TÉCNICO LISBOA

Imaging synthesized phosphatidylserine incorporation in single liposomes

Adriana Calaça Serrão

Thesis to obtain the Master of Science Degree in

Biological Engineering

September 2018

Supervisor(s)

Prof. Christophe Danelon

Prof. Gabriel António Amaro Monteiro

Examination Committee

Chairperson: Prof. Arsénio do Carmo Sales Mendes Fialho

Supervisor: Prof. Gabriel António Amaro Monteiro

Member of the Committee: Prof. Fábio Monteiro Fernandes

The work presented in this thesis was performed at the Department of Bionanoscience of Technische Universiteit Delft (Delft, Netherlands), during the period February-September 2018, under the supervision of Prof. Christophe Danelon, Duco Blanken and Dr. David Foschepoth, and within the frame of the Erasmus programme. The thesis was co-supervised at Instituto Superior Técnico by Prof. Gabriel Monteiro.

Abstract

Reconstituting lipid synthesis inside self-assembled lipid vesicles is fundamental to achieve growth and division of a bottom-up minimal cell. Detection of cell-free lipid synthesis encapsulated in liposomes has been successfully performed in bulk. However, the incorporation of the synthesized lipids in the membrane of single-vesicles has not been demonstrated yet. With this work, we aim to detect phospholipid incorporation within the liposome membrane. To test this, we encapsulated the genes of the kennedy pathway for lipid synthesis in *Escherichia coli* except *psd*, together with the cell-free gene expression protein synthesis using recombinant elements (PURE) system. The liposomes were incubated with the enhanced green fluorescent protein (eGFP)-C2 domain of lactadherin (lactC2) protein probe, which has shown to bind specifically to phosphatidylserine (PS). By image analysis using confocal microscopy, we showed that there is clear colocalization of the eGFP signal with the membrane in the majority of liposomes and that, on average, 55% of lipid vesicles are enriched in PS. The time tracking of enriched liposomes allowed to have a rough estimate of the lipid synthesis and incorporation kinetics: between 3.1 and 5.7 hours. The quantification of the incorporated phospholipid levels on the membrane was not possible, since the characterization of the method revealed that there is not a linear correlation between fluorescence intensity and PS concentration. We believe this work demonstrates the membrane incorporation of intravesicular synthesized phospholipids and poses as an advance for future studies regarding the growth and division of a liposome-based artificial cell.

Keywords: bottom-up synthetic biology, minimal cell, lipid synthesis, confocal fluorescence microscopy, PURE system

Resumo

Reconstituir a síntese de lípidos no interior de vesículas lípidicas é fundamental para atingir o crescimento e divisão de uma célula artificial com os componentes mínimos para ser considerada viva, numa abordagem *bottom-up*. A deteção da síntese de lípidos no interior de lipossomas foi, até à data, apenas realizada por métodos que analisam a população total de vesículas. No entanto, a incorporação de lípidos síntetizados na membrana de vesículas individuais ainda não foi demonstrada. Este projeto tem o objetivo de detetar a incorporação de fosfolípidos na membrana de lipossomas. Para isto, os genes da via metabólica Kennedy, responsáveis pela síntese de fosfolípidos em *Escherichia coli* e o sistema transcrição-tradução *in vitro* PURE foram encapsulados. Os lipossomas foram incubados no exterior com a proteína eGFP-lactC2, que demonstrou especificidade para a fosfatidilserina (PS). Utilizando microscopia confocal de fluorescência, a colocalização do sinal de eGFP com a membrana da maioria dos lipossomas foi observada. Adicionalmente, demonstrou-se que 55% das vesículas estão enriquecidas em PS. Através do seguimento temporal dos lipossomas enriquecidos, foi possível obter uma estimativa da cinética de síntese e incorporação de fosfolípidos, no sistema desenvolvido: entre 3.1 a 5.7 horas. A quantificação dos níveis de fosfolípidos incorporados não foi possível, dado que a caracterização do método revelou uma relação não linear entre a intensidade de fluorescência na membrana e a concentração de PS. O presente trabalho comprova a incorporação de fosfolípidos sintetizados no lúmen de lipossomas na membrana das respetivas vesículas. Estes resultados são considerados pioneiros, possibilitando estudos futuros com respeito ao crescimento e divisão de uma célula artificial.

Keywords: biologia sintética *bottom-up*, célula artificial, síntese de lípidos, microscopia confocal de fluorescência, sistema PURE

Contents

List of Tables

List of Figures

List of Acronyms and Abreviations

18:1 CL 1',3'-bis[1,2-dioleoyl-sn-glycero-3-phospho]-glycerol

CL cardiolipin

CDP cytidine diphosphate

CDP-DAG cytidine diphosphate - diacylglycerol

CTP cytidine triphosphate

CdsA cytidine diphosphate - diacylglycerol synthase A

ClsA cardiolipin synthase A

CMP cytidine monophosphate

CoA coenzyme A

BSA bovine albumin serum

DOPC 1,2-dioleoyl-sn-glycero-3-phosphocholine

DOPE 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine

DOPG 1,2-dioleoyl-sn-glycero-3-phosphoglycerol

DOPS 1,2-dioleoyl-sn-glycero-3-phospho-L-serine

DHPE-TexasRed Texas Red 1,2-Dihexadecanoyl-sn-Glycero-3-Phosphoethanolamine

DPPS 1,2-dipalmitoyl-sn-glycero-3-phospho-L-serine

DSPE-PEG(2000)-biotin 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[biotinyl(polyethylene glycol)-2000

eGFP enhanced green fluorescent protein

EGF epidermal growth factor

FadD fatty acyl-CoA synthetase

fV factor V

fVIII factor VIII

G3P glycerol 3-phosphate

GUV giant unilamellar vesicles

GPAT glycerol 3-phosphate acyl transferase

IPTG isopropyl beta-D-1-thiogalactopyranoside

IVTT *in vitro* transcription translation

lactC2 C2 domain of lactadherin

LC liquid chromatography

LPA lysophosphatidic acid

LPAAT lysophosphatidic acid acyltransferase

Ni-NTA nickel-nitrilotriacetic acid

NBD nitrobenzo-2-oxa-1,3-diazole

o-CoA oleoyl-Coenzyme A

p-CoA palmitoyl-Coenzyme A

PA phosphatidic acid

PC phosphatidylcholine

PE phosphatidylethanolamine

PG phosphatidylglycerol

PGP phosphatidylglycerolphosphate

PgpA phosphatidylglycerophosphatase A

PgsA phosphatidylglycerophosphate synthase A

POPS palmitoyl-oleoyl-sn-glycero-3-phospho-L-serine

PS phosphatidylserine

PssA phosphatidylserine synthase A

Psd phosphatidylserine decarboxylase

PEG polyethylene glycol

PURE protein synthesis using recombinant elements

SUV small unilamellar vesicles

T_m melting temperature

MS mass spectrometry

TP terminal protein

YFP yellow fluorescent protein

Chapter 1

Introduction

Life, as we know it, is extremely complex. However, is this complexity essential for life? Which is the minimal set of components that a system should have to be considered *alive*? Is it possible to experimentally create a system with the minimum complexity possible? The minimal cell project attempts to tackle these questions, engaging in one of the greatest challenges of the synthetic biology field. Particularly, the Christophe Danelon lab focuses on a bottom-up approach, constructing a contemporary-based living organism by assembling it from the minimal amount of biochemical parts. The creation of a minimal cell will certainly elucidate about the principles of life. Moreover, numerous applications can be envisioned for a functional and tunable cell scaffold, for instance as a starting point to higher complexity organisms with specific functions. Throughout the next section, the notion of minimal cell and the current strategies to reach its construction will be addressed.

1.1 The minimal cell

A minimal cell is a system that has the minimal albeit sufficient structural components to be considered alive. The construction of a minimal artificial living entity relies on a specific definition of being *alive*. In order to define this notion, it has been postulated that the minimal cell should be a self-sustained system and capable of reproduction and evolution [**Luisi2002a**]. Considering this definition, the essential characteristics of living organisms can be further detailed (Fig. 1.1): **compartmentalization** of cellular function, presence of a **flow of information** and capacity of **proliferation** [**Caschera2014**]. It is important to note that this definition is deduced from the empirical analysis of contemporary organisms, not excluding the possibility of other minimal cell setups either with artificial or natural components not yet described. In contemporary living cells, the flow of information relies in nucleic acids serving as coding and memory and in proteins as functional agents. The compartment, consisting of a phospholipid membrane in all known organisms, is responsible for providing cells with a boundary and selectively allow nutrients to enter the cell and allow communication with the environment. Lastly, the proliferation comprises all the functions regarding growth, reproduction and evolution. This includes for instance lipid synthesis, DNA replication and cell division.

There are currently several lines of action being researched to tackle the construction of the simplest living cell under the clarified concept, which can overall be grouped in two approaches, the bottom-up and the top-down approach [**Luisi2002**].

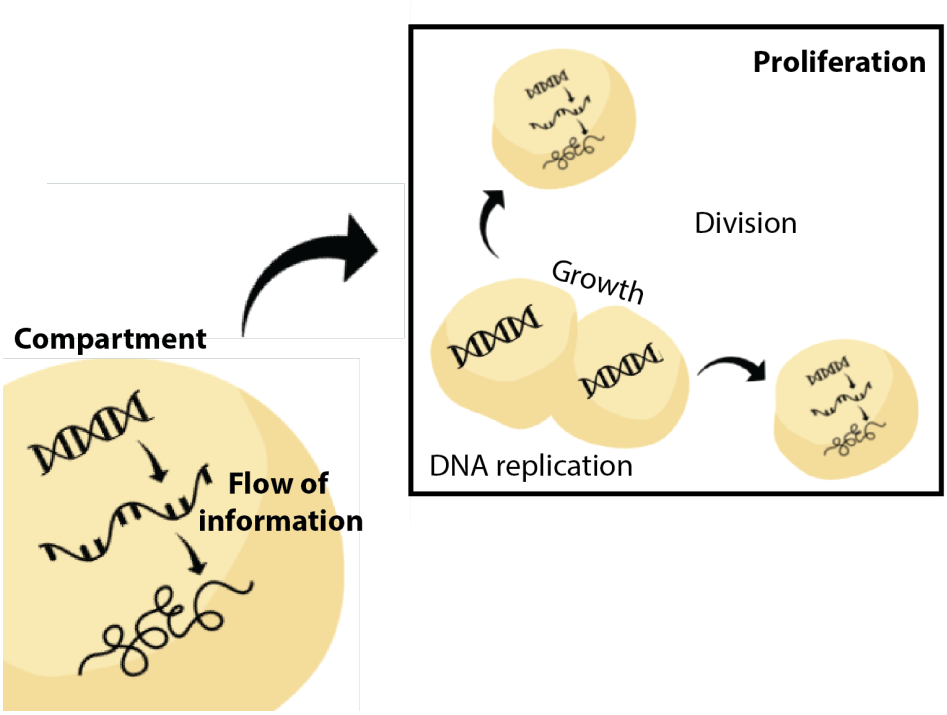


Figure 1.1: The essential features of the minimal cell.

1.1.1 Approaches to the minimal cell

The top-down approach consists of simplifying an already existing living system, e.g. removing non-essential genes, while maintaining viability. Most of the efforts in this field are focused on achieving a minimal genome commonly performed by successive deletion, while using the chassis of the starting-point organism [**Jewett2010**]. There have been successful attempts to reconstitute a minimal genome by leaving only the essential or quasi-essential genes present in a small microorganism chromosome. For this, species of the genus *Mycoplasma* have been recently used as they are the simplest species known capable of autonomous growth. In the work of Venter and colleagues, the genome of *Mycoplasma mycoides* (which originally had 985 putative genes) was sequentially reduced until it reached 473 genes, less than any known natural genome of a self-replicating free-living organism. From these 149 cannot be assigned a biological function [**Hutchison2016**]. This reflects the major disadvantage of the top-down approach: the fact that it does not necessarily allow for complete comprehension of each of the parts of the minimal organism obtained.

Alternatively, the bottom-up approach consists in pooling together essential purified macromolecules within a compartment. This way, all the components are of known function, easily manipulated and defined. In order to tackle the bottom-up minimal cell, two lines of action have been explored: encapsulation of either externally purified overexpressed proteins or of a minimal genome and a cell-free expression system. In the latter, protein expression occurs within the compartment, resembling more closely natural organisms. In summary, the bottom-up approach poses a fully synthetic methodology for the construction of a minimal cell, as it is assembled from the minimal set of separate parts. The knowledge of the current living organisms restricts the possibilities for a bottom-up approach. There are several possible configurations of a minimal cell and potentially unknown organisms with simple and efficient biochemical parts that could be applied. Presently, the

bottom-up efforts are mostly inspired by well described organisms so that each of the parts is fully understood.

1.1.2 Implementation of the minimal cell

The approach developed in our lab is inspired by contemporary living organisms and consists of building a DNA-based cell compartmentalized by a phospholipid membrane. For this, we use cell-free expression systems, encapsulated inside liposomes, to ultimately construct a bottom-up minimal cell. The vesicles should also entrap a minimal genome which allows mainly for functions related with proliferation and maintenance of the flow of information.

Compartment Compartmentalization is important to define the border of the cell and control the concentration of intracellular components. Ultimately, the compartment allows the cell to achieve homeostasis. The liposome, composed of phospholipids, is a synthetic analog of the contemporary bacterial cell membranes. For this reason, it is the chosen compartment for our minimal cell. Liposomes are vesicles composed of one or more phospholipid bilayers. Due to the amphiphilic nature of phospholipids, when hydrated in aqueous solution they self-assemble as vesicles, which allows encapsulation of components present in the solution. This bilayer conformation, with the hydrophobic tail pointing inwards and at each other and the hydrophilic head pointing outwards at the aqueous solution (in the lumen or the external medium) is more entropically favourable than a single layer, since it reduces the interaction of non-polar tails with the polar solution. Liposomes can be classified regarding their size - small (R<200nm) , large (200nm<R<1μm) and giant (>1μm) - and their lamellarity (uni- or multi-) [**Akbarzadeh2013**]. For this project, giant unilamellar vesicles (GUV) were used to facilitate encapsulation of all the components.

Flow of information The flow of information assures that the genetic information is translated to proteins that exert desired function, as stated on the central dogma of molecular biology. To perform transcription and translation within the liposome, a cell-free gene expression system should be encapsulated. These systems have been assembled both by purification of synthesized proteins or crude cell extracts [**Jewett2010**], based mostly on *E. coli*. Even though crude extracts have higher yield and reduced costs, their main disadvantage is the contamination with poorly characterized cytoplasmic components that might interfere with protein synthesis. The system for *in vitro* PURE solves this problem by having the purified enzymes and co-factors, with well-defined components and concentrations [**Nourian2014**].

The PURE system, an IVTT system, includes mostly elements purified from *E. coli* and the T7 RNA polymerase. These are the minimum elements necessary to allow protein synthesis from a template DNA in cell-free environment. Besides DNA, precursor molecules (i.e. aminoacids, nucleotides), translation factors, enzymes, energy sources, tRNA and ribosomes are the main components that should also be present in adequate buffer concentrations to recreate the current living systems flow of information [**Shimizu2005**]. Parallel to the protein synthesis reactions, there are also processes focused on recycling the nutrients, such aminoacylation of tRNAs and energy source regeneration. These four functions (transcription, translation, aminoacylation and energy restoration) are the main processes carried by the PURE system components and are represented in Fig. 1.2.

