, as previously reported[6]. b, CBB stain image of the gel shown in a. Background proteins from the PURE system are visible, whereas the yield of synthesized proteins is too low to be directly observed.
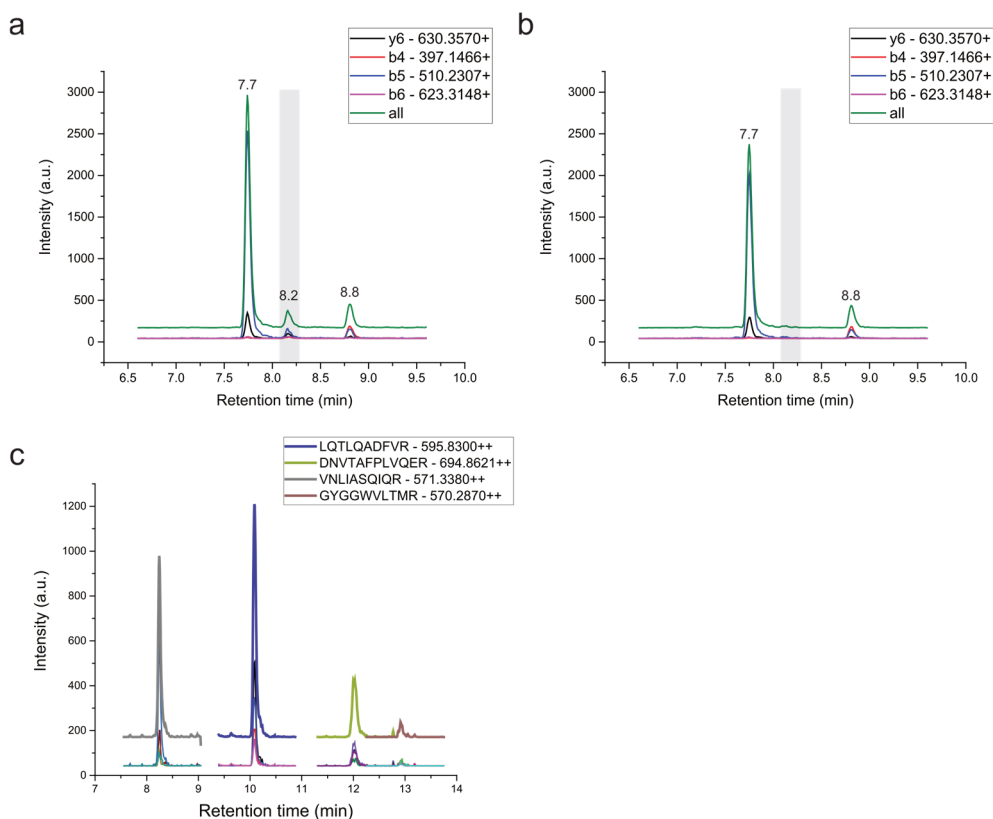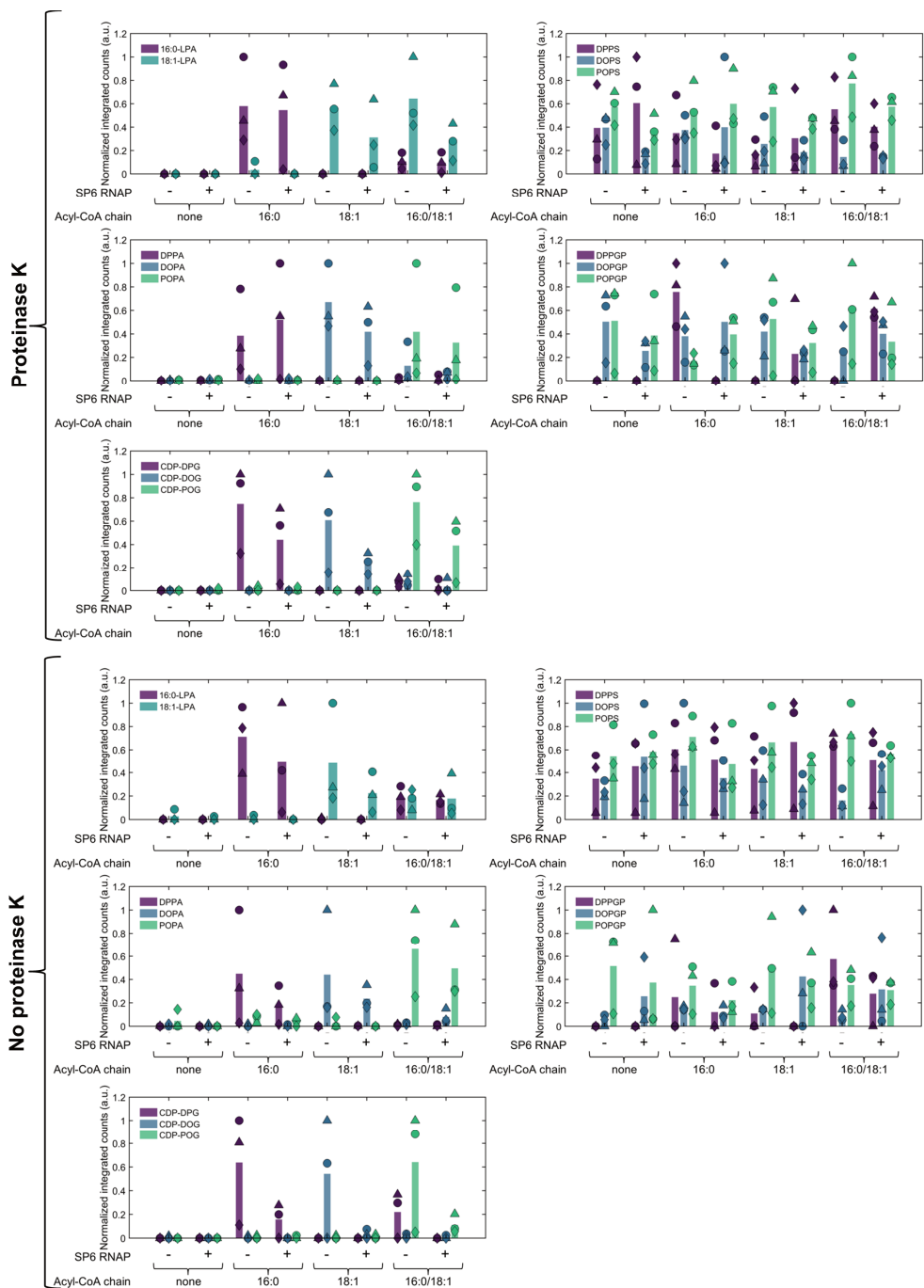
Supplementary Figure 3.7: LC-MS detection of CdsA by a single specific peptide DSGHLIPGHGGILDR and PgpC by four peptides. Peaks corresponding to tryptic peptides of the other enzymes were clearly identified and the chromatograms are not shown here. Their specifications can be found in **Supplementary Table 3.4**. **a**, Plotted intensities of four fragments of the CdsA peptide together with the combined intensities of all fragments. The overlaid gray area indicates the expected position at 8.2 min of the specific peak in the chromatogram. This specific peak is absent in a control cell-free protein expression reaction incubated without DNA (**b**). The visible peaks at 7.7 min and 8.8 min retention times can be attributed to background signals from the PURE*frex*2.0 proteins themselves or other contaminants. **c**, Chromatograms of all transitions and the combined measured ion current of each peptide of PgpC that were detectable after trypsin digestion. Source Data are available for panels **a** and **b**.
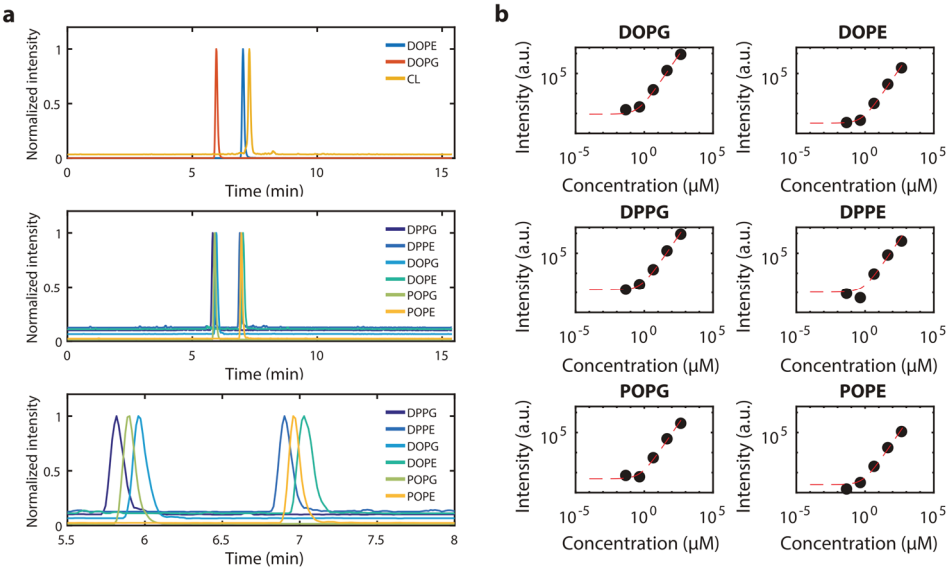
Supplementary Figure 3.8: LC-MS detection of phospholipids. a, Representative LC chromatograms of the studied phospholipids. Experimental conditions are as described in **Fig. 3.1c**. Two AU of SP6 RNAP, along with a 1:1 mix of oleoyl- and palmitoyl-CoA were supplemented to LUVs. Peak intensity was normalized to aid visualization. (top) Chromatograms of the lipids present in the liposome matrix. (middle, bottom) Chromatograms of synthesized phospholipids. The lower panel is a zoom-in graph, showing the typical elution pattern. Species with longer (oleoyl) chains elute later than species with shorter (palmitoyl) chains. **b,** Representative calibration curves for quantification of phospholipid concentrations. Five different known concentrations of the indicated phospholipids were injected, from low to high concentration, and plotted versus the resulting total integrated counts (TIC) of the corresponding peak labelled as 'Intensity' on the graphs. A linear fit was computed to obtain a calibration curve for each lipid (overlaid red dashed line). TIC values from cell-free synthesized lipids were then converted into absolute concentrations. Source Data are available for panel **b**.
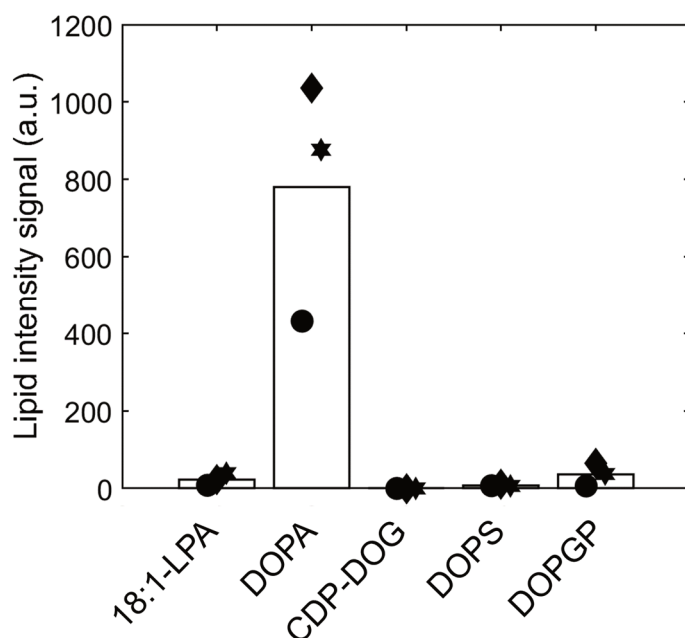
Supplementary Figure 3.9: Accumulation of lipid synthesis intermediates. Signal corresponding to all five lipid synthesis intermediate species. The measurements correspond to the rightmost data points shown in **Fig. 3.1c**, that is expression of pGEMM7 in the presence of LUVs and 4 AU of SP6 RNAP. Only DOPA accumulates in significant amounts. Other species have a total integrated peak intensity below 100 a.u.; therefore, the signals cannot be distinguished from noise. Source Data are available.

3

Supplementary Figure 3.10: LC-MS detection of synthesized phospholipids and accumulated intermediate products in giant liposome experiments. Phospholipid intermediates LPA, PA, and CDP-DAG, but not PGP and PS, were detected. Data are plotted as total integrated counts normalized to the highest value per species, for all intermediate lipid species. Dataset was obtained from the same samples as shown in main text **Fig. 3.3d,e**. Bars

indicate mean values. Symbols indicate values from individual experiments and correspond to data shown main text **Fig. 3.3d,e**. Precursors palmitoyl-CoA (16:0 acyl-CoA) and oleoyl-CoA (18:1 acyl-CoA) were used as indicated. Addition of proteinase K in the external medium confines gene expression to the interior of liposomes (upper panels). For LPAs, PAs and CDP-DAGs, signals were observed as expected, i.e. di-palmitoyl (DP) or di-oleoyl (DO) products when starting with 16:0 or 18:1 acyl-CoA, respectively, and both DP, DO and mixed chain products (PO) when starting with a mixture of the two precursors. For PS and PGP species, no clear pattern was detected and only very low peak intensities (< 100 integrated counts) were measured. Because successful detection of PS was confirmed in other experiments (Supplementary Fig. 16), the results indicate that PS species do not accumulate under these conditions. In contrast, PGP signal was never observed, suggesting that its detection requires optimized protocols. Accumulation of LPAs, PAs, and CDP-DAGs was lower in the presence of SP6 RNAP when the PG synthesis pathway is active. Presumably, having two downstream pathways instead of one might result in a higher flux through the upstream (pre-branchpoint) part of the pathway. Source Data are available.
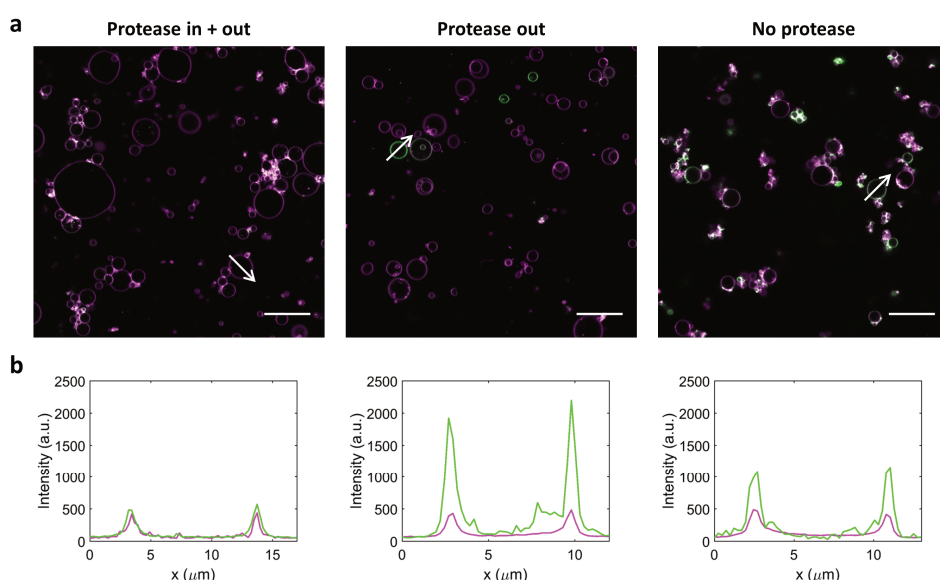

Supplementary Figure 3.11: Confocal fluorescence images of giant liposomes with synthesized NBD-labelled phospholipids. Experimental conditions are as indicated in main text **Fig. 4d**. Briefly, liposomes encapsulating pGEMM7 and PUREfrex2.0 were prepared and incubated in the presence of NBD-palmitoyl-CoA and palmitoyl-CoA (1:9 mol. fraction). **(a)** Representative images of different liposome samples are shown. Proteinase K was co-encapsulated inside liposomes, preventing gene expression and thus lipid synthesis (left). Proteinase K was added to preformed liposomes, preventing lipid synthesis from occurring outside the liposome lumina (middle). When no proteinase K was supplemented, gene expression and lipid synthesis can take place both inside and outside liposomes (right). NBD fluorescence is shown in green and the Texas Red-dyed liposome membrane in magenta. Scale bars represent 20 μm. **(b)** Line profiles of Texas Red (red) and NBD (green) signals along the arrows appended in a. a.u., arbitrary unit.
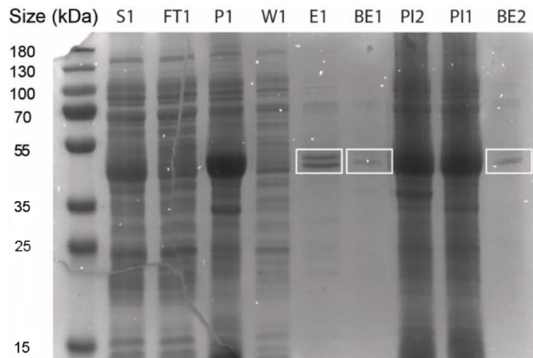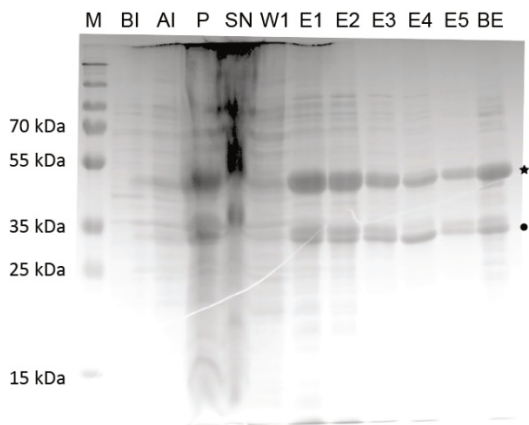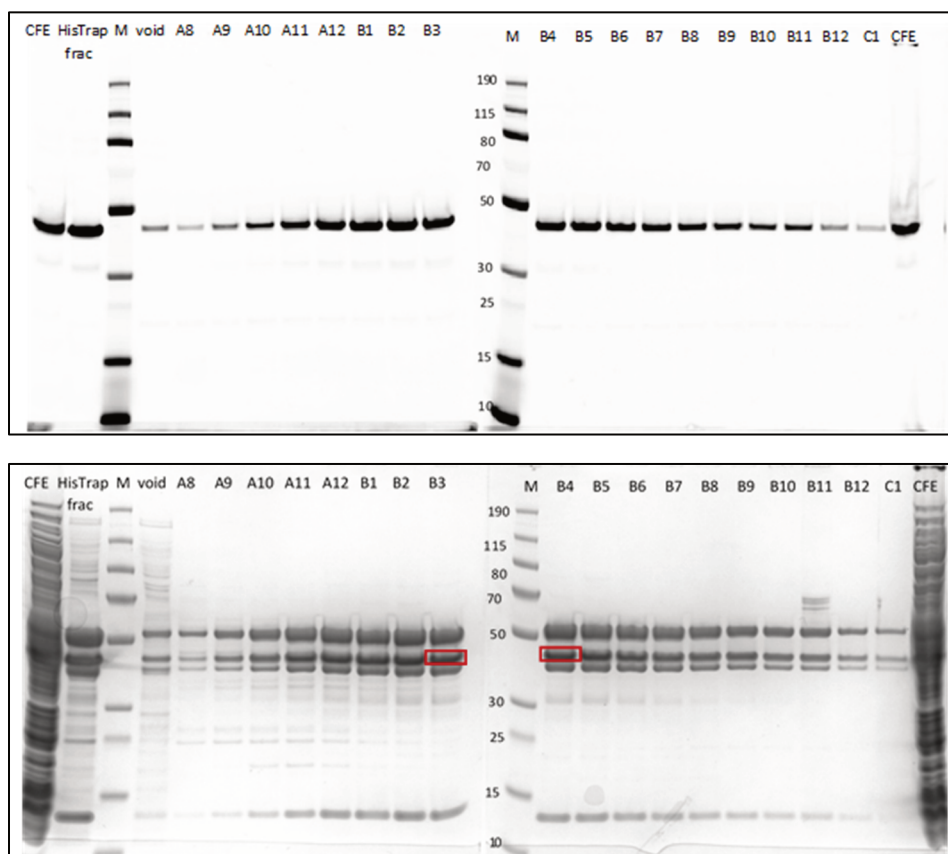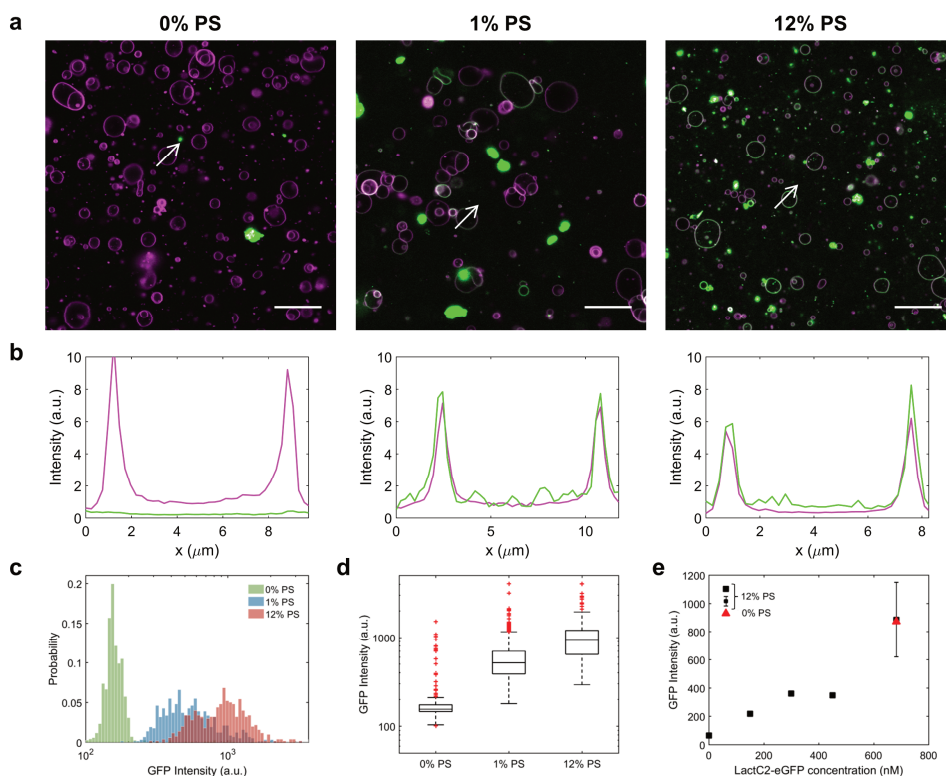
Supplementary Figure 3.12: Purification of LactC2-eGFP and LactC2-mCherry. Images of SDS-PAGE gels to visualize the purification steps of LactC2-eGFP (**a**) and LactC2-mCherry (**b**). **(a)** Recombinant LactC2-eGFP was expressed in *E. coli* and purified using Ni-NTA chromatography column. "1" stands for the protein produced by strain Rosetta ER2566, "2" by strain Rosetta 2; "PI" stands for post-induction with IPTG, "S" for the supernatant and "P" for the pellet, both after cell lysis by sonication, "FT" for the flow-through of the membrane binding step, "W" for the flow-through of the column washing step, "E" for eluate and "BE" for the final purified protein after buffer exchange. The white boxes indicate the expected region for the protein band. Double bands occurring in lane E1 can hint at incomplete reduction or denaturation of the protein. **(b)** In lane "M" 5 μL of Page ruler plus ladder (Promega) were loaded. In the following lanes a sample of the expression strain before induction (BI), after overnight induction at RT (AI), the pellet after lysis and centrifugation (P), the supernatant after lysis (SN), the first washing step (W1), elutions 1 to 5 (E1-E5) and a sample after buffer exchange (BE). The expected product size of the fusion protein is 47 kDa corresponding to the bands marked by a star. The dot marker indicates a set of two bands roughly corresponding to a monomeric mCherry protein including the T7 tag (30 kDa).

Supplementary Figure 3.13: Analysis of gel filtration fractions of purified LactC2-mCherry on 4-12% Bis-Tris SDS-PAGE gels. Samples from different purification steps were loaded on the gels. The upper panel shows a fluorescence image of the gels to localize the folded, active mCherry. The bottom panel shows the same gels stained with Coomassie to visualize the total protein content. The red boxes indicate the fluorescent band of the native mCherry fusion protein. The other two prominent bands above and below very likely correspond to differently denatured states of the mCherry, similar to what Mestrom et al. observed [3]. CFE, cell-free extract; HisTrap frac, pooled HisTrap fractions corresponding to the starting material for gel filtration; M, PageRuler[TM] Plus Prestained protein ladder (ThermoFisher); void, void volume of column; A8 to C1, gel filtration fractions.

3
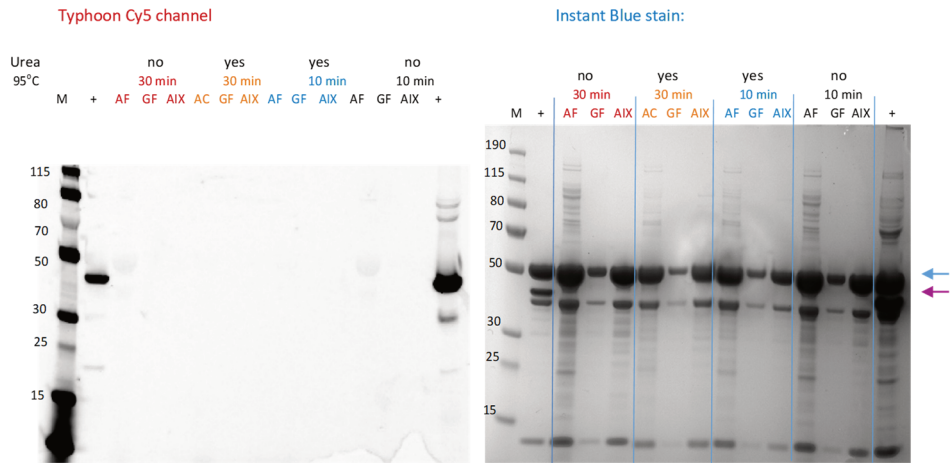
Supplementary Figure 3.14: SDS-PAGE analysis of LactC2-mcherry samples subjected to different denaturing conditions. The left panel shows a fluorescence image, where only the folded, active mCherry can be seen. The right panel shows the same gel stained with Coomassie to visualize the total protein content. AF, affinity chromatography; GL, gel filtration chromatography; AIX, anion exchange chromatography. The time in minutes indicates the duration of the treatment at 95 °C in Laemmli sample buffer supplemented with 10 mM DTT. The 'yes' or 'no' refers to the addition, or not, of 2 M urea in the sample loading buffer. The blue arrow points to the fully denatured LactC2-mCherry. The magenta arrow indicates the native protein, of which the mCherry fluorescence disappears after heat treatment. The 'M' lane is the ladder in kDa. The lane '+' corresponds to gel filtration sample in Bolt 4xLDS sample buffer (ThermoFisher) incubated for 10 minutes at 75 °C.
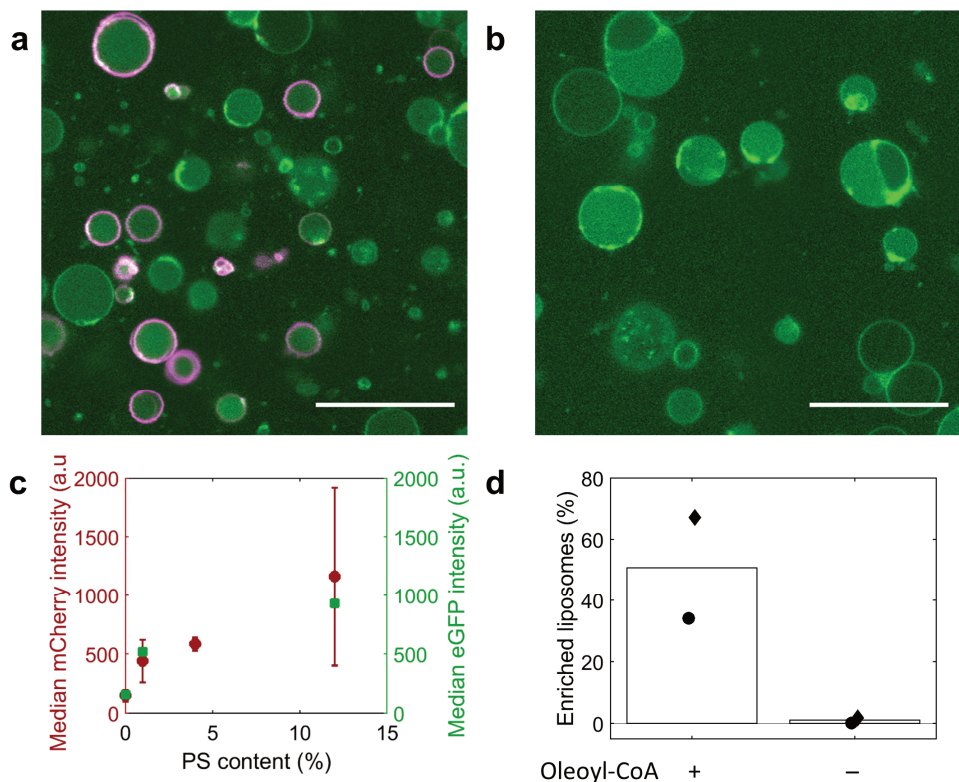
Supplementary Figure 3.15: LactC2-eGFP specifically binds to PS-containing liposomes. (a) Confocal fluorescence images of liposomes (membrane in magenta) containing 0%, 1%, or 12% DOPS in the membrane. The liposomes were incubated with 150 nM of LactC2-eGFP (green signal) before imaging. Green spots are LactC2-eGFP aggregates, which were typically observed in all conditions and do not colocalize with vesicles. Scale bars represent 20 μm. **(b)** Fluorescence intensity line profiles corresponding to the arrows appended in **a**. Texas Red membrane dye signal is in magenta and LactC2-eGFP signal is in green. Clear signal colocalization can be seen in the presence of 1% and 12% PS. **c**, Average LactC2-eGFP rim intensity for 0%, 1% and 12% PS. The distributions of LactC2-eGFP fluorescence at 1% and 12% PS significantly overlap. Number of analyzed liposomes N = 675, 560, 496, respectively. **(d)** Box plot of the data displayed in **c**. N = 675, 560, 496 (from left to right). The box represents all data point between the 25th and 75th percentile, with the line indicating the median. Whiskers extend the most extreme data points not considered outliers. Outliers are plotted individually as red '+' signs. The relationship between LactC2-eGFP fluorescence intensity and PS concentration is nonlinear. **(e)**, Average LactC2-eGFP rim intensity for liposomes with 0% or 12% PS and varying concentrations of LactC2-eGFP. For 12% PS, N = 503, 273, 93 and 403 liposomes (from left to right), all from one experiment. The error bar corresponds to the mean ± SD of two 12% PS samples, with N = 314 and N = 71. The red triangle represents the average GFP intensity of 60 liposomes, from one experiment, for liposomes containing 0% PS. Clear non-specific binding of the LactC2-eGFP probe to the membrane can be observed at high concentration. To ensure PS-binding specificity, a working concentration of LactC2-eGFP of 150 nM was employed in all measurements with expressed pGEMM7. a.u., arbitrary unit. Source Data are available for panels **c-e**.