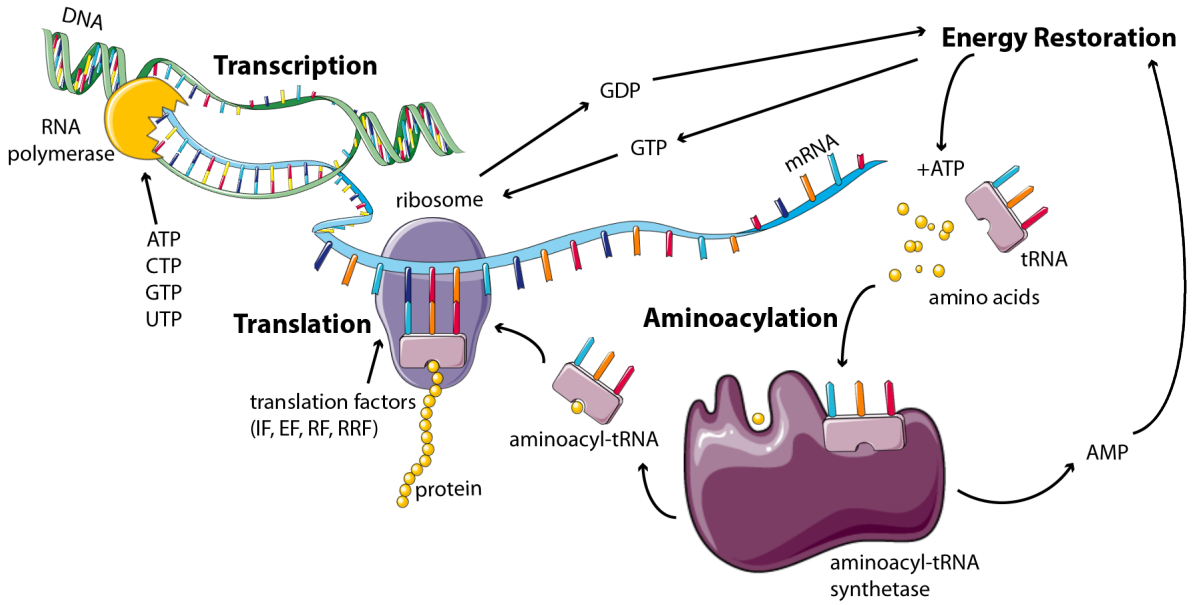


Figure 1.2: The four main reactions of cell-free protein synthesis. Image based on the PURE sytem description by Shimizu *et al.*, 2005 [**Shimizu2005**].

Proliferation Ultimately, a minimal cell should not only be able to perform gene expression but also to divide and generate offspring. In order to do this, the compartment should expand so that the parent cell has enough volume to create two daughter cells, by synthesizing and incorporating lipids in the membrane. This expansion could be the trigger of the division by unbalanced surface-to-volume ratio and shape deformation leading to scission. On the contrary, the division can be provoked by proteins known to be implicated in this process in contemporary organisms. In particular, the FtsZ and the ESCRT-III system are being studied for application in the minimal cell [**Nourian2014**]. Moreover, other processes like DNA replication and synthesis of the PURE system components are necessary to for independent cell proliferation. The growth and division should be coupled with component regeneration otherwise the dilution of the molecules with each cell cycle would inhibit any metabolism in subsequent generations, unless the nutrients are constantly fed in the medium.

Experimental approach Liposomes are generated by natural swelling. This method offers advantages when comparing to other preparation methods (e.g. water-in-oil emulsion transfer, microfluidic devices) to create large or giant vesicles. The most considerable advantage of the natural swelling technique is its flexibility, permitting the production of vesicles with almost unrestricted membrane composition. This allows the introduction of dye-conjugated lipids for fluorescence detection, the tuning of physical properties of liposomes by manipulation of their composition (e.g. permeability), the increase of their stability by insertion of polymer-conjugated lipids and, moreover, their immobilization in glass surfaces for imaging [**Nourian2012**].

After a dry lipid film is deposited, giving rise to a thin layer, the water molecules of the swelling solution penetrate the lipid bilayer. Attracted by the polar head group of the phospholipids, the molecules accumulate, causing swelling and eventually the budding off of the liposome [**Akbarzadeh2013**]. In our lab the original protocol was modified to increase liposome and encapsu-

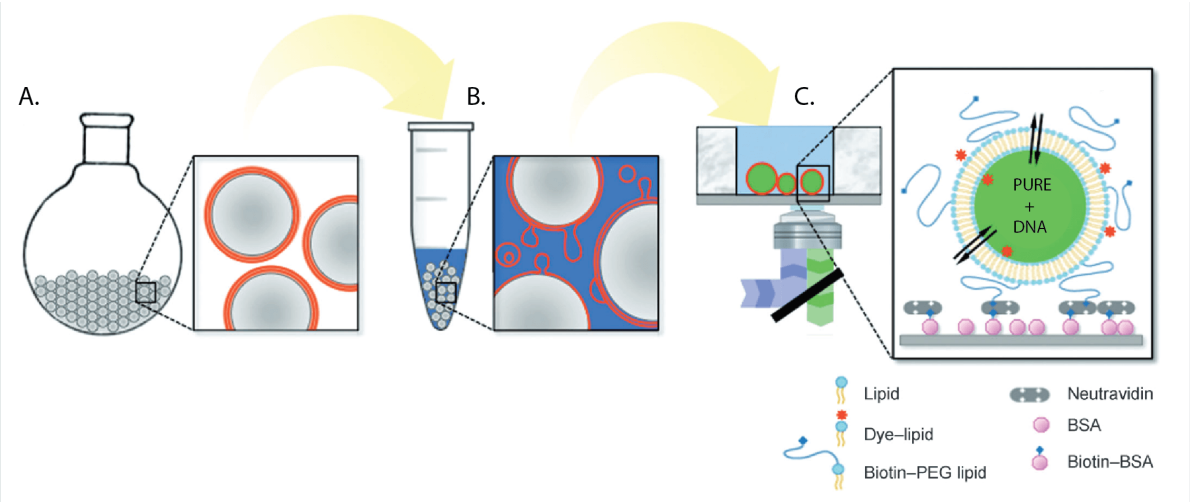


Figure 1.3: Experimental setup to generate liposomes and encapsulate PURE system. A. Lipid-coated beads are assembled in a round-bottom flask with the desired lipid composition. **B.** Beads are submerged on the swelling solution that should contain the components to be encapsulated. Liposomes are formed. **C.** The liposomes are transferred to a chamber, immobilized with the biotin-steptavidin method and imaged with microscopy. Image adapted from Nourian *et al.*, 2012 [**Nourian2012**]

lation yield. Using porous submillimeter glass beads instead of round bottom flasks to deposit the lipid film, increases the surface area and leads to a higher concentration of liposomes and better entrapment of macromolecules. Moreover, encapsulation of PURE system is now economically feasible, since this setup allows for microliter volume applications [**Nourian2012**]

In this project, the lipid film included lipids conjugated with a fluorescent dye and lipids bound to both polyethylene glycol (PEG) and biotin. These lipids allow liposome stabilization through the PEG spacer and immobilization to the glass chamber surface with the biotin-streptavidin approach. The encapsulated solution is composed of the DNA template, PURE system and other precursors. The extravesicular solution can also contain PURE system buffer, to maintain osmotic pressure similar both inside and outside and nutrients that are able to diffuse through the membrane.

1.2 Lipid synthesis and membrane growth *in vitro*

One of the milestones for creating a synthetic cell is the production and incorporation of the membrane constituents through internally synthesized enzymes. The presence and diversity of these constituents - the phospholipids, which compose the membranes of all known organisms - is essential for many cellular processes. This project focuses on the detection of lipids incorporated in the liposome membrane, leading to the growth of the vesicle. Therefore, bacterial phospholipids and the routes for lipid synthesis that are being reproduced in the minimal cell will be detailed.

1.2.1 Phospholipids: role, morphology and diversity

Phospholipids play not only an important function defining the permeable barrier of cells, but constitute also the matrix that influences the properties of the proteins associated with the membrane [1]. The chemical nature of the lipid constituents, as the headgroup and the carbon chains, can affect protein binding and the physical structure of the membrane.

Lipids can assemble in various structures, and can be considered bilayer-forming (when they have a cylindrical shape) or nonbilayer-forming (when having an inverse conical shape that imposes negative curvature on the membrane), a phenomenon called lipid polymorphism [2]. The ratio between the two types of lipids is an important factor in membrane functionalization: the nonbilayer-forming lipids can create discontinuities that may play a role in membrane fusion and vesicle formation, movement of macromolecules through the membrane, stabilization of protein complexes, lateral lipid movement and non-lipid component integration [1].

Although the variety of headgroups is large, the key precursor for all glycerol-based phospholipids is phosphatidic acid (PA), which consists of a glycerol esterified with two fatty acids and a phosphoric acid. This lipid is a nonbilayer-forming lipid and, therefore, the membranes could not be completely formed of PA. The headgroup modifications that PA suffers give rise to the diversity of lipids that constitute the bacterial membranes and that can be classified as zwitterionic or anionic.

Zwitterionic phospholipids For most of the biological membranes, zwitterionic phospholipids (mainly phosphatidylethanolamine (PE)) comprise the majority of the membrane composition (>70%). However, phosphatidylcholine (PC), also a zwitterionic phospholipid can be found in some organisms. Whereas the phosphate group is combined with choline in PC, or it is combined with ethanolamine in PE, the lipid backbone is PA. PC has been preferentially used in *in vitro* studies due to its facility in forming vesicles possibly due to its cylindrical shape. On the other hand, PE has a smaller head group and is capable of forming hydrogen bonds, but has an inverse cone shape that does not allow the formation of a lipid bilayer on its own. These lipids are known to play an important role in the folding and assembly of proteins and to support active transport by *E.coli* permeases [1].

Anionic phospholipids Present in almost all bacterial cell membranes, anionic phospholipids (phosphatidylglycerol (PG) and cardiolipin (CL)) have similar structures. While PG has the PA structure with a glycerol headgroup, CL has two phosphatidic acid moieties connected with a glycerol backbone in the center to form a dimeric structure. This means it has four alkyl groups and two potential negative charges. Although under normal physiological conditions (wherein pH is around 7), the molecule may carry only one negative charge, as its second pKa is much higher than the first. PS, another anionic lipid that has a L-serine as a headgroup which is less common in bacterial membranes [3]. It is believed that the presence of these anionic lipids influence the replication initiation and affect the translocation of proteins across membranes.

Fatty-acid diversity Fatty-acids, the molecular precursors that, esterified to the glycerol's carbons, give rise to the alkyl chains in phospholipids also contribute to the phospholipid variety. These carbon chains can be diverse and its chemical nature can vastly influence the physical characteristics of the lipid, in particular the melting temperature (T_m). The T_m can be defined as the temperature from which the bilayer changes from lipid-ordered (where the carbon tails are fully extended and packed) to the lipid-disordered phase (a more flexible structure). In both phases the lipids are constrained to the two dimensional plane of the membrane, but in liquid phase bilayers the molecular diffusion happens more freely.

The liquid phase behaviour is affected by the carbon-chain length and the level of unsaturation. Regarding the former, the longer the lipid is, the stronger the Van der Waals interaction are with neighbouring lipids, leading to less flexibility of movement and higher T_m. The presence of an unsaturation has a contrary effect, producing a disruption in the structure which creates free space in the bilayer and leads more freedom of movement. For this reason, the higher the number of unsaturations, the lower the T_m.

Lipid flip-flop Phospholipids can not only move within each leaflet of the membrane, but can also exchange leaflets, in a process called flip-flop. Moreover, the distribution of lipids across the membrane is commonly asymmetrical, and this asymmetry is known to play an important role in a series of functions, mostly regarding membrane mechanical stability and modulation of the activity of membrane proteins [4]. The flip-flop can be spontaneous or catalyzed by proteins: flippases, floppases and scramblases [5]. In the case of spontaneous flip-flop, the kinetics of the process has been studied and is strongly dependent upon the nature of the lipid, the composition and the physical state of the membrane, ranging from the millisecond (small lipids) to hours or days [6]. It is believed membrane defects that possibly arise from medium conditions (temperature, osmotic pressure, pH) can speed up the process of spontaneous lipid flip-flop.

The diversity of both phospholipid morphology (on the headgroup level or alkyl chain composition) and distribution is concluded to influence many key cellular processes. A minimal cell should be able to synthesize the necessary yet sufficient number of phospholipids, and for this, have the genes encoding for the necessary pathways encapsulated.

1.2.2 Kennedy pathway: *E.coli* phospholipid metabolism

As in most microbiology studies, *E.coli* has also been the standard model for the study of bacterial membrane lipids. In its membrane, three major lipids are accumulated: PE (about 75%), PG (about 20%) and CL (about 5% or lower). Although, the relative quantities depend on the growth stage of the bacterium [7]. Herein, the Kennedy pathway [8] for the formation of these phospholipids will be described.

Synthesis of PA and CDP-DAG The first step in phospholipid synthesis is the conversion of free fatty-acids in acyl-CoA, needed for subsequent acylation steps. This reaction is catalyzed by the cytosolic enzyme fatty acyl-CoA synthetase (FadD). The glycerol backbone of the universal lipid precursor PA is derived from one molecule of glycerol 3-phosphate (G3P) [10]. PA synthesis is done in two sequential steps, starting with the acylation of the *sn-1* position of G3P by glycerol 3-phosphate acyl transferase (GPAT), using acyl-CoA (coenzyme A (CoA)) as a substrate, giving rise to lysophosphatidic acid (LPA). A second acylation is performed in the *sn-2* position, by lysophosphatidic acid acyltransferase (LPAAT)[11]. GPAT, coded by the gene *plsB* is an integral membrane protein whereas LPAAT, coded by *plsC* is a membrane-bound protein [12]. The lipid cytidine diphosphate - diacylglycerol (CDP-DAG), considered the common intermediate for the synthesis of several phospholipids, is synthesized by the simultaneous dephosphorylation of cytidine triphosphate (CTP) to cytidine diphosphate (CDP) and attachment to PA as a headgroup, a reaction catalized by the membrane-bound cytidine diphosphate - diacylglycerol synthase A (CdsA). This lipid does not accumulate, as it believed the CdsA enzyme is

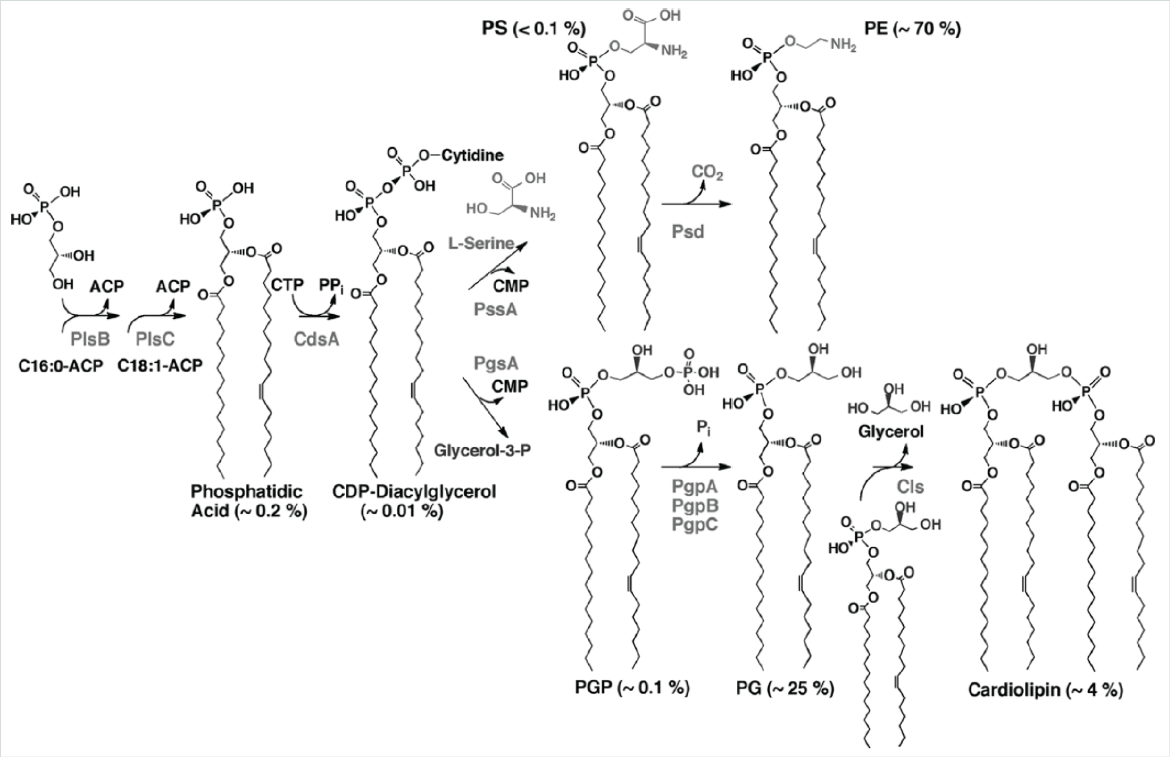


Figure 1.4: Kennedy pathway for phospholipid synthesis in *E.coli*. The genes that catalyze each step are in grey. The relative amount of each phospholipid in the *E.coli* membrane is in brackets. Image extracted from Lu *et al.* [9]

product inhibited [13], and can follow two synthetic routes. One of the derived pathways leads to the synthesis of PS and PE and the other to the synthesis of PG and CL.