3
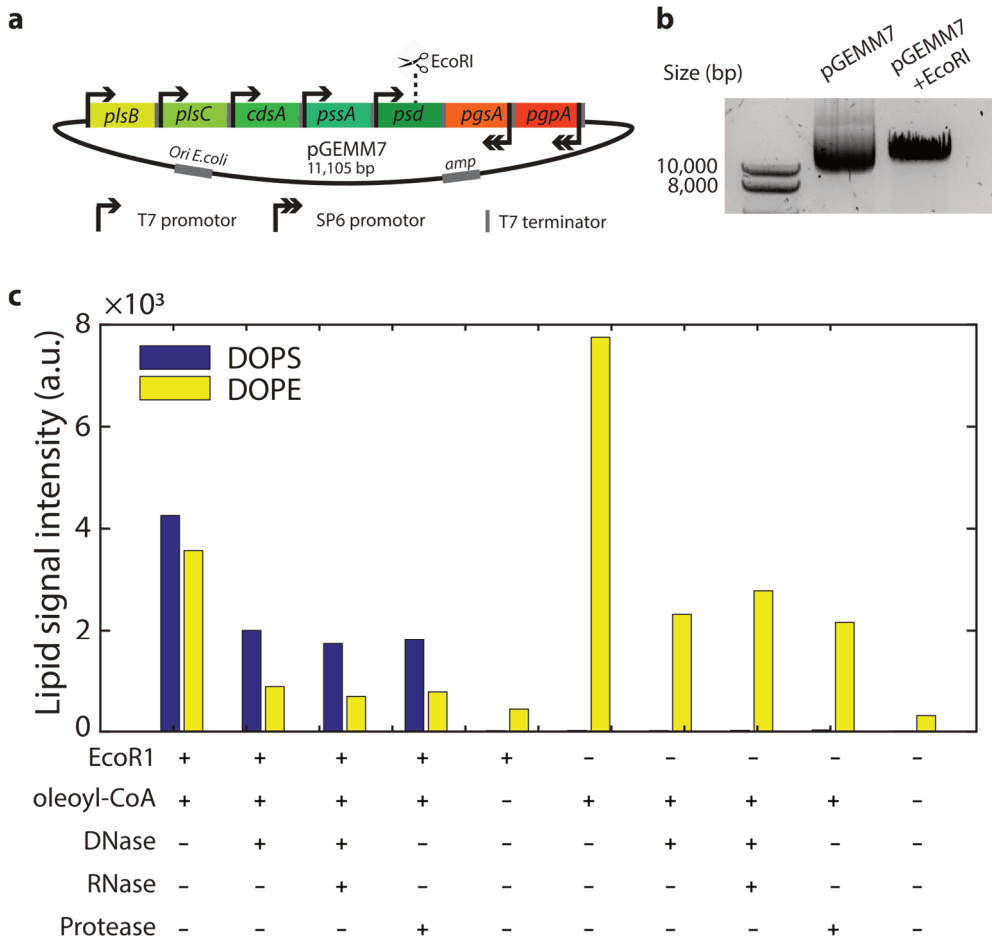
Supplementary Figure 3.16: Expression of EcoRI-restricted pGEMM7 results in accumulation of phosphatidylserine. (a) Schematic representation of the pGEMM7 plasmid map. The EcoRI site in the *psd* gene is depicted. **(b)** 1% agarose gel showing unrestricted pGEMM7 and pGEMM7 cleaved with EcoRI. **c,** LC-MS analysis of synthesized DOPS and DOPE with circular pGEMM7 (−EcoRI) or linearized pGEMM7 (+EcoRI) expressed inside giant liposomes. No SP6 RNAP was added. The total integrated counts for DOPS and DOPE are shown. Accumulation of PS is only observed with the linear DNA template, i.e. when expression of the Psd enzyme which converts PS into PE is inhibited. Some residual PE synthesis is observed, probably due to incomplete enzymatic digestion of pGEMM7. Various methods to confine gene expression to the inside of vesicles have been tested. Using proteinase K or DNase gave equivalent results, corroborating the findings presented in the main text **Fig. 3.3a**. Source Data are available for panel **c**.
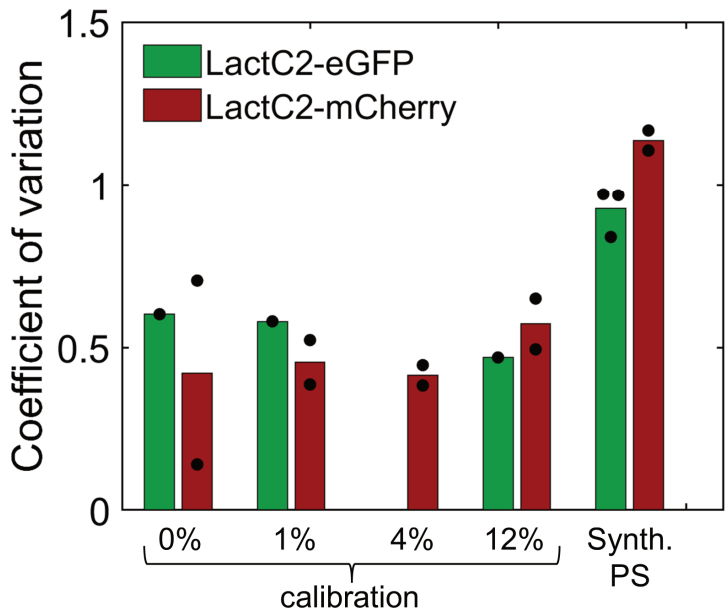
Supplementary Figure 3.17: Liposome-compartmentalized PS synthesis results in a higher coefficient of variation in LactC2-eGFP signal than when PS is directly included in the vesicle membrane. The coefficient of variation was defined as σ/μ, where σ the standard deviation and μ the mean, of the LactC2-eGFP or -mCherry signal. It provides a good measure of the excess liposome-to-liposome heterogeneity of membrane-incorporated PS caused by cell-free protein and lipid synthesis. The coefficient of variation was calculated for control samples with liposomes containing a fixed amount of PS (0, 1, 4 or 12 mol. %, as indicated) and for in-liposome synthesized PS samples. Liposome-confined synthesis of PS results in higher values of the coefficient of variation with both fluorescent probes, providing a quantitative measure of the liposome-to-liposome heterogeneity in the yield of synthesized PS. When an error bar is displayed, data represents the mean ± SD of two different biological samples, except for synthesized PS with LactC2-eGFP, where the mean ± SD of three biological replicates is shown. Calibration experiments with LactC2-eGFP have been performed once. Source Data are available.

Supplementary Figure 3.18: LactC2-mCherry is also a compatible PS-specific binding probe. Confocal fluorescence images of liposomes encapsulating all necessary components for pGEMM7Δ*psd*-directed lipid synthesis, in the presence (**a**) or absence (**b**) of oleoyl-CoA. LactC2-mCherry signal is displayed in magenta. Since mCherry is not spectrally compatible with the Texas Red membrane dye, liposomes were alternatively stained by addition of acridine orange (AO, green). AO is commonly used as a nucleic acid dye, but we have previously observed that the hydrophobic molecule partitions also into liposome membranes[7]. Compared with the eGFP fusion protein, less prominent clusters of LactC2-mCherry were observed. Scale bar represents 20 μm. **(c)** Median LactC2-mCherry intensity values for liposomes containing various amounts of PS (indicated in mol. %). Data are the mean ± SD from three repeats. Data obtained with LactC2-eGFP (**Supplementary Fig. 3.15**) are superimposed for comparison. **(d)** Percentage of PS-enriched liposomes after expression of pGEMM7Δ*psd* in the presence or absence of oleoyl-CoA. Bars represent the mean from two independent repeats. Individual data points are marked with a different symbol. Results are similar as with LactC2-eGFP (main text **Fig. 3.5e**), demonstrating that the nature of the fluorescent protein does not influence the membrane binding ability, despite the fact that LactC2-eGFP is more prone to aggregation. Source Data are available for panels **c** and **d**.
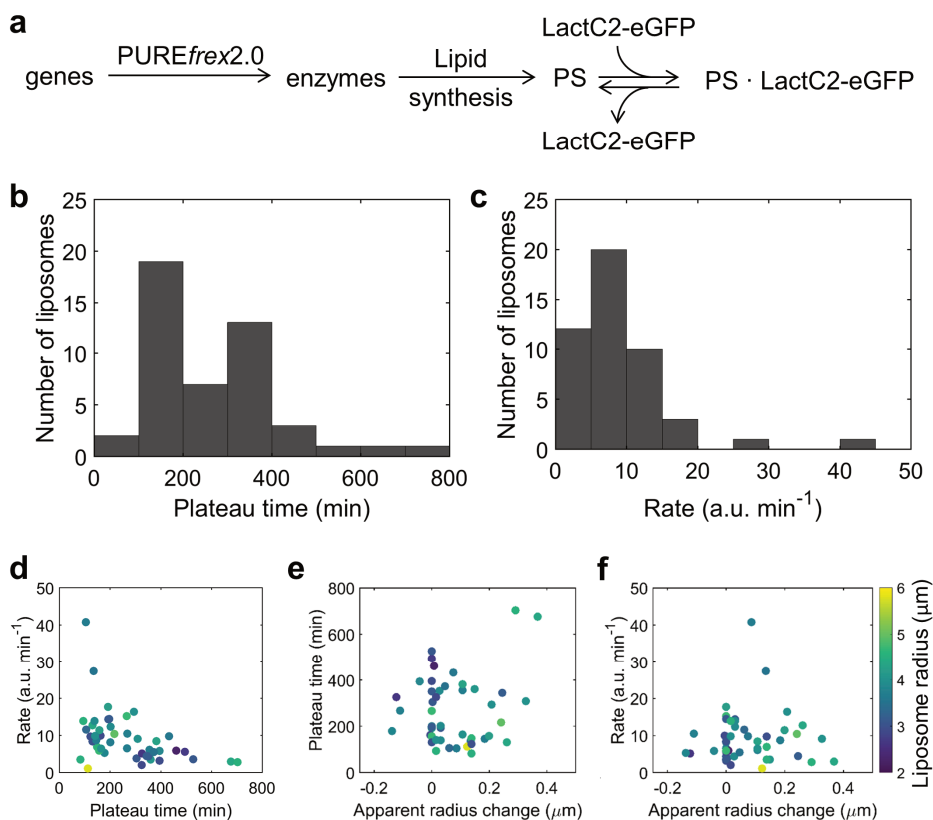
Supplementary Figure 3.19: Kinetic analysis of combined enzyme expression, PS synthesis, and LactC2-eGFP binding. (a) Overview of the main kinetic steps resulting in binding of LactC2-eGFP to synthesized PS, as observed in main text **Fig. 3.5.b,c.** Histograms, for the 47 traces shown in **Fig. 3.5g**, of plateau time (**b**) and apparent rate (**c**), as defined in Eq. 2 and Eq. 3, respectively. **(d)** Scatter plot of the plateau time and apparent rate as shown in **b** and **c**. Color coding corresponds to the (apparent) liposome radius[4]. No clear size dependence can be observed. Plateau time and rate are somewhat negatively correlated (Pearson's correlation coefficient $\rho = -0.44$). **(e, f)** The plateau time (**e**) and rate (**f**) have been plotted against the apparent radius change. Color coding and the analyzed liposomes are the same as in **d**. Source Data are available for panels **b-f**.

SUPPLEMENTARY REFERENCES

1. Van Nies, P., Canton, A. S., Nourian, Z. & Danelon, C. Monitoring mRNA and protein levels in bulk and in model vesicle-based artificial cells. *Methods Enzymol* **550,** 187-214 (2015).

2. He, B., Rong, M., Lyakhov, D., Gartenstein, H., Diaz, G., Castagna, R., et al. Rapid mutagenesis and purification of phage RNA polymerases. *Protein Expression and Purification* **9,** 142-151 (1997).

3. Mestrom, L., Marsden, S.R., Dieters, M., Achterberg, P., Stolk, L., Bento, I., Hanefeld, U., Hagedoorn, P.L. Artificial fusion of mCherry enhances trehalose transferase solubility and stability. *Appl. Environ. Microbiol.* **85,** e03084-18 (2019).

4. Blanken, D., Van Nies, P., & Danelon, C. Quantitative imaging of gene-expressing liposomes reveals rare favorable phenotypes. *Phys. Biol.* **16,** 045002 (2019).

5. Godino, E., et al. De novo synthesized Min proteins drive oscillatory liposome deformation and regulate FtsA-FtsZ cytoskeletal patterns. *Nat. Commun.* **10,** 4969 (2019).

6. Scott, A., et al. Cell-free phospholipid biosynthesis by gene-encoded enzymes reconstituted in liposomes. *PLoS One* **11,** e0163058 (2016).

7. Van Nies, P., et al. Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. *Nat. Commun.* **9,** 1583 (2018).

Chapter 4: A cell-free expressed membrane protein synthesizes lipids and remodels giant vesicles

Synthesis of lipid membrane constituents is a milestone in the development of a self-replicating synthetic cell. Phospholipid N-methyltransferase (Pmt) proteins catalyze the triple methylation of phosphatidylethanolamine to form phosphatidylcholine (PC), a popular membrane phospholipid in liposome research. Furthermore, the membrane remodeling activity of some Pmts might endow Pmt with a dual role for in a synthetic cell. Here, we demonstrate the cell-free expression of Rhodobacter sphaeroides *PmtA. Reconstitution of catalytic activity is demonstrated by liquid chromatography-coupled mass spectrometry detection of synthesized PC, including when PmtA was integrated downstream the Kennedy pathway, together forming a six-enzyme cascade reaction. Moreover, internal expression of PmtA is found to remodel giant unilamellar vesicles on the micrometer scale, as observed by fluorescence microscopy. Membrane remodeling yields septumed liposomes, a process that is accompanied by a decrease in average liposome size. The morphological transformations might be indicative of liposome fission, although the exact nature of the membrane septum remains to be clarified (single or double bilayer), as well as the localization of PmtA. Further investigations into* R. sphaeroides *PmtA are therefore of great interest for research aiming to reconstitute cell growth and division* in vitro.

4.1 INTRODUCTION

The ability to proliferate, that is autonomously generating offspring, is one of the key distinguishing features of every living organism[1]. Therefore, bottom-up reconstitution of cell proliferation has garnered considerable interest in the field of synthetic cell research[2–4]. The synthesis of molecules forming the cell compartment, such as phospholipids[5–7], synthetic phospholipid analogues[8], or amphiphilic peptides[9], is a key milestone towards synthetic cell proliferation. We have recently reconstituted the seven-gene *Escherichia coli* Kennedy pathway for synthesis of the phospholipids phosphatidylethanolamine (PE) and phosphatidylglycerol (PG) inside liposomes[7] (chapter 3). Expanding the repertoire of membrane-forming lipids produced internally would aid in the construction of liposomes with improved stability, and membrane signaling and mechanical properties. An interesting candidate is the non-*E. coli* lipid phosphatidylcholine (PC), which is commonly used for synthetic membrane research. A family of membrane proteins called phospholipid N-methyltransferase proteins (Pmts) catalyze the triple methylation of PE to form PC[10]. Interestingly, these proteins are also shown to have *in vitro* and *in vivo* membrane remodeling activity (**Fig. 4.1**)[11,12], a property that could be exploited for liposome division.
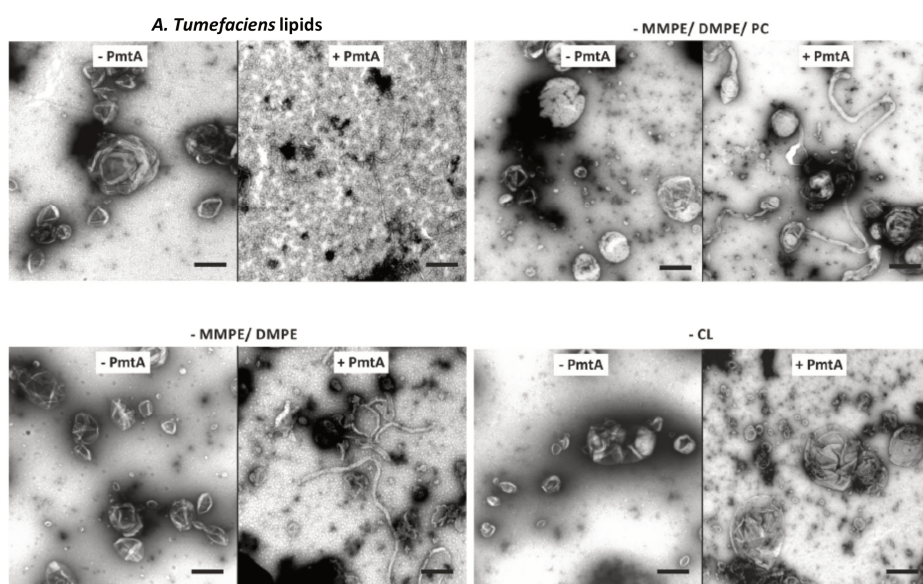

Figure 4.1: *A. tumefaciens* PmtA remodels liposome membranes *in vitro*. Transmission electron microscopy micrographs of 10 µM recombinant *A. tumefaciens* PmtA, incubated with 0.5 mg/mL of liposomes of indicated lipid composition and an approximate diameter of 250 nm. *A. tumefaciens* has a native lipid composition of 40% PE, 15% MMPE, 5% DMPE, 20% PC, 15% PG, and 5% CL. In this native composition, liposomes are vesiculated by PmtA (top left). If MMPE and DMPE are left out (top right, bottom left), the remodeling phenotype is changed to tubulation. Without CL, PmtA causes vesiculation but the number of resulting small vesicles is modest. Scale bars are 100 nm. Figure taken from Danne et al.[12]

In this study, we investigated the lipid-synthesizing and membrane-remodeling properties of *Rhodobacter sphaeroides* PmtA in a synthetic cell context (**Fig. 4.2**). Synthesis of PC by cell-free expressed PmtA was demonstrated by liquid chromatography-coupled mass spectrometry (LC-MS). Moreover, PmtA could work in tandem with the cell-free expressed PE biosynthesis pathway, demonstrating the functional coupling between PmtA and the upstream PE-synthesizing enzymes. Expression of PmtA inside giant unilamellar vesicles led to membrane remodeling on the micrometer scale. Both the formation of septums and a decrease in liposome size was observed as a result of PmtA expression. Cardiolipin and methylated PC synthesis intermediates were found to stimulate septum formation. Size decrease and septum formation might be indicative of liposome fission. This interesting phenomenon will require further study in the nature of the septumed liposomes and the positioning of the PmtA protein.

4.2 RESULTS AND DISCUSSION

4.2.1 Cell-free expressed PmtA can convert PE to PC

We started by cell-free expression of the *pmtA* gene using the PURE*frex*2.0 *in vitro* transcription translation (IVTT) system. Expression was performed in the presence of 400 nm large unilamellar vesicles (LUVs) consisting of DOPC, DOPE, DOPG, and cardiolipin (CL), to serve as a membrane scaffold for the newformed enzyme. Upon catalytic activity of the expressed PmtA, DOPE is methylated using heavy isotope-labelled S-adenosyl-methionine (SAM) as a methyl donor. Production of PC was confirmed by LC-MS. The resulting DOPC can be distinguished from the native DOPC by the 9 Da mass shift due to incorporation of isotopically labelled methyl groups. (**Supplementary Table 1, Fig. 4.3a**).
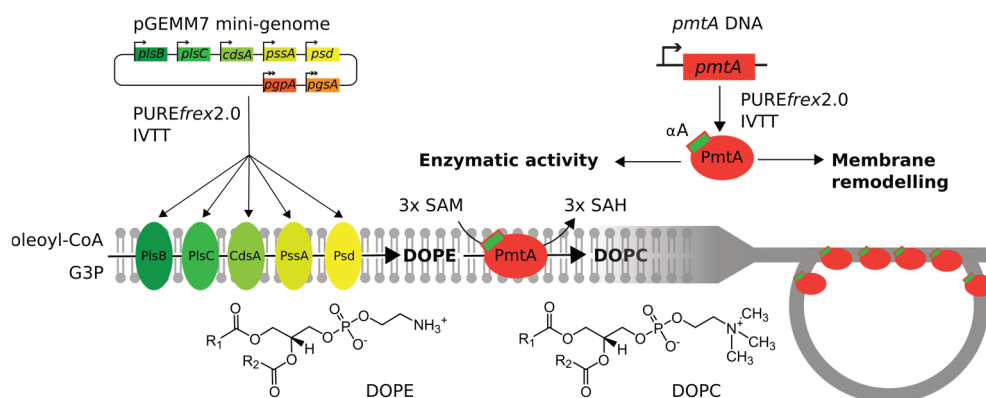
Figure 4.2: Cell-free expressed *R. sphaeroides* PmtA has PC synthesis and membrane remodeling activity. Schematic of the dual activity of PmtA. DNA encoding for PmtA is expressed using PURE*frex*2.0. The resulting protein, including αA helix, is membrane-associated. Using SAM as a cofactor, it catalyzes the triple methylation of DOPE to form DOPC and SAH (S-adenosylhomocysteine). R_1, R_2 are oleoyl chains. DOPE is synthesized, starting from oleoyl-CoA and G3P, by five enzymes (PlsB, PlsC, CdsA, PssA, and Psd) that are encoded on the pGEMM7 mini-genome. The genes *pgpA* and *pgsA* are silenced. Alternatively, DOPE is included in the lipid membrane *a priori*. Next to its catalytic activity, PmtA can also promote membrane remodeling, resulting in the formation of septums.

In the presence of both the *pmtA* gene and SAM, 4.3 µM ± 0.8 µM DOPC was formed. Based on the initial concentrations of DOPE and SAM, the upper limit of DOPC that can be synthesized is 167 µM; the yield observed here is ~2.5%. When *pmtA* and/or SAM was omitted, no synthesis of DOPC was observed (**Fig. 4.3b**). These results indicate that *R. sphaeroides* PmtA can be expressed in the PURE system, and that its catalytic activity can be reconstituted, although the percentage of converted precursors is modest.
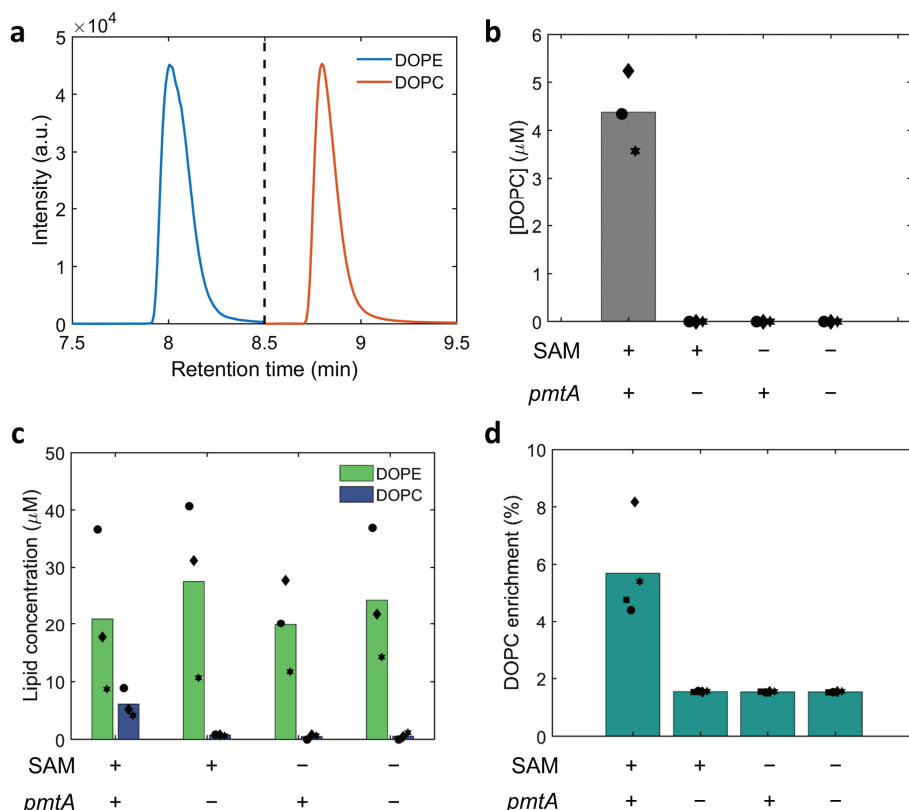
Figure 4.3: LC-MS measurements show that PmtA can synthesize PC alone and in conjunction with pGEMM7. (a) A representative chromatogram of a standard sample containing purified DOPE and DOPC is shown. Isotopically labelled analogues of DOPE and DOPC elute according to an identical pattern. The dashed line indicates the switch of polarity of the detector, from negative to positive mode. This switch allows for the detection of negatively ionized PE and positively ionized PC in the same run. **(b)** Concentration of DOPC, synthesized from DOPE present in the scaffold LUVs, in the presence and absence of the *pmtA* gene and heavy isotope-labelled SAM. Bar height represents the mean from three independent experiments, each represented by a different symbol. **(c)** DOPE is synthesized by expression of pGEMM7 in the presence of all necessary precursors and LUVs that did not contain any DOPE initially. Synthesized DOPC is only detected in the presence of both the *pmtA* gene and SAM. Bar height represents the mean from three independent experiments, each represented by a different symbol. **(d)** The intensity signal of synthesized DOPC expressed as a percentage of the total DOPC intensity signal, i.e. the enrichment of the amount of DOPC upon synthesis. Individual sample points correspond to those in (c) with similar symbol. One extra data set, represented by a square (□), is included. Bar height represents the mean from four independent experiments. The DOPC enrichment of ~1.5% observed in negative controls corresponds to the fraction of native DOPC molecules that carry heavy isotopes resulting in a 3 Da mass shift similar to the one used to distinguish native and synthesized lipids.

4.2.2 Cell-free expressed PmtA functions in conjunction with a reconstituted PE synthesis pathway

Subsequently, we set out to reconstitute *R. sphaeroides* PmtA in conjunction with our previously described[7] five-enzyme *E. coli*-based PE synthesis pathway encoded on a DNA plasmid called pGEMM7. Co-expression was performed with PURE*frex*2.0 in the presence of all necessary reaction substrates, including precursor molecules oleoyl-CoA and isotopically labelled glycerol-3-phosphate (^{13}C-G3P), and 400 nm LUVs containing DOPC, DOPG, and cardiolipin, but not DOPE.

In the presence of the *pmtA* gene and SAM, 6.1 µM ± 2.5 µM DOPC was synthesized **(Fig. 4.3c)**, similar to the synthesized DOPC concentration when only *pmtA* was expressed in the presence of DOPE-containing liposomes. The synthesized DOPC represents 4.1 ± 2.0% of the total amount of DOPC in the LUV membranes **(Fig. 4.3d)**. Without SAM and/or *pmtA*, no DOPC synthesis was observed. The amount of synthesized DOPE (measured DOPE plus measured DOPC) was similar in the presence of SAM and the *pmtA* gene (27 ± 17 µM) to the situation where *pmtA* (28 µM ± 15 µM), SAM (20 µM ± 8 µM) or both (26 µM ± 10 µM) were left out **(Fig. 4.3c)**. Of synthesized PE, 25% ± 7% is converted to PC by PmtA. Summarizing, PmtA can synthesize PC in conjunction with the PE biosynthesis pathway, resulting in the functional reconstitution of a six-enzyme enzymatic cascade *in vitro*, based on genes from two separate species.

The concentration of DOPC that is synthesized when *pmtA* is expressed alone, with a large DOPE precursor pool (~200 µM), is similar to when it is expressed in conjunction with pGEMM7, with much less DOPE available (~25 µM). The concentration of the PmtA enzyme is expected to be lower when co-expressed with pGEMM7, because of IVTT resource sharing[13]. This indicates that the low DOPC yield is not a result of either limited substrate availability, or a low concentration of the PmtA enzyme. Therefore, our experimental conditions are likely suboptimal for PmtA catalytic activity. This could be a result of product inhibition by S-adenosylhomocysteine (SAH) and DOPC[10]. Optimizing the membrane scaffold lipid composition by reducing the amount of PC or increasing the amount of PG to stimulate PmtA membrane association[10,14] could improve the headgroup conversion catalyzed by PmtA.

4.2.3 Expression of PmtA in GUVs results in the formation of septumed liposomes

Now we confirmed that cell-free expressed *R. sphaeroides* PmtA is enzymatically active, we wondered if it would also display an effect on membrane morphology, as reported for the purified *Agrobacterium tumefaciens* PmtA reconstituted in large unilamellar vesicles (**Fig 4.1**). To this end, we prepared giant unilamellar vesicles (GUVs), consisting of DOPC, DOPE, DOPG, and CL, encapsulating PURE*frex*2.0, with or without the *pmtA* gene. Using GUVs instead of LUVs[12] allowed us to use fluorescence laser scanning confocal microscopy (LSCM) in place of electron microscopy. After overnight incubation, we observed, in the sample expressing *pmtA*, a large number of structures we will refer to as 'septumed liposomes' **(Fig. 4.4a)**.

Septumed liposomes appear significantly more when *pmtA* is expressed (17% ± 6%, in the presence of SAM) than when it is not expressed (3% ± 2% without SAM, 2% ± 2% with SAM). Perhaps surprisingly, septums are also generated when *pmtA* is expressed in the absence of SAM (10% ± 6%), indicating that their formation is not solely driven by conversion of DOPE to DOPC. The expression of another protein such as YFP does not result in the formation of septumed liposomes (2% ± 1%) (**Fig. 4.4b**). Time-lapse microscopy of *pmtA*-expressing liposomes shows that the percentage of septumed liposomes gradually increases from <5% at the start of incubation, plateauing after ~10h with 20% of

liposomes displaying a septum (**Fig. 4.4d**). The formation of septumed liposomes therefore occurs on a timescale similar to in-liposome protein expression[13,15] and phospholipid synthesis[5,7].
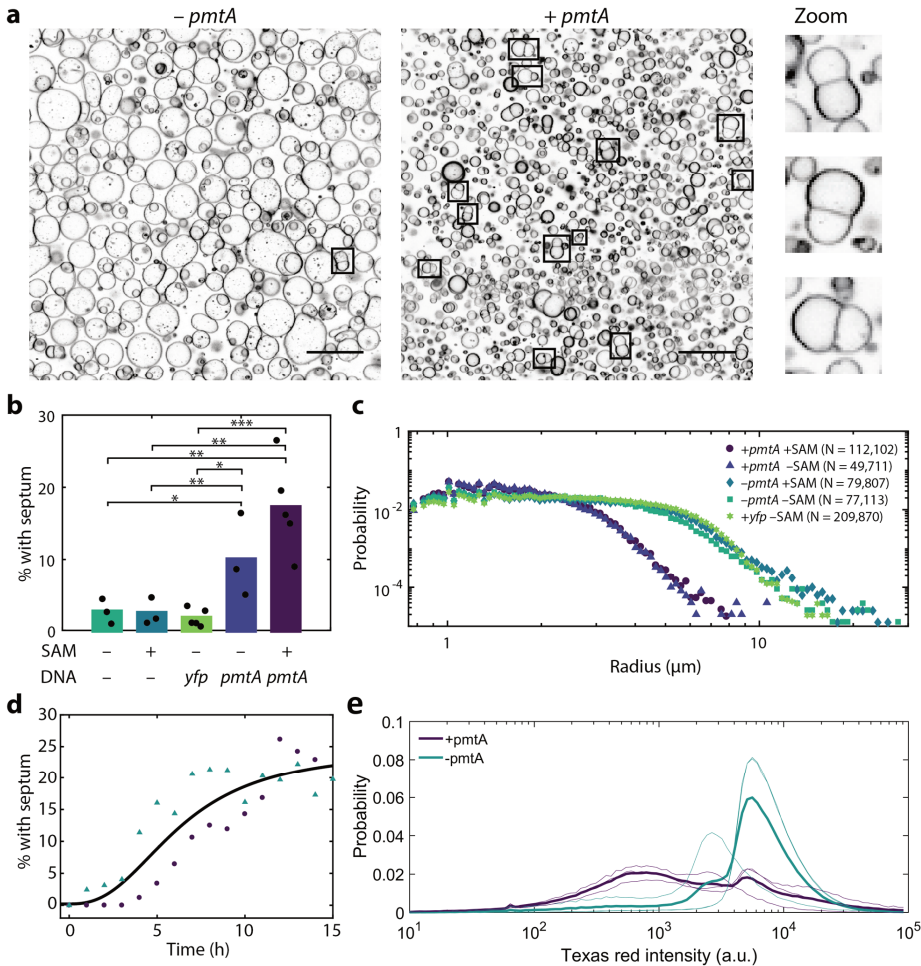
Figure 4.4: Expression of PmtA inside GUVs leads to septum formation and decrease in liposome size. (a) LSCM micrographs of GUVs without *pmtA* DNA and expressing *pmtA* DNA. Framed liposomes exhibit a septum, with three septumed liposomes shown in close-up. Scale bar is 20 μm. **(b)** The percentage of septumed liposomes in a random field-of-view, for different conditions, for multiple repeats (from left to right: $N = 3$, $N = 3$, $N = 5$, $N = 3$, $N = 5$). A two-sample *t*-test was applied between conditions, with *p*-values indicated as stars, with * for $p < 0.05$, ** for $p < 0.01$ and *** for $p < 0.001$. **(c)** Size distributions for liposomes under different conditions. Experiments represented in (b) are pooled. The displayed number is the total number of liposomes analyzed. **(d)** Percentage of septumed liposomes over time, for two independent experiments, with SAM and *pmtA* DNA. The black line is drawn to guide the eye. **(e)** Histograms of Texas Red membrane fluorescence with and without the *pmtA* gene, as determined with flow cytometry. Thin lines represent three individual measurements for each condition, bold lines represent the average. Objects with low Texas Red intensity, which are observed more often in the presence of *pmtA*, are indicated by an arrow.

The exact nature of septumed liposomes is unknown: they could represent two liposomes attached to each other (separated by a double bilayer), or two liposomes separated by a septum (single bilayer). Cryo-electron microscopy could be applied to elucidate the septum nature.

Septum formation represents a morphological transition on the micrometer scale, as opposed to the vesiculation and tubulation previously observed with *A. tumefaciens* PmtA, that display typical length scales of ~10 nm[12]. The important role of the αA helix of *A. tumefaciens* PmtA in membrane remodelling[11], and the finding that this helix and its activity are conserved over a wide range of Pmts[16], suggest that it would also be a main player in the membrane remodeling observed here. However, since the observed morphological change is strikingly different, additional mechanisms should not be excluded, necessitating investigation into the three-dimensional structure of (membrane-bound) *R. sphaeroides* PmtA. Moreover, fluorescent labelling of PmtA might give new insights by informing on its location on the membrane.

4.2.4 Cell-free expression of PmtA reduces the size of liposomes

Cell-free expressed PmtA also influences the size of the liposomes (**Fig. 4.4c**). When PmtA is expressed, liposomes are on average significantly smaller, both in the presence (mean/median radius: 1.63 µm / 1.46 µm) and absence (1.62 µm / 1.46 µm) of SAM, than without gene expression (2.43 µm / 1.91 µm with SAM, 2.29 µm / 1.87 µm without SAM) or when expressing YFP (2.60 µm / 2.14 µm). The size reduction in the absence of SAM corroborates a morphological process not solely dependent on the conversion of DOPE to DOPC. Flow cytometry data of *pmtA*-expressing liposomes display a peak at low Texas Red intensity that is much smaller in the absence of PmtA, demonstrating the presence of small membranous objects (**Fig. 4.4e**). This peak might correspond to small liposomes as observed by LSCM but might also represent small vesicles with a size below the resolution of our microscopy setup (~250 nm). The number of events per volume is markedly higher when *pmtA* is expressed (~3×10^5 events per µL, considered volume is prior to dilution) than when it is not (~1×10^5 events per µL), suggesting vesiculation or liposome fission, presumably with septumed liposomes as an intermediate step. Alternatively, the size decrease could be explained by tubulation, as observed by Danne et al.[12], with the resulting tubes below the diffraction limit of the microscopy set-up. This mechanism would however not explain the formation of septums or the increase in event concentration. The formation of septumed liposomes by (hemi-)fusion would result in a lower number of larger liposomes, and is therefore in the light of these findings not a likely mechanism. This could be further confirmed by performing a similar experiment with two populations of liposomes with a different membrane dye.
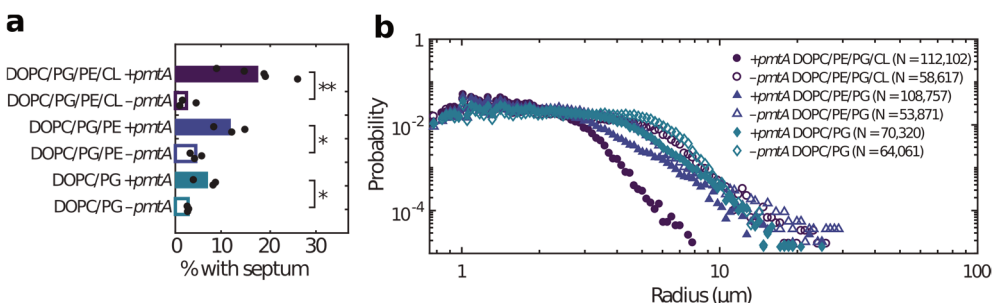
4.2.5 The presence of cardiolipin and methylated PE intermediates stimulates septum formation

The ability of *A. tumefaciens* PmtA to remodel membranes was shown to be strongly dependent on lipid composition[12] (**Fig. 4.1**). Particularly, anionic lipids like PG are needed for membrane binding and cardiolipin was found to be essential for membrane deformation. Moreover, the presence of monomethyl PE (MMPE, the first intermediate in PC synthesis) was found to change the remodeling type from tubulation to vesiculation. To investigate if lipid composition affects septum formation by *R. sphaeroides* PmtA, we also prepared GUVs containing PURE*frex*2.0 without cardiolipin and without cardiolipin and DOPE (so no PE-methylated species could be synthesized), all in the presence of SAM (**Fig. 4.5a**). In the absence of cardiolipin, but with DOPE, a slightly lower percentage of septumed liposomes was observed upon *pmtA* expression (12% ± 3%). This percentage is however still significantly higher than the percentage of septums observed in similar liposomes in the absence of PmtA (4% ± 1%). Expressing *pmtA* in liposomes composed of solely DOPC and DOPG, preventing catalytic activity of the protein and therefore the formation of MMPE and the second intermediate

dimethyl PE (DMPE), results in an even lower number of septumed liposomes (7% ± 3%), but again higher than in similar liposomes without PmtA (2.8% ± 0.2%). We observed that GUVs lacking cardiolipin display a size reduction upon PmtA expression (**Fig. 4.5b**, mean/median radius with PmtA: 1.87 µm / 1.55 µm, without PmtA: 2.34 µm / 1.96 µm). Liposomes consisting of only DOPC and DOPG are very large in the absence of PmtA (2.78 µm/ 2.26 µm), but display sizes comparable to DOPC/PE/PG/CL liposomes upon *pmtA* expression (2.26 um / 1.85 um). The decrease of the mean radius upon PmtA expression without cardiolipin (20% decrease) and without cardiolipin and DOPE (19% decrease) is lower than when both are included (33% decrease), but still significant. Concluding, septum formation and size decrease are stimulated by – but not strictly dependent on – MMPE and DMPE synthesis and cardiolipin.

If the morphological transformations observed really constitute liposome fission, the activity of *R. sphaeroides* PmtA would be the most minimal mechanism to reconstitute cell division, that is not dependent on a varying environment, such as temperature cycles[17,18] or osmolarity changes[19]. The most popular system under investigation with regard to synthetic cell division, the Z-ring, needs at least two proteins (FtsZ and FtsA/ZipA)[20], and a further three for positioning in the mid-cell (MinC, MinD, and MinE)[21]. Other candidates such as the ESCRT system also require multiple gene products[22]. None of these systems have been successfully employed to realize cell division yet. Therefore, further investigations into PmtA as a minimal cell division protein could represent an exciting new avenue in synthetic cell research.

Figure 4.5: Septum formation and size reduction by PmtA is promoted by, but not strictly dependent on, cardiolipin and methylated intermediates. **(a)** The percentage of septumed liposomes for different lipid compositions, with and without expression of the *pmtA* gene, for multiple repeats (top: $N = 5$, others: $N = 3$). A two-sample *t*-test was applied between conditions, with *p*-values indicated as stars, with * for $p < 0.05$, and ** for $p < 0.01$ **(b)** Size distributions of liposomes of different lipid composition, with (filled symbols) and without (open symbols) *pmtA* DNA. Experiments represented in (a) are pooled. The displayed number is the total number of liposomes analyzed.

4.3 MATERIALS AND METHODS

All buffers and solutions were made using Milli-Q grade water with 18.2 MΩ resistivity (Millipore, USA). Reagents were purchased from Sigma-Aldrich unless otherwise indicated.

4.3.1 DNA constructs

A plasmid containing *R. sphaeroides pmtA* in a pUC57 backbone, with codon usage optimized for *E. coli* and therefore the PURE system, was ordered at GenScript. Polymerase chain reaction with forward primer ChD300 (5'-GCGAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTCCCTCTAGAAATAATTTTG-3') and reverse primer ChD91 (5'-AAAAAACCCCTCAAGACCCGTTTAGAGG-3') was applied to generate a 748 bp linear DNA fragment encoding for *pmtA* under a T7 promoter. The construction of the pGEMM7 plasmid is described in Blanken & Foschepoth et al.[7].

4.3.2 Preparation of LUVs

Large unilamellar vesicles (LUVs) were prepared by extrusion of large multilamellar vesicles (LMVs). A film of 2 mg of phospholipids was made by mixing the lipids, dissolved in chloroform, in a 2 mL glass vial, and subsequently applying a gentle argon flow for 10 min followed by an hour of desiccation. For experiments involving only PmtA, the lipid composition was DOPC/DOPE/DOPG/CL (50 mol%/36 mol%/12 mol%/2 mol%). When *pmtA* and pGEMM7 were co-expressed, the lipid composition was DOPC/DOPG/CL (51 mol%/47 mol%/2 mol%). All lipids were purchased from Avanti Polar Lipids. The lipid film was then resuspended in 250 µL buffer A (20 mM HEPES, 180 mM potassium glutamate, 14 mM magnesium acetate, pH 7.6). LMVs were made by vortexing, and subsequently extruded with the Avanti mini-extruder equipped with a 400 nm membrane, according to instructions provided by the manufacturer. LUVs were aliquoted, snap-frozen in liquid nitrogen, and stored at −20 °C.

4.3.3 Lipid synthesis experiments

For experiments where only the *pmtA* gene was expressed, PURE*frex*2.0 (GeneFrontier) was assembled according to instructions by the supplier, and supplemented with 500 µM (RS)-S-Adenosyl-L-methionine-d$_3$ (S-methyl-d$_3$) tetra(p-toluenesulfonate) salt (isotopically labelled SAM, CDN Isotopes), 0.4 mg/mL LUVs, 0.75 U/µL SUPERase (Invitrogen), and 22 nM of the *pmtA* gene. When *pmtA* was expressed in conjunction with pGEMM7, isotopically labelled SAM was replaced by an equal amount of non-labelled SAM, and 5 mM β-mercaptoethanol, 500 µM ^{13}C-labelled glycerol-3-phosphate (^{13}C-G3P), 1 mM cytidine triphosphate, 500 µM L-serine, 0.4 mg/mL LUVs, and 1 nM of pGEMM7 plasmid were added to the aforementioned PURE*frex*2.0 reaction mix which was subsequently added to a dried oleoyl-Coenzyme A (oleoyl-CoA) film. The oleoyl-CoA concentration upon rehydration was 100 µM. Dried films of oleoyl-CoA were prepared in a 0.2 mL reaction tube by evaporation of the storage solution chloroform:methanol:water (80:20:2) under ambient pressure. Reactions were incubated overnight at 37 °C in a BioRad C1000 Touch Thermal Cycler.

4.3.4 Liquid chromatography-coupled mass spectrometry

The amount of (synthesized) lipids was measured using liquid chromatography-coupled mass spectrometry (LC-MS), according to the method previously described in chapter 3 (Blanken & Foschepoth et al.[7]). Pre-ran PURE*frex*2.0 samples were diluted ten times in methanol containing 10 mM EDTA and 2 mM acetylacetone. After 10 min of sonication and 5 min of centrifugation at 16,000 g, the lipid-containing supernatant was transferred to Agilent LC-MS vials with low-volume inset,

flushed with argon and stored at −20 °C. Liquid chromatography was performed with a ACQUITY UPLC® Peptide CSH™ C18 column using a gradient of two mobile phases (A: water with 0.05% ammonium hydroxide and 2 mM acetylacetone, and B: 80% 2-propanol, 20% acetonitrile, 0.05% ammonium hydroxide and 2 mM acetylacetone) at a flow rate of 300 µL/min and a column temperature of 60 °C. To equilibrate the column, a ratio of mobile phase A to mobile phase B of 70:30 was used. Upon injection, this ratio was gradually changed to 100% mobile phase B over the course of 8 min and then kept constant for 2 min. Subsequently, over the course of 1 min, the initial 70:30 ratio of mobile phase A and B was reset, which was then used for the last 4 min of the run. The built-in autosampler of the LC-MS system was used to inject 1 µL of sample solution.

Mass spectrometry was performed with an Agilent 6460 Triple Quad MS system coupled to the liquid chromatography set-up. Transitions were based on our previous work[5,7]. Synthesized lipids were distinguished by a 3 Da mass shift resulting from the incorporation of heavy G3P (*pmtA* with pGEMM7 experiments), or a 9 Da mass shift resulting of the incorporation of three methyl groups from isotopically labelled SAM (*pmtA*-only experiments). New transitions were determined for DOPC and its isotopically labelled analogues (**Supplementary Table 4.1**). In contrast to the other lipid species, DOPC bears a positive charge upon ionization and it could not be detected simultaneously. Making use of the fact that DOPC elutes after all other lipid compounds, a dynamic multi reaction monitor (dMRM) acquisition protocol was set up. First, 8.5 min of detection in negative mode was performed, scanning for negatively ionized lipids. Then, for 1 min, the detector was switched to positive mode to detect DOPC and analogues (**Fig. 4.2a**). All samples were injected twice, in random order. Quantification was performed by injecting a dilution series of purified lipids of interest in methanol with 10 mM EDTA and 2 mM acetylacetone, before and after measurement.

4.3.5 Shape deformation assay

Giant unilamellar vesicles (GUVs) encapsulating the *pmtA* gene and PURE*frex*2.0 were prepared as described previously in chapters 2 and 3[7,13]. In short, 212-300 µm glass beads were coated with a mixture of phospholipids (DOPC:DOPE:DOPG:CL 50:32:16:2 supplemented with 0.5 mass% DHPE-Texas Red and 1mass% DSPE-PEG-biotin, unless noted otherwise). All lipids were from Avanti, except DHPE-TR that was from Thermo Fisher Scientific. A PURE*frex*2.0 reaction was assembled according to instructions provided by the supplier, with 22 nM *pmtA* DNA and 500 µM SAM added. Resuspension of the lipid film with the PURE system mixture results in the formation of liposomes after two hours of swelling. Four freeze-thaw cycles were applied by dipping the sample into liquid nitrogen to increase unilamellarity and encapsulation efficiency. For microscopy experiments, after freeze-thawing, 2 µL of liposomes were transferred to a 5.5 µL of feeding solution consisting of PURE*frex*2.0 Solution I and Milli-Q (4:7 vol/vol) with RQ1 DNase (0.07 U/µL, Promega)) to prevent gene expression in the extra-liposomal solution. Liposomes were subsequently immobilized on the surface of a custom-made glass imaging chamber, using biotin-streptavidin interaction. For flow cytometry experiments, 0.25 µL DNase I solution was added to 4.75 µL of liposomes. Samples were incubated overnight at 37 °C before analysis.

4.3.6 Microscopy and image analysis

Liposomes were imaged using a Nikon A1R laser scanning confocal microscope, using an SR Apo TIRF 100× oil immersion objective. The 561 nm laser line was used to image the Texas Red fluorescence of the liposome membranes. The height of the focal plane above the surface of the chamber was adjusted manually to equatorially dissect as many liposomes as possible. For each sample, multiple montages of six-by-six fields-of-view were acquired, containing thousands of liposomes.

To obtain the size distribution of liposomes, we applied our previously described automated image analysis MATLAB script (see chapter 2). To investigate the occurrence of the septumed liposomes, we opted for a manual image analysis procedure. A 500×500 pixel field-of-view was randomly (while assuring sufficient number of liposomes) selected from the data. The number of septumed liposomes in this field of view was then manually counted independently by two people. To prevent bias, the counters were not informed of the conditions of the particular experiment. The total number of liposomes in the field-of-view was determined using the aforementioned image analysis algorithm and used to determine the percentage of septumed liposomes. All percentages reported in the text are the mean of those determined by the two counters.

4.3.7 Flow cytometry

To facilitate flow cytometry, the liposome mixture was diluted 100× in buffer A after overnight incubation. A BD FACSCelesta flow cytometer (BD Biosciences) was used to measure 200 µL of diluted liposome solution. Texas Red fluorescence was obtained with the pre-set mCherry setting. Data was analyzed using MATLAB and the online Cytobank application (https://community.cytobank.org/cytobank/experiments).

REFERENCES

1. Luisi, P. L., Ferri, F. & Stano, P. Approaches to semi-synthetic minimal cells: A review. *Naturwissenschaften* **93,** 1–13 (2006).

2. Nourian, Z., Scott, A. & Danelon, C. Toward the assembly of a minimal divisome. *Syst. Synth. Biol.* **8,** 237–247 (2014).

3. Caspi, Y. & Dekker, C. Divided we stand: splitting synthetic cells for their proliferation. *Syst. Synth. Biol.* **8,** 249–269 (2014).

4. Schwille, P. et al. MaxSynBio: Avenues towards creating cells from the bottom up. *Angew. Chemie Int. Ed.* **57,** 2-13 (2018).

5. Scott, A. et al. Cell-free phospholipid biosynthesis by gene-encoded enzymes Reconstituted in Liposomes. *PLoS One* **11,** e0163058 (2016).

6. Exterkate, M., Caforio, A., Stuart, M. C. A. & Driessen, A. J. M. Growing membranes in vitro by continuous phospholipid biosynthesis from free fatty acids. *ACS Synth. Biol.* **7,** 153–165 (2018).

7. Blanken, D., Foschepoth, D., Serrão, A. C. & Danelon, C. Genetically controlled membrane synthesis in liposomes. *bioRxiv* 2020.03.29.013300 (2020) doi:10.1101/2020.03.29.013300.

8. Bhattacharya, A., Brea, R. J., Niederholtmeyer, H. & Devaraj, N. K. A minimal biochemical route towards de novo formation of synthetic phospholipid membranes. *Nat. Commun.* **10,** 300 (2019).

9. Vogele, K. et al. Towards synthetic cells using peptide-based reaction compartments. *Nat. Commun.* **9,** 3862 (2018).

10. Aktas, M. & Narberhaus, F. In vitro characterization of the enzyme properties of the phospholipid N-methyltransferase PmtA from Agrobacterium tumefaciens. *J. Bacteriol.* **191,** 2033–2041 (2009).

11. Danne, L. et al. Membrane-binding mechanism of a bacterial phospholipid N-methyltransferase. *Mol. Microbiol.* **95,** 313–331 (2015).

12. Danne, L. et al. Membrane remodeling by a bacterial phospholipid-methylating enzyme. *MBio* **8,** 1–15 (2017).

13. Blanken, D., van Nies, P. & Danelon, C. Quantitative imaging of gene-expressing liposomes reveals rare favorable phenotypes. *Phys. Biol.* **16,** 045002 (2019).

14. Shibuya, I. Metabolic regulations and biological functions of phospholipids in Escherichia coli. *Prog Lipid Res*. **31,** 245–299 (1992).

15. Godino, E. et al. De novo synthesized Min proteins drive oscillatory liposome deformation and regulate FtsA-FtsZ cytoskeletal patterns cytoskeletal patterns. *Nat. Commun.* **10,** 4969 (2019)

16. Danne, L. et al. Dissection of membrane-binding and -remodeling regions in two classes of bacterial phospholipid N-methyltransferases. *Biochim. Biophys. Acta - Biomembr.* **1859,** 2279–2288 (2017).

17. Sakuma, Y. & Imai, M. Model system of self-reproducing vesicles. *Phys. Rev. Lett.* **107,** 1–5 (2011).

18. Jimbo, T., Sakuma, Y., Urakami, N. & Imai, M. Role of inverse-cone-shape lipids in temperature-controlled self-reproduction of binary vesicles. *Biophys. J.* **110,** 1551–1562 (2016).

19. Döbereiner, H. G., Käs, J., Noppl, D., Sprenger, I. & Sackmann, E. Budding and fission of vesicles. *Biophys. J.* **65,** 1396–403 (1993).

20. Errington, J., Daniel, R. A. & Scheffers, D.-J. Cytokinesis in bacteria. *Microbiol. Mol. Biol. Rev.* **67,** 52–65 (2003).

21. Rowlett, V. W. & Margolin, W. The Min system and other nucleoid-independent regulators of Z ring positioning. *Front. Microbiol.* **6,** 1–10 (2015).

22. Makarova, K. S., Yutin, N., Bell, S. D. & Koonin, E. V. Evolution of diverse cell division and vesicle formation systems in Archaea. *Nat. Rev. Microbiol.* **8,** 731–741 (2010).

SUPPLEMENTARY INFORMATION

Supplementary Table 4.1: Mass spectrometry settings for phospholipid detection. Asterisk (*) indicates species incorporating ^{13}C-G3P, resulting in a 3 Da mass shift with respect to the regular species. Two asterisks (**) indicate species that have incorporated d_3-methyl groups for isotopically labelled SAM, resulting in a 9 Da mas shift with respect to the regular species. Fragmentor voltage and collision energy were adapted from our previous works[5,7].

Compound name	Precursor Ion (*m/z*)	Product Ion (*m/z*)	Fragmentor (V)	Collision energy (eV)	Retention time window (min)	Polarity
DOPC	786.4	184.0	135	50	8.5 - 9.5	Positive
DOPC*	789.4	184.0	135	50	8.5 - 9.5	Positive
DOPC**	795.4	193.0	135	50	8.5 - 9.5	Positive
DOPE	742.5	281.2	140	25	0 - 8.5	Negative
DOPE*	745.5	281.2	140	25	0 - 8.5	Negative

Chapter 4: A cell-free expressed membrane protein synthesizes lipids and remodels giant vesicles

4

Chapter 5: Towards 100% membrane growth: improving lipid synthesis yield

The yield of phospholipid synthesis by cell-free expressed Kennedy pathway proteins reconstituted inside liposomes is currently limited by the low solubility of the acyl-CoA precursor. In this chapter, several approaches are considered to ameliorate this problem. Using the more soluble palmitoyl-acyl carrier protein complex as an alternative acyl donor proved ineffective. Upstream pathway extension with fadD *facilitated the use of the more soluble precursor oleic acid, but the higher substrate concentration was offset by a lower conversion yield, resulting in phospholipid yields similar to what was achieved with oleoyl-CoA. Circumventing the Kennedy pathway, native chemical ligation-based synthesis of membrane-forming lipid analogues, as developed by the Devaraj lab, was investigated. In an optimized reaction buffer, lipidic structures could be synthesized* de novo, *but the reactive precursors significantly inhibited PURE system gene expression, making this approach incompatible with our synthetic cell framework. Continuous delivery of low concentrations of acyl-CoA with a flow reactor set-up is proposed as an alternative approach to realize the 100% lipid increase needed for synthetic cell proliferation.*

5.1 INTRODUCTION

The role of lipid synthesis within the synthetic cell framework is to increase the membrane area to sustain proliferation cycles. To grow and proliferate without shrinking, a doubling of the membrane surface per cycle is necessary, either prior to or following fission. Therefore, the amounts of lipids synthesized should equal the amount of lipids present in the initial membrane of the parent liposome; in other words, the amount of lipids should increase by 100%. When division results in more than two daughter cells (i.e. via budding off small vesicles), the amount of lipids synthesized by parent and offspring should be even higher.

In chapters 3 and 4, we describe the synthesis of dioleoylphosphatidylglycerol (DOPG), dioleoylphosphatidylethanolamine (DOPE), and dioleoylphosphatidylcholine (DOPC), which together make up 98% of our initial liposome membrane. However, the total concentration of synthesized lipids was in the order of ~30 µM. This stands in contrast to the millimolar amounts of lipids already present at the start of the reaction. The increase in lipid content and consequently membrane area is therefore in the order of <5%, very far removed from the desired 100%.

The main culprit limiting the lipid synthesis yield is the solubility of the acyl-CoA precursor. Low acyl-CoA solubility in the presence of Mg^{2+} limits the initial concentration to 100 µM[1,2]. Since every synthesized lipid molecule contains two acyl chains, even with perfect conversion the maximum expected yield is 50 µM. This problem can be circumvented in two ways: (1) continuous resupply of acyl-CoA and (2) utilizing (much) more soluble precursors. The first approach will involve the use of flow chambers or microfluidic systems. The second approach, which is the focus of this chapter, involves upstream pathway extension or implementing an alternative route to produce membrane lipids from non-natural substrates.

In nature, acyl chain precursors are fed to the Kennedy phospholipid synthesis pathway in two distinct ways. Firstly, exogenous fatty acids are activated by the enzyme FadD that synthesizes acyl-CoA from fatty acids and coenzyme A[3]. *In vitro* reconstitution of purified FadD in combination with the Kennedy pathway has been realized by the group of Driessen[4]. Using high amounts of fatty acids as a precursor, they could synthesize millimolar amounts of DOPE and DOPG, resulting in putative membrane growth.

Secondly, fatty acids are synthesized *de novo* via the fatty acid synthesis (FAS) pathway. The central protein cofactor in the FAS system is the acyl carrier protein (ACP) that carries all intermediates as well as the acyl chain end product[5]. Acyl-ACP can also serve as an acyl-chain donor to PlsB and PlsC[6–9], and is, in fact, the dominant precursor *in vivo* when fatty acids are synthesized endogenously[9]. In bacteria, plants, and algae, fatty acid synthesis starts with the formation of malonyl-CoA from acetyl-CoA, catalyzed by acetyl-CoA carboxylase. The CoA moiety is subsequently exchanged for ACP by FabD. FabH then condensates malonyl-ACP with acetyl-CoA, after which the FA synthesis cycle, consisting of the proteins FabG, FabA/FabZ, FabI, and FabB/FabF, elongates the molecule two carbons at a time to a full length product of 16 or 18 carbons in length[5]. This pathway **(Fig. 5.1a)** is referred to as type II fatty acid synthesis (FASII). In non-plant eukaryotes, these functionalities are provided by a single multi-domain megasynthase referred to as FAS type I[10] **(Fig. 5.1b)**.

Impressively, the group of Khosla has reconstituted the full FASII pathway from fourteen purified proteins, synthesizing up to 50 µM of palmitic acid[11,12]. Murtas has published results on *in vitro* palmitic acid synthesis using a purified type I fatty acid synthesis megaprotein from *Brevibacterium ammoniagenes*[13]. Whereas using fatty acids as lipid synthesis precursor requires only a single extra gene (*fadD*) to be expressed, reconstituting either type I or II fatty acid synthesis by cell-free gene

expression will be severely complicated by respectively the size or amount of enzyme(s) needed. Reconstitution of the FAS pathway in the context of an autonomous synthetic cell will therefore only become attainable when major advances in the performance of cell-free gene expression systems are realized.
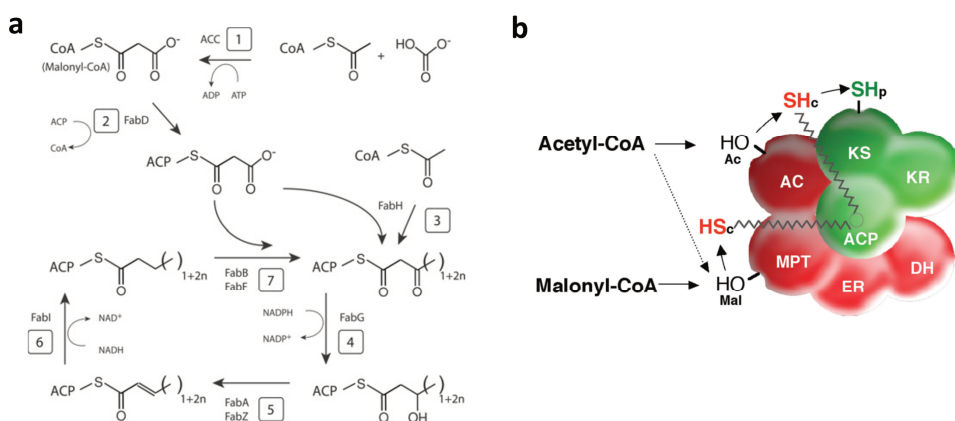
Figure 5.1: Fatty acid synthesis is performed by two distinct systems in nature. (a) Type 2 fatty acid synthesis pathway as described in the main text. Numbers indicate reaction order. Taken from Chan & Vogel[5]. **(b)** Type 1 fatty acid synthesis is performed via a megaprotein with multiple domains: ac(et)yltransferase (AC), malonyl/acetyl- or malonyl/ palmitoyl-transacylase (AT, MPT), ketoacyl synthase (KS), ketoacyl reductase (KR), dehydratase (DH), enoyl reductase (ER), and ACP. ACP guides the nascent fatty acid through the pathway, here indicated by the zigzag line. Taken from Schweizer & Hofmann[14].
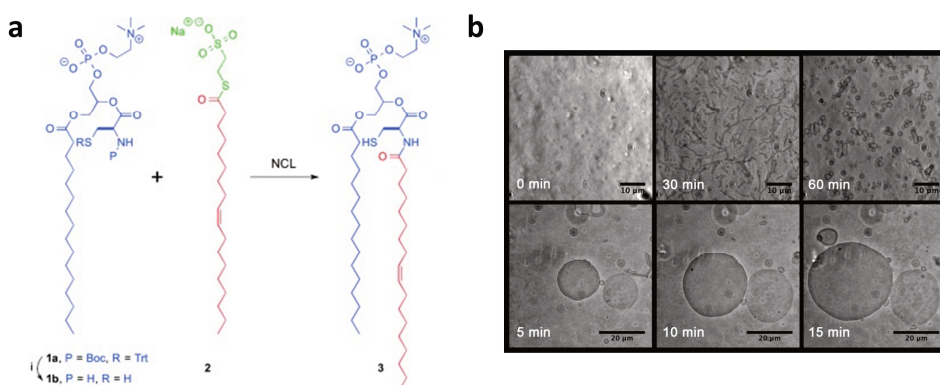
Figure 5.2: Vesicle synthesis and growth by lipid analogue production via native chemical ligation. (a) 1-palmitoyl-2-(L-Cys)-*sn*-glycero-3-phosphocholine (p-LPC, **1b**) and sodium 2-mercaptoethane-sulfonate (MESNA) oleate (**2**) form lipid analogue **3** via native chemical ligation (NCL). Substituting the palmitoyl chain in **1b** yields 1-oleoyl-2-(L-Cys)-*sn*-glycero-3-phosphocholine (o-LPC). **(b)** Phase contrast microscopy at different times of *in situ* vesicle formation (top) and vesicle growth (bottom) in the presence of 1 mM **1b** and **2**. Figure taken from Brea, Cole, & Devaraj[15].

A completely different approach to lipid synthesis from soluble precursors has been investigated by the group of Devaraj. They have demonstrated fast synthesis of high concentrations of membrane-forming lipid analogues from soluble precursors by native chemical ligation (**Fig. 5.2**). This click-chemistry-based approach is attractive for its simplicity, (presumed) biocompatibility, and since no enzymes are needed for catalysis[15–19]. They have demonstrated coupling of chemical lipid synthesis to an enzymatic reaction catalyzed by cell-free expressed FadD10 (a FadD analogue)[21]. Coupling to a genetically encoded element is essential when integrating chemical lipid synthesis into an autonomous synthetic cell.

Here, we will discuss phospholipid synthesis pathway extension by reconstitution of FadD, the use of acyl-ACP as a lipid synthesis precursor and attempts to combine chemical lipid synthesis with PURE system gene expression. We find that FadD can be functionally expressed in conjunction with the Kennedy pathway enzymes in the PURE system and can catalyze the formation of acyl-CoA from fatty acids and CoA. However, conversion yield was low and subsequent phospholipid yield from the downstream pathway was not significantly higher than when directly using acyl-CoA as a precursor, regardless of the addition of millimolar amounts of fatty acid. Acyl-ACP was found to be an ineffective acyl chain donor to PlsB and PlsC in our system, likely a result of its low purity. Chemical lipid synthesis resulted in the formation of large amounts of membrane-forming lipid analogues in a PURE system background; however, gene expression was dramatically decreased in the presence of chemical lipid synthesis precursors. Concluding, the three approaches towards higher lipid yield investigated here will require significant further development to achieve 100% membrane growth. We propose further investigations into continuous supply of acyl-CoA as the most promising next step.

5.2 RESULTS & DISCUSSION

### 5.2.1 Upstream lipid synthesis pathway extension with FadD

Co-expression of *fadD* and the pGEMM7 plasmid encoding for the pathway to synthesize PE (chapter 3), in the presence of large unilamellar vesicles (LUVs), resulted in DOPE synthesis from oleic acid, as detected by liquid chromatography-coupled mass spectrometry (LC-MS). This result indicates that FadD can both be expressed in the PURE system and is enzymatically active in combination with the Kennedy pathway. It therefore represents successful upstream pathway extension. However, the resulting DOPE concentration of ~30 µM was not significantly higher than can be reached starting with oleoyl-CoA, notwithstanding the ten- or twenty-times higher precursor concentration (**Fig. 5.3**). A microscopy assay was performed using giant unilamellar vesicles (GUVs) with 50 mol% oleic acid added to the membrane. Time lapse microscopy of these GUVs, encapsulating PURE*frex*2.0, the *fadD* gene, 500 µM CoA, and pGEMMΔ*psd* displayed growth of the GUVs at the expense of the oleic acid blob (**Fig. 5.4a**). However, this process was also observed in the absence of CoA (**Fig. 5.4b**), suggesting a dewetting process independent of acyl-CoA synthesis by FadD.