Synthesis of PS and PE To form PS, phosphatidylserine synthase A (PssA) condenses L-serine with CDP-DAG and releases cytidine monophosphate (CMP) and a proton [12]. Although PS is capable of accumulating and incorporating in the membrane in some species , normally the subsequent conversion, from PS to PE, is very efficient. This conversion is catalyzed by phosphatidylserine decarboxylase (Psd) and consists in the decarboxylation of the L-serine headgroup, by its reaction with a proton and the subsequent release of carbon dioxide [3] [12]. The Psd enzyme is a heterodimer (with subunits alpha and beta) located at the inner membrane, produced from a single peptide cleaved after translation [12]. This corresponds to the top branch of the pathway presented in Fig. 1.4.

Synthesis of PG and CL On the other hand, CDP-DAG can also condense with G3P leading to the synthesis of phosphatidylglycerolphosphate (PGP). This reaction is catalyzed by phosphatidylglycerophosphate synthase A (PgsA) and releases CMP and a proton [12]. PgsA is an integral membrane protein located in the inner membrane and completely magnesium dependent [12]. PGP is then dephosphorylated by phosphatidylglycerophosphatase A (PgpA) to yield PG. CL can be synthesized with two molecules of PG, by the enzyme cardiolipin synthase A (ClsA). This corresponds to the bottom branch of the pathway presented in Fig. 1.4.

1.2.3 Implementation of phospholipid synthesis in a minimal cell

To date some attempts have been made to stimulate phospholipid vesicle growth. In early attempts, PC was produced inside liposomes, encapsulating purified enzymes (GPAT, LPAAT, and two more specific enzymes for PS synthesis), isolated from pig liver [14]. Recently there was also the encapsulation of all the purified enzymes needed for the production of PE and PG (described in subsection 1.2.2), from the fatty-acid precursors [13]. Both of these efforts were done by feeding the externally produced proteins, whereas a truly synthetic cell should be able to produce its own components from within.

In opposition to the approach that relies exclusively on the use of purified proteins, a paradigm based on cell-free gene expression seems more viable. The production of phospholipids using an IVTT system encapsulated inside liposomes was first done by Kuruma *et al.*, 2009 [15]. In this work the *plsB* and *plsC* genes were encapsulated together with the PURE system, each at once, with the objective of producing the enzymes GPAT and LPAAT, leading to PA. After the proteins were expressed, the extracts were incorporated in liposomes and the precursors were fed to the system, having achieved active enzymes, which, however, failed to work in the same environment. Moreover, gene expression and lipid synthesis were not integrated and PA, although being the universal phospholipid precursor, is not commonly present in bacterial membranes. As a nonbilayer-forming lipid, it could contribute to incompatibilities with membrane-proteins if it were the only lipid synthesized to stimulate liposome growth.

Recently, Scott *et al.* [12] reported the production of all the Kennedy pathway enzymes in bulk starting from G3P and acyl-CoA precursors with PURE system. The PE and PG lipids were detected when synthesized in a pathway-selective manner. Moreover, for the first time, the membrane incorporation of *in vesiculo* synthesized phospholipids was detected. The cell-free expression with PURE system of both GPAT and LPAAT was performed inside liposomes and the membrane was enriched in PA. Since the publication of this work, a minimal genome with the seven kennedy pathway genes was constructed (pGEMM 7.0) and the PE and PG could be simultaneously detected with mass spectrometry (MS) (unpublished work by Duco Blanken and Dr. David Foschepoth).

However, the verification through real-time observation of lipid incorporation and potential membrane expansion has not been done yet. The project presently described is a follow-up of the previously done work, using imaging techniques that allow to follow single liposomes and investigate the incorporation of the produced lipids in the membrane.

1.3 Lipid analysis techniques

1.3.1 Bulk quantification

Thin-layer chromatography

Thin-layer chromatography (TLC) is a widely used technique for the detection of lipids in a mixture. This technique allows the separation of components in non-volatile mixtures of smaller and apolar compounds [16].

This technique is an based on the principles of affinity chromatography. After the sample is transferred to a plate, its different components migrate in different rates due to the differences in

their attraction to the stationary phase comparing with the mobile phase. The lower the affinity of the compound is with the stationary phase, based on its polarity, the higher it migrates. Since the plates normally contain a fluorescent indicator, the localization of the compounds can be observed using ultra-violet light. When the molecules absorb in this range of wavelenghts, the fluorophore is quenched and dark spots appear in the correspoding area [17].

TLC is a robust, convenient and simple classification method for phospholipids, separating the molecules according to the polar groups. Despite disadvantages of TLC as a high time demand, low sensitivity, and limited quantification precision, it is still routinely used as the first step for the analysis of a lipid mixture [18].

Mass spectrometry

Mass spectrometry (MS) is an analytic technique used for the quantification of materials, identification of unknown compounds within a sample, and to elucidate the structure and chemical properties of different molecules. Overall the process involves the sample conversion into gaseous ions, their fragmentation, and subsequent separation and analysis [19]. In the ionization step, gaseous radical ions are created which normally leads to molecular fragmentation. The products can, in turn, undergo further fragmentation, and so on. These molecules are separated based on their mass-to-charge (m/z) ratios and detected in proportion to their abundance, giving rise to a mass spectrum. [19].

The generated ions provide information concerning the structure of their precursor molecule. When analysing a pure compound, the molecular ion appears at the highest value of m/z (followed by ions containing heavier isotopes) and gives the molecular mass of the compound [19].

To allow the analysis of a complex mixture, as in lipidomic approaches, it is convenient to couple a chromatography (usually liquid chromatography (LC) for this purpose) prior to the sample injection, resulting in the separation of the components of the mixture. This strategy allowsto separate or concentrate different classes of compounds according to their physicochemical properties [20].

Overall, MS is useful for lipidomic assays, leading to quantitative analysis of lipid classes and the distinction of different molecular species, by the differences of carbon chains or headgroups. When dealing with liposome encapsulated gene expression, this technique enables to analyze the lipids produced in the bulk of the population. However, the heterogeneity of phenotypes between liposomes produced by natural-swelling is immense, resulting from the amount of compounds involved in gene expression necessary to encapsulate [**Nourian2012**]. For this, it is important to analyze the lipid production on the single liposome level.

1.3.2 Microscopy: molecular probes to visualize lipid location

Microscopic visualization of lipids poses as a powerful technique that is complementary to mass-spectrometry-based analysis. Since most of the lipids are not intrinsically fluorescent, many molecular probes that allow visualization have been described [21].

Probes for lipid imaging

Probes can be divided in two different categories: (1) Fluorophore-conjugated lipids, (2) Lipid-binding probes (which can be genetically encoded) [21].

Using fluorophore conjugated lipids is an adequate strategy to follow the movement and localization of the lipids. These fluorophores can, for instance, be attached to the molecular precursors (either the headgroup precursor or the acyl-CoA) for lipid synthesis which are, in turn, added to the growth medium. One example is nitrobenzo-2-oxa-1,3-diazole (NBD), which is a mildly polar molecule that can be attached to the acyl chain or the headgroup. The major concern with this approach is that these attached molecules can lead the lipids to have different chemical characteristics than the natural counterparts, biasing the analysis [21].

Lipid-binding probes consists of a group of fluorescently-tagged lipid-binding proteins and (conjugated) antibodies. In the case of antibodies, there are only a few that recognize phospholipids, but with some disadvantages: firstly, as the antibody molecule is much larger than the lipid, it recognizes more than one molecule, and, moreover, a small number of lipids have known and described antibodies [21]. This can be explained by the majority of lipids having an endogenous presence in the membranes of the animals used to harvest antibodies, lacking antigenicity [22].

There are, however, a number of protein domains, known for interacting with phospholipids, and that can be expressed fused to a fluorophore to track the position of specific lipids and the dynamic changes when subjected to specific conditions or stimuli. The major setbacks of this method is the varying affinity and specificity of the binding. For the lipids synthesized by *E.coli*, there are described protein probes for PA (like Sos1-PH [23] and Spo20p [24]), PE (like duramycin and cinnamycin peptides [25]) and PS (like Annexin V and the lactC2 [26]).

All of the above mentioned lipids are synthesized in our lab's approach to the minimal cell. Imaging PA would mean encapsulating only the *plsB* and *plsC* genes, since this lipid has a short life and is rapidly consumed. Moreover, the headgroup consists only of a phosphate, which is smaller and less complex than the one of *E.coli* synthesized membrane phospholipids (PE and PG), being a less good model to inquire about the incorporation of the lipids of interest.

In the case of PE, the available probes (duramycin and cinnamycin) are small-peptides that induce morphological changes in the membrane and that are curvature-dependent [25], meaning the essays for its detection would require the usage of smaller vesicles like small unilamellar vesicles (SUV) instead of GUV. However, with such small vesicle radii it would be impossible to encapsulate all the PURE system components.

This work focuses on imaging PS, encapsulating all the genes that allow the conversion of G3P to PS and PG, boycotting only the production of PE, so that PS accumulates. Even though PS is not a desired end product for our lipid synthesis module, studying its incorporation can give insights about the remaining lipids synthesized.

Lactadherin-C2 domain binding PS

Lacthadherin, a 47 kDa glycoprotein, is known to bind PS in a stereospecific way. This has been used to image PS in phospholipid membranes, by fusing the protein domain with a fluorescent probe. In this work, this strategy is employed to access the incorporation of lipids in the membrane of liposomes.

Lactadherin structure Lactadherin is present in cow milk fat globules where it is responsible for surrounding the surface of the phospholipid bilayer and stabilizing it. The C2 domain is responsible for binding phosphatidylserine through the phospho-L-serine motif present in the head group of this phospholipid. The protein comprises two N-terminal epidermal growth factor (EGF) homology domains (EGF1 and EGF2) and has two C-terminal regions (C1 and C2 domains), that share homology with the lipid binding “C” domains of blood coagulation factor VIII (fVIII) and factor V (fV). The presence of three analogous loops that are likely to take part in membrane binding, called "spikes", is common between lactC2 and fV and fVIII. The most evident difference between lactC2 and fV and fVIII structures is in spike 1 where lactC2 has a longer and more complex loop. Specifically, the presence of hydrophobic residues in both this and the spike 3, colored in cyan in Fig. 1.5 **A**, are thought to have an influence in the binding.

Binding characteristics The lactC2 PS ligation is performed in a stereospecific (to the L-form of serine motif) and calcium independent manner [27]. The binding has affinitiy for highly curved membranes, but it is not curvature-dependent, as it was shown to happen in both SUV and GUV. It is believed this is a two-step mechanism, in which the first and faster step is the binding event and the second a slower possible conformational change in the protein-lipid complex or lipid insertion in the membrane (shown in Equation 1.1).

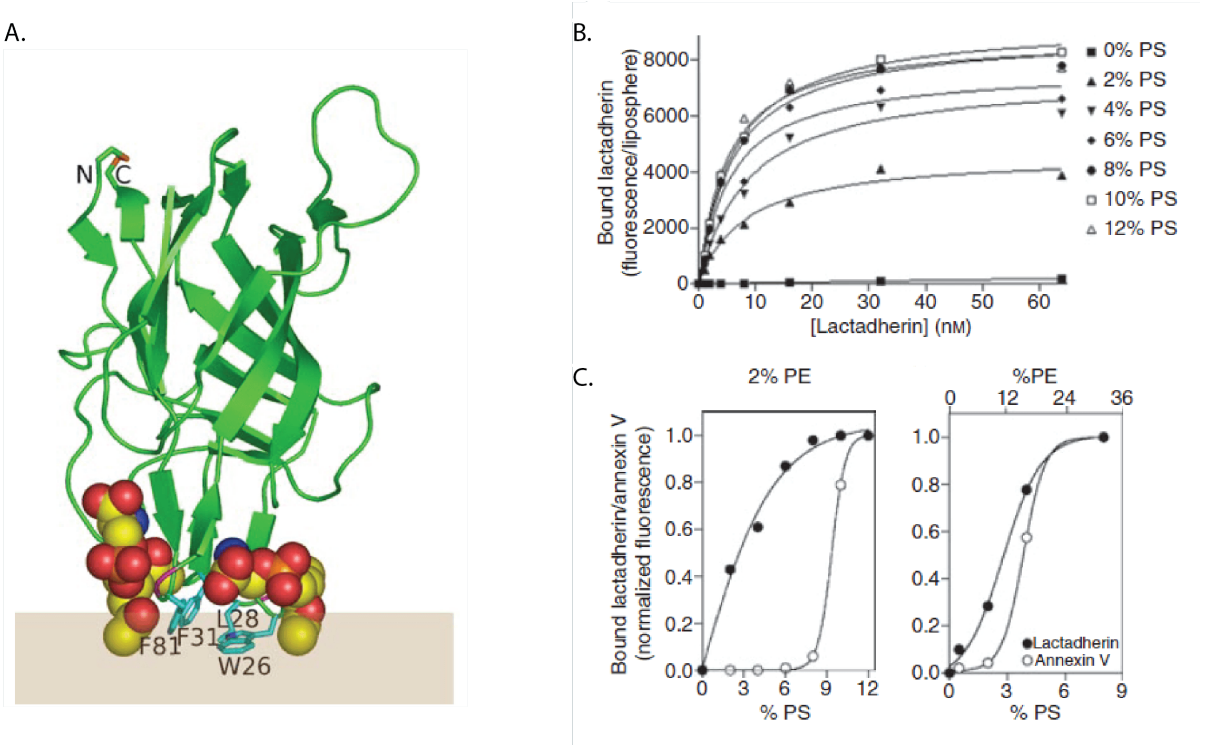


Figure 1.5: Binding of lactC2 domain to L-serine of PS and comparison with Annexin V. A. Proposed model of lact2 (green ribbon diagram) binding to the L-serine motifs (space-filling model) of the membrane (light brown). Important protein motifs belonging to the spikes are in blue. Image from Shao *et al.*, 2008 [28]. **B.** Relationship of relative amount of membrane PS with the fluorescence intensity, for increasing quantity of labelled lactC2. **C.** Binding threshold for lactC2 and annexin V probes and its relation with PE relative content. Both images extracted from Shi *et al.*, 2008 [29].

$$\mathrm{LactC_2} + \mathrm{PS} \underset{\text{k-1}}{\overset{\text{k1}}{\rightleftharpoons}} (\mathrm{LactC_2}\cdots\mathrm{PS}) \xrightarrow{k_2} \mathrm{LactC_2}\cdots\mathrm{PS} \tag{1.1}$$

The sensitivity of this method is extremely high, as quantities of PS as low as 0.03 mass% have been detected [27]. However, the binding is only linear (in regards to the concentration of PS) for PS < 4 mol%, reaching a plateau at around 8 mol% [29]. The experimental data supporting these conclusion is present in Fig. 1.5 **B**.
The lactC2 binding is thought to be specific to PS, as no binding was observed in studies done with both zwitterionic and anionic lipids (PC, PE and PA) [30]. This also suggests the nature of the binding is not charge-dependent, since other anionic lipids were also tested. Furthermore, it is also not dependent of the presence of other lipids in the membrane, like PE [29].