Implementing FadD to realize a significant improvement in lipid synthesis yield will require extensive optimization of reaction conditions for FadD, while retaining conditions that are compatible with PURE system gene expression and the other enzymatic steps in the lipid synthesis pathway. A potential workaround could involve the use of FadD10, which instead of acyl-CoA synthesizes fatty acyl adenylates[20] that potentially also can be used as acyl donor for PlsB and PlsC. A more innovative, albeit challenging, approach would be the use of directed evolution methods to create a FadD that can better function in the current experimental conditions. To that end, FadD activity needs to be coupled to PS synthesis, and selection pressure can be applied based on PS concentration as assessed by

LactC2-eGFP fluorescence (chapter 3). Moreover, the method by which fatty acids are supplied to the reaction could be optimized. For example, fusion of fatty acid vesicles with the FadD-containing liposomes might increase substrate conversion. Using fatty acids of different chain lengths than oleic acid might positively affect yield, but will result in the synthesis of phospholipids that do not match the initial liposome membrane.
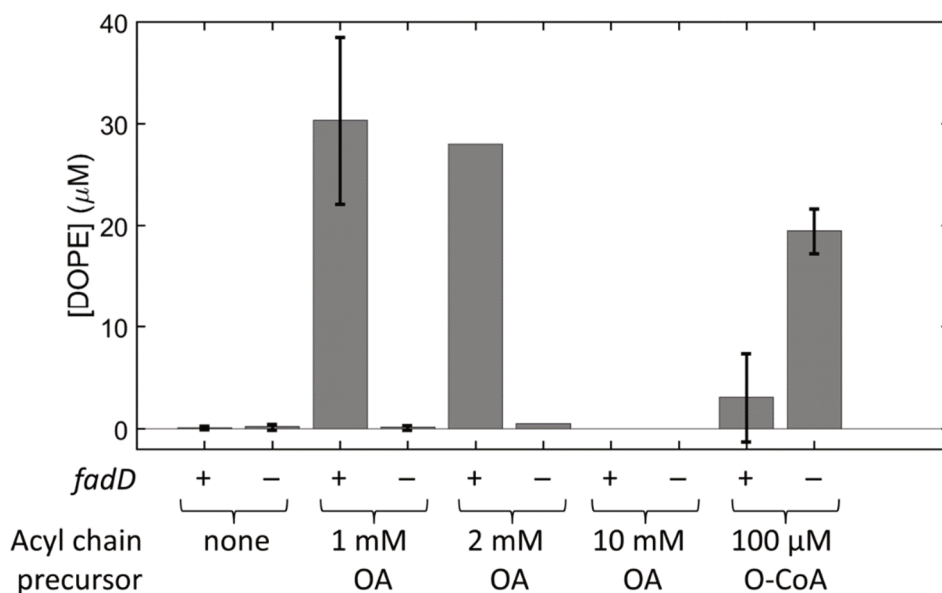

Figure 5.3: Oleic acid can be utilized as a phospholipid precursor by co-expressing pGEMM7 and the *fadD* gene. Lipid synthesis enzymes encoded by pGEMM7 were produced in the presence of LUVs and various amounts of oleic acid (OA) or oleoyl-CoA (O-CoA), with or without *fadD* DNA. In all conditions, 50 µM CoA was added. Data are absolute concentrations of synthesized DOPE as measured by LC-MS. When two independent experiments were performed, an error bar indicating the standard deviation is shown. Conditions without error bar are assayed once. FadD can partially convert 1 mM or 2 mM OA into oleoyl-CoA (O-CoA), which can subsequently be used for synthesis of DOPE. In the absence of *fadD* template, OA cannot be processed by the enzymatic pathway encoded in pGEMM7. Addition of 10 mM OA did not result in lipid synthesis, very likely because at such a high concentration, OA solubilizes the LUVs. Supplying the OA in a carrier, such as mixed PC-OA vesicles, might be explored to prevent this. Lipid synthesis from O-CoA is less efficient in the presence of FadD which presumably can also catalyze the reverse reaction, that is the conversion of O-CoA into CoA and OA, reducing the pool of acyl donor precursor for PlsB and PlsC.
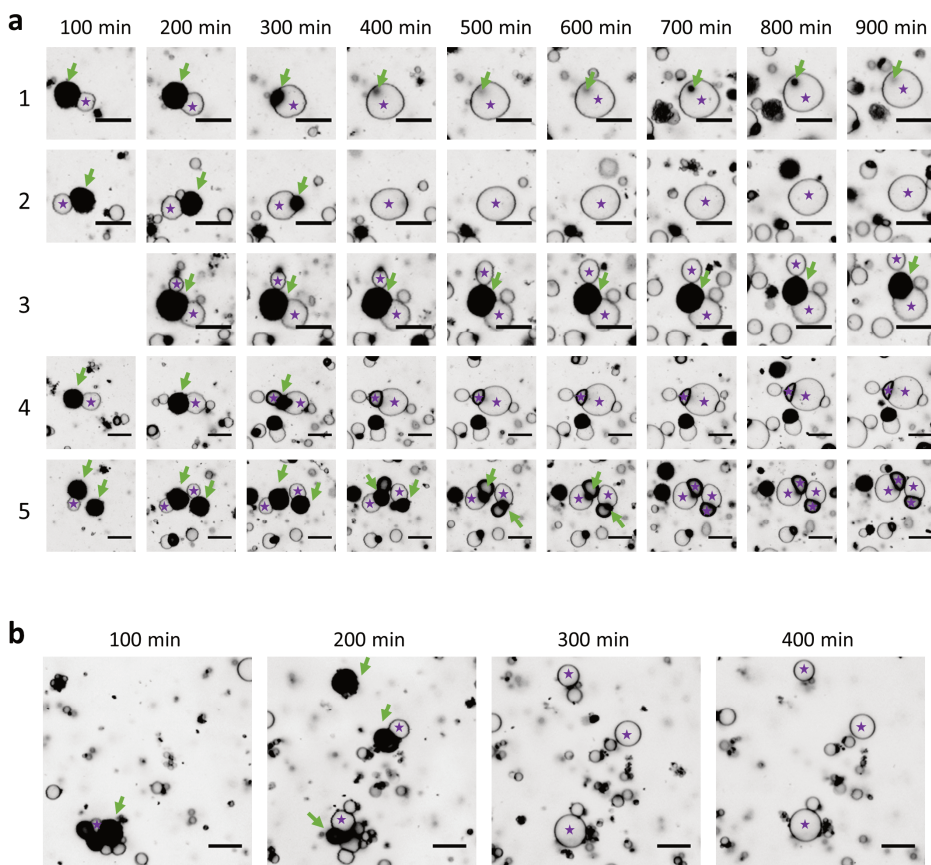
Figure 5.4: Time lapse micrographs of *fadD*-expressing liposomes are suggestive of growth but represent dewetting. (a) Time-lapse micrographs of five fields of view with liposomes containing PURE*frex*2.0, the *fadD* gene, 500 µm CoA, and pGEMM*Δpsd*. The black oleic acid blobs, indicated by green arrows, shrink, and the attached liposomes, indicated by purple stars, grow. Scale bar is 10 µm. **(b)** The same experiment, but in the absence of CoA, excluding FadD enzymatic activity. Blob shrinkage and liposome growth still occur. Scale bar 20 µm.

5.2.2 Palmitoyl-ACP is an inefficient acyl chain donor for pGEMM7-directed lipid synthesis

DPPE was synthesized by expressing pGEMM7 in the presence of LUVs and all necessary precursors (chapter 3). As acyl-chain donor, palmitoyl-acyl carrier protein (palmitoyl-ACP) and palmitoyl-CoA were used. Palmitoyl-ACP was taken from a stock solution of 1.65 mg/mL in 10 mM MgSO$_4$, kindly provided by Marek Noga. Biosynthesis yield of DPPE from 10 µM palmitoyl-ACP was more than one order of magnitude lower than the yield using an equal concentration of palmitoyl-CoA. Final concentration of input palmitoyl-ACP was limited by the stock concentration. In contrast, palmitoyl-CoA could be used at up to 100 µM, leading to a yield of synthesized DPPE that is over two orders of magnitude higher compared to that with 10 µM palmitoyl-ACP **(Fig. 5.5)**.
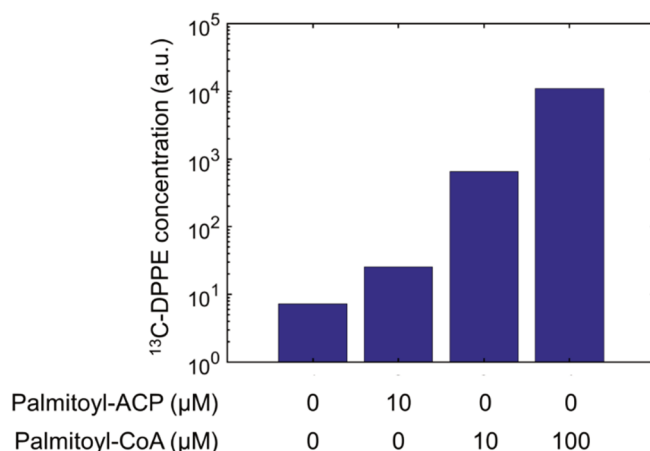
Figure 5.5: DPPE synthesis with palmitoyl-ACP or -CoA as acyl donor. Lipid synthesis enzymes encoded by pGEMM7 were produced in the presence of LUVs and all necessary precursors. Palmitoyl-ACP was supplied in solution, in contrast to palmitoyl-CoA, which was supplied as a dried film. Y-axis is total integrated counts. Experiment was performed once.

Purification of acyl-ACP to higher concentration and activity will be paramount to use it as a viable precursor for lipid synthesis. Alternatively, ACP could be cell-free expressed *in situ.* Although ACP is a small 9 kDa protein, it requires a posttranslational modification, where a phosphopantetheine group from CoA is attached to a serine residue by a protein called ACP synthase[5]. Moreover, fatty acids need to be bound to the ACP, requiring the expression of yet another enzyme such as FabD. Using *in situ* generated acyl-ACP will therefore be exceedingly challenging, and not very minimal.

5.2.3 Chemical lipid synthesis is incompatible with PURE system gene expression

We aimed to couple click chemistry-based lipid synthesis, as demonstrated by the Devaraj group, to gene-encoded fatty acid activation with the FadD enzyme. This approach could combine the advantages of gene-encoded processes for the synthetic cell (regulation, autonomy) with the high yields of chemistry-based methods for lipid synthesis.

We have received the precursors sodium 2-mercaptoethane-sulfonate (MESNA) oleate, 1-oleoyl-2-(L-Cys)-*sn*-glycero-3-phosphocholine (o-LPC), and 1-palmitoyl-2-(L-Cys)-*sn*-glycero-3-phosphocholine (p-LPC) from the Devaraj lab **(Fig. 5.2a)**. Mass spectrometry signal was found for MESNA oleate and o-LPC, but not for their synthesis product when incubated together (2.5 mM each) for 1 h at 25 °C in a 90 mM HEPES buffer with 20 mM TCEP. Confocal microscopy of these samples, spiked with DSPE-Texas Red, showed lipidic structures, in the form of large blobs, only when both precursors were present **(Fig. 5.6a)**. This could indicate successful formation of a lipidic product; however, emulsification of precursors cannot be ruled out. When the reaction was performed in a microscopy chamber coated with BSA, liposomes were observed **(Fig. 5.6b)**. However, these could be consisting solely of DSPE-Texas Red. Besides, BSA might sequester the precursors, inhibiting the synthesis reaction.

For this chemical lipid synthesis scheme to fit in the synthetic cell framework of cell-free gene expression of liposomes, it has to be compatible with PURE system gene expression. However, PURE*frex*2.0 expression of YFP inside liposomes in the presence of MENSA oleate and o-LPC is severely inhibited **(Fig. 5.6c)**. This could be a result of unspecific detrimental reactions of o-LPC with PURE

system proteins. Addition of 20 mM TCEP attenuates the inhibitory effect, but protein yields are still significantly lower than observed without lipid synthesis precursors **(Fig. 5.6c)**. Moreover, the addition of 20 mM TCEP has a detrimental effect on PURE*frex*2.0 YFP expression in bulk in the absence of chemical lipid synthesis precursors **(Fig. 5.6d)**. This could potentially be attributed to the very acidic nature of the TCEP solution, surpassing the buffering capacity of the PURE system.

The incompatibility of PURE*frex*2.0 gene expression and chemical lipid synthesis provides a major bottleneck for this project. In Bhattacharya et al.[20] , which formulated similar project goals as here, no gene expression in the presence of chemical lipid synthesis precursors was shown. Instead, they opted for a sequential process, where FadD10 is allowed to express before addition of lipid synthesis precursors. This suggests that PURE system inhibition is not particular to our specific experimental conditions. It should be assessed if the reduced gene expression in the presence of precursors is sufficient to express catalytically active amounts of FadD. This could be examined by expressing the *fadD* gene in the presence of oleic acid, CoA, o-LPC, and TCEP, and measuring the resulting product concentration by mass spectrometry. If no lipidic product is synthesized, a major revision of the envisioned reaction scheme is necessary. Even when lipidic products could be synthesized, it is expected that the reduction in gene expression will pose major problems when integrating this process into the wider synthetic cell framework.
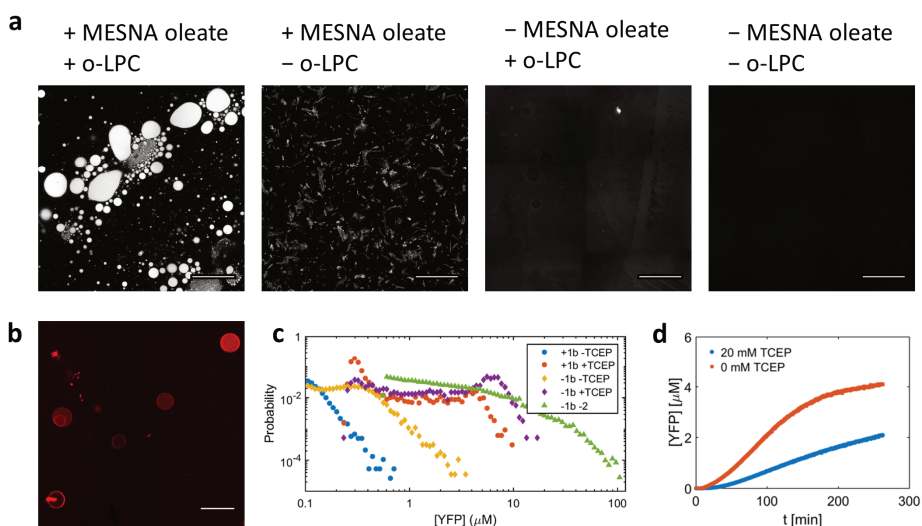
Figure 5.6: Chemical lipid synthesis results in the formation of lipidic blobs and possibly liposomes but causes inactivation of PURE system. (a) Confocal micrographs of samples after 1 h of chemical lipid synthesis, with indicated precursors. Scale bar is 50 μm. **(b)** Image of liposomes possibly resulting from *in situ de novo* lipid synthesis. Scale bar is 20 μm. **(c)** Histogram of YFP intensities in individual liposomes in the presence of MESNA oleate (2), with and without o-LPC (1b), and with and without 20 mM TCEP. Data from chapter 2 are superimposed in green. **(d)** PURE*frex*2.0 bulk YFP expression in the presence and absence of 20 mM of TCEP.

5.3 CONCLUSION & OUTLOOK

For synthetic cell growth and proliferation, a significant increase in lipid synthesis, compared to the yields reported in chapter 3 and 4, needs to be realized. In this chapter, we explored upstream pathway extension by FadD, the use of acyl-ACP as an alternative substrate, and a click-chemistry based alternative lipid synthesis route. None of the investigated methods could be straightforwardly applied to increase phospholipid synthesis to achieve 100% membrane growth in a manner that would fit within the synthetic cell framework established in chapter 1. Acyl-CoA is, regardless of the limits imposed by its limited solubility, the most effective precursor for phospholipid synthesis.

Methods to deliver higher amounts of acyl-CoA to the system that circumvent precipitation or detergent activity should be investigated. Potentially, an extra-liposomal solution with no or less Mg^{2+} or other cations could be developed. This would increase acyl-CoA solubility, allowing for a higher starting concentration. However, high initial concentrations of acyl-CoA will likely destabilize liposome membranes, affecting their permeability and structural integrity. Moreover, an ionic composition of the extravesicular solution that is very different from that of the liposome lumen will need to be carefully balanced to prevent high osmotic stress on the membrane. An approach based on a modified feeding solution furthermore excludes LUV experiments, where reactions take place in the extravesicular space, limiting the experimental feasibility of scanning a large parameter space.

Alternatively, continuous flow reactors hold promise to constantly supply small amounts of acyl-CoA. When inflow balances acyl-CoA consumption by the Kennedy pathway, it will not accumulate to high concentrations, thereby circumventing the associated problems. Developing a continuous flow reactor compatible with our synthetic cell scaffold (PURE system in GUVs produced with the natural swelling method) has value beyond increasing the synthesis of phospholipids (**Fig. 5.7**). All reactions currently reconstituted in liposomes by the group are fed by nutrients (NTPs, amino acids, creatine phosphate, specific reaction precursors) co-encapsulated in the liposomes with the PURE system[2,21–24]. Furthermore, nutrients can feed liposomes from the outside by passive diffusion across the liposome membrane (potentially facilitated by pores such as α-hemolysin or connexin-43) or, for hydrophobic compounds, partitioning into the bilayer. The resource pools both inside and outside of the liposomes are static, and slowly depleted. Evidently, with prolonged reactivity of the reactions inside the liposomes, the demand for inflow of nutrients from outside increases. Continuous flow reactors can supply a constant influx of nutrients and will also allow outflow of potentially toxic side products such as phosphate. Furthermore, it could be exploited to apply external chemical cues, like changing osmolarity.

Flow reactor systems have already been designed to improve bulk cell-free gene expression reactions. Steady-state transcription and translation in the PURE system was realized for 30 h in a microfluidic reactor by reagent exchange[25]. Such prolonged expression times facilitate the cell-free recapitulation of genetic oscillators[25,26]. Nutrient inflow has also been demonstrated to prolong compartmentalized gene expression in micro-chambers separated from a nutrient flow by a capillary[27] or a phospholipid membrane[28]. However, batch-produced liposomes have not been combined with flow reactor schemes. Often, the extravesicular solution can be considered an infinite source of resources, since in many set-ups, its volume is orders of magnitude higher than the total volume of all liposome lumina, obviating the need for its replenishment. However, this consideration will be challenged in the case of increased reactivity, or a higher volume fraction of liposomes, which would be required to create populations of a size amendable for *in vitro* evolution[29]. Liposomes produced with microfluidics are

by principle combined with flow reactors. However, microfluidic flow is often not used to create environmental changes in favor of (simpler) bulk methods[30].
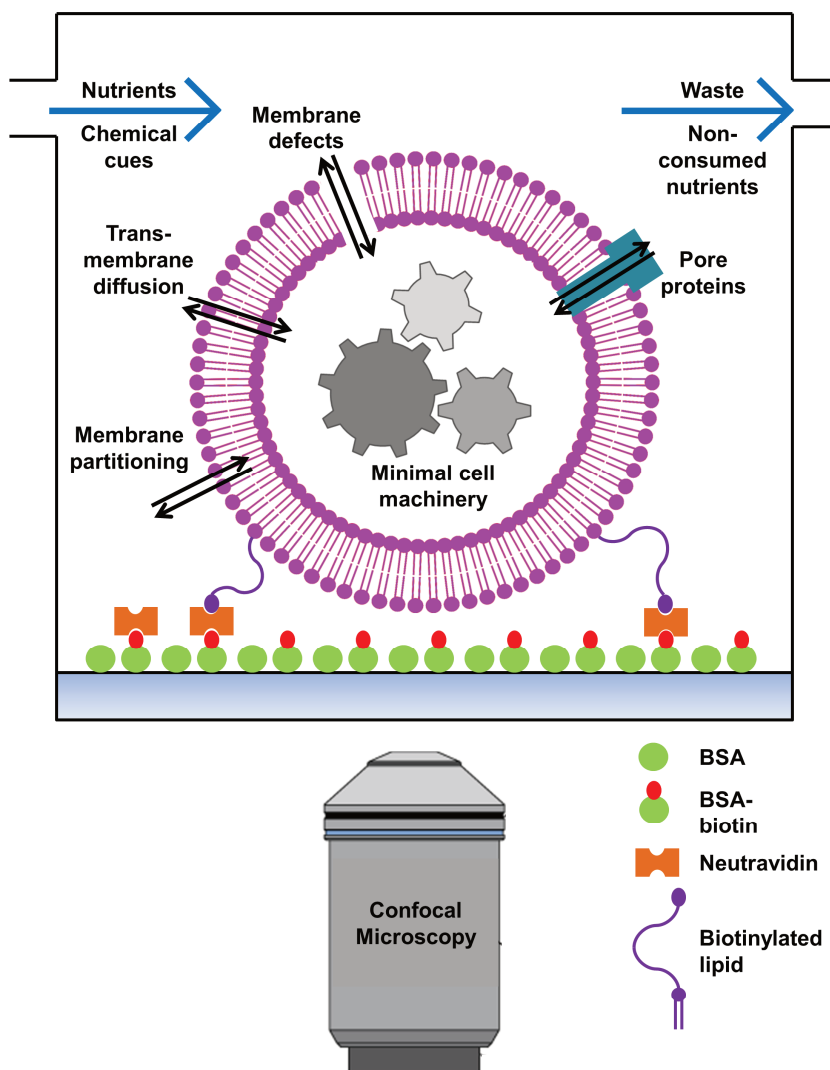
Figure 5.7: Schematic displaying batch-produced liposomes immobilized in a continuous flow reactor. Liposomes, encapsulating various machineries of interest for minimal cell research, are immobilized using biotin-neutravidin interactions in a flow chamber suitable for confocal microscopy imaging. Nutrients and/or small molecules representing chemical cues are flushed into the reactor and can cross the lipid bilayer in various ways. Waste and non-consumed nutrients are flushed out at a similar rate. Using a continuous flow reactor could be a particularly relevant strategy to circumvent the problem of acyl-CoA solubility.

5.4 METHODS

5.4.1 Preparation of *fadD* DNA

The *fadD* gene was amplified from genomic DNA of *E. coli* K12 with overhangs for Gibson assembly using primers ChD873 and ChD874. The vector backbone was amplified from pUC57 with primers ChD507 and ChD535. The resulting PCR fragments were purified with the PCR Clean-Up System kit and used for a Gibson assembly reaction, which was subsequently transformed into TOP10 *E. coli* cells. The correct assembly was verified by Sanger sequencing. Linear *fadD* DNA was regularly prepared by PCR with primers ChD73 and ChD181.

5.4.2 LUV experiments with FadD

A PURE*frex*2.0 reaction was assembled according to instructions by the manufacturer, and supplemented with 50 μM CoA, 5 mM β-mercaptoethanol, 500 μM ^{13}C-G3P, 1 mM CTP, 500 μM L-serine, 1 nM pGEMM7 DNA, 7 nM *fadD* DNA, and 0.4 mg/mL LUVs. LUVs were prepared as described in chapter 3. Oleic acid or oleoyl-CoA films were prepared in 0.2 mL tubes by evaporating the chloroform solvent at ambient temperature and pressure. The PURE*frex*2.0 mixture was subsequently added to the films, and incubated overnight at 37 °C.

Samples were prepared for mass spectrometry by diluting 100-fold in methanol with 5 mM EDTA and 2 mM acetylacetone, followed by 10 min of sonication and 5 min of centrifugation at 16,000 g. The resulting supernatant was subsequently harvested and added to Agilent vials with low-volume inset. Mass spectrometry was performed as described in chapter 3.

5.4.3 GUV experiments with FadD

Lipid-coated beads containing 50 mol% oleic acid (OA) were prepared as described in chapter 2, with a modified lipid composition of DOPC:DOPE:DOPG:CL:OA: DHPE-Texas Red/DSPE-PEG-Biotin (25 mol%: 18 mol%: 6 mol%: 1 mol%: 50 mol%: 0.5 mass%: 1 mass%). A PURE*frex*2.0 reaction was assembled according to instructions by the manufacturer, and supplemented with 50 μM CoA, 5 mM β-mercaptoethanol, 500 μM ^{13}C-G3P, 1 mM CTP, 500 μM L-serine, 1 nM pGEMM7Δ*psd* DNA, and 7 nM *fadD* DNA. Per 1 μL of swelling solution, 1 mg of lipid-coated beads was added. After 2 h of swelling on ice, with regular tumbling, four freeze-thaw cycles were performed. Then, 2 μL of supernatant was harvested and diluted in 5.5 μL of a feeding solution consisting of PURE*frex*2.0 Solution I and Milli-Q (3:7). The sample was then immobilized in an imaging chamber as described in chapter 2. Time-lapse confocal microscopy was performed with a Nikon A1R Laser scanning confocal microscope using the excitation/emission wavelengths 561/595 nm to visualize the Texas Red membrane dye. The sample height was adjusted manually in order to equatorially dissect as many liposomes as possible.

5.4.4 Acyl-ACP experiment

Palmitoyl-ACP was kindly provided by Marek Noga as a stock solution of 1.65 mg/mL in 10 mM MgSO$_4$. A PURE*frex*2.0 reaction was assembled according to instructions by the manufacturer, and supplemented with 5 mM β-mercaptoethanol, 500 μM ^{13}C-G3P, 1 mM CTP, 500 μM L-serine, 1 nM pGEMM7 DNA, 0.4 mg/mL LUVs, and 10 μM palmitoyl-ACP. Controls without palmitoyl-ACP were supplied with acyl chains from palmitoyl-CoA films, prepared as described above. After overnight incubations at 37 °C, samples were prepared for mass spectrometry analysis as described above. Mass spectrometry was performed as described in chapter 3.

5.4.5 Chemical lipid synthesis precursors

The precursors sodium 2-mercaptoethane-sulfonate (MESNA) oleate, 1-oleoyl-2-(L-Cys)-*sn*-glycero-3-phosphocholine (o-LPC), and 1-palmitoyl-2-(L-Cys)-*sn*-glycero-3-phosphocholine (p-LPC) were kindly supplied by Roberto Brea of the Devaraj Lab at UCSD in powdered form. LPC's were dissolved in methanol, MESNA oleate in Milli-Q water, and all were aliquoted and stored at −80 °C.

5.4.6 Chemical lipid synthesis reactions

Lipid synthesis with o-LPC and MESNA oleate was performed by mixing 2.5 mM of the compounds 1:1 in a 90 mM HEPES buffer with 20 mM TCEP, and incubating for 1 h at 25 C. Mass spectrometry transitions could be determined for o-LPC (522.3 $\rightarrow$ 184.0, positive ion mode) and MESNA oleate (405.1 $\rightarrow$ 297.2 and 405.1 $\rightarrow$ 141.0, negative ion mode), but not their reaction product. For visualization, 10 µg/mL DSPE-Texas Red dissolved in chloroform was added to the reaction mixtures, which were subsequently dried to a lipid cake by 30 min of vacuum concentration. Then, the film was resuspended in Milli-Q, followed by 1 h of continuous tumbling at room temperature. Microscopy chambers were coated with 2 mg/mL BSA before being filled with the liposome solution. For *in situ de novo* chemical lipid synthesis, the reaction mixture was spiked with Texas Red before incubation and directly added to the imaging chamber. Confocal microscopy was performed with a Nikon A1R Laser scanning confocal microscope using the excitation/emission wavelengths 561/595 nm to visualize the Texas Red membrane dye.

5.4.7 Combined chemical lipid synthesis and PURE system reactions

PUREfrex2.0 reaction mixture was assembled according to manufacturer's instructions, and supplemented with 7 nM of YFP DNA, 2.5 mM MESNA oleate, and 20 mM TCEP when indicated. Lipid-coated beads, prepared as described in chapter 2, were added, followed by 2 h of swelling at ice, subjected to regular tumbling. After four freeze-thaw cycles, 4.75 µL of supernatant was harvested, supplemented with 0.25 µL DNase I, and added to a dried film of o-LPC. The concentration of resuspended o-LPC was 2.5 mM. The liposome mixture was then immobilized in an imaging chamber as described in chapter 2. Confocal microscopy was performed with a Nikon A1R Laser scanning confocal microscope using the excitation/emission wavelengths 514/535 nm (YFP) and 561/595 nm (Texas Red). The sample height was adjusted manually in order to equatorially dissect as many liposomes as possible. Image analysis was performed as described in chapter 2.

REFERENCES

1. Constantinides, P. P. & Steim, J. M. Solubility of palmitoyl-Coenzyme A in acyltransferase. *Arch. Biochem. Biophys.* **250,** 267–270 (1986).

2. Scott, A. et al. Cell-free phospholipid biosynthesis by gene-encoded enzymes reconstituted in liposomes. *PLoS One* **11,** e0163058 (2016).

3. Kameda, K. & Nunn, W. D. Purification and Characterization of Acyl Coenzyme A synthetase from Escherichia coli. *J. Biol. Chem*. **256,** 5702–5707 (1981).

4. Exterkate, M., Caforio, A., Stuart, M. C. A. & Driessen, A. J. M. Growing membranes in vitro by continuous phospholipid biosynthesis from free fatty acids. *ACS Synth. Biol.* **7,** 153–165 (2018).

5. Chan, D. I. & Vogel, H. J. Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochem. J.* **430,** 1–19 (2010).

6. Ray, T. K. & Cronan, J. E. Acylation of sn-glycerol 3-phosphate in *Escherichia coli*. Study of reaction with native palmitoyl-acyl carrier protein. *J. Biol. Chem.* **250,** 8422–8427 (1975).

7. Rock, C. O., Goelz, S. E. & Cronan, J. E. Phospholipid synthesis in Escherichia coli. Characteristics of fatty acid transfer from acyl-acyl carrier protein to sn-glycerol 3-phosphate. *J. Biol. Chem.* **256,** 736–742 (1981).

8. Coleman, J. Characterization of the Escherichia coli gene for 1-acyl-sn-glycerol-3-phosphate acyltransferase (plsC). *MGG Mol. Gen. Genet.* **232,** 295–303 (1992).

9. Rottig, A. & Steinbuchel, A. Acyltransferases in Bacteria. *Microbiol. Mol. Biol. Rev.* **77,** 277–321 (2013).

10. Beld, J., Lee, D. J. & Burkart, M. D. Fatty acid biosynthesis revisited: structure elucidation and metabolic engineering. *Mol. BioSyst.* **11,** 38–59 (2015).

11. Yu, X., Liu, T., Zhu, F. & Khosla, C. In vitro reconstitution and steady-state analysis of the fatty acid synthase from Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 18643–18648 (2011).

12. Xiao, X., Yu, X. & Khosla, C. Metabolic flux between unsaturated and saturated fatty acids is controlled by the FabA:FabB ratio in the fully reconstituted fatty acid biosynthetic pathway of *Escherichia coli*. *Biochemistry* **52,** 8304–8312 (2013).

13. Murtas, G. Internal lipid synthesis and vesicle growth as a step toward self-reproduction of the minimal cell. *Syst. Synth. Biol.* **4,** 85–93 (2010).

14. Hofmann, J. & Schweizer, E. Microbial type I fatty acid synthases (FAS): Major players in a network of cellular FAS systems. *Microbiol Mol Biol Rev.* **68,** 501–517 (2004).

15. Brea, R. J., Cole, C. M. & Devaraj, N. K. In situ vesicle formation by native chemical ligation. *Angew. Chemie - Int. Ed.* **53,** 14102–14105 (2014).

16. Hardy, M. D. et al. Self-reproducing catalyst drives repeated phospholipid synthesis and membrane growth. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 8187–8192 (2015).

5

17. Brea, R. J., Rudd, A. K. & Devaraj, N. K. Nonenzymatic biomimetic remodeling of phospholipids in synthetic liposomes. *Proc. Natl. Acad. Sci. U.S.A.* **113,** 201605541 (2016).

18. Bhattacharya, A., Brea, R. J. & Devaraj, N. K. De novo vesicle formation and growth: An integrative approach to artificial cells. *Chem. Sci*. **8,** 7912–7922 (2017).

19. Devaraj, N. K. In Situ Synthesis of Phospholipid Membranes. *J. Org. Chem.* **82,** 5997–6005 (2017).

20. Bhattacharya, A., Brea, R. J., Niederholtmeyer, H. & Devaraj, N. K. A minimal biochemical route towards de novo formation of synthetic phospholipid membranes. Nat. Commun. 10, 300 (2019).

21. Nourian, Z., Roelofsen, W. & Danelon, C. Triggered gene expression in fed-vesicle microreactors with a multifunctional membrane. *Angew. Chemie - Int. Ed.* **51,** 3114–3118 (2012).

22. Van Nies, P., et al. Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. Nat. Commun. **9,** 1583 (2018).

23. Godino, E. et al. De novo synthesized Min proteins drive oscillatory liposome deformation and regulate FtsA-FtsZ cytoskeletal patterns cytoskeletal patterns. *Nat. Commun.* **10,** 4969 (2019)

24. Godino, E. et al. Cell-free biogenesis of bacterial division proto-rings that can constrict liposomes. *bioRxiv* 2020.03.29.009639 (2020) doi:10.1101/2020.03.29.009639.

25. Niederholtmeyer, H., Stepanoèa, È. & Maerkl, S. J. Implementation of cell-free biological networks at steady state. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 15985–15990 (2013).

26. Yelleswarapu, M. et al. Sigma factor-mediated tuning of bacterial cell-free synthetic genetic oscillators. *ACS Synth. Biol.* **7,** 2879–2887 (2018).

27. Karzbrun, E., Tayar, A. M., Noireaux, V. & Bar-Ziv, R. H. Programmable on-chip DNA compartments as artificial cells. *Science* **345,** 829–32 (2014).

28. Izri, Z., Garenne, D., Noireaux, V. & Maeda, Y. T. Gene Expression in on-chip membrane-bound artificial cells. *ACS Synth. Biol.* **8,** 1705–1712 (2019)

29. Abil, Z. & Danelon, C. Roadmap to building a cell: An evolutionary approach. *under review* (2020).

30. Deshpande, S. et al. Spatiotemporal control of coacervate formation within liposomes. *Nat. Commun.* **10,** 1800 (2019).

Chapter 6: Thermal deformation and division of gene-expressing liposomes

Phase transition in phospholipid membranes is a well-studied phenomenon in biophysics with applications in drug delivery and vesicle bioreactors. Crossing the phase transition temperature, the molecular surface area of the membrane drastically increases, which for closed vesicles leads to a sudden change in surface-to-volume ratio, and thereby liposome morphology. This phenomenon has previously been exploited to divide liposomes suspended in aqueous solution, triggering fission by the incorporation of inverse cone-shaped lipids. Here, we try to recapitulate this process in the context of PURE system gene-expressing liposomes. To form liposomes with an experimentally attainable phase transition temperature required modifications of our previously established protocol and resulted in liposome populations with many aggregates and multilamellar vesicles. Post-swelling encapsulation of DNA was needed to prevent expression during formation of liposomes. Liposomes expressing the gfp *gene elongate upon crossing the phase transition temperature and stay elongated and impermeable for several hours. Elongated gene-expressing liposomes could serve as a scaffold to reconstitute cellular processes that require the rod-like shape of many bacteria. Moreover, when subject to temperature cycles, gene-expressing liposomes could elongate and relax repeatedly in a reversible manner. In the PURE system, morphology changes occurred on the minute timescale, in contrast to transitions occurring on the second timescale reported previously in aqueous solution. Introduction of inverse cone-shaped lipids did not lead to fission of gene expressing liposomes, potentially as a result of slow deformation kinetics and poor liposome quality.*

6.1 INTRODUCTION

In synthetic cell research, including previous chapters in this thesis, liposomes are mostly regarded as static containers for biochemical processes. However, liposomes are dynamic assemblies capable of various shape transformations, triggered by environmental factors[1–6]. In this chapter, we will subject gene-expressing liposomes to temperature cycles to initiate deformation and division. Deformation to more elongated shapes breaks the spherical symmetry of liposome assemblies, bringing their morphology more in line with rod-shaped bacteria such as *Escherichia coli.* Division is the most important and dramatic structural rearrangement in the cell cycle and reconstituting it would be a main milestone in the synthetic cell project. Temperature-driven shape transitions have the potential to obviate the need for protein complexes that provide elongation (e.g. microtubules[7]) or realize division (divisome proteins, e.g. FtsZ[8–10], ESCRT[11]), resulting in further minimization of the synthetic cell design.

Phospholipid membranes exist in two distinct phases: the gel, or liquid-ordered (L_o), phase, and the fluid, or liquid-disordered (L_d), phase. Transition between the two phases occurs at the transition- or melting temperature T_m. The melting temperature of a membrane is mostly governed by phospholipid chain length and degree of saturation. Long chains result in many Van der Waals interactions between lipids and consequently high melting temperatures. Unsaturations (double bonds) introduce 'kinks' in the acyl chain, spacing the phospholipid molecules farther apart, thereby decreasing the Van der Waals interactions and consequently the melting temperature[12]. Since melting temperatures of membranes composed of liposome-forming lipids can be found within the biologically relevant temperature range, phospholipid membrane phase transition is a phenomenon of interest to synthetic cell research.
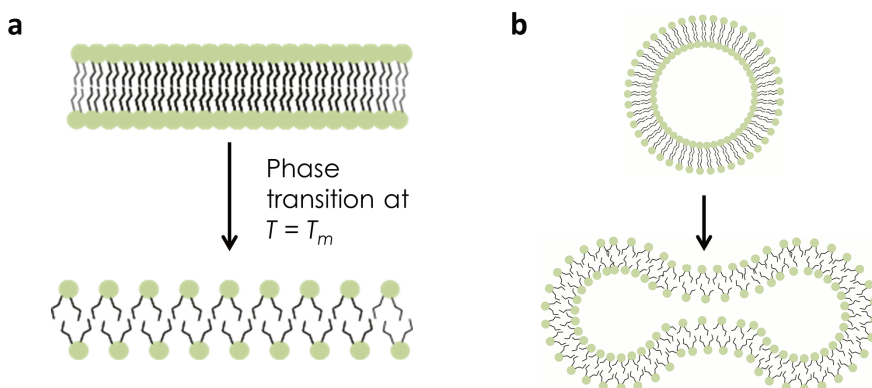

Figure 6.1: Crossing the phase transition temperature increases the molecular surface area of membrane lipids. **(a)** In the fluid phase (bottom), lipids are spaced further apart, the bilayer is thinner, resulting in an increase in membrane area. **(b)** Membrane area increase with constant volume will result in a change in *S/V* and therefore in liposome morphology.

The molecular surface area of a phospholipid is significantly higher (~30%, depending on lipid species) in the fluid phase than it is in the gel phase[13]. Consequently, a liposome that undergoes a transition from gel to fluid phase will have an increased membrane surface area (**Fig. 6.1a**). This results in an increase of surface-to-volume ratio (S/V), provided that surface increase occurs quicker than volume equilibration (**Fig. 6.1b**). Increased S/V will result in a shape transition of a spherical vesicle to a more energetically favorable state, such as an elongated prolate shape. Assuming constant volume, the following relationship between area increase factor α and liposome elongation λ can be derived (for a derivation, see supplementary information):

$$
\begin{cases}
\lambda > 1 \text{ (prolate)} \quad \alpha^{-\frac{3}{2}} = \dfrac{2\sqrt{2}\,\lambda}{\left(1+\frac{\lambda^2}{\sqrt{\lambda^2-1}}\sin^{-1}\left(\sqrt{\frac{\lambda^2-1}{\lambda^2}}\right)\right)^{\frac{3}{2}}} \\[3em]
\lambda < 1 \text{ (oblate)} \quad \alpha^{-\frac{3}{2}} = \dfrac{2\sqrt{2}\,\lambda}{\left(1-\frac{\lambda^2}{\sqrt{1-\lambda^2}}\tanh^{-1}(1-\lambda^2)\right)^{\frac{3}{2}}}
\end{cases}
\qquad \text{(Eq. 1)}
$$

Here, the elongation λ is defined as the ratio between the length of the major axis and the minor axis. According to this relationship, plotted in **Fig. 6.2**, a 30% area increase resulting from phase transition can cause a shape transition to a prolate with an elongation of 4.23. Hence, applying a small temperature difference to a liposome can result in a drastic morphological transition. Elongated liposomes mimic the rod-shape of the model organism *E. coli*, facilitating reconstitution of processes such as Min-system waves that require a non-spherically symmetric geometry[9], and are therefore of interest in minimal cell research.
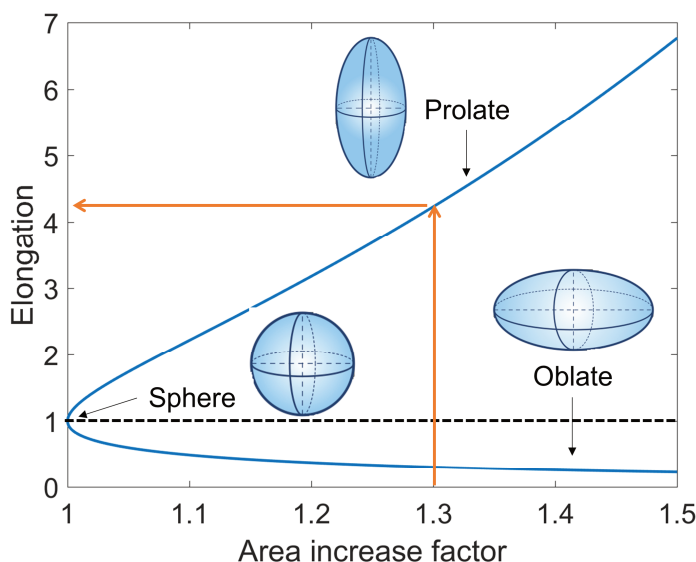
Figure 6.2: Graphical representation of the relationships in equation 1. Spheroids with an elongation larger than one (above dashed line) are prolate. Elongations lower than one (below dashed line) correspond to oblate spheroids. Spheres have an elongation of one. Superimposed in red arrows, a typical area increase factor of 30% corresponds to a prolate elongation of 4.23. Spheroid schematics adapted from https://py.checkio.org/en/mission/humpty-dumpty/.

Work by the group of Imai has demonstrated that, when a fraction of inverse-cone shaped lipids is added to the membrane, liposomes undergoing phase transition can not only deform but also divide[14–16]. Inverse-cone shaped lipids, such as dilauroylphosphatidylethanolamine (DLPE, **Fig. 6.3a**), prefer to partition into the inner leaflet of a spherical liposome due to their geometry that stabilizes positive curvature. This asymmetric lipid distribution causes dramatic shape transformations upon phase transition, including the formation of a neck, where DLPE accumulates. Inverse-cone lipids partition into the neck through a mechanism called curvature-mediated sorting, contracting and eventually destabilizing it, resulting in vesicle fission (**Fig. 6.3b**)[17]. This mechanism, mediated solely by lipid composition and temperature, could be used to realize synthetic cell division without divisome proteins.
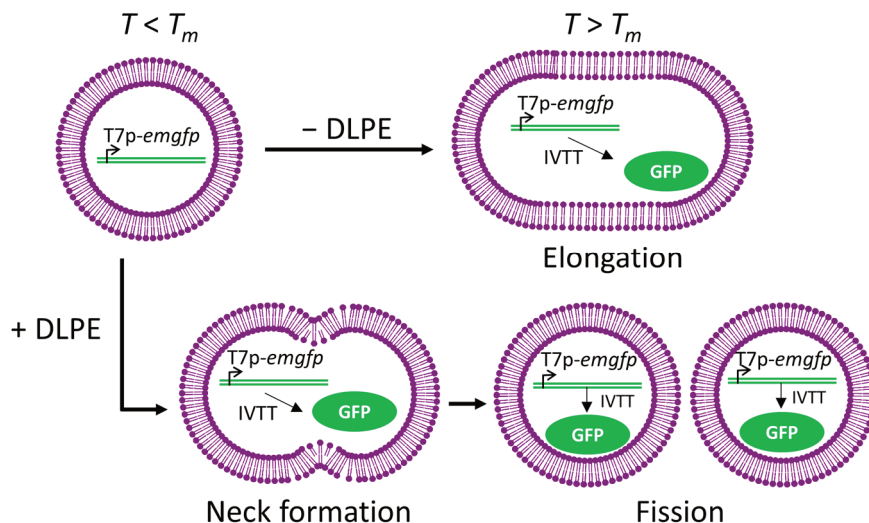
Temperature-driven liposome deformation and division has only been demonstrated in water or simple buffers[14,16]. In this chapter, we will attempt to reconstitute these processes within our minimal cell framework of gene expressing liposomes (**Fig. 6.4**). This will require several modifications of the liposome preparation protocol described in previous chapters. Firstly, a lipid composition with a melting temperature higher than the freezing point of water but low enough to prevent protein denaturation is needed. Therefore, we will use lipids with myristoylic acyl chains ($T_m \approx 25$ °C) instead of oleic acid chains ($T_m \approx -20$ °C) used elsewhere in this thesis. Since lipid film swelling necessitates the membranes to be in the fluid phase, a swelling temperature >25 °C is needed. At this elevated temperature, PURE system gene expression will commence before vesicles are fully formed[18]. Therefore, we explored various mechanisms to prevent gene expression during swelling and trigger it afterwards. Encapsulating DNA post-swelling using freeze-thaw cycles was found to be the most successful approach. We observed that with myristoyl lipids the yield of unilamellar liposomes is much lower than with oleoyl lipids, posing a significant experimental challenge.

Crossing the melting temperature resulted in the elongation of myristoyl lipid vesicles encapsulating expressed green fluorescent protein (GFP), as observed with laser scanning and spinning disc confocal microscopy. Liposomes maintained their elongated shape for several hours, while retaining GFP in the lumen. Temperature cycling across the melting temperature reversibly elongated liposomes containing GFP, i.e. the liposomes reverted back to spherical shape upon cooling. Selected liposomes were subjected to 3D image analysis, demonstrating that both surface and volume increased upon elongation, but returned to their initial values upon cooling. It was observed that shape transformations occurred on the minute time scale, compared to transformations on the second timescale observed by Sakuma and Imai[14]. We attribute this to the complex ionic composition of the PURE system interacting with (charged) lipid membranes.

Incorporation of DLPE to the lipid film before swelling yielded liposomes that underwent elongation but not division. Liposomes, fused after swelling with DLPE-containing small unilamellar vesicles (SUVs), underwent more varied deformations upon phase transition. However, division events as described in Jimbo et al.[16] were not observed.
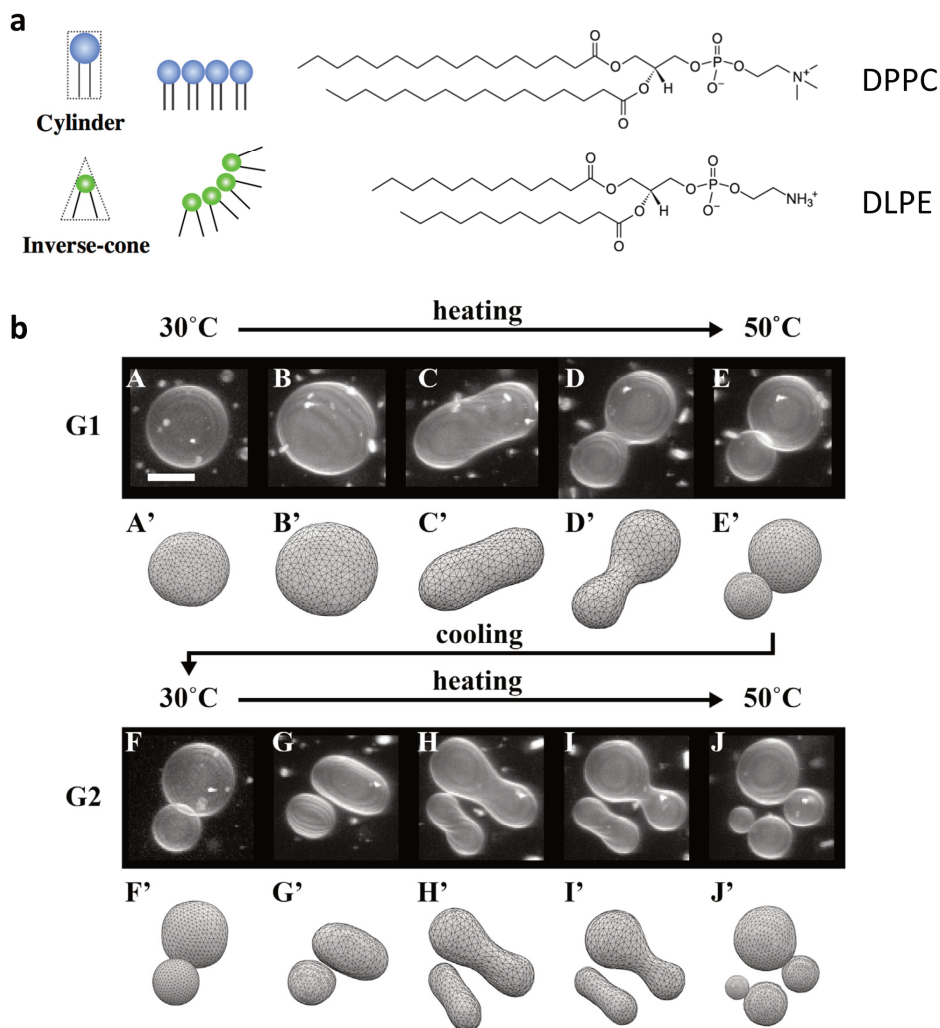
Figure 6.3: Liposomes containing a fraction of inverse-cone lipids can divide as a result of temperature cycling. **(a)** DPPC is a cylinder-shaped lipid. The size of its tails is offset by its big choline headgroup. DLPE is an inverse-cone shaped lipid. Although it has rather small tails, its minute ethanolamine headgroup gives rise to an inverse-cone shape. Lipid shape schematic taken from Sakuma and Imai[15] and lipid structures are taken from www.avantilipids.com. **(b)** Two generations of liposomes (DPPC/DLPE 8:2) deforming and dividing when temperature cycles are applied. Scale bar is 10 μm. Taken from Jimbo et al.[16]

Figure 6.4: Temperature-driven liposome deformation and division combined with PURE system gene expression. Without inverse-cone lipids like DLPE, liposomes are expected to elongate when crossing the melting temperature T_m. Moreover, the temperature increase will trigger expression of the encapsulated *emgfp* gene, resulting in synthesis of GFP. When DLPE is included, this will trigger the formation of a neck, followed by liposome fission.

6.2 RESULTS

6.2.1 Encapsulating DNA after swelling by freeze-thaw cycles triggers gene expression in myristoyl liposomes

Liposomes with myristoyl lipids (79 mol% 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC), 18 mol% 1,2-dimyristoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (DMPG), and 3 mol% 1',3'-bis[1,2-dimyristoyl-sn-glycero-3-phospho]-sn-glycerol (C14:0-cardiolipin, C14:0-CL)) are prepared using the glass beads methods, applying a swelling temperature of 37 °C. Addition to the swelling solution (PURE*frex* or PURExpress) of DNA encoding for the enhanced monomeric green fluorescent protein (emGFP) whose expression is under control of a T7 promoter (T7p-*emgfp*) results in protein expression prior to and during liposome formation. Consequently, emGFP is observed both in the liposome lumen and in the extravesicular solution (**Fig. 6.5a**). The lower fluorescence intensity in the extravesicular solution is a result of dilution in feeding solution before imaging.

Several approaches have been investigated to trigger gene expression after liposome formation, ensuring protein expression only takes place inside the liposome. Firstly, DNA was encapsulated in the liposomes using freeze-thaw cycles after swelling. This results in appreciable GFP expression localized inside the lumen of liposomes, although the fluorescent signal is significantly lower than when DNA is added before swelling (**Fig. 6.5b**). Higher GFP fluorescence was observed when using PURExpress than PURE*frex*, in agreement with the higher gene expression yield of PURExpress observed previously in chapter 2.
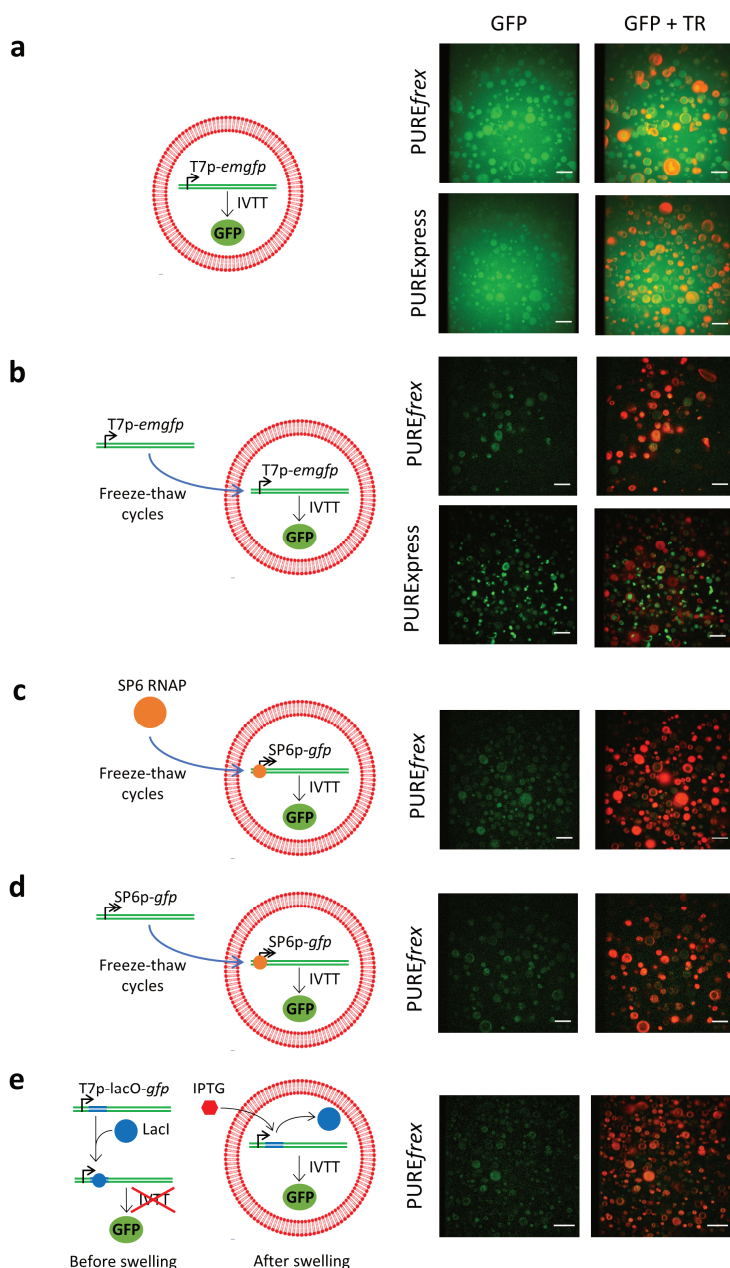
Figure 6.5: Various approaches to trigger gene expression in myristoyl liposomes after vesicle formation. (a) Addition of 7 nM T7p-*emgfp* to a swelling solution consisting of either PURE*frex* or PURExpress results in GFP signal (green) inside and outside the lumen of liposomes. Liposome membranes are stained with Texas Red (TR, red). **(b)** Encapsulation of 7 nM T7p-*emgfp* after formation of liposomes, using a swelling solution of either PURE*frex* or PURExpress, with six freeze-thaw cycles, results in GFP expression localized inside liposome lumina. **(c)** Encapsulation of 2 U/µL SP6 RNAP with six freeze-thaw cycles into liposomes formed with a swelling solution containing PURE*frex* and 7 nM SP6p-*gfp* DNA results in GFP signal in liposome lumina that is not clearly

distinguishable from crosstalk of the TR staining of the membranes. **(d)** Encapsulation of 7 nM of SP6p-*gfp* DNA with six freeze-thaw cycles into liposomes formed with a swelling solution of PURE*frex* and 2 U/µL SP6 RNAP gives results similar to panel (c). **(e)** After preparation of liposomes with PURE*frex*, 1 nM of T7p-lacO-*gfp* DNA and 1 µM of LacI, triggering gene expression by the addition of 200 µM IPTG before applying four freeze-thaw cycles does not result in appreciable GFP expression. All micrographs were obtained using spinning disk confocal microscopy with 50 ms exposure time and a gain of 200 a.u.. Scale bars represent 10 µm.

Secondly, a DNA encoding for GFP under an SP6 promoter (SP6p-*gfp*) was added to the swelling solution. Since the associated RNA polymerase (SP6 RNAP) is not present in the PURE system, gene expression could be initiated by encapsulating the SP6 RNAP with freeze-thaw cycles after swelling. Some GFP fluorescence could be observed; however, it is hardly distinguishable from the Texas Red membrane dye crosstalk signal (**Fig. 6.5c**).

Co-encapsulating the SP6 RNAP during swelling and adding the SP6p-*gfp* DNA with freeze-thaw cycles yielded similar results (**Fig. 6.5d**). Finally, we used the LacI protein to repress gene expression during swelling from a DNA with the Lac operon (lacO) between the T7 promoter and the GFP gene. After liposomes were formed, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to release the LacI repression and trigger gene expression. This approach did not result in GFP fluorescence in the liposome lumina that significantly exceeded the Texas Red crosstalk signal (**Fig. 6.5e**). Concluding, post-swelling encapsulation of T7p-*emgfp* DNA by freeze-thaw cycles is the most effective approach to confine gene expression to the lumen of myristoyl liposomes. Therefore, this strategy will be applied in the remainder of this chapter, with PURExpress as gene expression system, since it was found to yield higher GFP fluorescence than PURE*frex*.

For all conditions, a high fraction of multilamellar vesicles is observed, and few unilamellar liposomes are imaged per field of view. Liposome quality is significantly lower compared to oleoyl liposomes investigated in previous chapters. Attempts to remove multilamellar structures by centrifugation were not successful (data not shown).

6.2.2 Crossing the bilayer melting temperature leads to stably elongated impermeable liposomes

GFP expression in myristoyl liposomes containing PURExpress was triggered by addition of T7p-*emgfp* DNA with freeze-thaw cycles. After overnight incubation at 37 °C, which is above the phase transition temperature for these liposomes, laser scanning confocal microscopy was used to observe elongated liposomes containing GFP (**Fig. 6.6a**). This observation demonstrates that liposomes elongate when being heated from ambient temperature to 37 °C and they can maintain this elongation for 16 h. The fact that liposomes do not revert back to a spherical shape, and that gene expression can take place in the lumen, implies that they remain impermeable during and after phase transition.

To further test this hypothesis, we performed time lapse microscopy on a myristoyl liposome containing PURExpress but no DNA (**Fig. 6.6b**). The small fluorescent molecule Cy5 (~790 Da) was added to the feeding solution as a permeability probe. Elongation was triggered by increasing the temperature to 37 °C. Elongation was observed to slowly decrease during imaging, from a major versus minor axis ratio of >4 to ~2 (**Fig. 6.6c**). Integrated Cy5 fluorescence intensity was found to be similar in the liposome lumen and in the extravesicular solution, suggesting a permeabilizing effect occurring before the start of imaging, most likely when crossing the phase transition temperature. During imaging, Cy5 fluorescence intensity inside the liposome was observed to decrease faster than the Cy5 fluorescence intensity of the extravesicular solution (**Fig. 6.6d**); as described in chapter 2, this is a strong indicator of impermeability to Cy5.

Summarizing, liposomes elongated by crossing the phase transition temperature maintain an elongated shape for multiple hours while being impermeable to small molecules such as Cy5. A permeabilizing event when crossing the phase transition temperature allows for inflow of Cy5 but does not prevent gene expression. Temperature-driven liposome elongation can therefore be used as a tool to study reconstituted systems in a non-spherically symmetric environment.

6.2.3 Temperature cycling leads to repeated reversible elongation and contraction cycles of myristoyl liposomes

To further investigate the shape deformations that myristoyl liposomes undergo when crossing the phase transition temperature of their membranes, we applied heating and cooling cycles to a population of liposomes (see **Supplementary Fig. 1**). Increasing the temperature from 20 °C to 30 °C, thereby crossing the phase transition temperature, we observed liposome elongation, as before. Subsequent cooling from 30 °C to 20 °C resulted in another deformation, where the liposomes reverted back to their initial spherical shape (**Fig. 6.7a**). Analysis of the mean elongation of the imaged population of liposomes (see **Supplementary Fig. 2** for three independent repeats) revealed distinct features of the elongation/contraction process (**Fig. 6.7b**). A lag time of 2 min is observed between the start of the temperature ramp and the increase/decrease in elongation. Moreover, the liposome population requires 30-60 s to transition from one shape into the other. Approximately half of the observed liposomes undergoes elongation and contraction (**Fig. 6.7c**). After cooling, the mean elongation approaches the initial value, indicating a reversible process (**Fig 6.7b**). Indeed, ~90% of liposomes are spherical again at the end of the temperature cycle (**Fig. 6.7c**).

The reversibility of the deformation cycle encouraged us to apply repeated cycles. At least five cycles could be applied, which resulted in elongation and subsequent contraction (**Fig. 6.7d**). Interestingly, the average population elongation increased from ~1.2 to 2 in every cycle, indicating that repeated cycles do not influence the properties of deformation. Overlaying elongation data from another population of liposomes demonstrates very similar characteristics, suggesting the observed behavior is general across experiments.

Exploiting the very fast acquisition rate of spinning disk confocal microscopy, we acquired time-lapsed z-stacks of individual liposomes undergoing a deformation cycle. The 3D manager plugin of FIJI[19] was used to reconstruct three-dimensional images at various time points (**Fig. 6.8a**) and extract (normalized) liposome surface area (**Fig. 6.8b**) and volume (**Fig. 6.8c**). In all cases, the surface area was found to increase over a period of >2 min, which is longer than the elongation time of 30-60 s observed from population measurements. Surprisingly, in cases 1 and 3 (**Fig. 6.8b**), a surface area increase significantly higher than expected based on the molecular surface area increase of main membrane constituent DMPC (20%)[20] was observed. Moreover, in these cases the measured volume also increased and decreased during the cycle. However, this volume change was not sufficient to maintain a static S/V (**Fig. 6.8d**); therefore, elongation still occurs. Due to low image quality, the hypothesis that the high surface area increase and unexpected volume increase are imaging artefacts cannot be excluded.
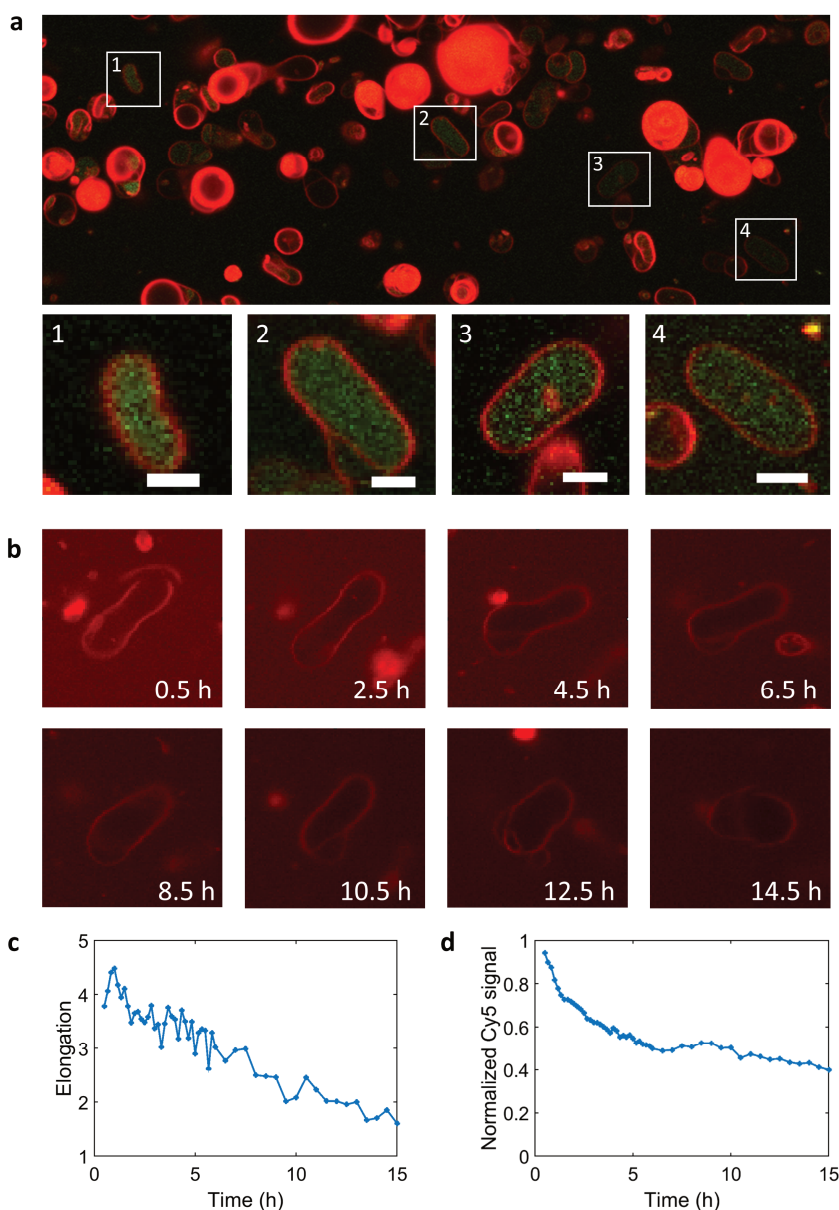
Figure 6.6: Myristoyl liposomes expressing GFP stably elongate upon melting temperature transition. (a) A population of liposomes encapsulating PURExpress and with T7p-*emgfp* DNA added during freeze-thaw cycles applied after swelling, was imaged after 16.5 h of incubation at 37 °C. Numbered boxes correspond to zoom-in images showing examples of elongated liposomes with expressed GFP fluorescence. Scale bars are 3 μm. **(b)** Time lapse microscopy of a myristoyl liposome prepared with PURExpress but without any DNA. Molecular dye Cy5 was added to the outside solution as a permeability probe. **(c)** Elongation of the liposome shown in (b), at various time points. **(d)** Cy5 intensity over time in the lumen of the liposome shown in (b), normalized to the extravesicular Cy5 intensity in the corresponding micrograph.
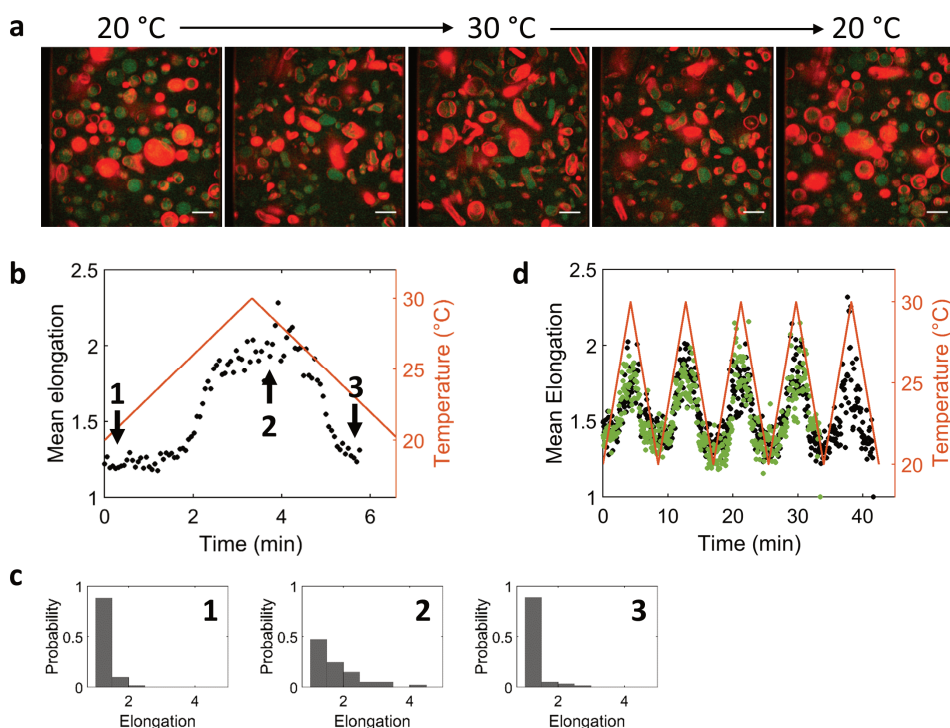
Fig 6.7: Temperature cycling reversibly elongates myristoyl liposomes. (a) Spinning disk confocal micrographs of a population of myristoyl liposomes containing PURExpress and T7p-*emgfp* DNA added during swelling, subject to a temperature cycle from 20 °C to 30 °C and back. Scale bar is 10 μm. **(b)** Mean elongation of the population of liposomes shown in (a), over time. The applied temperature ramp is overlaid in orange. The number of liposomes in the field of view varied during acquisition, with $26 \leq N \leq 68$ **(c)** Histograms of elongation for the population of liposomes shown in (a), at the time points indicated by the arrows in (b). Histogram 1: $N = 61$. Histogram 2: $N = 40$. Histogram 3: $N = 55$. **(d)** Mean elongation for two populations of liposomes (black: $7 \leq N \leq 57$, green: $7 \leq N \leq 45$) subject to repeated temperature cycles. The applied temperature ramp is overlaid in orange.