Alternative PS probes Annexin V is a PS binding protein used as the benchmark probe for imaging PS [26]. However, comparison studies between have shown that lactC2 has enhanced binding in low PS content vesicles, showing a steep binding increase between 0-2% of PS, whereas annexin V has a sigmoidal behaviour, having a threshold above 8% [26]. Moreover, the binding with annexin V is calcium-dependent, which imposes restrictions on the composition of the encapsulated solution, and is affected by the presence of PE, lowering the detection threshold to around 2.5% [29]. This is shown in Fig. 1.5 **C**.

1.4 Research goals

The ultimate goal of this project was to detect the incorporation of the lipids synthesized inside liposomes by the enzymes expressed with the PURE system. To do this, lactC2 will be used as a probe to bind PS and allow imaging of its incorporation. The work developed can be divided in a list of sub-goals:

1. Establishing and characterizing a method to use lactC2 as a probe for the presence of PS.
   - Overexpression and purification of lactC2 with bound eGFP, to allow fluorescence detection.
   - Assessment of the probe binding specificity to PS, in comparison to other phospholipids in the membrane.
   - Characterization of the binding, focusing on the investigation of the minimum PS threshold detected and the dependency of binding with PS concentration.
2. Detection of PS incorporation in liposomes with confined gene expression.
   - Development of a protocol to confine gene expression and analyze permeability of liposomes to the probe.
   - Analysis of the eGFP fluorescence signal in a population of liposomes expressing enzymes that synthesize PS.
   - Determination the time needed for lipid synthesis and incorporation in the membrane of liposomes.

In this work, by succeeding in the implementation of the aforementioned milestones, we detect the incorporation of intravesicular synthesized PS in the membrane of liposome-based synthetic cell models.

Chapter 2

Material and Methods

2.1 Plasmid cloning

The first step of the project was to produce the protein probe eGFP-LactC2, and, to do so clone these genes into a bacterial expression plasmid. The final sequence of the plasmid was designed with the aid of SnapGene software. The *egfp* and *lactC2* genes were inserted in the pET-11a plasmid backbone, downstream of the promoter for bacteriophage T7 RNA polymerase. Prior to the *egfp* gene, in the region corresponding the N-terminus of the desired protein, the sequence that codes for a polyhistidine-tag consisting of 6 histidine residues was also inserted, to allow protein purification with nickel-nitrilotriacetic acid (Ni-NTA) chromatography. Moreover, the plasmid backbone contains the *lacI* gene that is repressed in medium with isopropyl beta-D-1-thiogalactopyranoside (IPTG) and allows for protein induction and the *amp* gene that confers resistance to ampicillin, allowing colony selection.

PCR amplification The plasmid containing *egfp* and *lactC2* genes was kindly provided as a gift from Sergio Grinstein [30]. Regular PCR reactions were performed to amplify both the plasmid backbone and the gene construct. Adequate primers were chosen for each fragment (Appendix A.1): for the amplification of the pET-11a backbone primers ChD 471 (forward) and ChD 850 (reverse) were used; for the amplification of the *egfp-lactC2*, primers ChD 848 (forward) and ChD 849 (reverse) were used. The reaction was performed with 10 ng of template DNA, 1 unit of Phusion High-Fidelity DNA Polymerase (New England Biolabs) in HF buffer, present in the enzyme kit and supplemented with 0.2 mM of dNTPs, 0.2 μM of forward primer and 0.2 μM of reverse primer in a final volume of 50μL. After an initial heating step at 95°C for 5 min to allow denaturation of dsDNA, the PCR reaction consisted of 34 cycles of 30 s for melting DNA at 95°C, followed by the hybridization of the primers for 30s 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 more in order to allow the DNA polymerase to complete possible unfinished elongations, thus leading to less side products. Unless stated otherwise, the remaining PCRs described used the same conditions and concentrations.

Analysis of DNA on agarose gel The composition of the PCR products was verified with an TAE agarose gel (1% w/v) using SybrSafe staining (Thermo Fisher). The BenchTop 1-kb DNA Ladder from

Promega* was used to allow visualization of the size of fragments. The fragments corresponding to the adequate sequence length (1.3kb and 5.6kb) were excised from the gel and purified using the Promega Wizard SV Gel and PCR Clean-Up System kit. The concentration of the eluate was determined by spectrophotometry, measuring the absorbance at 260 nm with the Thermo Scientific NanoDrop 2000c spectrophotometer.

Gibson assembly The pET-11a backbone and eGFP-LactC2 genes were assembled into the desired plasmid using a Gibson assembly protocol. For 100 ng of backbone, the equimolar quantity of gene was used. In the 20 μL total mixture volume, besides the DNA, both the enzymes (1.3×10^{-3} U/µL of T5 exonuclease, 2.5×10^{-2} U/µL of Phusion polymerase, 4.8 U/µL of Taq ligase) and the ISO Buffer were added. The assembly reaction was incubated at 50°C for 60min. After this time, 20 U/µL of Dpn1 were added to digest possible methylated DNA left and the mixture incubated for an additional 15 min a 37°C.

Transformation by heat shock and cell culture In order to have plasmid propagation, 25 ng of the construct obtained with Gibson assembly was transformed into 500 µL of One Shot™ TOP10 chemically competent *E.coli* (ThermoFisher) cells using heat shock. For this, the cells were put in a water bath at 42°C for 45s and then back on ice for 2 min, to reduce cell damage. The cells were then incubated in 1 mL of LB medium (1:20 dilution of cells) for 20min at 37°C, after which 50µL of the suspension was spread in LB plates supplemented with ampicillin. The remaining volume, after being pelleted and re-suspended in 50 µL of LB medium, was also plated. All of the plates were incubated overnight at 37°C.

Colony PCR After incubation, six colonies were picked to perform colony PCR and a replica plate was made. A regular PCR reaction was performed with 0.5 units of GoTaq DNA Polymerase (Promega) in GoTaq Buffer supplemented primers and dNTPs. Adequate primers were chosen in order to amplify the gene region and part of the backbone sequence upstream and downstream of the gene (Appendix A): ChD 25 as a forward primer and ChD 310 as a reverse. As the colonies are mixed directly in the reaction, the initial denaturing step is important to lyse the cells and extract the genetic material. The DNA from the colonies that presented the band corresponding to the predicted length (6.9 kb) was extracted from the gel using the PureYield Plasmid Miniprep System (Promega) and its concentration was determined by spectrophotometry at a wavelength of 260 nm, corresponding to peak of absorption of DNA.

Restriction enzyme digestion The plasmids were further tested by a restriction enzyme digestion analysis, in which 2.5 units of DraI and 2.5 units of StuI were mixed with 500ng of DNA, in a final volume of 20µL (both enzymes by New England Biolabs). The mixture was then incubated at 37°C for 1h. The results of the verification techniques were visualized in TAE agarose gel (1%), with the same protocol as described previously. The expected digestion pattern was produced with the SnapGene software, knowing the cutting sites of this enzyme, and compared to the pattern obtained in gel.

Sequencing To infer the quality of the construct on the sequence level, the DNA extracted from the six colonies was sequenced externally by Sanger sequencing (Macrogen). To 300 ng of plasmid DNA, 0.25μM of adequate primer was added, in a final volume of 10μL. The primers, no further than 1000 kb apart, should allow for verification of the gene sequence. In this reaction ChD 288 and ChD 25 were used (Appendix A). By analyzing the quality of the sequencing and the possible mutations in the gene, using SnapGene, it was possible to infer which colonies had produced the desired plasmid.

2.2 Protein overexpression and purification

After having the designed plasmid with the correct sequence inserted, the next step is to transform competent cells, express the construct and purify it from the remaining cellular components. In this section these steps will be explained in further detail.

Cell culture and protein induction The plasmid was transformed by heat shock into both E. coli Rosetta ER2566 cells (New England Biolabs) and Rosetta 2 (Novagen) suited for protein overexpression. A preculture of these strains was incubated overnight at 37°C in LB medium supplemented with ampicillin. After, the cultures were diluted in the same medium, in a ratio of 1:1000, and incubated at 37°C with agitation (200 rpm) until achieving an OD_{600} of around 0.6, when the cells should be in exponential growth. In this moment, the protein production was induced with 1 mM of IPTG and the cells were incubated at 30°C for 3h with agitation (200 rpm).

Cell lysis by sonication After achieving the desired OD_{600} the cells were pelleted by centrifugation at 13000 rpm for 5 min. The pellet was ressuspended in Buffer A (150 mM NaCl, 20 mM Tris (pH 7.5), 20 mM imidazole), the cells were disrupted by sonication. The sonication was carried with ten pulses of 10 s and 30s of interval, using 30% of amplitude. After centrifugation at 4°C for 15min and 13000 rpm, the supernatant should be clear and contain cytosolic proteins in solution, while the cell debris form a pellet.

Chromatography with Ni-NTA column The protein purification was done using Ni-NTA Spin Columns (Qiagen). This chromatographic columns of Ni-NTA are selective for the affinity tag of six histidine residues present in the desired protein. The standard protocol of the supplier was followed. The column was equilibrated and washed with Buffer A and the protein was eluted with Buffer B (150 mM NaCl, 20 mM Tris, 500 mM imidazole; pH 7.5). In order to exchange the elution buffer for the storage buffer (10 mM Hepes-KOH; pH 7.5), Zeba Spin Desalting Columns (ThermoFischer) were used. 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 taken for subsequent visualization in polyacrylamide gel.

Gel verification by SDS-PAGE The 12% polyacrylamide resolving gel and the 4% stacking gel were done as indicated in Table 2.1 The solutions with all the components, except APS, was degassed in order to prevent bubbles in the gel. APS was added after to begin polimerization. To load in the gel, 15 μL of the protein sample was mixed with 1 μL DTT and 15 μL Laemmli 2× Concentrate Loading Buffer (Sigma-Aldrich) and the denatured at 95°C for 10 min. The gel was run at 100V while the

samples were in the stacking gel and at 180V in the resolving gel. The gel was run submerged in running buffer (250mM Tris-HCl, 200mM glycine, 1% w/v SDS; pH 8.3).

Table 2.1: Volumes of the resolving and stacking gels

Component	Resolving gel 12%	Stacking gel 4%
Milli-Q	6.3mL	3.2 mL
1.5 M Tris-HCl pH 8.8	1.25 mL	78
20% SDS w/v	0.075 mL	0.025 mL
AA/BAA 37.5%/0.8%	4.8 mL	0.53 mL
TEMED	0.075 mL	0.025 mL
10% APS	0.01 mL	0.005 mL

Protein concentration measurement The concentration of the protein was measured with a Bradford essay. The essay was done by dissolving the Bradford reagent (Sigma-Aldrich) 1:10 and mixing 500 µL of this dilution with 20µL of the protein. bovine albumin serum (BSA) was used as a standard in a range of seven concentrations from 0.25 mg/mL to 2 mg/mL. The standards were diluted with 1 mL of the Bradford reagent 1:10 and 5µL of the protein solutions. Each of the samples was done in triplicate, including a milli-Q sample, and the absorbance at a wavelength of 595nm was measured by spectrophotometry. The standards were diluted eight times comparing to the protein allowing for measuring low concentrations of protein within the linear region of absorption.

2.3 Lipid synthesis inside liposomes

In order to have protein-synthesizing liposomes, glass beads are coated in lipids with the desired composition, from which the liposomes are created by natural swelling when incubated with a solution containing the molecules to encapsulate. The swelling solution has an IVTT system containing all the components necessary for gene expression plus the DNA and precursors. After swelling, the liposomes are immobilized in a coverslip with biotin-neutravidin recognition sites and imaged through fluorescence confocal microscopy.

DNA coding lipid synthesis enzymes The genes for DOPS synthesis (*plsB*, *plsC*, *cdsA* and *pssA*) were used either in the form of individual PCR amplified constructs or in the linearized plasmid format (pGEMM 7.0). The pGEMM 7.0 plasmid consists of the pUC19 backbone containing *plsB*, *plsC*, *cdsA*, *pssA* and *psd* under the domain of the T7 polymerase promoter and the genes *pgpA* and under *pgsA* the domain of the SP6 polymerase promoter. The plasmid was already available in the lab, synthesized previously by Dr. David Foschepoth and Gemma van der Voort. A digestion was conducted with 5U of EcoRI and 10µg of plasmid DNA. The enzyme has a single cutting region in thex *psd gene*, responsible for converting PS into PE. This way, the linearized pGEMM 7.0 can be used to accumulate PS.

IVTT system The PUREfrex 2.0 system (GeneFrontier Corporation, Japan) was used for gene expression inside liposomes. The kit consists of three separate solutions. Solution I is the feeding

solution (containing aminoacids, NTPs, tRNAs etc.), Solution II contains the essential enzymes (T7 RNA polymerase, energy recycling system, etc.) and Solution III containg ribosomes. The PUREfrex 2.0 solution for a 20 μL reaction consists of 10 μL of Solution I, 1 μL of Solution II, 2 μL of Solution III and the linearized DNA. The supplemented DNA can be the individual genes (each in 1 nM), the pGEMM 7.0 linearized plasmid, or the gene synthesizing for the YFP protein, depending on the nature of the experiment. In the case of lipid synthesis experiments, the reaction was supplemented with the necessary precursors: 1 mM CTP, 0.5 mM G3P and 0.5 mM of L-serine. Beta-mercaptoetanol (5 mM) was added to provide a reducing environment for the GPAT enzyme and 0.75 U/μL of SUPERase (ThermoFischer) to inhibit possible contamination with RNases.

Precursor films Oleoyl-CoA (Avanti Polar Lipids) is kept in aliquots of 0.5 mg/mL, dissolved in chloroform:methanol:water (80:20:2) and stored under argon. With adequate pipettes for high pressure vapour liquids (MICROMAN pipettes, Gilson), the volume corresponding to 100μM of oleoyl-CoA, in the final reaction mix, was transferred to a PCR tube. The organic solvents were evaporated at ambient temperature and pressure for 2h, leading to a dried precursor film.

Lipid-coated beads To prepare lipid-coated beads, a mixture of 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) (50 mol%), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE) (36 mol%), 1',3'-bis[1,2-dioleoyl-sn-glycero-3-phospho]-glycerol (18:1 CL) (2 mol%), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[biotinyl(polyethylene glycol)-2000 (DSPE-PEG(2000)-biotin) (1 mass%) (all from Avanti Polar Lipids) and Texas Red 1,2-Dihexadecanoyl-sn-Glycero-3-Phosphoethanolamine (DHPE-TexasRed) (0.5 mass%) (Invitrogen), for a total mass of 2 mg, was assembled in a 10 mL round-bottom flask. The remaining 12% were variable between different combinations of PS and PG, depending on the nature of the experiments. The pipette tips were flushed with chloroform twice to assure efficient lipid removal. To the mixture 0.4 of 100 mM rhamnose (Sigma-Aldrich) in methanol were added, to promote lipid film swelling. Afterwards, 0.6g of 212-300 μm glass beads (Sigma-Aldrich) were added and the organic solvent was removed by rotary evaporation at 200 mbar for 2h, followed by overnight dessication. Both the lipid solutions in chloroform and the lipid-coated beads were stored under argon at -20°C. Before each use, the beads were redessicated for at least 30 min.

Preparation of GUVs by natural swelling In order to prepare GUVs, 1 mg of lipid coated beads is added per μL of the solution desired to encapsulate - denominated swelling solution. The swelling solution contained the PURE system the DNA template. In the experiments with the objective of verifying gene expression, the mYFP-LL-Spinach construct was encapsulated [31]. The GUVs were formed by natural swelling of the lipids in the surface of the beads, during 2h in ice and protected from light to avoid photobleaching. Regular tumbling was provided to promote detachment of the GUVs. Four freeze-thaw cycles were then applied by submerging the sample in liquid nitrogen and thawing on ice. Afterwards, 2 μL of the supernatant containing liposomes was transferred to 5.5μL of the solution desired to be outside the liposomes - denominated feeding solution. The manipulation of suspensions containing liposomes should be done with a cut pipette tip to avoid breakage of the vesicles. The feeding solution composition depended on the experiment, usually having 0.3 vol% of PUREfrex 2.0 Solution I and 150 nM of eGFP-LactC2 protein, unless stated otherwise. To confine

gene expression to the liposome lumen, RQ1 DNase (Promega) was added (0.07 U/μL of feeding solution). The remaining enzymes tested - Proteinase K and RNAse I (both from Thermo Fischer) - were also used in the same concentration. In the lipid synthesis experiments, the feeding solution was previously transferred to the oleoyl-CoA films and mixed by pipetting up and down.

2.4 LC-MS

Liquid chromatography of phospholipid samples was performed using a CSH C18 2.1*50mm, 1.7 μM column, mobile phase A (water with 0.05% ammonium hydroxide and 2 mM acetylacetone), and mobile phase B (% 2-propanol, 20% acetonitrile, 0.05% ammonium hydroxide and 2 mM acetylacetone).

A triple-quad mass spectrometer was used to detect PE and PS phospholipids. Transitions were established based on previous work by Scott *et al.*, 2016 [12], and scanning measurements of pure standards. Synthesized phospholipids were distinguished from phospholipids present at the start of the reaction (as part of the liposome matrix) by incorporation of ^{13}G3P, resulting in a 3 Da mass shift.

Mass spectrometry data was analysed using the Agilent Masshunter Quantitative analysis program, which automatically integrates peaks corresponding to the transitions set in the method. The integrated peak intensities were analyzed with MATLAB. For each transition, the average integrated counts signal of two injections was determined.

2.5 Fluorescence confocal microscopy

Imaging chamber preparation The sample is prepared in custom-made imaging chambers. To produce the chambers, two holes are drilled in a glass slide, a narrow (aprox. 3 mm) and a wide one (aprox. 5 mm). On one side, a cover slip was glued with NOA 61 glue (Norland Products), creating the two compartments. Before each use, the glass slide was cleaned with a protocol of sequential 10min sonication steps. Firstly a solution of chloroform and methanol (50 vol% of each) then Hellmanex III 2 vol% (Sigma-Aldrich), 1M KOH, milli-Q deionized water and finally, ethanol (100%). The narrow chamber was then incubated with 10μL of BSA:BSA-biotin (1mg/mL) for 10min and then Neutravidin (1mg/mL). After the immobilization, the surface was washed three times with milli-Q and the suspension containing both the liposomes and the feeding solution was transferred to this chamber. The wider chamber is filled with 15μL of water, to saturate the air inside the compartment and avoid evaporation of the sample. The compartment is closed placing a silicone square and a cover slip. The sample is incubated at 37°C overnight to allow for expression of the genes and binding of the protein probe.

Image Acquisition Image acquisition was performed using a Nikon A1R Laser scanning confocal microscope with the the SR Apo TIRF 100x oil immersion objective. The following excitation/emission wavelengths were used: 488/509 nm (lactC2-eGFP) 514/540 nm (YFP), and 561/595 nm (Texas Red). In order to allow comparison between experiments the acquisition settings were kept constant between experiments, for each of the fluorophores and are presented in Table **??**. The pinhole was

1.2 AU and the image resolution 512px×512px for all the experiments. The samples were always imaged in 6x6 stitched fields of view.

Table 2.2: Settings for each of the fluorophores used.

Fluorophore	Laser Power (a.u.)	Offset (a.u.)	Gain (a.u)
eGFP-LactC2	30	0	30
Texas Red	10	0	10
YFP	50	0	5

The sample height was adjusted manually in order to capture the liposomes immediately above the surface, corresponding to the immobilized ones. In the kinetic experiments, the sample is incubated in the microscope to allow for image acquisition in the desired time intervals.

Image analysis To determine lactC2-eGFP fluorescence intensity at the membrane automated image analysis has been applied. Firstly, an algorithm was used to determine liposome lumina, based on the texas red membrane signal and searching for circular components (rationale detailed in Appendix C). To determine the eGFP intensity along the membrane, first the centroid and radius were determined for every detected liposome. Then, line 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, corresponding to the membrane intersection, was recorded, and averaged to obtain the eGFP intensity of the membrane. It is important to note not all liposomes of the sample are detected with this methods but it is assumed the detected population represents the whole. The MATLAB scripts can be found in Appendix C. In the experiments where lumen pixel intensity was computed, the same rationale to find liposomes was applied.

Statistical analysis To compare distributions, a two-sample kolmogorov-kmirnov test was applied to the data. The test returns a decision for the null hypothesis that both samples belong to the same continuous distribution. The MATLAB inbuilt function *kstest2* was applied to perform this analysis. The output is 1 if the null hypothesis is rejected with a 5% significance and 0 otherwise.

Chapter 3

Results and Discussion

The experimental approach carried in this project was divided in three sequential steps. The first step of the project was to express and purify eGFP-lactC2, responsible for binding PS and facilitating fluorescence detection. An active probe that binds specifically PS is the foundation for the successful imaging of phospholipid incorporation. After the production of protein, the applicability of the method was tested. The production of PS by enzymes expressed with the PURE system was verified by MS and the probe functionality (e.g. specificity and threshold detection) was characterized, using liposomes that contained PS *a priori*. Finally, after characterizing the method, imaging of PS incorporation in liposomes was performed. The key objectives that inspired the experiments in this last section were: 1) observing PS detection with encapsulated gene expression and 2) determining the time-scale for lipid synthesis and incorporation.

3.1 Expression and purification of eGFP-lactC2 protein

3.1.1 Preparation of linear construct

The linear construct containing both the *egfp* and *lactC2* genes was amplified by PCR and successfully assembled into the expression vector pET11a. The PCR product was analyzed with an agarose gel (Fig. 3.1). The gel contains both the separate gene fragments and the final construct. The lanes with the expression vector (5.7 kb) and the *egfp-lactC2* construct (1.2 kb) have bands in the expected size range.

The strain responsible for expressing the plasmid was transformed with the Gibson assembly product and plated. After incubation, two clones were chosen for verification and the DNA was extracted. Both clones have an intense band in the 7 kb region, where the assembled plasmid is expected to be. *In vivo*, plasmid DNA is mostly in the supercoiled form. As a result, this DNA runs faster than its linear or nicked forms and the corresponding band is smeared, as can be seen for clones 1 and 2.

The *egfp-lactC2* genes were analyzed on the sequence level by Sanger sequencing. This was done to verify if the genes were inserted without mutation, for both clones analyzed. From this, it is concluded that the expression plasmid was successfully assembled and the genes have no point mutations.

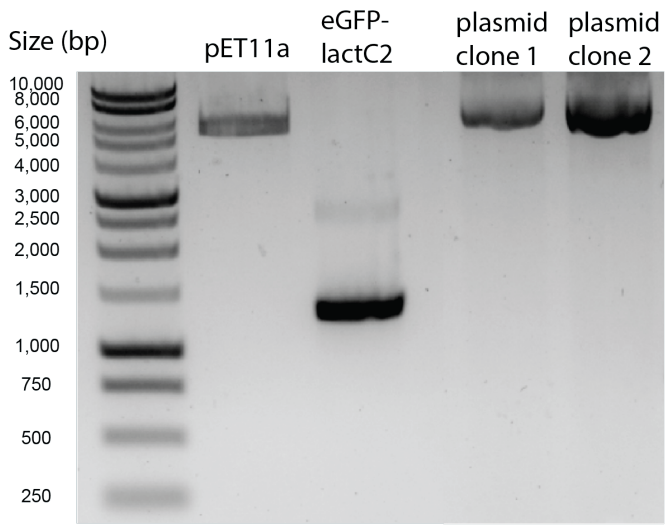


Figure 3.1: Gel electrophoresis (1 mass% of agarose) of separate DNA fragments: the cloning vector pET11a, the eGFP-LactC2 contruct and the assembled plasmid. The plasmid DNA from two clones (1 and 2) was extracted after transformation with the assembly product.

3.1.2 Protein production in both Rosetta 2 and ER2566 and purification by Ni-NTA chromatography

After DNA extraction from clones containing the successfully assembled plasmid, two strains of Rosetta cells (2 and ER2566) were transformed and eGFP-lactC2 expression was induced. Rosetta cells contain codons rare in *E.coli* and are generally used to overexpress eukaryotic proteins. For this reason, the two strains tested are adequate to produce the protein probe, since lactC2 is of bovine origin.

The cells were lysed and subsequent protein purification from the cell lysate was done with Ni-NTA column chromatography. Samples were taken for SDS-PAGE gel visualization (Figure 3.2), through each of the steps of induction, cell lysis and purification. Even though the production was tested with both strains, only for Rosetta ER2566 the intermediate steps were visualized. Bands corresponding to the full protein are predicted in the 45 kDa region.

A high intensity band is observed in the expected region after protein overexpression (lanes PI1 and PI2 of Figure 3.2). This indicates that cells of both strains produced eGFP-lactC2. With the cell lysis, both the soluble fraction (lane S1) and the insoluble fraction (lane P1) result in a band in the expected region. However, the band is more intense in lane P1, corresponding to the pellet. The insolubility of a fraction of the protein could be explained by protein aggregation due to misfolding [32].

The purification of the soluble fraction allowed to remove most of the undesired cytosolic proteins in the binding and washing steps, seen in lanes FT1 and W1. The purified protein, seen in lane E1, has two well defined bands in the region corresponding to 45 kDa, possibly due to the presence of another protein in the cells with affinity to the column. After buffer exchange (lanes BE1 and BE2), the bands are fainter, particularly the top one. This indicates loss of protein in the size exclusion column.

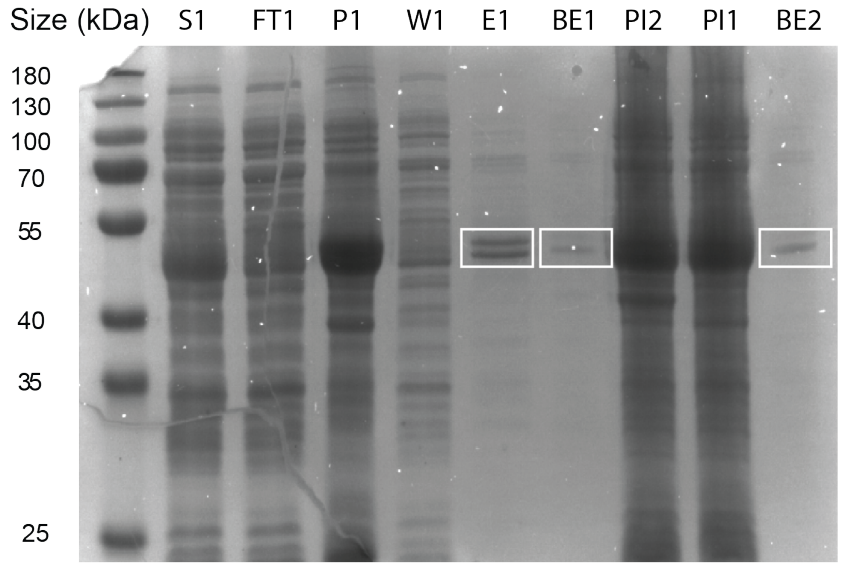


Figure 3.2: Visualization of the results and intermediate steps of the protein purification using Ni-NTA chromatographic column by SDS-PAGE gel. "1" stands for the protein produced by strain Rosetta ER2566, "2" by strain Rosetta 2; "PI" stands for post-induction with IPTG, "S" for the supernatant and "P" to the pellet, both after cell lysis by sonication, "FT" to the flow through of the membrane binding step, "W" to the flow through of the column washing step, "E" to eluate and "BE" to the final purified protein after buffer exchange. The white boxes indicate the expected region for the protein in the lanes with faint bands.

3.1.3 Concentration of purified protein

The concentration for the final protein solutions, overexpressed from both cell lines, was measured using a Bradford assay. Both the BSA standards and the eGFPlactC2 protein, used in higher concentrations (8x, comparatively to standards) were analyzed. The calibration curve (Equation 3.1) was obtained by measuring the absorbance at 595nm of BSA standards in triplicates. In the equation, C stands for concentration (mg/mL) and A for absorbance. The absorbance measured and concentrations obtained using the calibration are presented in Table 3.1. The final concentration obtained was 611 ± 78 and 787 ± 50 for Rosetta 2 and ER2566, respectively.

$$A = 0.254 \times C + 0.005 | (R^2 = 0.9924) \tag{3.1}$$

Table 3.1: Concentration of eGFP-lactC2 obtained using the Bradford essay. The concentration was obtained by applying the calibration curve equation to the absorbance measured and multiplying by a factor of eight. The final concentration corresponds to the average of the triplicates.

	Rosetta 2		Rosetta ER2566	
Measurement	Abs_{595}	C(nM)	Abs_{595}	C(nM)
1	0,069	711	0,079	824
2	0,059	600	0,079	824
3	0,052	521	0,069	712
Final concentration (nM)	611 ±78		787 ±50	

3.2 PS is synthesized and accumulated by PURE system: Bulk verification with SUV

PS was chosen as the model to study phospholipid incorporation because it can be imaged with the eGFP-lactC2 probe. However, this lipid is rapidly converted to PE in *E.coli*, not having a relevant membrane presence. Since PS is not an end product of the lipid synthesis pathway, its production and accumulation with PURE system had not been verified before. The success of the PS visualization in the membrane depends on its production in detectable levels.

In this experiment, PS was synthesized by enzymes expressed with the PURE system in the presence of SUV, responsible for providing a membrane support to the enzymes. Three samples were analyzed with LC-MS, in which either palmitoyl-Coenzyme A (p-CoA), oleoyl-Coenzyme A (o-CoA) or both were provided as precursors. There was also a negative control without any film precursor. The level of each possible PS carbon-chain variant (1,2-dioleoyl-sn-glycero-3-phospho-L-serine (DOPS), palmitoyl-oleoyl-sn-glycero-3-phospho-L-serine (POPS) and 1,2-dipalmitoyl-sn-glycero-3-phospho-L-serine (DPPS)) is presented in Fig 3.3.

In all samples except the negative control, PS was found to accumulate. As expected, the sample with o-CoA had only DOPS and the sample with p-CoA had only DPPS. The sample with both showed a majority of the mixed chain lipid, POPS and the negative control, in its turn, had no production of lipid, as acyl-CoA is an essential lipid precursor.

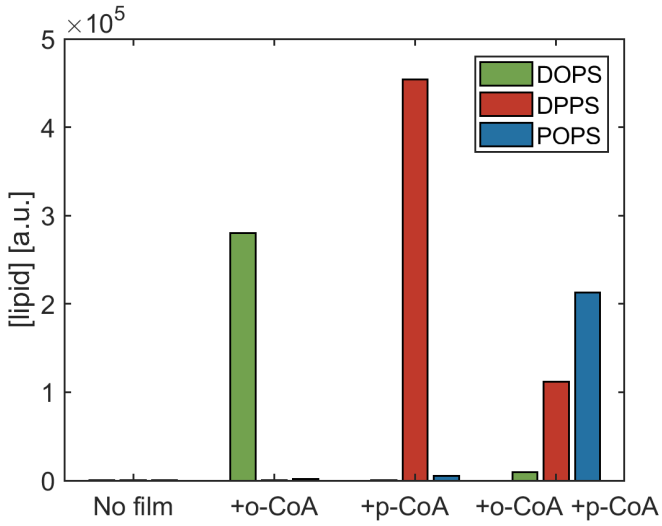


Figure 3.3: LC-MS analysis of the PS synthesis by PURE system in the presence of SUV. "No film" corresponds to the negative control without lipid film precursors. The precursors p-CoA and o-CoA were supplied in separate samples at a concentration of 100μM ("+o-CoA" and "+p-CoA") and in a equimolar mixture with 50μM of each ("+o-CoA +p-CoA").