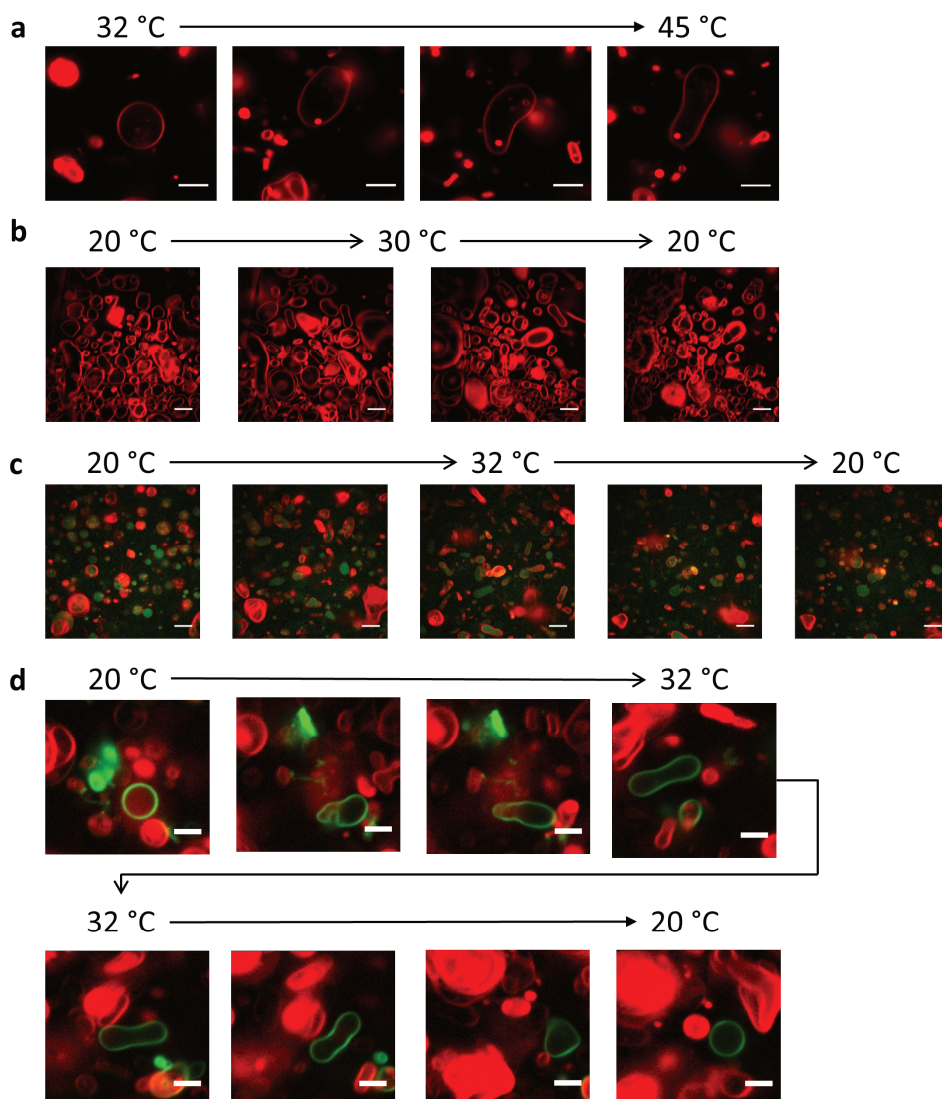## 6.2.4 Incorporation of inverse-cone shaped lipids does not result in division upon temperature cycling

Next, we sought to realize temperature-driven liposome division by incorporating the inverse-cone shaped lipid DLPE into the liposome membrane. DLPE has acyl chains of only twelve carbons, but its very small ethanolamine head group gives it the shape of an inverse cone, with corresponding morphological properties that can trigger division. We first attempted to replicate the results from Sakuma and Imai[14], by preparing liposomes containing dipalmitoylphosphatidylcholine (DPPC) and DLPE (80:20) liposomes in Milli-Q by traditional lipid film swelling, which, in their work, divided upon melting temperature ($T_m = 41$ °C) crossing. In our hands, elongation but no division was observed (**Fig. 6.9a**). However, imaging was complicated by liposome movement due to convective flows induced by the large temperature gradient in our set-up. Using DMPC instead of DPPC allowed to cycle around a lower temperature. Also here, elongation but not division was observed (**Fig. 6.9b**). In both cases, we observed that shape transformations occurred on the second time scale, much faster than in deformation assays in a PURE system background.

Subsequently, liposomes were prepared from a myristoyl lipid film including 10% DLPE, using PURExpress with T7p-*emgfp* DNA as a swelling solution. Yet again, these liposomes underwent elongation but not division upon cycling around the melting temperature (**Fig. 6.9c**). Hypothesizing that DLPE did not incorporate in the liposome membrane upon swelling, post-swelling DLPE inclusion was attempted. To this end, 30 nm DLPE SUVs were prepared and fused with liposomes after swelling by applying six freeze-thaw cycles. Successful membrane incorporation of DLPE was confirmed by colocalization of Texas Red and Oregon Green, which was added as a membrane dye to the SUVs. Increasing the temperature from 20 °C to 32 °C resulted again only in elongation; however, some intermediate structures that appeared to have a neck could be observed. This neck formation did nevertheless not result in a division event (**Fig. 6.9d**).
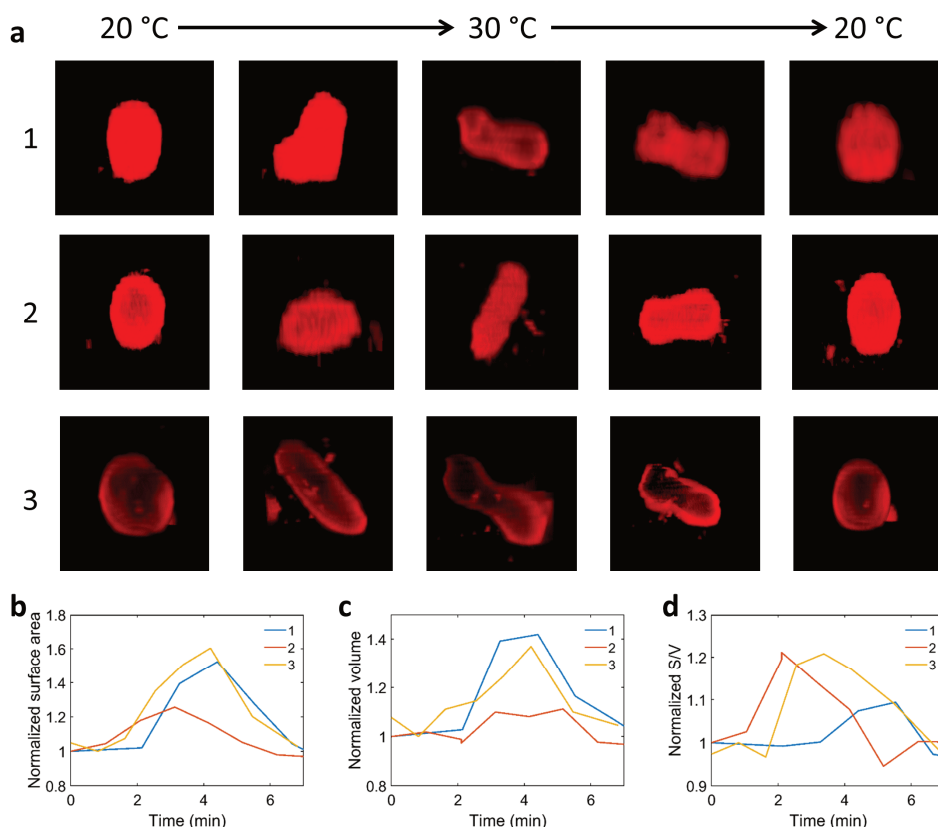
Figure 6.8: Three-dimensional reconstitution of liposome deformation reveals high surface area and volume increase. (a) Three-dimensional reconstruction of the temperature-driven deformation of three individual myristoyl liposomes. Liposome surface area **(b)** and volume **(c)**, normalized to the lowest value attained during the cycle, for the three liposomes displayed in (a). **(d)** Surface-to-volume ratio (*S/V*) for the three liposomes shown in (a). *S/V* was calculated based on normalized values displayed in (b) and (c).

Figure 6.9: Liposomes incorporating inverse-cone shaped lipids undergo deformation but not division. (a) A liposome with a lipid composition of DPPC:DLPE (80:20), prepared using Milli-Q as a swelling solution, undergoing a temperature increase, deforms but does not divide. **(b)** A population of liposomes with a lipid composition of DMPC:DLPE (80:20), prepared using Milli-Q as a swelling solution, undergoing a temperature cycle. As before with myristoyl liposomes, elongation occurs upon crossing the T_m. However, no division events are observed. **(c)** A population of myristoyl liposomes including 10% DLPE, encapsulating PURExpress and T7p-*emgfp* (added before swelling), subjected to a temperature cycle. As before, elongation and subsequent contraction can be observed. However, no division events are observed. **(d)** A myristoyl liposome containing PURExpress but no DNA, undergoing a temperature cycle. Small unilamellar vesicles containing DLPE and Oregon Green (green signal in micrographs) were fused to the myristoyl liposomes by applying freeze-thaw cycles. Although some intermediate states display a neck, no division is observed. Scale bars are 5 µm.

6.3 DISCUSSION

In the attempt to equip a synthetic cell with a very minimal division mechanism, we attempted to reconstitute temperature-driven division in gene expressing liposomes. Liposomes encapsulating a functional cell-free gene expression machinery, as demonstrated by expression of GFP, were prepared with myristoyl phospholipids with a melting temperature of ~25 °C. The spherical symmetry of liposomes was broken through elongation upon crossing the melting temperature. However, no division events were observed, even when including a significant fraction of inverse-cone shaped lipids that are known to promote temperature-driven vesicle fission[14,16].

Several factors that could potentially have inhibited liposome division can be identified. First, the quality of myristoyl liposomes was significantly lower than when preparing oleoyl liposomes with the same method. This is exemplified by the large amount of multilamellar structures present. We focused our investigations on liposomes that appeared unilamellar upon visual inspection; however, it cannot be excluded that these vesicles are in fact multilamellar, impacting their morphological response to phase transition. The reason liposome quality is low is as of yet unknown.

Secondly, during a large part of the temperature cycles applied with myristoyl liposomes, the inverse-cone lipid DLPE, that has a melting temperature of 29 °C, is in the gel phase. This might inhibit the curvature mediated sorting that is necessary to contract and stabilize the neck, which is the intermediate state towards division[16]. This would however not explain why no division was observed in DPPC:DLPE control experiments, where DLPE was in the fluid phase for the complete cycle.

Finally, it was observed that shape deformation kinetics are much slower when the transition takes place in the PURE system, compared to in Milli-Q water. This effect is likely caused by interactions between cations in the PURE system and negatively charged lipids in the liposome membrane. Moreover, the higher viscosity could impact deformation kinetics[21]. These slow kinetics might give the vesicle time to accommodate the vesicle volume to maintain a more stable surface-to-volume ratio. Three-dimensional shape analysis of individual vesicles indeed suggested that surface area increase is accompanied by a volume increase. Furthermore, slower deformation kinetics will allow for lipid redistribution between the inner and outer leaflet, provided the energy penalty of the misbalance is large enough to overcome the energy barrier of lipid redistribution; a misbalance between the two leaflets is identified as the driving force of vesicle fission[16]. Therefore, temperature driven division and high ionic solution might be functionally incompatible with the tested lipids.

Based upon the notion of synthetic cell autonomy, as elaborated in the introduction of this thesis, one could posit that temperature-driven division is no valid mechanism for a living system to proliferate. The idea of self-replication is compromised by the fact that division is triggered by the environment/the experimenter, and not by the cell itself, for example by a timing mechanism encoded in its genome. Therefore, direct coupling between genotype and (division) phenotype cannot be realized, excluding synthetic cell evolution. Potentially, this coupling can be restored by integrating lipid synthesis in the division scheme: timely synthesis of inverse-cone lipids could serve as gene-encoded division trigger. This would however not obviate the need for temperature cycles; although these could be provided by for example day-night cycles, I cannot envisage a scenario in which their necessity does not compromise synthetic cell autonomy.

Although temperature-driven division has not been realized, the stable elongation triggered by melting temperature crossing could be of use in the building of a synthetic cell. For many division-related processes, a deviation from the spherical symmetry is required. Liposome elongation creates

symmetry breaking, resulting in morphology mimicking the rod-like shape of bacteria like *E. coli*. It would therefore be interesting to investigate Min system oscillations, Z-ring positioning, or genome organization in thermally elongated liposomes, complementing other elongation techniques, such as applying osmotic stress[4], using microfluidic traps[22], or reconstitution of protein filaments that can deform liposomes[7]. While thermally elongated liposomes as prepared here are of low quality, they can be elongated without restricting their shape to a predefined morphology and applying a constant pressure on the vesicle, as with microfluidic traps. Elongation by protein filaments will require expression of extra proteins at rather high concentration, which might be very taxing on the cell-free gene expression machinery. Applying osmotic stress must be done very carefully to prevent liposome permeabilization or rupture.

The low liposome quality can potentially be mitigated by including in the membrane a fraction of DO lipids (i.e. lipids with two oleoyl acyl chains), which have been applied in the previous chapters to create high-quality liposome samples. Since DO lipids will always be in the liquid phase in liquid water, they will not contribute to elongation. However, replacing 50% of DM lipids (i.e. lipids with two myristoyl acyl chains) with DO lipids will still result in an overall area increase factor of 10% when crossing the T_m of the DM lipid, corresponding to an elongation of 2.22, sufficient to break the spherical symmetry of the liposome. The membrane composition of such multi-component vesicles might however become laterally inhomogeneous, resulting in domain formation[4]. How this effects liposome elongation in a PURE system background will need to be studied.

6.4 MATERIALS AND METHODS

6.4.1 DNA constructs

Three linear DNA construct encoding for the enhanced monomeric green fluorescent protein (emGFP) were used in this chapter. The construct T7p-*emgfp* consisted of the 720 bp *emgfp* gene, flanked by the T7 promoter and T7 terminator, and has a total length of 1090 bp. The T7 promoter was replaced by an SP6 promoter to generate the SP6p-*gfp* construct (1043 bp). The construct T7p-lacO-*gfp* contained the *emgfp* gene, with upstream the T7 promoter and a Lac Operon (lacO), and downstream the T7 terminator, for a total length of 1090 bp. DNA constructs were regularly prepared by polymerase chain reaction from a circular construct encoding for T7p-*emgfp*, with reverse primer ChD181 (5'-CAAAAAACCCCTCAAGACCCGTTTAGAGG-3') and forward primer ChD73 (5'-GCGAAATTAATACGACTCACTATAGGGAGACC-3') for T7p-*emgfp*, ChD202 (5'-AAATTTAGGTGACACTATAGAATACAAGCTTGGGCTGCAGCTAG-3') for SP6p-*gfp*, and ChD558 (5'-CGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAATAATTTTGTTT AACTTTAA-3') for T7p-lacO-*gfp*.

6.4.2 Liposome preparation

Liposomes were prepared according to the lipid-coated glass beads method[18,23], except when noted otherwise. A mixture of 1.33 mg phospholipids dissolved in chloroform was prepared in a round-bottom flask, and supplemented with 1 mass% DHPE-Texas Red and 0.5 mass% DSPE-PEG(2000)-biotin, and mixed with 100 mM rhamnose in methanol in a 2.5:1 ratio (v/v). 0.4 g of acid-washed glass beads with a 212-300 µm diameter (Sigma) were added to the lipid mixture, which was subsequently subjected to 2 h of rotary evaporation at 20 mbar, followed by overnight desiccation. Lipid-coated glass beads could be stored under argon for two months. In all experiments, the lipid mixture consisted of 79 mol% DMPC, DMPG, C14:0-CL, except in division assays, were 80 mol% 1,2- DPPC and

20 mol% DLPE, or 71 mol% DMPC, 16 mol% DMPG, 10 mol% DLPE, and 3 mol% C14:0-CL was used. All lipids were from Avanti Polar Lipids, except DHPE-Texas Red, which was from Invitrogen.

To trigger natural swelling, 10 mg of lipid-coated beads were added to 10 µL of swelling solution, which consisted of the PURE*frex (*GeneFrontier) or PURExpress (New England Biolabs) reaction mixture, assembled according to instructions of the manufacturer, with the addition of 7 nM of DNA and 0.75 U/µL SUPERase RNAse inhibitor (Invitrogen). When indicated, the PURE reaction mixture was replaced by an equal volume of Milli-Q water. Liposomes were allowed to swell for two hours at a temperature above the phase transition temperature of the membrane (37 °C for myristoyl liposomes, 48 °C for palmitoyl liposomes), being subjected to regular tumbling. Then, six freeze-thaw cycles were performed by dipping the sample in liquid nitrogen and thawing it at the swelling temperature (so above the T_m). When encapsulating compounds by freeze-thaw cycles, those compounds were added to the liposome solution right before the first freezing step. Subsequently, 1.5 µL of the supernatant of the liposome solution was taken up, with cut tip, and added to 5.5 µL of feeding solution (4:7 vol/vol of PURE*frex* Sol I, or PURExpress A, and Milli-Q supplemented with 83 mg/L Proteinase K).

To prepare 80:20 DPPC:DLPE and DMPC:DLPE liposomes in Milli-Q, traditional lipid film swelling was performed. A 200 µL solution of 10 mg/mL of the appropriate lipids in chloroform was evaporated in a 4 mL glass vial under gentle nitrogen flow, followed by 2 h of desiccation. The lipid film was subsequently rehydrated with 500 µL of Milli-Q and kept for 2 h at a temperature above the melting temperature of the lipid mixture. Then, 10 µL of liposomes was harvested with a cut pipette tip, added to the imaging chamber, and visualized without further manipulation.

6.4.3 Image acquisition

Experiments at constant temperature were imaged using a A1R Nikon Laser Scanning Confocal microscope with an SR Apo TIRF 100× oil immersion objective. The samples were installed in the set-up in a custom-made glass chamber, which contained an extra reservoir of water to slow down sample evaporation. Imaging was performed using the laser lines 488 nm (GFP), 561 nm (Texas Red) and 640 nm (Cy5), with appropriate emission filters.

When applying temperature cycles, an Olympus iX81 inverted fluorescence microscope including a Nipkow spinning disk (CSU-XI, Yokogawa) equipped with a 100x oil immersion objective and an iXon3 EMCCD camera was used. Imaging was performed with laser lines 491 nm (GFP) and 640 nm (Texas Red), with appropriate emission filters.

6.4.4 Temperature cycling and calibration

To perform rapid temperature cycles, a HCS60 water-cooled Peltier thermal stage (Instec) was installed in the spinning disk microscope. The temperature was controlled using a mK1000 high precision temperature controller (Instec). To accommodate heat transfer to the sample, an aluminum sample chamber was custom-made to fit exactly in the thermal stage. Temperature calibration was performed using a fine gauge thermocouple from Okolab. Applying temperature cycles resulted in significant drift and convection in the sample. To ameliorate this, the IX3ZDC2 Z-drift compensation module (Olympus) was used. In many cases, this was not sufficient to retain focus; thus, the focus was adjusted manually in real time.

6.4.5 Image analysis

To determine the elongation of (populations of) liposomes, a custom MATLAB (R2017b, MathWorks) script was used. Using the GFP fluorescence as a lumen marker, a segmentation algorithm was performed to obtain objects representing the liposomes. The function `regionprops.m` was then applied to extract morphological parameters. From these, the elongation was determined by dividing the major axis length by the minor axis length. Determination of liposome surface area and volume was done using the 3D manager plug-in of FIJI[19]. Z-stacks were prepared for analysis by manual labelling of the liposome lumen.

6.5 ACKNOWLEDGMENTS

REFERENCES

1.   Canham, P. B. The minimum energy of bending as a possible explanation of the biconcave shape of the human red blood cell. *J. Theor. Biol.* **26,** 61-81 (1970).

2.   Berndl, K., Käs, J., Lipowsky, R., Sackmann, E. & Seifert, U. Shape transformations of giant vesicles: Extreme sensitivity to bilayer asymmetry. *Eur. Lett,* **13,** 659–664 (1990).

3.   Käs, J. & Sackmann, E. Shape transitions and shape stability of giant phospholipid vesicles in pure water induced by area-to-volume changes. *Biophys. J.* **60,** 825–44 (1991).

4.   Döbereiner, H. G., Käs, J., Noppl, D., Sprenger, I. & Sackmann, E. Budding and fission of vesicles. *Biophys. J.* **65,** 1396–403 (1993).

5.   Seifert, U. & Lepowsky, R. Chapter 8 Morphology of vesicles. *Handb. Biol. Phys.* **1,** 403–463 (1995).

6.   Döbereiner, H.-G., Evans, E., Kraus, M., Seifert, U. & Wortis, M. Mapping vesicle shapes into the phase diagram: A comparison of experiment and theory. *Phys. Rev. E* **55,** 4458 (1997).

7.   Emsellem, V., Cardoso, O. & Tabeling, P. Vesicle deformation by microtubules: A phase diagram. *Phys. Rev. E* **58,** 4807–4810 (1998).

8.   Furusato, T. et al. De novo synthesis of basal bacterial cell division proteins FtsZ, FtsA, and ZipA inside giant vesicles. *ACS Synth. Biol.* **7,** 953–961 (2018).

9.   Godino, E. et al. De novo synthesized Min proteins drive oscillatory liposome deformation and regulate FtsA-FtsZ cytoskeletal patterns cytoskeletal patterns. Nat. Commun. 10, 4969 (2019)

10. Godino, E. et al. Cell-free biogenesis of bacterial division proto-rings that can constrict liposomes. *bioRxiv* 2020.03.29.009639 (2020) doi:10.1101/2020.03.29.009639.

11. Makarova, K. S., Yutin, N., Bell, S. D. & Koonin, E. V. Evolution of diverse cell division and vesicle formation systems in Archaea. *Nat. Rev. Microbiol.* **8,** 731–741 (2010).

12. Marsh, D. Structural and thermodynamic determinants of chain-melting transition temperatures for phospholipid and glycolipids membranes. *Biochim. Biophys. Acta - Biomembr*. **1798,** 40–51 (2010).

13. Heimburg, T. Mechanical aspects of membrane thermodynamics. Estimation of the mechanical properties of lipid membranes close to the chain melting transition from calorimetry. *Biochim. Biophys. Acta - Biomembr.* **1415,** 147–162 (1998).

14. Sakuma, Y. & Imai, M. Model system of self-reproducing vesicles. *Phys. Rev. Lett.* **107,** 1–5 (2011).

15. Sakuma, Y. & Imai, M. From vesicles to protocells: The roles of amphiphilic molecules. *Life* **5,** 651–675 (2015).

16. Jimbo, T., Sakuma, Y., Urakami, N. & Imai, M. Role of inverse-cone-shape lipids in temperature-controlled self-reproduction of binary vesicles. *Biophys. J.* **110,** 1551–1562 (2016)

17. Urakami, N., Jimbo, T., Sakuma, Y. & Imai, M. Soft Matter Molecular mechanism of vesicle division induced membrane curvatures. *Soft Matter* **14,** 3018-3027 (2018)

18. Nourian, Z., Roelofsen, W. & Danelon, C. Triggered gene expression in fed-vesicle microreactors with a multifunctional membrane. *Angew. Chemie - Int. Ed.* **51,** 3114–3118 (2012).

19. Schindelin, J. et al. Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9,** 676–682 (2012).

20. Needham, D. & Evans, E. Structure and mechanical properties of giant lipid (DMPC) vesicle bilayers from 20 °C below to 10 °C above the liquid crystal-crystalline phase transition at 24 °C. *Biochemistry* **27,** 8261–8269 (1988).

21. Fujiwara, K. & Yanagisawa, M. Liposomal internal viscosity affects the fate of membrane deformation induced by hypertonic treatment. *Soft Matter* **13,** 9192–9198 (2017).

22. Fanalista, F. et al. Shape and size control of artificial cells for bottom-up biology. *ACS Nano* **13,** 5439–5450 (2019).

23. Blanken, D., van Nies, P. & Danelon, C. Quantitative imaging of gene-expressing liposomes reveals rare favorable phenotypes. *Phys. Biol*. **16,** 045002 (2019).

SUPPLEMENTARY INFORMATION

Derivation of the relation between surface area increase and liposome elongation

We will consider the case of the prolate (grape-shape) and oblate (orange-shape) separately. For a prolate, the surface area can be expressed as[1]:

$$A_{prolate} = 2\pi a^2 \left(1 + \frac{c}{ae}\sin^{-1} e\right) \qquad \text{(Eq. 1)}$$

Where a is the minor axis, c the major axis, and e is the eccentricity, defined as:

$$e^2 = 1 - \frac{a^2}{c^2} \qquad \text{(Eq. 2)}$$

Plugging equation 2 in equation 1, we obtain

$$A_{prolate} = 2\pi a^2 \left(1 + \frac{\lambda^2}{\sqrt{\lambda^2-1}}\sin^{-1}\sqrt{\frac{\lambda^2-1}{\lambda^2}}\right) \qquad \text{(Eq. 3)}$$

Where we used the elongation λ as c/a. We now use the reduced volume, which represents the volume of the spheroid divided by the volume of a sphere with the same area A as the spheroid:

$$V_{reduced} = \frac{\frac{4}{3}\pi a^2 c}{\frac{4}{3}\pi\left(\frac{A}{4\pi}\right)^{\frac{3}{2}}} \qquad \text{(Eq. 4)}$$

Plugging in equation 3 in equation 4, we obtain the reduced volume of a prolate with elongation λ:

$$V_{reduced,prolate} = \frac{2\sqrt{2}\,\lambda}{\left(1+\frac{\lambda^2}{\sqrt{\lambda^2-1}}\sin^{-1}\left(\sqrt{\frac{\lambda^2-1}{\lambda^2}}\right)\right)^{\frac{3}{2}}} \qquad \text{(Eq. 5)}$$

For an oblate, the surface area can be expressed as[2]:

$$A_{oblate} = 2\pi a^2 \left(1 + \frac{1-e^2}{e^2}\tanh^{-1} e\right) \qquad \text{(Eq. 6)}$$

Here, a is the major axis and c the minor axis, and the eccentricity is defined as:

$$e^2 = 1 - \frac{c^2}{a^2} \qquad \text{(Eq. 7)}$$

In a similar fashion as for the prolate, the reduced volume can be derived, resulting in the following expression:

$$V_{reduced,oblate} = \frac{2\sqrt{2}\,\lambda}{\left(1-\frac{\lambda^2}{\sqrt{1-\lambda^2}}\tanh^{-1}(1-\lambda^2)\right)^{\frac{3}{2}}} \qquad \text{(Eq. 8)}$$

To now link the reduced volume to the area increase, we define the area increase factor:

$$\alpha = \frac{A_1}{A_0}$$ (Eq. 9)

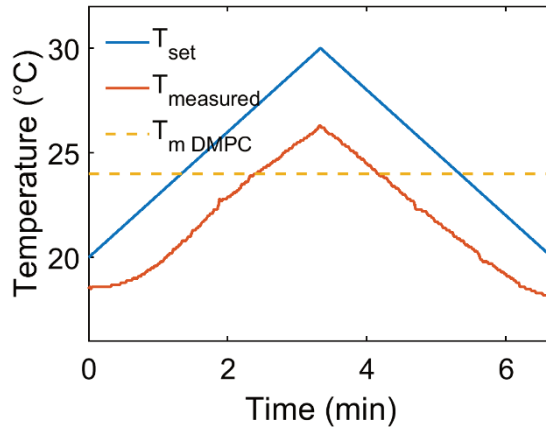Where A_0 is the area of the initial sphere of radius R:

$$A_0 = 4\pi R^2$$ (Eq. 10)

6

and A_1 is the area of the elongated spheroid. Combining with equation 4, we obtain:

$$V_{reduced} = \alpha^{-\frac{3}{2}}$$ (Eq. 11)

Filling in equations 5 and 8 in equation 11, we obtain the equation in the main text.

Supplementary figures

Supplementary figure 6.1: Calibration of temperature cycles. A temperature cycle was applied to a reaction chamber filled with water. The temperature of the water was probed with a thermocouple. An offset between the applied temperature and the measured temperature can be observed. However, the actual temperature of the sample also crosses the melting temperature of DMPC, allowing for shape deformations to take place. From the start of the cycle, it takes ~2.5 min for the sample to reach the T_m of DMPC. This accounts for the lag phase observed in elongation measurements.
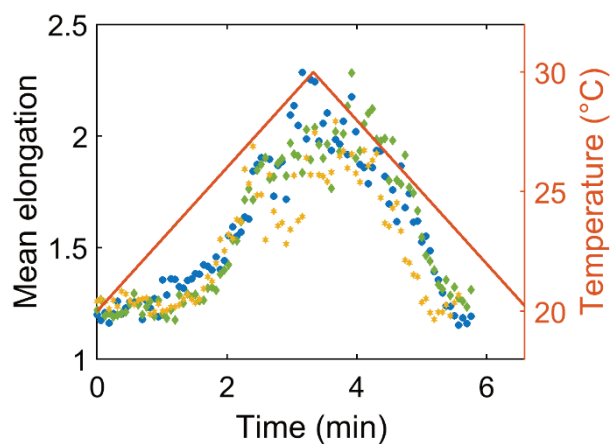
Supplementary figure 6.2: Mean elongation of myristoyl liposomes during one temperature cycle for three independent repeats. These data represent repeats of the data shown in **Fig. 6.3b**. The number of liposomes varied during acquisition. Yellow: $23 \leq N \leq 59$. Green: $26 \leq N \leq 68$. Blue: $15 \leq N \leq 62$.

Supplementary references

1. Weisstein, E. W. Prolate Spheroid. *MathWorld--A Wolfram Web Resource* https://mathworld.wolfram.com/ProlateSpheroid.html (2020).

2. Weisstein, E. W. Oblate Spheroid. *MathWorld--A Wolfram Web Resource* https://mathworld.wolfram.com/OblateSpheroid.html (2020).

6

Chapter 7: A perspective on module integration in artificial cell assembly

The various modules needed for a synthetic cell are mostly developed separately. Integrating these modules is a necessary step towards synthetic cell assembly and provides a major challenge. Within the synthetic cell framework of cell-free gene expression from DNA inside liposomes, DNA and its amplification have a central role. Therefore, DNA amplification should be taken as starting point for module integration. Here, we attempt to integrate the phospholipid synthesis pathway reconstituted in chapter 3 of this thesis with the φ29 DNA replication strategy previously developed in the lab. The genes required for both systems were included on a single DNA called pMAR. Lipid synthesis and modest DNA amplification could only be established by increasing the relative concentration of φ29 genes by including an auxiliary plasmid encoding for φ29 DNA replication machinery. Differences in optimal conditions, DNA occupancy, resource sharing, and interfacing between machinery from different organisms likely all contribute to the non-satisfactory DNA amplification results. The integration of these modules therefore needs to be further optimized. This optimization can be performed via the traditional approach of rational design, building, and testing. However, the large parameter space might be prohibitive. An evolutionary approach might be much more efficient, but it represents a new constructive paradigm of synthetic cell research. A hybrid approach, where evolution is used to make major strides in the fitness landscape, which are subsequently rationalized via the traditional approach, might provide an attractive middle ground.

7.1 INTRODUCTION

The work in previous chapters of this thesis is mostly concerned with synthetic cell compartment growth. This is just one of the modules needed to build a synthetic cell. As discussed in chapter 1, DNA replication, regeneration of the cell-free gene expression machinery, and compartment division will need to be reconstituted to realize an autonomously replicating synthetic cell. As these modules are studied more and more extensively individually, the next big challenge is their integration into a functioning whole[1].