It was concluded that PS accumulated and the precursors were being used by the enzymes in the pathway. DOPS was used in further studies since phospholipids containing 18-carbon chains are preferred for liposome construction. Liposomes formed with di-palmitoyl (DP) lipids were previously observed to have low expression levels. This can be explained by high $\mathbf{T}_m$**!** ($\mathbf{T}_m$**!**) which also constrains lipid film swelling to temperatures above 37°C, the optimal temperature for PURE system [**Nourian2012**]. Since di-oleoyl (DO) liposomes have worked for past experiments, displaying gene expression with encapsulated PURE system, these liposomes were used for the experiments in

this project.

3.3 Characterization of the detection method

Before using eGFP-lactC2 to detect PS synthesized in liposomes, it is indispensable to know the characteristics of this imaging method. With this objective in mind, the approach focused on 1) assuring binding specificity to PS, in comparison to other phospholipids and 2) studying how binding correlates with PS and lactC2 concentrations.

3.3.1 Similar radius distribution and gene expression levels in liposomes with 12% PG vs. 12% PS

In order to analyze the binding specificity of the probe to PS, it is important to rule out possible interference of the presence of PS in the membrane with physical characteristics of the liposomes and the gene expression levels.

Throughout this work, the standard bilayer composition will be considered to be 50 mol% DOPC, 36 mol% DOPE, 12 mol% 1,2-dioleoyl-sn-glycero-3-phosphoglycerol (DOPG) and 2 mol% 18:1 CL. In this experiment, liposomes with both the standard bilayer composition and a modified version, switching PG for PS in same molar quantity, were prepared. Confocal microscopy was used to image liposomes encapsulated PURE system and 4 nM *yfp* DNA. This construct was used to ensure direct visualization of translation activity.

The first observation was that both membrane compositions led to the production of several micrometer sized liposomes (labeled with a red membrane dye) as seen in the micrographs (Fig. 3.4 **A**). Moreover, the radius was analyzed on both populations of liposomes to understand possible impacts on the size distribution (Figure 3.4 **B**). Using both membrane compositions, the majority of vesicles has a radius ranging from <1 μm to 10μm. The probability density profiles overlap, demonstrating a equivalent size distribution.

The concentration of yellow fluorescent protein (YFP) produced was computed by extracting the average pixel intensity in the lumen of liposomes, for the YFP channel, and applying a calibration that was previously established (unpublished work by Duco Blanken). The liposomes successfully expressed YFP and the concentration distribution was similar for both samples, mostly ranging from 0.1 μM to 1 μM (Figure 3.4 **C**). It was therefore concluded that liposomes with 12% PS could be used to test binding with eGFP-lactC2 probe in an unbiased manner, that is, without significantly altering the properties of the system.

PS was chosen to be replaced by PG in order to keep the membrane charge unchanged, as both lipids have a -1 charge. The two lipids also share a cylindrical shape and are considered bilayer forming [33]. The data regarding the radius distribution confirms the hypothesis that these similarities possibly lead to identical membranes structures.

Since PS is not present in the membrane of *E.coli* and the encapsulated proteins (from the PURE system) are mostly purified from this bacterium it was hypothesized that it could interfere with gene expression. However, the membrane composition does not influence the production of YFP. This is likely because the enzymes responsible for gene expression are cytosolic and do not interact directly with the membrane.

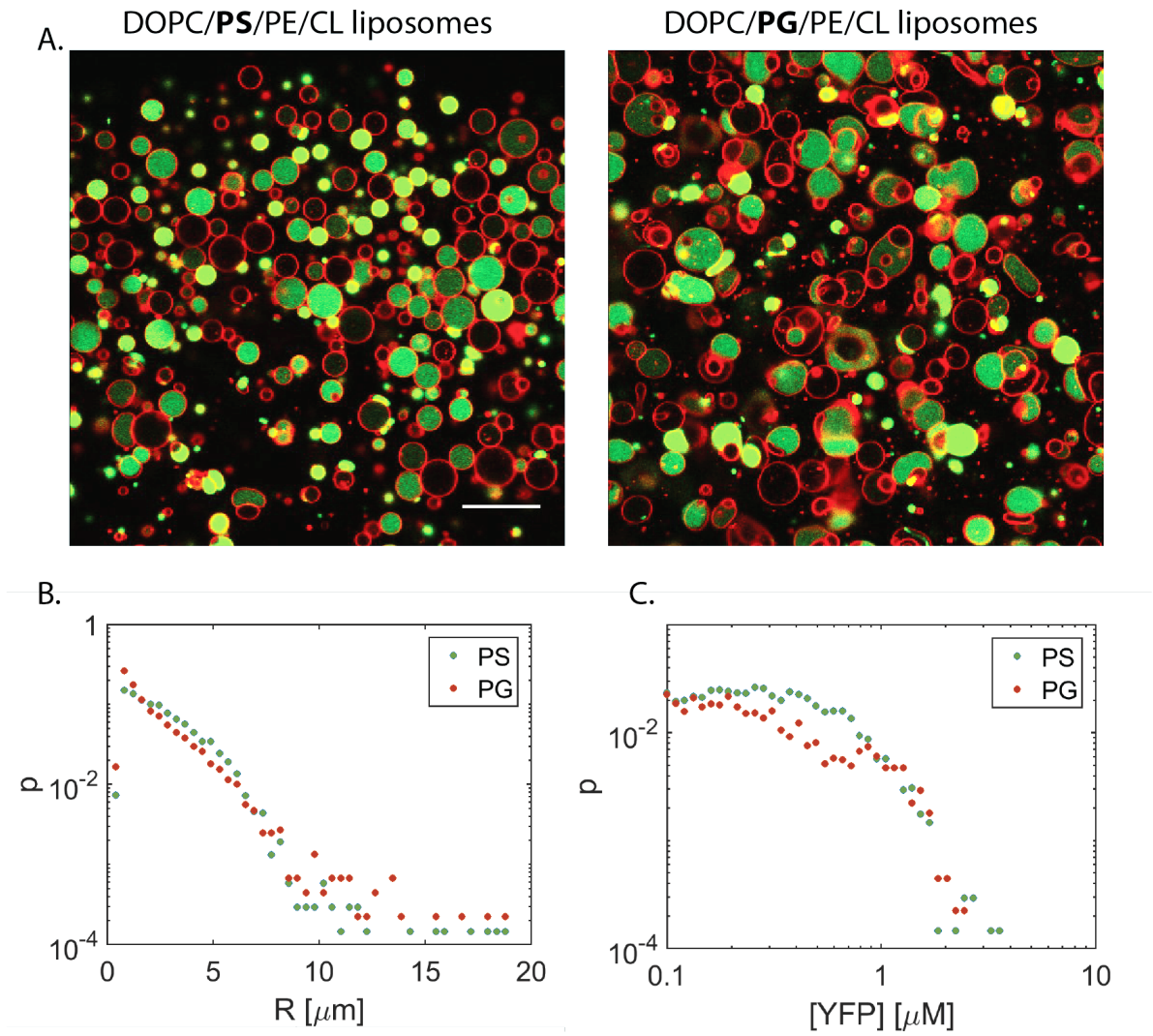


Figure 3.4: Morphology comparison of liposomes done with 12% PG vs. 12% PS. A. Micrographs of both samples obtained using confocal microscopy. The red channel corresponds to the liposome membrane and the green to the expressed YFP. Data obtained processing confocal microscopy images. Histograms showing the distribution of **B.** liposome radius (μm) and **C.** YFP expression (μM).

3.3.2 Protein probe binds specifically to PS-containing liposomes

After verifying that liposomes with PS instead of PG have an equivalent morphology, the specificity of lactC2 was tested. The probe needs to bind specifically to PS in comparison to other phospholipids to allow detection of membrane incorporation.

Liposomes were prepared with both the standard and modified composition (12% PS or 12% PG, respectively). Although there was no gene expression, PURE system was also encapsulated to assure coherence with the remaining experiments. This allows similar osmotic pressure, pH and divalent ion concentration. In the feeding solution 150nM of eGFP-lactC2 was added. The samples were imaged with fluorescent confocal microscopy (Fig. 3.5 **A**) and the average pixel intensity across the rim of liposomes was extracted (Appendix C).

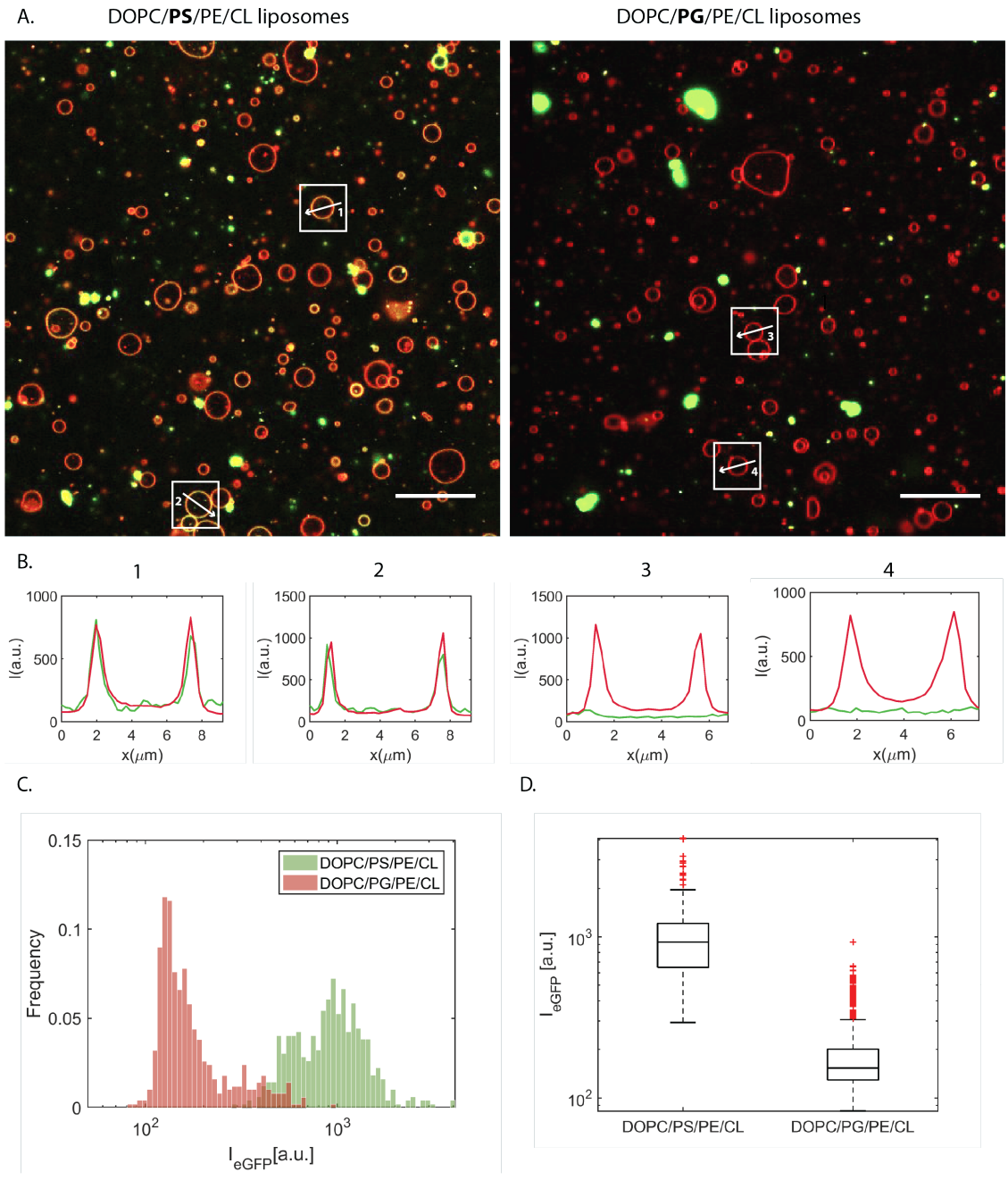


Figure 3.5: Comparison of the lactC2 specificity between PS and PG. A. Confocal micrographs of liposomes (labelled in red) of different membrane compositions and incubated with 150 nM of eGFP-lactC2 outside (labelled in green). Scale bar is 20 μm. **B.** Line intensity profile for four liposomes (framed in **A.**), given that each of the colors represent the corresponding channel. **C.** Histogram and **D.** box plot displaying the distribution of average intensities across the rim of liposomes in the eGFP channel. Several liposomes (n= [496, 500], respectively for the samples in the presented order) were taken into account.

From the analysis of the micrographs, it is clear the green fluorescence signal of eGFP colocalizes with the membrane in the sample with liposomes containing PS. Moreover, when focusing on specific liposomes, the two fluorescence peaks of the line profile (Fig. 3.5 **B**), corresponding to the membrane, completely overlap in both channels. The sample without PS, however, has no colocalization of the eGFP signal with the membrane. The signal throughout the line profile is approximately constant and similar to the background value. This demonstrates the eGFP-lactC2 probe accumulates on the membrane of liposomes containing PS but not on the absence of PS.

Furthermore, the rim intensity of the population of liposomes, for eGFP channel, in different fields of view was analysed (Fig. 3.5 **C** and **D**). The histogram shows a distribution of the intensities centered on 1000 for the liposomes containing PS, close to a order of magnitude higher than for liposomes that did not contain PS. In addition, there is little overlap of both distributions. Examining the box plot of the same data, the difference in the populations is also clear (see Appendix D for a detailed explanation of this plot type). The average rim intensity for the PS-containing liposomes analyzed is higher than the majority of liposomes of the standard composition.

To compare both distributions, a two-sample kolmogorov-kmirnov test was applied to the data. The test returns a decision for the null hypothesis that both samples belong to the same continuous distribution. The null hypothesis was rejected for this data, confirming the observations. These results show that the protein probe binds specifically to liposomes containing PS. The binary behaviour validates this method as a promising tool for examining incorporation of synthesized PS. Moreover, the possible charge-dependent nature of the binding ([27]) was not observed with this experiment. PG, equally an anionic lipid, did not show colocalization of the eGFP signal with the membrane.

3.3.3 Detection threshold of the method below 1 mol% of PS

With the objective of quantifying the amount of PS incorporated in the liposomes, it is important to know the limit of detection of the method. Three samples of liposomes with varying composition of both PS and PG were made (0%+12%, 1%+11% and 12%+0%, respectively). Again, the vesicles encapsulated PURE system and were incubated with 150nM of eGFP-lactC2. The distribution of the average rim intensity for each sample is shown in Fig. 3.6.

From the analysis of the histogram (Fig. 3.6 **A**), the distinction between both sample with liposomes containing PS and liposomes of the standard composition is very clear, with almost no overlap of the distributions. However, even though the sample with 12% PS has a distribution shifted to higher rim intensities, most of it overlaps with the 1% PS sample. The box plot (Fig. 3.6 **B**) reinforces this result, as it shows that the median of average rim intensities across the sample (aprox. 800 for 1% PS and 1000 for 12% PS) is not proportional to the concentration of PS in the membrane, for this percentage. Otherwise, for the 12% PS sample, the box would be shifted to intensities twelve times higher.

Previous studies have shown that it is possible to detect binding between PS and lactC2 for membrane PS content above 0.03 mass% [27]. Additionally, the binding is only linear at low concentrations of PS, achieving saturation at 8 mass%. The results obtained confirm this, as the detection threshold was found to be lower than 1 mol% ($\simeq$ 1 mass%) and non-linear behaviour was observed between samples with 1% and 12% PS. It was concluded eGFP-lactC2 can be applied to detect as low as $\sim$1% of PS but not to quantify the relative amount.

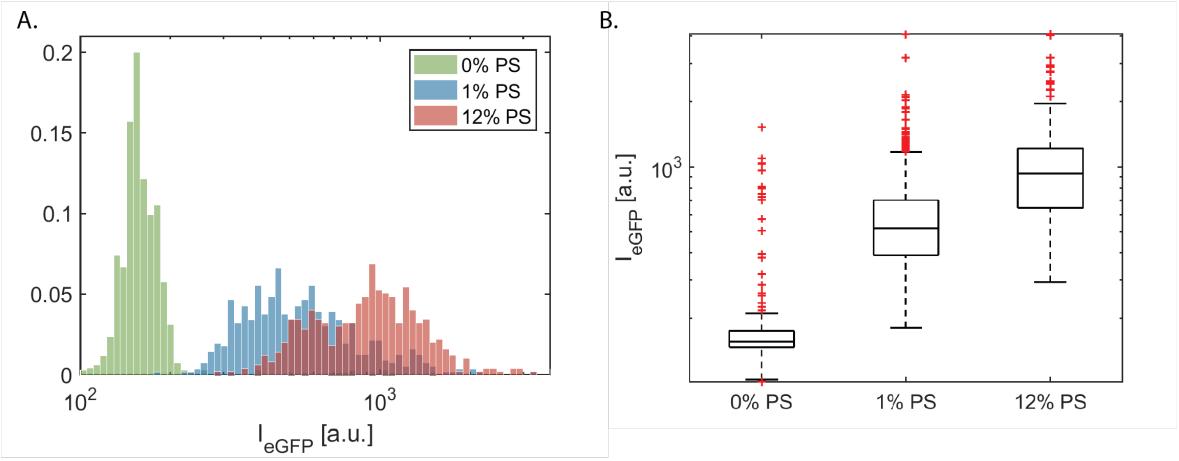


Figure 3.6: Binding behaviour for different PS concentrations. **A.** Histogram and **B.** box plot of the distribution of average rim intensities for liposomes with 0%, 1% and 12% of PS. Several liposomes (n=[675, 560, 496], respectively for the samples in the presented order) were taken into account.

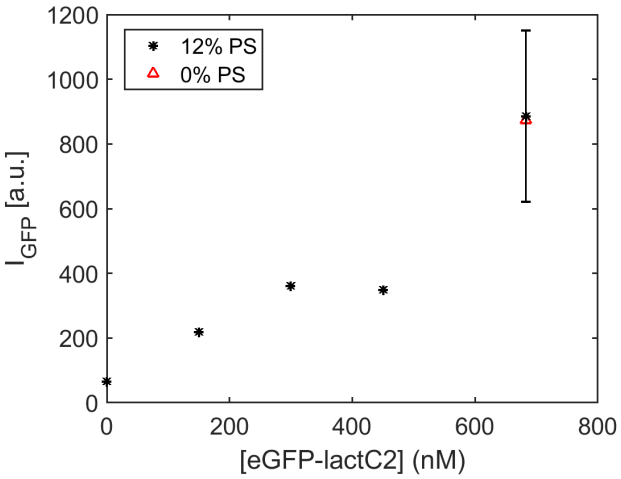


Figure 3.7: Non-specific binding for high concentrations of eGFP-lactC2. Averaged rim intensities across the sample, plotted against the concentration of probe used. For the highest concentration, two repeats were done hence the error bar with the standard deviation. Several liposomes (n = [503, 273, 93, 403, 314, 71, 60], respectively for the samples in ascending order of concentrations) were taken into account.

3.3.4 Non-specific binding verified for high concentrations of eGFP-lactC2

To apply successfully a fluorescence imaging technique, a high enough concentration that allows both detection of fluorescence signal above background noise and saturation of all the molecular targets should be used.

To test the saturation concentration of the probe, liposomes with 12% PS were made and incubated with increasing quantities of eGFP-lactC2. The average rim intensity of the population was then analyzed (Fig 3.7). For liposomes containing PS, the eGFP signal in the membrane increased with the concentrations of the probe. Moreover, liposomes of the standard composition were also incubated with the highest concentration of probe. This sample showed similar average rim fluorescence to the 12% PS.

The results obtained indicate non-specific binding at high concentrations. This could explain why, contrary to expectation, a saturation plateau was not reached. We reasoned that, with the

increasing concentrations of probe, the non-specific binding also increases. The fraction of each cannot be discerned, as the method only allows for the detection of eGFP and not the eGFP—PS complex.

It should be taken into consideration that when using a concentration that allows binary response (that is, concluding if PS is present or not), saturation is not necessarily guaranteed. However, even though theoretically the binding requires a ratio of one molecule of probe to one of PS, not all the lipid is necessarily exposed to the extravesicular solution. Furthermore, after the protein binds the headgroup of one PS molecule, it may hinder the binding of other nearby PS molecules. Considering these limitations, all the following experiments will continue to be done with 150nM of eGFP-lactC2 unless otherwise stated, since it has been proven to have a binary response (section 1.3.2).

3.4 Imaging incorporation of synthesized PS in the membrane

The key aim in this chapter is to visualize the incorporation of PS produced with the enzymes synthesized by PURE system, confined inside liposomes. To do so, firstly the PS synthesized both intra- and extra-vesicle was detected. After, the impact of both confinement approaches and eGFP-lactC2 placements on the detection of signal colocalized with the membrane were studied. Lastly, the time scale of lipid synthesis and incorporation was studied with kinetic experiments.

3.4.1 PS expressed by IVTT system is successfully incorporated the membrane of liposomes

Motivated by the results demonstrating that the method is specific to PS and has a high sensitivity, our next step was to investigate the incorporation in the membrane of PS synthesized by enzymes expressed with the PURE system. As a first step, liposomes containing 1nM of each of the *plsB*, *plsC*, *cdsA* and *pssA* separate constructs, the lipid precursors (G3P, L-serine and CTP) and PURE system, were incubated in the presence of the 150nM of eGFP-lactC2 and 100μM of o-CoA. It is important to note that at this stage no experimental measures were taken to confine the expression to the inside of liposomes. This means that there is also gene expression outside of liposomes, generated by the molecules that were not encapsulated during lipid film swelling.

Three different negative controls were designed. In the first, to cease lipid synthesis, no o-CoA was supplemented to the sample. However, acyl-CoA molecules also incorporate in the membrane. In the unlikely possibility that the eGFP-lactC2 probe binds to o-CoA, a second negative control sample was tested. In this sample, o-CoA was supplemented, but *plsB* (the gene coding for GPAT, responsible for converting G3P to LPA, the first step in lipid synthesis) was not. It was reasoned that, even in the absence of GPAT, another enzyme could have catalyze the same reaction, with a residual activity. Particularly LPAAT, as both enzymes are responsible for catalyzing acylation reactions. To tackle this, a third negative control was performed, in which none of the gene constructs were supplemented. Instead, phage ϕ29 terminal protein (TP) DNA in the same concentration was added to this sample [34]. In its absence, it will just maintain the gene expression in the sample to similar levels. For the same reason, the gene concentration in the second negative control were increased accordingly.

In the positive sample clear probe binding was observed indicating successful synthesis and membrane incorporation of PS. From the analysis of the micrographs for the four samples described,

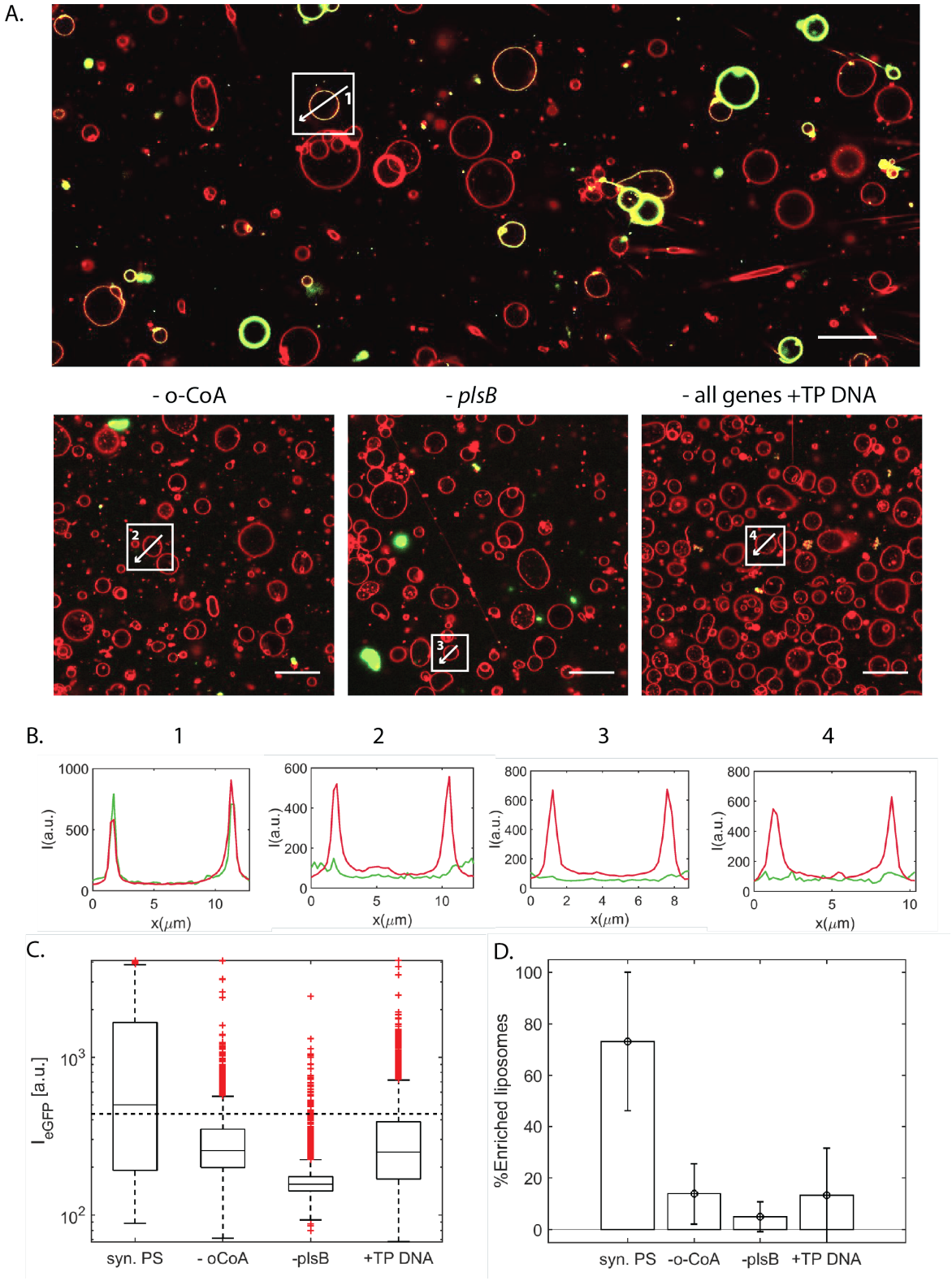


Figure 3.8: Synthesized PS incorporates in the membrane of liposomes. **A.** Confocal micrographs of liposomes (labelled in red) encapsulating PURE system, L-serine and CTP and incubated with 150 nM of eGFP-lact2 (labelled in green) and 100μM of o-CoA outside . Scale bar is 20 μm. **B.** Line intensity profile for four liposomes (framed in **A.**), given that each of the colors represent the corresponding channel. **C.** Distribution of average intensities across the rim of liposomes in the eGFP channel. The dashed line indicates the threshold used to consider a liposome enriched in PS. **D.** Percentage of liposomes enriched in PS for each sample. Four repeats of each sample were made, from which several liposomes (n= [767, 3539, 1244, 3742], respectively, in the presented order) were taken into account. Error bars represent the inter-sample standard deviation.

present in Fig. 3.8 **A**, the colocalization of the eGFP signal with the membrane is only present in the sample containing all genes and precursors. This is further supported by the line profile of individual liposomes (Fig. 3.8 **B**). This positive sample is the only one with overlapping peaks in the green and red channel. The threshold to compute the percentage of PS-enriched liposomes (Fig. 3.8 **D**) was obtained assuming a normal distribution of average rim intensities to the liposomes with 0% PS (Fig. 3.5 **C**). A liposome was considered enriched, if the average rim intensity was above the intensity of 95% of the liposomes in the 0% PS sample (section 1.3.3). By the definition of normal distribution, the threshold is given by $\mu + 2\sigma$, where μ is the mean and σ the standard deviation. The threshold obtain corresponded to 422 a.u. of average rim eGFP fluorescence intensity. The amount of enriched liposomes obtained is three times higher for the positive samples, where it was observed to be above 40% in all the four repeats performed. From the box plot is very clear that all the negative controls have most of the liposomes below the threshold, when for the case of the positive sample the median is above this value.

In conclusion, the PS is produced by the enzymes synthesized with the cell-free PURE system and detected with the eGFP-lactC2 fluorescence imaging method. The difference between the positive sample and the negative controls designed was clear, both on the level of probe fluorescence colocalization with the membrane but also on the number of liposomes that displayed a considerable signal (denominated PS-enriched). Since the expression occurred both on the outside and inside of liposomes and the probe was fed to the outside solution, no conclusions can be drawn related to possible lipid flip-flop.

Furthermore, a high level of heterogeneity was observed, both between liposomes in the same repeat and between repeats of the same sample (day-to-day variability). Liposome to liposome variability is not dependent of encapsulation restrictions in this case, since there is expression outside. However, spontaneous crowding of some molecules (e.g. proteins and ribosomes) may occur [35], leading to concentration heterogeneity across the sample. Moreover, diffusion constraints can slow down and limit the molecule interactions needed for expression. The higher the number of components in the system, the more prominent is the heterogeneity. Day-to-day variability is not entirely understood. Different batches of reagents, in particular lipid-coated beads, could be a cause. It was observed often that older beads led to less liposome quality. In the experiments that had less liposomes in solution rim intensities were sometimes brighter (always using the same eGFP concentration). This was hypothesized to be due to a higher ratio of eGFP molecules per lipid molecule. However, these observations were not further tested in independent experiments. Moreover, PURE system has a lot of day-to-day variability regarding the expression levels (unpublished work by Anne Schwabe).

3.4.2 Strategies to encapsulate gene expression

Having four separate gene constructs might contribute to the variability within liposome population. To prevent variability due to inefficient encapsulation, a minimal genome with the genes involved in the lipid synthesis (pGEMM 7.0 - 11kb) was used. The plasmid contains the seven genes of the lipid synthesis pathway responsible, starting from G3P as a precursor (see Fig. 3.9 **A**). Since *psd* is also included, conversion of PS in PE occurs. To circumvent this, as it would prevent PS accumulation, the plasmid was linearized using EcoRI, an enzyme with an unique restriction site in

the *psd* gene.

The circular and the linearized plasmid were run in an agarose gel (Fig 3.9). Both of the bands are in the expected region: above 10kb. The band corresponding to the circular construct presents a smear and ran faster, given its likely supercoiled structure.

Using the circular plasmid led to no production of PS, indicating a rapid conversion of PS to PE (see Fig. 3.9 **C**). Moreover, there is PE present even in samples with linearized plasmid. This can be justified by possible residual production, due to the incomplete digestion of pGEMM and resulting presence of active Psd. On the contrary, the linearized plasmid showed PS accumulation, which validates that it can be used in further PS confined synthesis experiments.

Confining the lipid synthesis to the lumen of the liposomes is also an important step for a self-growing minimal cell, as all the components and functions should be within the compartment. The most efficient way to confine gene expression was investigated by incubating liposomes with three different combinations of enzymes: DNAse, DNAse and RNAse or proteinase. Each of the enzymes is responsible for arresting gene expression at different levels. Both DNase and proteinase prevent transcription, the former by digesting the DNA template and the later by digesting proteins and in particular RNA polimerase. RNAse on the other hand prevents translation, digesting the mRNA transcripts.