This chapter will provide a perspective on module integration. First, the central role of the DNA replication module is discussed. This will be accompanied by the description of a testcase: the integration of φ29 DNA replication and phospholipid synthesis. The findings of this testcase will inform us about the challenges of module integration. Two approaches to overcome these challenges and realize module integration are discussed: the 'traditional' approach, and the 'evolutionary' approach. Finally, I will propose a strategy where both approaches are used in conjunction, inspired by the Lévy flight model.

7.2 DNA REPLICATION IS THE STARTING POINT FOR MODULE INTEGRATION

Liposomes, containing an *in vitro* transcription translation (IVTT) system, expressing genes encoding on DNA, are the most promising scaffold on which to build a minimal cell, as argued in chapter 1. Modules that are developed in a different framework (e.g. polymersome compartment, purified proteins instead of cell-free gene expression) will need to be converted to fit within the liposome-DNA-IVTT paradigm before they can be integrated into the synthetic cell. Within this paradigm, the whole design of the synthetic cell is encoded in the DNA. Therefore, replication and expression of the DNA are at the very core of successful self-replication of the system. Consequently, DNA replication should be the starting point for module integration. Starting by integrating other DNA-encoded modules, such as compartment growth and compartment division, already requires decisions to be made on DNA design. Should this design, for any reason, be incompatible with DNA replication, integration will have to be repeated. On the other hand, later integration of e.g. compartment division will not necessitate redesign of the DNA replication module, since this module does not interact with the DNA itself. Changes in the DNA to be replicated as a result of including a new module might necessitate further optimization in the DNA replication module, but this cannot be circumvented by taking a different module than DNA replication as a starting point.

In the Danelon lab, DNA replication has been reconstituted using the machinery of the bacteriophage φ29[2]. The φ29 DNA replication strategy is well-suited for reconstitution in a synthetic cell for two reasons. Firstly, the DNA replication is protein-primed, and therefore avoids loss of sequence information or subsequent DNA processing[3]. Secondly, the full DNA replication can be recapitulated by only four proteins: a DNA polymerase (DNAP, p2), a terminal protein (TP, p3) that binds to the origins of replication (ori) on both ends of a linear DNA template and serves as a primer[3], and auxiliary proteins p5[4] and p6[5].

PURE*frex* gene expression of a DNA encoding for p2 and p3, and flanked by oris (oriLR-p2p3), results in the synthesis of functional DNAP and TP, which can subsequently replicate the DNA they were expressed from. The replicated DNA can serve as a template for further rounds of DNA replication,

leading to significant DNA amplification (**Fig. 7.1a**). Addition of p5 and p6 improves DNA amplification but is not essential in bulk reactions[2]. This process is also recapitulated inside liposomes, where DNA amplifies with the help of p5 and p6, and forms aggregates, or blobs, that were visible by staining with the nucleic acid dyes acridine orange[2] (AO) (**Fig. 7.1b**) or dsGreen (**Fig. 7.2a**). The latter was found to bind more specifically to double-stranded DNA than AO. Staining with dsGreen revealed three distinct DNA amplification phenotypes: (1) background dsGreen signal, (2) homogeneous dsGreen signal, and (3) dsGreen signal in the form of a blob, as could also be observed with AO. Without dNTPs, no DNA can be synthesized, and only phenotype 1 is observed. The observed background dsGreen fluorescence comes from non-amplified DNA and RNA. No blobs (AO and dsGreen) or homogenous dsGreen signal were detected when dNTPs were omitted (**Fig. 7.1b,7.2b**). Phenotypes 2 and 3 are therefore an indication of successful DNA amplification. Taking the maximum dsGreen fluorescence intensity of the negative control (150 a.u.) as a threshold, 37% of the liposomes in **Fig. 7.2a** can be considered to have successfully amplified DNA.
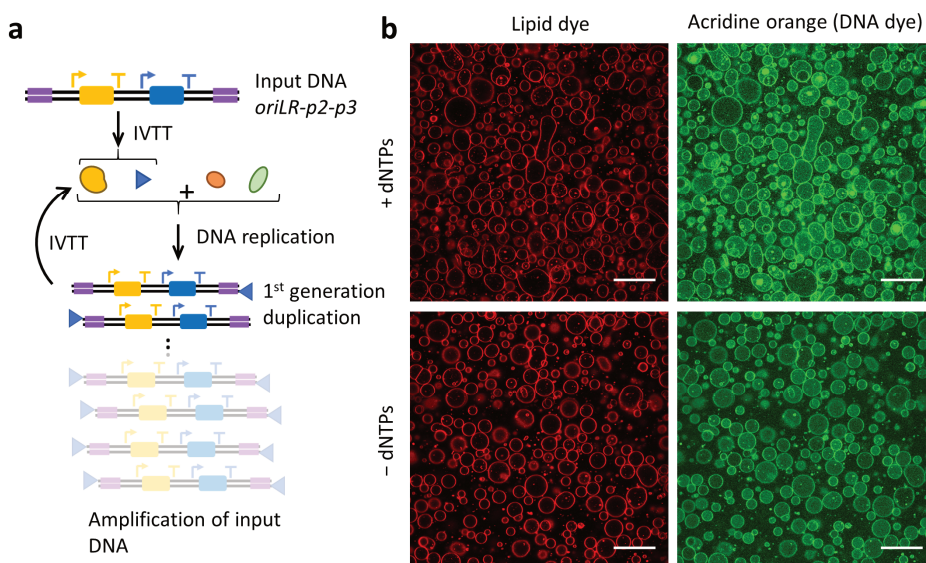
Figure 7.1: Amplification of DNA by its encoded proteins in liposomes. (a) The input DNA, encoding for p2 and p3, is expressed by PURE*frex* in vitro transcription translation (IVTT). When assisted by purified p5 and p6, the expressed p2 and p3 can replicate their parental DNA, kicking off an autocatalytic cycle of DNA amplification. **(b)** Liposomes encapsulating PURE*frex*, oriLR-p2-p3, purified p5 and p6, with or without dNTPs, visualized in the lipid dye channel (left), and acridine orange (AO, a DNA-staining dye that due to its hydrophobic aromatic groups also stains the lipid membrane) channel (right). With dNTPs, bright green blobs are visible in the AO channel that do not colocalize with aggregates in the lipid dye channel. These blobs are not observed when no DNA amplification takes place (−dNTPs). Hence, the blobs likely represent high density aggregates of amplified DNA. Scale bars are 20 µm. Figure is adapted from Van Nies et al.[2]
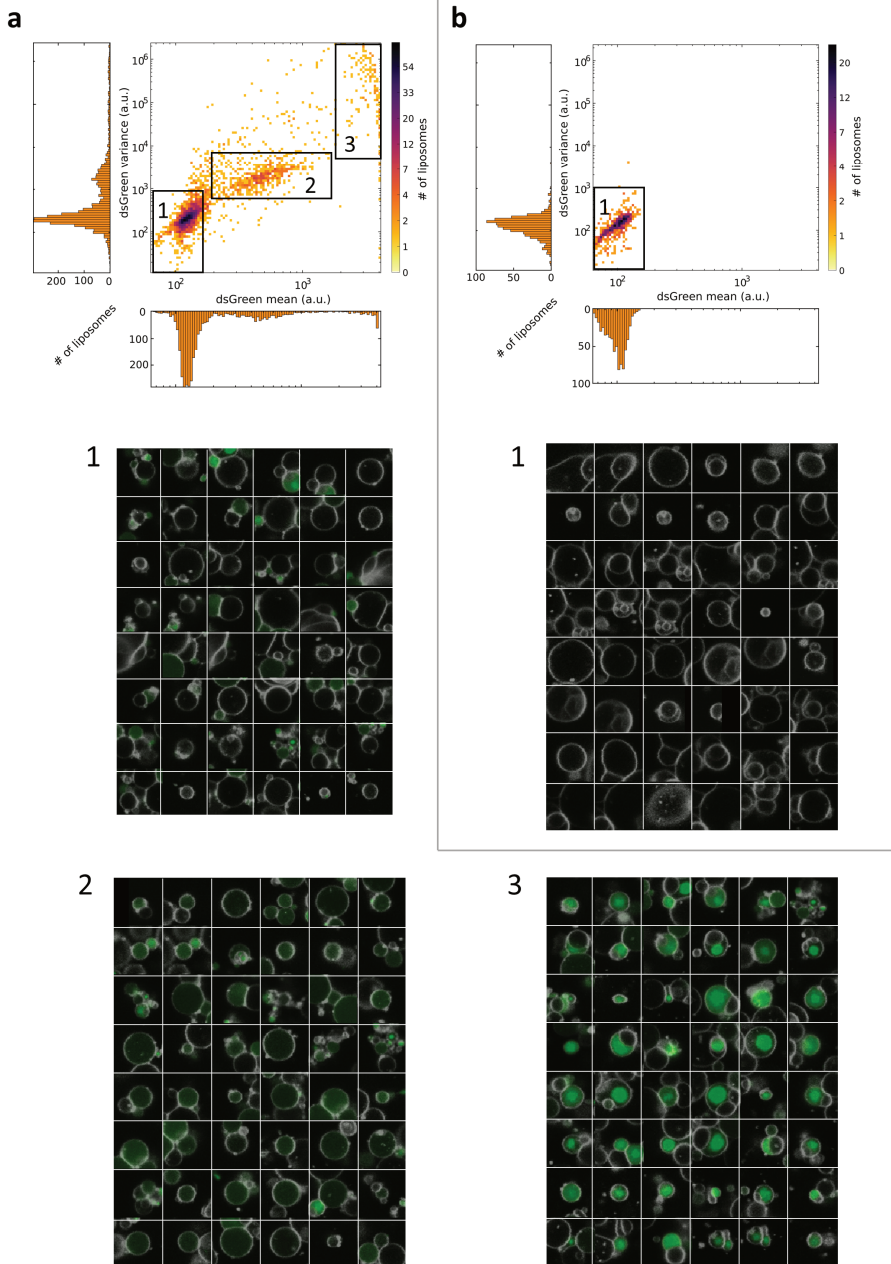
Figure 7.2: φ29 DNA replication inside liposomes results in three distinct phenotypes. (a) Cy5-stained liposomes encapsulating PURE*frex*2.0, oriLR-p2-p3, purified p5 and p6, with dNTPs and dsGreen as a DNA dye. Micrographs of liposome populations were acquired using laser scanning confocal microscopy, from which individual liposome micrographs and corresponding dsGreen signal were extracted using a custom MATLAB program. A heatmap of the mean and variance of the dsGreen signal of individual liposomes reveals three distinct phenotypes as defined in the main text. Lower panels show example micrographs of each phenotype. **(b)** Idem, but without dNTPs. Only phenotype 1 is observed, indicating no DNA amplification took place.

Applying the φ29 DNA replication machinery to copy the genetic information of the synthetic cell requires the synthetic cell genome to be linear and flanked by origins of replication. The DNA can either be present in the form of a single full-genome piece of DNA, or as multiple smaller DNA molecules. Using multiple pieces of DNA is attractive since it introduces the ratio between DNAs as an optimizable parameter. Genes that require high expression could be on a high copy number DNA, and vice versa. Moreover, having multiple smaller DNA molecules will be less exacting on the processivity of the φ29 DNAP. However, since all genetic information needs to be copied for the synthetic cell to proliferate, more pieces of DNA will require more initiations of DNA replication. Consequently, DNAP and TP will likely need to be expressed at higher concentrations, which will introduce a heavy burden on the IVTT system. The aim should therefore be to limit the number of different DNA molecules that comprise the synthetic cell genome.

7.3 TEST CASE: INTEGRATING DNA REPLICATION AND LIPID SYNTHESIS

Since both the φ29 DNA replication machinery and the phospholipid synthesis machinery have been reconstituted successfully within the same framework, they are prime candidates for a first attempt at module integration. A single linear DNA is designed, called pMAR, recapitulating both modules (**Fig. 7.3a**). Four phospholipid synthesis genes, *plsB, plsC, cdsA,* and *pssA*, under control of T7 promoters, are included. Expression of these genes yields the four-enzyme phosphatidylserine (PS) synthesis pathway. The genes encoding for p2 and p3 are included to reconstitute the DNA replication module. The expression of both genes is under control of an SP6 promoter. The orthogonality of the SP6-T7 promoter pair allows us to control and tune the expression of the different modules independently, as previously done with the pGEMM7 construct in chapter 3. Moreover, the lipid synthesis and DNA replication modules are placed on opposite strands of the DNA, to prevent unwanted expression as a result of transcription read-through.

DNA amplification is detected by staining with dsGreen, as discussed above. Liposomes are considered to have successfully amplified DNA when the mean dsGreen fluorescence intensity of the lumen is higher than the highest mean dsGreen fluorescence intensity of the lumen of liposomes of a negative control performed in parallel. Phospholipid synthesis is visualized by addition of the LactC2-mCherry probe, according to the procedure established in chapter 3. Liposomes are considered to synthesize lipids when mean mCherry fluorescence at the liposome membrane exceeds the threshold value, which was established as the mean plus two standard deviations of the mCherry rim intensity of a population of liposomes without PS or PS synthesis machinery. Cy5-labelled phospholipids are used to visualize the liposome membrane, yielding a spectrally non-overlapping set of three fluorophores.

Expression of pMAR in PURE*frex*2.0 with T7 and SP6 RNAP inside liposomes in the presence of dNTPs, purified p5 and p6, and all lipid synthesis precursors resulted in a negligible amount (0.1 %) of liposomes displaying elevated dsGreen fluorescence (**Fig. 7.3b**). When no SP6 RNAP is added, no liposomes amplified DNA, as expected (**Fig. 7.3b**). We reasoned that the low percentage of DNA-amplifying liposomes was a result of insufficient expression of p2 and p3. We added the plasmid G340, which contains the genes for p2 and p3, both under a T7 promoter, but cannot be amplified, with the aim of increasing the p2 and p3 concentrations and stimulate DNA amplification. Indeed, addition of G340 results in successful DNA amplification in 4.2% of the liposomes, in one iteration of the experiment (**Fig. 7.3b**). The heatmap of mCherry fluorescence versus dsGreen fluorescence revealed

that all DNA-amplifying liposomes also successfully synthesized PS (**Fig. 7.3c**). Liposomes from the high dsGreen, high mCherry region displayed both uniformly distributed dsGreen signal and blobs (**Fig. 7.3d**). In repeats, the percentage of liposomes exhibiting DNA amplification was very low (0.2% and 0%). This indicates that even when assisted by expression from G340, DNA amplification is unreliable. In contrast, PS enrichment is invariably high under the same conditions, with 88 % ± 16% of liposomes exhibiting mCherry fluorescence at the membrane that surpasses the threshold (**Fig. 7.3b**).

7.4 THE CHALLENGES OF MODULE INTEGRATION

The φ29 DNA replication module and the phospholipid synthesis module were previously independently reconstituted within the same experimental framework, but when combined, DNA replication was severely inhibited. This can be attributed to a variety of factors, which will be discussed here. Since these problems are not specific for the integration of the DNA replication and lipid synthesis module, their discussion will be extended to module integration in general.

7.4.1 Differences in optimal conditions between modules

To start, although the general experimental framework is similar, minor differences in optimal experimental conditions exist between the modules. Whereas PURE system gene expression and phospholipid synthesis have a temperature optimum of 37 °C[6], φ29 DNA replication works best at 30 °C[2]. The experiments described in the testcase were performed at 37 °C; studying the integration of the modules at a lower temperature is an obvious next step. Moreover, NTPs and tRNAs, needed for PURE system gene expression, are found to inhibit φ29 DNA replication[2,7,8]. Lowering their concentration may improve DNA replication, but might result in lower gene expression. Their concentrations therefore need to be carefully tuned. Other experimental factors, such as the inclusion of lipid synthesis precursors and β-mercaptoethanol required for lipid synthesis, did not seem to inhibit DNA replication with the oriLR-p2-p3 construct in preliminary experiments . Vice versa, addition of dNTPs and ammonium sulfate did not seem to inhibit phospholipid synthesis, exemplified by the high PS enrichment observed in **Fig. 7.3b**. It may however not be excluded that such additions could impact the performance of other modules to be incorporated. Therefore, integrating new modules will require repeated optimization of temperature and concentrations of NTPs, tRNAs, and other reaction components.
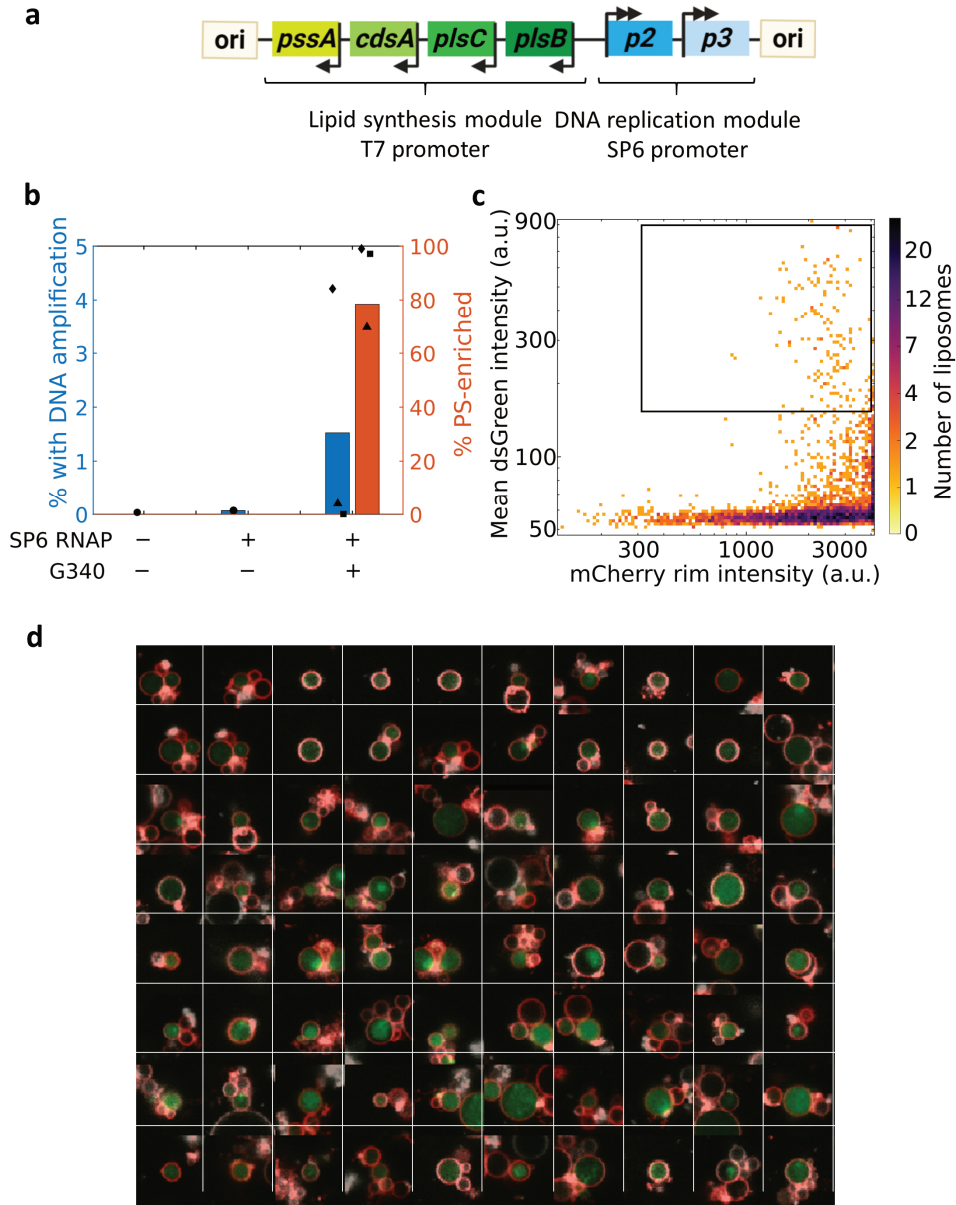
Figure 7.3: pMAR expression inside liposomes results in combined DNA replication and PS synthesis when assisted by G340. (a) Design of the pMAR DNA construct. The lipid synthesis module, consisting of the genes *plsB, plsC, cdsA,* and *pssA* are placed under T7 promoter. On the other strand, the DNA replication module, consisting of the genes encoding for p2 and p3, is placed under an SP6 promoter. Encoding the modules on opposing strands excludes non-specific expression due to transcription read-through (see chapter 3). **(b)** Percentage of liposomes that have a dsGreen fluorescence higher than the maximum fluorescence in a negative control acquired in parallel (blue, left axis) and that have a mCherry fluorescence higher than mean mCherry fluorescence plus two standard deviations of a control without PS synthesis (orange, right axis). All liposomes that have dsGreen fluorescence

above the threshold also have mCherry fluorescence above the threshold. Data points correspond to individual experiments, with similar data point markers in different conditions indicating experiments performed in parallel. Bars indicate the mean of indicated measurements. For conditions without G340, no mCherry measurement was performed. **(c)** Heatmap of the mCherry fluorescence intensity at the liposome rim (a measure for PS synthesis) versus the dsGreen fluorescence intensity in the liposome lumen (a measure of DNA amplification) for a population of 3,302 liposomes, corresponding to experiment indicated by diamonds in panel (b). The box indicates liposomes whose fluorescence intensity surpasses both the dsGreen and mCherry threshold (150 a.u. and 317 a.u., respectively), indicating both high PS synthesis and DNA amplification (4.2% of the total population). **(d)** A selection of liposomes corresponding to the data points within the box in panel (c). The liposome under investigation is the one in the center of the panel. Note that in two experimental repeats, the percentage of liposomes exhibiting DNA amplification was not significant, indicating that further improvement is needed to increase robustness.

7.4.1 Inhibition of DNA replication by high DNA occupancy

Another factor that could inhibit DNA replication is the occupancy of the DNA in module integration experiments. As pMAR contains six promoter sites, of two types, on both strands of the DNA, many RNA polymerases will be processing the DNA. Moreover, nascent RNA chains will bind ribosomes to form large transcription-translation complexes on the DNA. Collision with these complexes might cause the DNAP to stall or drop off, reducing the DNA replication rate[9]. *In vivo*, temporal regulation of transcription and DNA regulation is thought to limit the effect of collisions[9]. In our current *in vitro* system however, transcription and DNA replication occur simultaneously, with potentially inhibitory effects. Using multiple replicable DNAs encoding for a lower number of genes will decrease the chances of polymerase collision. For example, one could include the lipid synthesis and DNA replication genes on separate replicable DNAs. Expressed p2 and p3 will replicate their parental DNA, as well as the other DNA molecule that included the lipid synthesis genes. This will also facilitate tuning the relative gene concentrations (see below), but will require higher p2 and p3 concentrations.

7.4.3 Resource sharing and lack of regulation prevents optimal protein concentration

When expressing multiple genes in the PURE system, the final concentration of protein per gene drops (chapter 2). This is attributed to the limited availability of gene expression resources in the PURE system. What exactly is the limiting factor, and if the limiting factor is the same in all situations, is unclear, although depletion of initiation factors and tRNAs is suggested[10]. Regardless of which factor is limiting, when multiple genes are expressed, less PURE system machinery and nutrients are available per gene. Increasing the number of genes to be expressed will strengthen this effect. How the functioning of a module is affected by lower protein concentrations depends on the nature of the proteins. Enzymes, like the proteins of the lipid synthesis pathway, can catalyze many reactions at low concentrations, whereas processive machinery like the φ29 DNAP needs to be present in rather high concentration to successfully initiate DNA replication. The terminal protein is even consumed during DNA replication. Therefore, a lower expression per gene will much more strongly affect DNA replication than phospholipid synthesis. Indeed, even expressing the seven enzyme-encoding genes on pGEMM7 concurrently enabled the synthesis of phospholipids (chapter 3). Introduction of modules that include structural proteins (e.g. FtsZ, microtubules) will exaggerate the resource sharing problem, since these proteins require high concentrations to assemble in the many-protein complexes that provide their function.

Major improvements in PURE system protein yield will be needed to facilitate the expression of multiple modules in parallel. Regardless of the total gene expression yield, the widely varying optimal

concentrations of different types of proteins will require (fine)tuning of the relative expression levels. For this, applying the *E. coli* transcription machinery for IVTT[11], which is regulated via sigma factors, might be more suitable than depending on poorly regulated single subunit RNAPs like T7 RNAP or SP6 RNAP. In the testcase, expression is tuned by addition of the G340 DNA, encoding for extra p2 and p3. Hereby, the relative concentration of the p2 and p3 genes with respect to the lipid synthesis genes is increased, resulting in a higher relative expression level, and better DNA amplification. However, introduction of genetic material that cannot be transmitted to the next generation, such as the non-replicable G340, will result in dead-by-dilution of this information. It therefore does not represent a viable design strategy for the synthetic cell.

### 7.4.4 Interfacing modules from different organisms

All proteins of the lipid synthesis module, and most proteins of the PURE system, are derived from *E. coli*. However, the DNA replication machinery comes from the bacteriophage φ29, and the T7 RNAP, one of the key proteins in the PURE system, is derived from the T7 bacteriophage. Interfacing modules derived from different organisms might give rise to a range of problems. First, codon usage differs per organism, and needs to be optimized to match the availability of tRNAs in the PURE system. When properly expressed, the different proteins still need to be compatible (not inhibit each other) and orthogonal (not get in the way of each other's functions)[12]. Even when these requirements are met, interfacing machinery from different organisms will not recapitulate highly evolved cross-regulation mechanisms. For example, in the PURE system, the tight coupling between transcription and translation in *E. coli*[13] is lost by substitution of the transcription machinery by the much faster T7 RNAP. This results in the formation of inhibitory RNA structures and inefficient usage of mRNA, which have an inhibitory effect on gene expression[10]. Potentially, integrated lipid synthesis and DNA replication is similarly affected by an unknown inhibition or lack of coupling.

7.5 DIFFERENT PATHS TOWARDS SUCCESSFUL MODULE INTEGRATION

To realize module integration, the challenges detailed above will need to be overcome. There are two general approaches to do this, which I will refer to as the 'traditional approach' and the 'evolutionary approach'. Both come with clear advantages and drawbacks, and therefore I will here propose a hybrid approach that will combine the best of both methods.

7.5.1 The traditional approach

The traditional approach is exemplified by the testcase described above. It entails the combination of two modules at a time, taking into account a large amount of *a priori* knowledge on the different modules. As observed in the testcase, likely the first combination attempt will not be successful, but will yield new knowledge, which can be used in a new integration design. This kicks of a design-build-test (DBT) cycle[14], where the knowledge of the two integrated modules is expanded step-by-step (**Fig. 7.4a**). Every new module will be integrated via a series of DBT cycles.
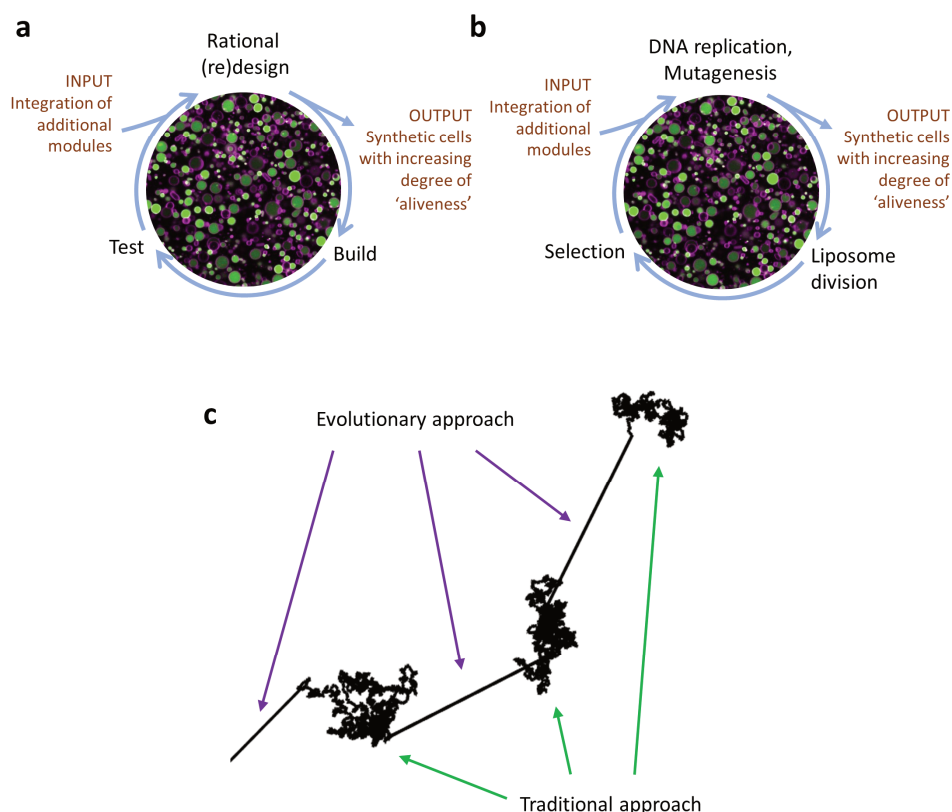
Fig 7.4: Module integration can be pursued via three different approaches. (a) In the traditional approach, a new module is introduced via rational design. This design is subsequently built and tested, with the outcome informing a new design. **(b)** In the evolutionary approach, DNA encoding for integrated modules is diversified via mutagenesis, spread over liposomes, and the best-performing liposomes are selected. In continuous mode, the first two steps can be achieved without liposome disruption via error-prone DNA amplification and liposome division. **(c)** Levy flight model of a hybrid approach. The evolutionary approach is applied to make big steps, when introducing a new module. The evolutionary solutions will inform rational design of the integrated modules (small steps). The trajectory is simulated in MATLAB as a 2D random walk with normally distributed steps, with manually defined 'jumps'. Panel (a) is adapted from and panel (b) is taken from Abil & Danelon[12].

The traditional approach is the clearest representation of the constructive paradigm of synthetic cell research: learning by building[6,15]. DBT cycles will provide a lot of information on the various modules and their interplay, which will aid research beyond the synthetic cell field. Moreover, since every possible parameter of the system will be designed, build, and tested, the end result will be a fully characterized cell with a complete blueprint: the goal of synthetic cell research.

However, the traditional approach might easily become 'prohibitively difficult'[12]. Even in the testcase, the parameter space in which optimization has to be performed is huge. As described above, the concentration of many components affects the module performance, the sequence and number of DNA molecules is an important parameter, the expression levels of proteins need to be tuned, and interplay between parts taken from different organisms needs to be optimized. Likely, auxiliary modules, for example responsible for regulating gene expression in time, will need to be included to

get DNA replication and phospholipid synthesis to work together. Concluding, only combining two relatively simple modules via the traditional approach is an incredibly complex project. With further integration of more modules, the parameter space and associated workload will explode to Herculean proportions.

7.5.2 The evolutionary approach

The alternative is the evolutionary approach[12]. Here, the complex parameter space arising from integration of two modules is not scanned via a DBT cycle, but evolution is applied to reach a local optimum (**Fig. 7.4b**). This requires much less *a priori* knowledge of the systems to be integrated and of their interplay. Evolution can be applied discontinuously and continuously. In discontinuous mode, the DNA(s) encoding for the integrated modules is amplified and mutagenized. Subsequently, it is encapsulated in liposomes, expressed by IVTT, and selection is applied based on the phenotype. The selected liposomes will be disrupted, the DNA will be extracted, amplified, and mutagenized to start a new cycle. In the continuous mode, liposomes will not be disrupted, but DNA amplification and diversification will take place inside the vesicle, and new genotypes will be distributed over its offspring. This approach becomes particularly attractive if DNA replication and liposome proliferation are reconstituted to a considerable extent and will allow for many evolutionary cycles with little experimental interference.

The evolutionary approach harnesses the power of *in vitro* evolution to efficiently scan the fitness landscape. Moreover, using large DNA libraries will result in large heterogeneous populations of liposomes, which each individually can be considered an optimization experiment. Coupled to efficient, high-throughput techniques of phenotype interrogation and selection (such as fluorescence activated cell sorting (FACS)[16] or intelligent image activated cell sorting (IACS)[17]), the evolutionary approach is well-suited to dramatically speed up module integration. Furthermore, integration of evolution, one of the key characteristics of life (chapter 1), into the synthetic cell design strategy is exceedingly elegant and will yield a wealth of information on evolutionary processes.

The traditional constructive paradigm of synthetic cell research is abandoned when applying the evolutionary method. Instead of realizing module integration via conscious design decisions, this is left to the randomness of evolution. Therefore, applying this method will not result in what, in my opinion, are the most important goals of synthetic cell research: a fully characterized synthetic cell and its blueprint. Of course, such information can be interrogated after the fact, but this *a posteriori* interrogation of the system, the default research mode of traditional biology, is what synthetic cell research attempts to circumvent. Therefore, it is likely that the evolutionary approach will yield a living synthetic cell (much) earlier than the traditional approach, but it will, on its own, not achieve some of the goals of synthetic cell research.

7.5.3 The hybrid approach

Since both the traditional and evolutionary approach have very clear advantages and drawbacks, a hybrid approach, taking 'the best of both worlds', might prove interesting. It should combine the power of evolution to efficiently find its way through the fitness landscape, with the constructive paradigm of the traditional approach. I envision the following strategy (**Fig. 7.4c**): at the start of an integration project, the modules are combined in rather simplistic fashion, and the evolutionary approach is applied to optimize fitness to a point where the two modules can be considered to function together. This would require little *a priori* knowledge of the system and would rather efficiently result in big steps in the fitness landscape. However, it will not reveal the blueprint of how

these two modules can be integrated. Therefore, in the second step, the evolutionary solutions of module integration will be studied, and serve as the *a priori* knowledge needed for traditional module integration. With this knowledge, traditional module integration can be more targeted, and thereby more efficient. In this way, the outcome of the evolutionary approach will be rationalized, and integrated in the constructive paradigm of synthetic cell research.

In many ways, this hybrid approach represents a Lévy flight[18]: a random walk in which the step-lengths have a probability distribution that is heavy-tailed. A Levy flight characteristically consists of many small steps, interspersed with rare long jumps to new regions (**Fig. 7.4c**). In this analogy, the evolutionary approach is responsible for the long jumps. It will serve to quickly bring the system to a new region, say a region where it consists of two modules instead of one. In that region, the traditional approach, which is much slower (much smaller steps) will be applied to rationalize the position of the system in this new region. At a certain point, this can be followed by another long evolutionary jump, and subsequent constructive biology experiments, etcetera.

7.6 CONCLUSION

The integration of different biological modules into a functioning whole is perhaps the biggest challenge to be faced by synthetic cell research. The different optimal conditions for each module, resource sharing in cell-free gene expression, DNA occupancy, and the interfacing of machinery from different organisms are all complicating factors for module integration. The extremely vast parameter space to be covered by optimization might not be covered by a traditional design-build-test approach. Therefore, harnessing the power of in vitro evolution could provide a solution. However, this means a deviation from the conventional constructive paradigm of synthetic cell research. Therefore, a hybrid between traditional, rational design and an evolutionary approach might be the best way forward. Here, evolution will allow us to make big leaps, which are subsequently rationalized according to the traditional approach.

7.7 METHODS

7.7.1 DNA constructs

The synthesis of the oriLR-p2-p3 DNA was previously published[2]. The G340 plasmid is a modified version of the oriLR-p2-p3 DNA, where one of the G10 leader sequences was replaced by a sequence with similar secondary structure but different primary structure.

The plasmid pTU1 (Addgene) was used as backbone for pMAR. The six individual genes, and a region containing the oris with a Pme1 site, were amplified by PCR. All pieces were combined using stepwise Golden Gate assembly. The final plasmid was sequence verified and stored in a glycerol stock. Before use, the plasmid was linearized by digestion with Pme1 to form the amplifiable pMAR.

7.7.2 Liposome experiments

A PURE*frex* (only Fig. 7.1) or PURE*frex*2.0 (both GeneFrontier) reaction was assembled according to instructions by the manufacturer, and supplemented with 20 µM of purified p5, 8 µM of purified p6, 20 mM of ammonium sulfate, 0.75 U/µL SUPERase (Ambion, only Fig. 7.1), and 300 µM of PCR Nucleotide mix (Promega). For experiments involving the synthesis of lipids, 500 µM G3P, 1 mM CTP, 500 µM L-serine, 11.25 mM β-mercaptoethanol (all from Sigma-Aldrich) were added. When indicated,

1 U/µL SP6 RNAP (Promega) and 2 nM G340 DNA was added. Either 2.3 nM of oriLR-p2-p3 DNA or 2 nM pMAR DNA was added.

Lipid-coated beads, with a lipid composition of DOPC:DOPE:DOPG:CL:DSPE-PEG(2000)-Biotin:DHPE-Texas Red (only Fig. 7.1)/DOPE-Cy5 (50 mol% : 36 mol% : 12 mol%: 2 mol% : 1 mass% : 0.5 mass%), were prepared as described in chapter 2. All lipids were from Avanti Polar Lipids, except DHPE-Texas Red which was from Invitrogen. For Fig. 7.1, 12 mg of beads was added to 10 µL of PURE system mixture. In all other experiments, 5 mg was added to 12.5 µL. Swelling was performed by incubating on ice for two hours, followed by four freeze-thaw cycles. For Fig. 7.1, 2 µL of supernatant was then harvested and diluted in a feeding solution consisting of PURE*frex* Solution 1 and Milli-Q (4:7), with 91 µg/mL proteinase K, to a final volume of 7.5 µL. In other experiments, 12 µL was harvested, and supplemented with 0.5-0.75 µL DNase 1 (Promega) solution, and, when indicated, 1.5 µL of LactC2-mCherry (for details on purification, see chapter 3). For lipid synthesis experiments, the sample was added to a dried film of oleoyl-CoA (Avanti Polar Lipids), for a final concentration of 100 µM.

After overnight incubation at 30 °C (DNA replication only) or 37 °C (DNA replication and lipid synthesis), the sample was immobilized as described in chapter 2. For Fig. 7.2b, 62.5 µM acridine orange (Sigma Aldrich) was added to the sample. For other experiments, 13 µL of a staining solution containing 10,000× diluted dsGreen (Lumiprobe), LactC2-mCherry, and buffer (20 mM HEPES, 180 mM potassium glutamate, 14 mM magnesium acetate, pH 7.6) was added to 2 µL of liposomes. Imaging was performed after 30 min of incubation at 37 °C.

Microscopy was performed using a Nikon A1R laser scanning confocal microscope with a SR Apo TIRF 100× oil immersion objective, with 488 nm (acridine orange/ dsGreen), 561 nm (Texas Red, mCherry), and 647 nm (Cy5) laser lines with appropriate emission filters.

7.7.3 Image analysis

Image analysis was performed with a custom MATLAB program. Liposomes are detected in large montages in the membrane dye channel, following the approach described in chapter 2. Subsequently, fluorescence and morphological parameters are extracted for, and a 30×30-pixel crop is made of, each liposome. Data were then investigated with a custom graphical user interface, in which all liposomes in a data set can be plot as a heatmap of two selected parameters. By indicating a region of interest in the heatmap, a sample of liposome image crops corresponding to that region are displayed.

The percentage of liposomes that had successful DNA amplification was determined by thresholding with respect to the maximum value of a negative control (minus dNTPs, or minus amplifiable DNA) performed in parallel. This threshold was found to vary significantly from day-to-day, excluding the use of a general threshold. The mCherry signal was more stable, and a general threshold could be determined as the mean plus two standard deviations of the mCherry signal of a population of liposomes without PS.

7.8 ACKNOWLEDGEMENTS

REFERENCES

1. Caschera, F. & Noireaux, V. Integration of biological parts toward the synthesis of a minimal cell. *Curr. Opin. Chem. Biol.* **22,** 85–91 (2014).

2. Van Nies, P. et al. Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. *Nat. Commun.* **9,** 1583 (2018).

3. Blanco, L., Lazaro, J. M., Vega, M. D. E., Bonnin, A. N. A. & Salas, M. Terminal protein-primed DNA amplification. *Proc. Natl. Acad. Sci. U. S. A.* **91,** 12198–12202 (1994).

4. Esteban, J. A., Blanco, L., Villar, L. & Salas, M. In vitro evolution of terminal protein-containing genomes. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 2921–2926 (1997).

5. Blanco, L., Gutièrrez, J., Làzaro, J. M., Bernad, A. & Salas, M. Replication of phage φ29 DNA in vitro: Role of the viral protein p6 in initiation and elongation. *Nucleic Acids Res.* **14,** 4923–4937 (1986).

6. Nourian, Z., Roelofsen, W. & Danelon, C. Triggered gene expression in fed-vesicle microreactors with a multifunctional membrane. *Angew. Chemie - Int. Ed.* **51,** 3114–3118 (2012).

7. Blanco, L. et al. Highly efficient DNA synthesis by the phage φ29 DNA polymerase. Symmetrical mode of DNA replication. *J. Biol. Chem.* **264,** 8935–8940 (1989).

8. Sakatani, Y., Ichihashi, N., Kazuta, Y. & Yomo, T. A transcription and translation-coupled DNA replication system using rolling-circle replication. *Sci. Rep.* **5,** 1–5 (2015).

9. Elías-Arnanz, M. & Salas, M. Bacteriophage φ29 DNA replication arrest caused by codirectional collisions with the transcription machinery. *EMBO J.* **16,** 5775–5783 (1997).

10. Doerr, A. et al. Modelling cell-free RNA and protein synthesis with minimal systems. *Phys. Biol.* **16,** 025001 (2019).

11. Shin, J. & Noireaux, V. An E. coli cell-free expression toolbox: Application to synthetic gene circuits and artificial cells. *ACS Synth. Biol.* **1,** 29–41 (2012).

12. Abil, Z. & Danelon, C. Roadmap to building a cell: an evolutionary approach. *Under Review* (2020).

13. Castro-Roa, D. & Zenkin, N. In vitro experimental system for analysis of transcription-translation coupling. *Nucleic Acids Res.* **40,** e45 (2012).

14. Hutchison, C. A. et al. Design and synthesis of a minimal bacterial genome. Science. 351, aad6253 (2016).

15. Altamura E, Carrara P, D'Angelo F, Mavelli F and Stano P. Extrinsic stochastic factors (solute partition) in gene expression inside lipid vesicles and lipid-stabilized water-in-oil droplets: a review *Synthetic Biology* **3,** ysy011 (2018).

16. Fujii, S., Matsuura, T., Sunami, T., Kazuta, Y. & Yomo, T. In vitro evolution of α-hemolysin using a liposome display. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 16796–16801 (2013).

17. Nitta, N. et al. Intelligent Image-Activated Cell Sorting. *Cell* **175,** 266-276 (2018).

18. Viswanathan, G. M. et al. Lévy flight random searches in biological phenomena. *Phys. A Stat. Mech. its Appl.* **314,** 208–213 (2002).

7

Summary

Cells, the building blocks of life, are vastly complex. This complexity confers to every living organism the ability to maintain oneself, reproduce oneself, and evolve. Creating a minimal system from non-living components that is capable of self-maintenance, self-reproduction, and evolvability, will greatly increase our understanding of life. Essential features of every cell, synthetic or otherwise, are the compartment, a form of information transfer, and the ability to proliferate. In **chapter 1**, I argue that autonomy, that is self-governance, is another key characteristic of life. Therefore, to create a synthetic cell, aforementioned features should be recapitulated in an autonomous manner. This informs the synthetic cell design that is adhered to in this thesis. Phospholipid vesicles, so-called liposomes, serve as a compartment. Reconstitution of the central dogma of molecular biology by cell-free gene expression with the PURE system enables the use of a genetic program encoded on DNA. By encoding its working instructions on its own DNA, the cell will be autonomous. Proliferation is carried out by a set of modules encoded on the DNA. These modules will serve to replicate the DNA itself, to replenish the gene expression machinery, and to stimulate growth and trigger division of the compartment. The goal of this thesis is to reconstitute compartment growth by cell-free gene expression of DNA encoding for phospholipid synthesis machinery.

The scaffold on which all work in this thesis is built, PURE system gene expression inside liposomes, is the object of study in **chapter 2**. The lipid-coated glass beads protocol for liposome preparation, established previously in our lab, is improved to increase the yield and quality of giant unilamellar liposomes. The yellow fluorescent protein (YFP) is expressed inside these liposomes, its fluorescence intensity serving as a direct measure of the amount of protein expressed. High-content fluorescence confocal microscopy is applied to determine the YFP concentration in tens of thousands of liposomes. A small fraction of liposomes exhibits favorable phenotypes, with protein yield and/or expression time much higher than observed in bulk PURE system reactions. Membrane permeability is probed with the small molecular dye Cy5. Transient permeability events are observed, and stable permeabilization to small molecules can be established by the expression of pore-forming protein connexin-43. Increased permeability does not increase gene expression yield or time, indicating that depletion of nutrients is not limiting protein synthesis.

Chapters 3-5 form the core of this thesis and are concerned with the growth of the synthetic cell compartment by the synthesis of phospholipids. In **chapter 3**, the seven-enzyme Kennedy pathway of bacterial phospholipid synthesis is encoded on a single DNA construct called pGEMM7. The genes of the five enzymes needed to synthesize phosphatidylethanolamine (PE) are under control of a T7 promoter, whereas the two-enzyme pathway branch leading to the synthesis of phosphatidylglycerol (PG) is encoded under an SP6 promoter. Via this scheme, the ratio of PE versus PG can be tuned genetically. Moreover, metabolic regulation by activation of the PssA enzyme by anionic lipids is recapitulated *in vitro*. Expressing pGEMM7 inside liposomes results in the synthesis of up to six phospholipid end products in a one-pot reaction, as detected with liquid chromatography-coupled mass spectrometry (LC-MS). Two fluorescence microscopy methods are applied to study phospholipid synthesis and membrane incorporation at the single-liposome level. The cell-free expressed phospholipid synthesis pathway was able to convert acyl chains with a fluorescent NBD moiety into a novel fluorescent lipid. Membrane incorporation of NBD-labelled lipids was demonstrated by an increase in fluorescence colocalizing with the membrane upon lipid synthesis. Assays with the LactC2 probe, fused to a fluorescent protein, that binds specifically to the lipid phosphatidylserine (PS), corroborate the findings with NBD. Moreover, it demonstrated heterogeneity I n PS enrichment

between liposomes and could be used to obtain kinetics of the complete process of gene expression, lipid synthesis, membrane incorporation, and probe binding. Concluding, the synthesis of phospholipids and their incorporation in the liposome membrane demonstrate autonomous DNA-encoded growth of liposomes. However, the amount of synthesized lipids, and thereby the relative volume increase of the liposomes, is severely limited by the low solubility of the acyl-CoA precursor. Therefore, physical growth was not significant enough to be directly observed with optical microscopy methods.

The main lipid constituent of the liposomes used in this research is phosphatidylcholine (PC). In **chapter 4**, the pathway investigated in chapter 3 is extended by the enzyme PmtA, which synthesizes PC from PE. LC-MS measurements determined that cell-free expressed PmtA from *Rhodobacter sphaeroides* could synthesize PC, both alone and in conjunction with the *Escherichia coli*-based PE synthesis pathway encoded on pGEMM7. Moreover, it was found that cell-free expressed PmtA could remodel liposomes, resulting in the formation of septums, and a decrease in average liposome size. In contrast to previously observed nanoscale remodeling of liposomes with PmtA from *Agrobacterium tumefaciens*, we observed remodeling taking place on the micrometer scale that could be identified with confocal microscopy. This unexpected and drastic remodeling property could be of interest for reconstituting synthetic cell compartment division.

The low solubility of acyl donor acyl-CoA limits the amount of phospholipids that can be synthesized with the pathway reconstituted in chapter 3. Therefore, in **chapter 5**, the use of more soluble precursors is explored. Using the acyl-acyl carrier protein complex resulted in significantly lower lipid yields. Oleic acid could be used as a more soluble precursor by upstream pathway extension with cell-free expressed FadD. However, the higher precursor concentration was offset by a much lower conversion efficiency. Therefore, the phospholipid yield was not higher than when using acyl-CoA as precursor. Circumventing the Kennedy pathway altogether, native chemical ligation-based synthesis of membrane-forming lipid analogues, as developed by the Devaraj lab, was explored. The chemical lipid synthesis precursors turned out to be incompatible with PURE system gene expression. Therefore, this reaction scheme cannot be reconstituted within the synthetic cell framework established in chapter 1. Continuous supply of low concentrations of acyl-CoA in a flow reactor setup is suggested as an alternative method of increasing lipid yield.

Phospholipid membranes drastically increase in area upon phase transition. In **chapter 6**, triggering phase transition by crossing the melting temperature is applied to deform gene-expressing liposomes. Producing gene-expressing liposomes with an experimentally attainable phase transition temperature required extensive modification of the protocol developed in chapter 2 and resulted in a decrease in liposome quality. Nevertheless, GFP-expressing liposomes could be stably elongated for multiple hours by crossing the melting temperature. Stably elongated liposomes could be applied to study the reconstitution of cell systems that depend on a non-spherical geometry. Moreover, temperature cycling resulted in reversible elongation and contraction of liposomes. Addition of inverse cone-shaped lipids did not result in liposome division, in contrast to the results of Sakuma and Imai in aqueous solution. This is likely caused by the lower rate of deformation in the PURE system.

To build a synthetic cell, modules developed independently, such as the phospholipid synthesis machinery reconstituted in this thesis, should be integrated in a functioning whole. In **chapter 7**, I argue that module integration for a DNA-based autonomous synthetic cell should start with a DNA replication module. A testcase is presented, where the φ29 DNA replication module, previously developed in our lab, is combined with the phospholipid synthesis module developed in chapter 3. Although based on a similar framework, DNA replication is severely inhibited when combined with

phospholipid synthesis. Minor differences in experimental conditions, DNA occupancy, resource sharing, and interfacing between modules from different evolutionary contexts are possible complicating factors. Further module integration can proceed according to the traditional approach, where design-build-test cycles are applied to combine modules. However, this requires a large amount of *a priori* knowledge and might become prohibitively difficult. The evolutionary approach is an attractive alternative but presents a deviation from the conventional constructive paradigm of synthetic cell research. A hybrid approach, where evolutionary solutions are subsequently rationalized by traditional design, might provide an attractive middle ground.

Samenvatting

Cellen, de bouwstenen van het leven, zijn enorm complex. Door deze complexiteit is elk levend organisme in staat zichzelf te onderhouden, zichzelf te reproduceren en te evolueren. Het creëren van een levend systeem, van niet-levende onderdelen, dat in staat is tot zelfonderhoud, zelfreproductie en evolutie zal ons begrip van het leven spectaculair vergroten. Essentiële onderdelen van elke cel, zowel synthetisch als natuurlijk, zijn het compartiment, een vorm van informatieoverdracht en het vermogen tot vermenigvuldiging. In **hoofdstuk 1** beargumenteer ik dat autonomie, het onafhankelijk functioneren, ook een kerneigenschap van het leven is. Daarom moeten, om een synthetische cel te maken, voornoemde eigenschappen op autonome wijze gerecreëerd worden. Deze eis ligt ten grondslag aan het celontwerp dat gevolgd wordt in deze thesis. Blaasjes van fosfolipiden, zogenaamde liposomen, dienen als compartiment. Celvrije genexpressie met het PURE-systeem recapituleert het centrale dogma van de moleculaire biologie en stelt ons in staat een genetisch programma, gecodeerd in DNA, te gebruiken. Het coderen van zijn werkinstructies in zijn eigen DNA maakt een cel autonoom. Vermenigvuldiging wordt uitgevoerd door een aantal in het DNA gecodeerde modules. Deze modules voeren de replicatie uit van het DNA zelf, verversen de machinerie voor genexpressie, en stimuleren de groei en deling van het compartiment. Het doel van deze thesis is om groei van het compartiment te realiseren door de synthese van fosfolipiden door middel van de celvrije genexpressie van machinerie die gecodeerd is op DNA.

Het fundament waarop al het werk in deze thesis is gebouwd, genexpressie met het PURE-systeem in liposomen, wordt bestudeerd in **hoofdstuk 2**. De methode om liposomen te maken door middel van met lipiden gecoate glazen korreltjes, zoals eerder ontwikkeld in ons lab, is verbeterd om tot een groter aantal en een betere kwaliteit van grote liposomen met een enkel membraan te komen. Het *yellow fluorescent protein* (YFP, geel fluorescent eiwit) is tot expressie gebracht binnen deze liposomen, waarbij de fluorescentie-intensiteit een directe maat is voor het aantal tot expressie gebrachte eiwitten. Door middel van confocale fluorescentiemicroscopie is de concentratie van YFP in tienduizenden liposomen bepaald. Bijzonder gunstige fenotypen, met eiwitproductie of expressietijd veel hoger dan in bulkexperimenten, zijn waargenomen in een kleine fractie van liposomen. Met behulp van de kleine moleculaire kleurstof Cy5 is de permeabiliteit van het membraan bestudeerd. Kortdurende momenten van permeabiliteit zijn waargenomen, en door expressie van het eiwit connexin-43, dat een porie vormt, kan het membraan stabiel permeabel gemaakt worden voor kleine moleculen. De opbrengst of duur van genexpressie wordt niet verhoogd door de toegenomen permeabiliteit, wat een aanwijzing is dat het opraken van voedingsstoffen geen limiterende factor is in eiwitsynthese.

Hoofdstukken 3-5 vormen het hart van deze thesis en hebben betrekking op de groei van het compartiment van de synthetische cel door de synthese van fosfolipiden. In **hoofdstuk 3** is de zogenaamde *Kennedy pathway* voor de synthese van fosfolipiden, die bestaat uit zeven bacteriële enzymen, gecodeerd op een enkel stuk DNA genaamd pGEMM7. De expressie van de genen voor de vijf enzymen die nodig zijn voor de synthese van phosphatidylethanolamine (PE) wordt gecontroleerd door een T7-promoter. De expressie van de genen die coderen voor de twee enzymen die samen een andere tak van de reactieketen vormen die leidt tot de synthese van phosphatidylglycerol (PG) wordt daarentegen gecontroleerd door een SP6-promoter. Hierdoor kan de ratio van PE tot PG op genetische wijze afgestemd worden. Bovendien is de metabolische regulatie van deze ratio door middel van de activatie van het PssA-enzym door anionische lipiden *in vitro* gerecreëerd. Het tot expressie brengen van pGEMM7 binnen liposomen resulteert in de synthese van tot wel zes fosfolipide eindproducten in een enkele reactie, zoals gedetecteerd met *liquid chromatography-*

coupled mass spectrometry (LC-MS, vloeistofchromatografie gekoppeld aan massaspectrometrie). Twee methoden gebaseerd op fluorescentiemicroscopie zijn toegepast om de synthese van lipiden en hun incorporatie in het membraan te bestuderen op het niveau van individuele liposomen. De celvrij tot expressie gebrachte reactieketen voor de synthese van fosfolipiden was in staat acylketens met een fluorescente NBD-groep in een nieuwe fluorescente lipide te incorporeren. Een toename van het fluorescentiesignaal aan het membraan als gevolg van lipidensynthese toonde aan dat deze NBD-gelabelde lipiden in het membraan opgenomen worden. Proeven met LactC2, dat specifiek bindt aan de lipide phosphatidylserine (PS) en gefuseerd was met een fluorescent eiwit, bevestigden de bevindingen met NBD. Bovendien lieten deze proeven zien dat de verrijking met PS verschilt per liposoom, en kon de LactC2-methode gebruikt worden om de kinetiek te bepalen van het volledige proces van genexpressie, lipidensynthese, membraanopname en het binden van LactC2. De synthese van fosfolipiden en hun opname in het liposoommembraan tonen tezamen de autonome, DNA-gecodeerde groei van liposomen aan. De hoeveelheid gesynthetiseerde lipiden, en daarmee de relatieve toename in volume van de liposomen, wordt echter ernstig beperkt door de lage oplosbaarheid van acyl-CoA, één van de precursors van de lipiden. De fysieke groei van de liposomen was daarom niet groot genoeg om direct te detecteren met behulp van methoden gebaseerd op lichtmicroscopie.

Het belangrijkste lipide bestanddeel van de liposomen die gebruikt zijn in dit onderzoek is phosphatidylcholine (PC). In **hoofdstuk 4** is de in hoofdstuk 3 bestudeerde reactieketen uitgebreid met het enzym PmtA, dat PC maakt van PE. Door metingen met LC-MS kon worden vastgesteld dat celvrij tot expressie gebracht PmtA van *Rhodobacter sphaeroides* in staat was PC te synthetiseren, zowel alleen als in samenwerking met de van *Escherichia coli* afkomstige reactieketen voor PE gecodeerd op pGEMM7. We vonden bovendien dat celvrij tot expressie gebracht PmtA liposomen kan remodelleren, resulterend in het ontstaan van septums en een afname in de gemiddelde grootte van liposomen. In tegenstelling tot de eerder geobserveerde remodelering op de nanoschaal door PmtA van *Agrobacterium tumefaciens,* observeerden wij vervorming die plaats vond op de microschaal en dus met behulp van confocale microscopie kon worden geïdentificeerd. Deze onverwachte en spectaculaire eigenschap zou van belang kunnen zijn voor de reconstitutie van compartimentsdeling in een synthetische cel.

De slechte oplosbaarheid van acyl-CoA, het molecuul dat acylketens levert voor lipidensynthese, beperkt de hoeveelheid fosfolipiden die gevormd kan worden met behulp van de reactieketen beschreven in hoofdstuk 3. Daarom is in **hoofdstuk 5** het gebruik van beter oplosbare precursors bestudeerd. Het gebruik van het *acyl-acyl carrier protein complex* resulteerde in significant lagere productie van lipiden. *Oleic acid* (oliezuur) kon als beter oplosbare precursor gebruikt worden door de reactieketen uit te breiden met celvrij tot expressie gebracht FadD. De hogere precursorconcentratie werd echter tenietgedaan door een lagere productie-efficiëntie. De fosfolipidenproductie was dientengevolge niet hoger dan wanneer acyl-CoA als precursor wordt gebruikt. De synthese van membraanvormende lipiden-analogen gebaseerd op *native chemical ligation,* zoals ontwikkeld door de groep van Devaraj, is verkend als alternatief om de *Kennedy pathway* volledig te omzeilen. De precursors voor chemische lipidensynthese bleken onverenigbaar met genexpressie in het PURE-systeem. Dit reactieschema kan daarom niet worden toegepast binnen het kader van de synthetische cel zoals uitgewerkt in hoofdstuk 1. De continue toevoer van lage concentraties van acyl-CoA in een stroomreactor is gesuggereerd als alternatief om de lipidenproductie op te schroeven.

De oppervlakte van een fosfolipide membraan verandert dramatisch wanneer het een fasetransitie ondergaat. In **hoofdstuk 6** is het teweeg brengen van een fasetransitie door de smelttemperatuur over te steken toegepast om liposomen te vervormen terwijl er genexpressie in plaats vindt. De methode om liposomen te maken, zoals beschreven in hoofdstuk 2, moest drastisch aangepast worden om liposomen te genereren die zowel in staat waren tot genexpressie als een fasetransitietemperatuur hadden die experimenteel haalbaar was, hetgeen resulteerde in een afname van de kwaliteit van de liposomen. Desondanks konden liposomen waarin GFP tot expressie was gebracht voor meerdere uren stabiel uitgerekt worden door de smelttemperatuur over te steken. Deze stabiel uitgerekte liposomen zouden toegepast kunnen worden om de reconstitutie van systemen die een niet-sferische geometrie nodig hebben te bestuderen. Het cyclisch variëren van de temperatuur rond het smeltpunt resulteerde bovendien in reversibele uitrekking en samentrekking van de liposomen. De toevoeging van invers-conische lipiden resulteerde niet in deling van het liposoom, in tegenspraak met de resultaten van Sakuma en Imai in water. Dit wordt vermoedelijk veroorzaakt door de lagere snelheid van vervorming in het PURE-systeem.

Om een synthetische cel te bouwen moeten onafhankelijk ontwikkelde modules, zoals de machinerie voor fosfolipidensynthese in deze thesis, geïntegreerd worden tot een functionerend geheel. In **hoofdstuk 7** beargumenteer ik dat de integratie van modules voor een op DNA gebaseerde autonome synthetische cel de DNA-replicatiemodule als uitgangspunt moet nemen. Om de proef op de som te nemen is de φ29 DNA-replicatiemodule, eerder ontwikkeld in ons lab, gecombineerd met de fosfolipidensynthesemodule ontwikkeld in hoofdstuk 3. Hoewel de twee modules ontwikkeld zijn binnen hetzelfde kader neemt DNA-replicatie enorm af wanneer het gecombineerd wordt met fosfolipidensynthese. Mogelijke complicerende factoren zijn kleine verschillen in experimentele condities, de bezetting van het DNA, het delen van grondstoffen en de koppeling tussen twee modules uit een verschillende evolutionaire context. Een verdere integratie van de modules kan gerealiseerd worden via de traditionele aanpak waar cycli van ontwerpen, bouwen en testen worden toegepast om modules te combineren. Voor een dergelijke aanpak is echter een grote hoeveelheid *a priori* kennis nodig; bovendien wordt het mogelijk buitensporig ingewikkeld. De evolutionaire aanpak is een aantrekkelijk alternatief, maar hiermee wijkt men af van het conventionele constructieve paradigma van het onderzoek naar synthetische cellen. Een hybride aanpak, waarbij evolutionaire oplossingen gerationaliseerd worden via het traditionele ontwerpproces, biedt mogelijk een aantrekkelijke tussenoplossing.

Acknowledgments

As much as I would like to pretend that the thesis in front of you is immaculately conceived from my brain alone, of course it is not. So many people have had an influence, small or large, on this research and on me during my time as a PhD, that thanking them all will become 'prohibitively difficult'. If I missed you, I'm sorry!

The only reasonable person to start with is of course **Christophe**. It's more than six years ago since I joined your Origins of Life course because it 'sounded cool' and because BN courses were rumored to be easier than condensed matter physics. Over those years, you were a profound influence on my development as a scientist and as a person. You've been the main counterforce against my impatience, triggering me to consider again, do more controls, and change the phrasing of sentences yet again. I will never forget the discussions we had where both of us seemed incapable to agree, but then, after a good night sleep, we found out that one of us (usually you) was right all along. I always felt free to not agree ('high-content' in a paper title, come on...) and trying to convince you is what triggered me to become a proper scientist. Moreover, you always left space for me to find a proper work-life balance, stressing that a PhD is a marathon, not a sprint. Thanks for that.

Then come my lab mates, starting with **Pauline.** As my master thesis supervisor, you introduced me to the lab, to the synthetic cell, and inspired me to be a scientist. As a colleague, you continued to be some kind of mentor to me. **David,** most of my PhD and your postdoc have been all about our shared project. I loved working with you, and I've truly come to learn what a funny and kind man is hiding behind those super-nerdy glasses. **Elisa**, you're maybe the most talented scientist I know, and a very nice person too. I can only stand in awe of your work ethic. **Johannes**, you're a cynic, but in a good way. I loved having you as a lab mate. **Jonas**, el rey del futbolín. **Alicia**, you were weird enough to survive the male-dominated phase of the lab, and you always make me laugh. **Zhanar**, your dedication to science is unrivalled and very inspiring. **Ana,** you're always surrounded by a cloud of fun. Ik vind het heel knap hoe snel je Nederlands leert. **Anne**, if I've learned anything these last five years, it's to listen to you. Writing the thesis, I realized how many awesome ideas actually came from you. **Ilja**, you managed to ground our super-international lab firmly in Rotterdam. **Essie**, I love how you still think I'm a nice guy. **Fab, Huong, Zohreh, Andrew**, although we spend little time together in the lab, you've all been very nice.

My favorite part of the PhD was working with students. **Miranda,** you dared to become my very first student and even getting hit by a car didn't stop you from performing. **Mees**, last of the physicists! I am still sorry your project was so short; I would have loved to have worked longer with you. **Jurjen**, when you walked into class in your Feyenoord shirt, I couldn't image we would be working together so nicely for almost a year after that. **Adriana**, your kind words for me in the acknowledgments of your (awesome) thesis were a highlight of my PhD. I'm sure you'll inspire a whole new generation of students in Munich, with your intelligence, self-deprecating humor, and love of pure science. **Mats**, due to this whole corona situation, you've outlived me in the lab. Your ability to think outside of the box is a rare gift. **Celine, Ilias, Jard**, I might not have supervised you, but you've been such an integral part of my time in the lab that I have to thank you.

Most of my social life at BN revolved around the foosball table. **Thijs, Luuk, Sam, Alberto, Alessio, Antony, George,** all the others I might forget, thank you for all the fun. Special thanks to **Fede**, how you addressed mental health so publicly was very supportive.

Curriculum vitae

Duco Martijn Blanken

19-12-1991   Born in Harderwijk, The Netherlands

Education

2003-2009   VWO
                   Augustinianum, Eindhoven, The Netherlands

2009-2013   Bachelor Applied Physics
                   TU Delft, Delft, The Netherlands

2013-2016   Master Applied Physics
                   TU Delft, Delft, The Netherlands

2016-2020   PhD in synthetic biology
                   TU Delft, Delft, The Netherlands
                   Thesis: *Genetically encoded phospholipid production for autonomous synthetic cell proliferation*
                   Promotor: Dr. C.J.A. Danelon
                   Co-Promotor: Prof. Dr. A.M. Dogterom

List of publications

Van Nies, P., Westerlaken, I., **Blanken, D**., Salas, M., Mencia, M. & Danelon, C. Self-replication of DNA by its encoded proteins in liposome-based synthetic cells. *Nat Commun* **9,** 1583 (2018).

Blanken, D., van Nies, P. & Danelon, C. Quantitative imaging of gene-expressing liposomes reveals rare favorable phenotypes. *Phys. Biol.* **16**, 045002 (2019).

Blanken, D., Foschepoth, D., Serrão, A.C. & Danelon, C. Genetically controlled membrane synthesis in liposomes. *Nat Commun* **11,** 4317 (2020).