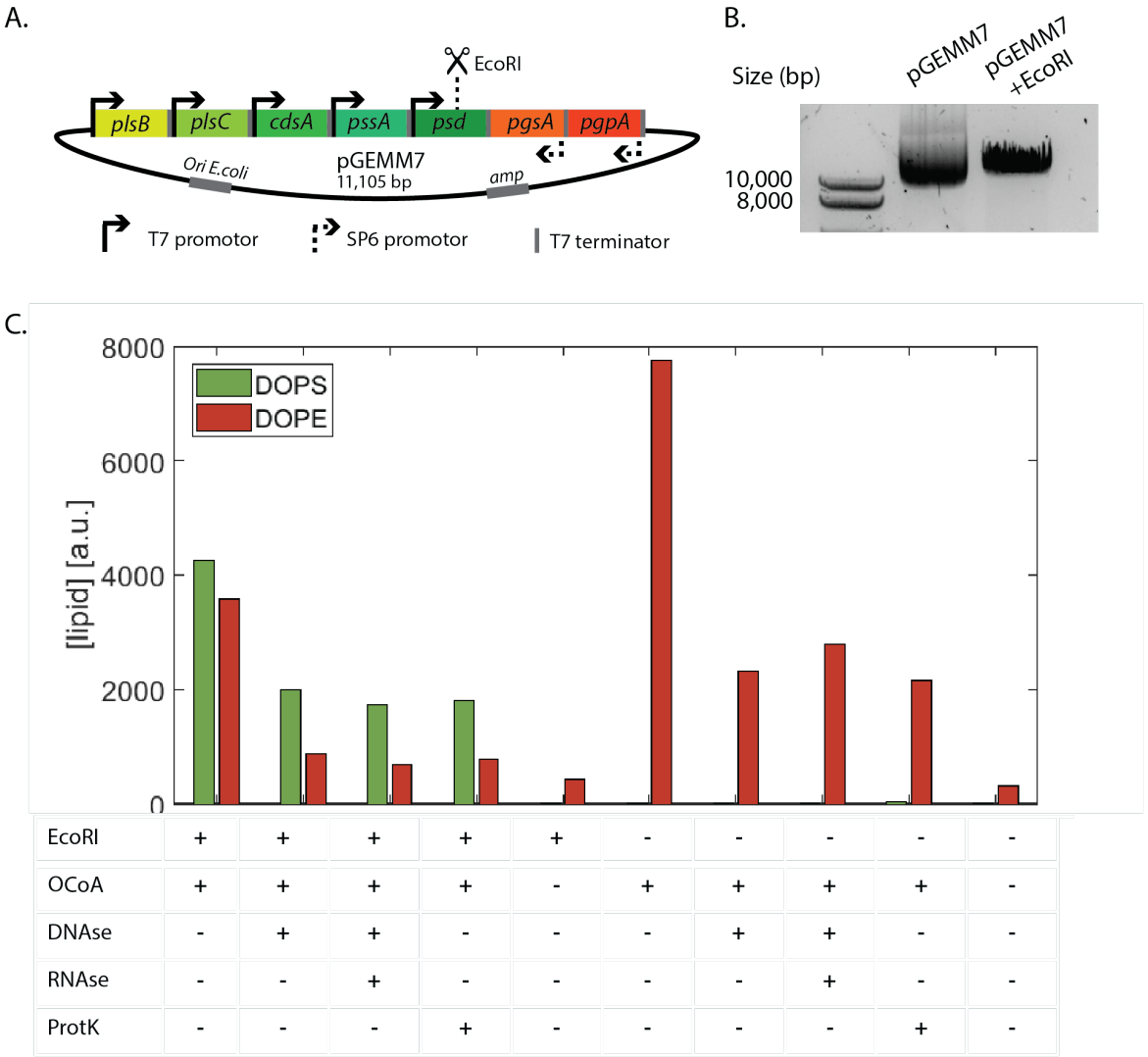
Using proteinase is the most powerful approach of the three described: it does not only arrest the gene expression but also all the other enzyme catalyzed reactions present in the PURE system. The drawback of this approach is that the protein probe used for imaging would also be digested. For this reason, the results obtained with the other enzymes were compared with proteinase. It is important to note that DNAse and RNAse were combined because RNAse by itself would potentially allow residual expression, since transcription is still carried out.

To test the effect of these proteins on the gene expression outside, liposomes with the standard composition encapsulated 1nM of the *yfp* construct and PURE system and were subsequently incubated with one of the enzyme combinations. Each of the samples was imaged through confocal microscopy, in addition to a sample that had no restriction enzymes added in the medium, to allow comparison (Fig. 3.10 **C**). To understand the effect on extra-vesicular gene expression, the background signal for the YFP channel was measured.

As expected, the highest background signal was displayed by the sample incubated without any enzymes (549.4 a.u.). Both the sample with DNAse and the sample with proteinase had a similar background (about 80 a.u.). In its turn, the sample containing DNAse and RNAse had a slightly higher background (about 150 a.u.). This could be due to the high liposome density observed.

Lipid production levels were further analyzed and quantified in bulk using MS. Samples with liposomes encapsulating circular or linearized pGEMM 7.0 and PURE system were incubated with the three enzyme combinations described. The level of PS and PE produced is shown in Fig 3.9 **C**. The level of PS produced without incubation with restriction enzymes is about double as for the three samples with confined expression. These samples have similar PS levels, reinforcing the conclusion that the confinement efficiency is equivalent for the strategies designed.

In conclusion, using DNAse has proven to be as effective as proteinase to stop gene expression outside liposomes without degrading the protein probe. Adding RNAse did not show to enhance the effect of DNAse. For this reason, in the following experiments, DNAse will be used to confine gene expression.



EcoRI	+	+	+	+	+	-	-	-	-	-
OCoA	+	+	+	+	-	+	+	+	+	-
DNAse	-	+	+	-	-	-	+	+	-	-
RNAse	-	-	+	-	-	-	-	+	-	-
ProtK	-	-	-	+	-	-	-	-	+	-

Figure 3.9: Accumulation of PS with encapsulation of linearized pGEMM 7.0. A. Structure of pGEMM 7.0 indicating the EcoRI restriction site in the *psd* gene. **B.** Agarose gel (1 mass%) showing circular and linearized pGEMM 7.0. **C.** Lipid production in bulk using PURE system, 1nM of pGEMM 7.0 and lipid precursors, in the presence of SUV, measured with LC-MS for different confinement conditions. "EcoRI +" stands for the plasmid linearized with this enzyme and "EcoRI-" for the circular plasmid. For the remaining molecules "+" stands for its presence in the medium and "-" for the absence.

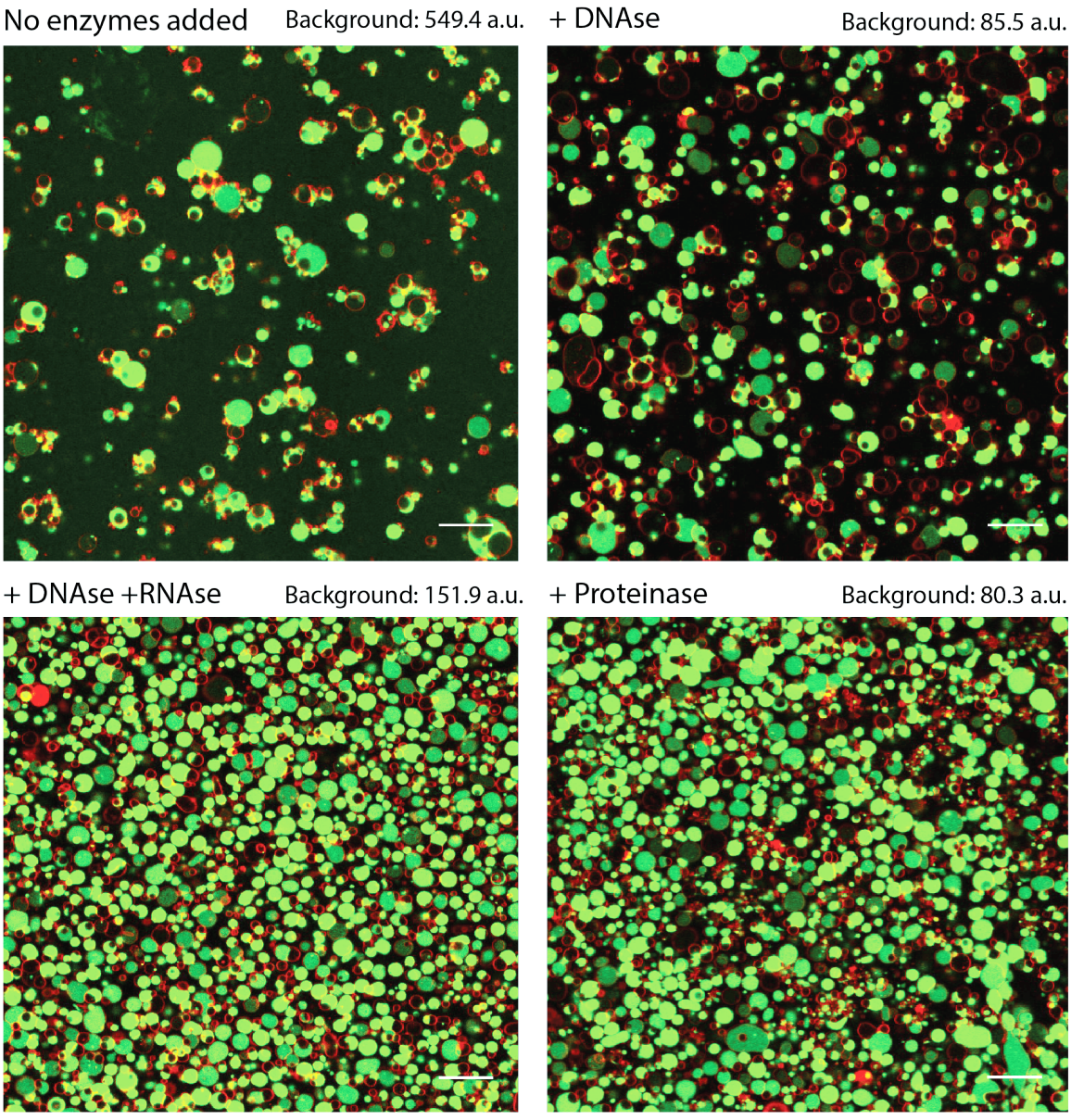


Figure 3.10: Comparison of confinement of gene expression using different enzymes. Confocal micrographs of liposomes (labelled in red) encapsulating PURE system and 4nM of *yfp*. The synthesized YFP is responsible for the green signal. The liposomes were incubated with 0.07U/μL of each enzyme. Scale bar is 20 μm.

3.4.3 Probe location does not influence the distribution of average rim intensities.

In this experiment, different positions for the probe, with respect to the liposome, were tested. The aim was to understand if the fluorescence intensity differs with the probe location. We reasoned that this could give insights about possible single leaflet staining and set the foundation for investigating spontaneous flip-flop in our system.

Liposomes of the standard membrane composition were made , encapsulating PURE system and 1nM of linearized pGEMM 7.0, together with all the lipid precursors previously described. The three samples consisted of encapsulating the probe in the swelling solution (inside), incubating liposomes with eGFP in the feeding solution (outside) and both. After incubation, the samples were imaged through confocal fluorescence microscopy (Fig 3.12) and the average rim intensities analyzed (Fig 3.11).

The experiment was designed with the rationale that having the probe outside, only the outer leaflet would be stained whereas, with the probe both inside and outside, both leaflets would be stained. For this reason, incubating simultaneously inside and outside was anticipated to lead to higher rim intensities, followed by encapsulating the probe inside and lastly, incubating with it outside. Contrary to expectation, the distribution of rim intensities did not show significant differences between locations of the probe. The median is located between 100 and 200 a.u. for the three samples, all of them containing a few liposomes with a high intensity (classified as outliers in the box plot of Fig 3.11). The non-linear behaviour of the binding, demonstrated in Section 1.3.3, can explain why the higher eGFP concentration did not lead to an increased membrane signal, regarding the sample incubated both inside and outside. Moreover, given the results for the sample incubated with eGFP outside, we hypothesized that either flip-flop of PS is occurring or potential membrane permeability to the probe might be preventing the distinction between leaflets.

To test this rationale, we investigated if the membrane is permeable to eGFP-LactC2 by analyzing the aforementioned micrographs. The pixel intensity for the eGFP channel was intentionally over saturated (Fig. 3.12, right micrographs) to conclude about the presence of the probe in the lumen of liposomes. The results were as expected: the sample with the probe inside, displayed higher pixel intensities in the lumen and the sample with probe both inside and outside, displayed similar pixel intensities in both locations. Understandably, for the former, there is signal outside. This happens because, despite the fact the probe is included in the swelling solution, the non-encapsulated components remain in solution.

The sample incubated with probe outside had liposomes with a variety of phenotypes. There were liposomes that had a pixel intensity in the lumen similar to the background (Fig. 3.12, liposomes 1 and 4), implying membrane permeability to the probe. Most of the liposomes that show colocalization of the green signal with the membrane had this behaviour. In these cases, it is possible that the membrane is stained because the probe is binding PS in the inner leaflet. On the other hand, there are also potentially impermeable liposomes, with a average lumen intensity lower than the background (Fig. 3.12, liposomes 2 and 3). Interestingly, some of the impermeable liposomes show signal colocalization with the membrane (liposome 3), proposing that in some cases, exclusive staining of the outer leaflet may happen.

The heterogeneity observed in the permeability throughout the liposome population could be due to morphological differences. With the production and incorporation of lipids, the surface tension

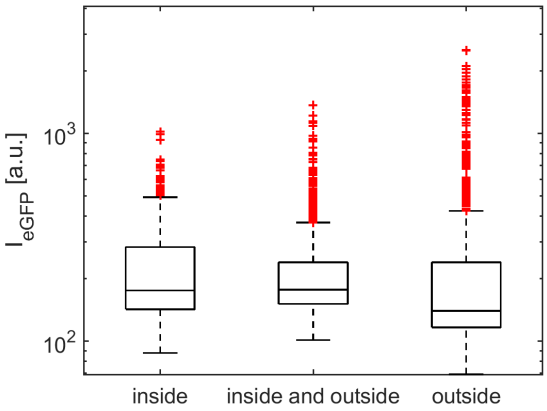


Figure 3.11: Average rim intensity does not vary substantially with probe location. Comparison of average rim intensity, extracted from micrographs of liposomes encapsulating PURE system, 1nM linearized pGEMM 7.0 and lipid precursors incubated with 150nM eGFP-lactC2 in different locations: swelling solution (inside), both swelling and feeding solution (inside and outside) and only feeding solution (outside). Several liposomes were taken into account (n=[355, 1252, 786], respectively for the samples in the presented order).

of the membrane increases due to an area difference between the inner and outer leaflets. When the tension reaches a critical value, one of the possible responses is the opening of a pore, allowing the mixture of the inner and outer solutions [36]. This mechanism explains how the eGFP-LactC2 could enter the liposome and stain the inner leaflet. Moreover, these transient defects might facilitate lipid flip-flop [37] which can further explain the similar levels of staining in the different probe locations.

3.4.4 PS expressed *in vesiculo* is incorporated in the membrane.

The key aim of this work is to detect the membrane incorporation of PS, synthesized inside liposomes by the enzymes expressed with the PURE system. The synthesis and incorporation of PS was previously verified (Section 1.4.1), however, in this experiment, expression was not confined to the interior of the liposomes. Motivated by previous experiments, demonstrating a successful encapsulation of gene expression with DNAse, we confined lipid synthesis and investigated its incorporation in the membrane from within.

Liposomes of the standard composition, encapsulating PURE system, lipid precursors and 1nM of linearized pGEMM 7.0 were incubated with DNAse and 150nM eGFP-lactC2 in the feeding solution. The negative controls were equivalent to the ones used in Section 1.4.1: no o-CoA and TP DNA substituting the linearized plasmid. As now pGEMM 7.0 is used, the negative control removing *plsB* is not feasible.

The samples were imaged with confocal microscopy (micrographs in Fig. 3.13 **A**) and the average rim intensity was extracted for a large number of liposomes (Fig. 3.13 **C**). The percentage of enriched liposomes (Fig. 3.13 **D**) was computed knowing the enrichment threshold previously obtained (Section 3.4.1). The positive sample shows a bright signal colocalized with the membrane of liposomes. The negative controls had no signal colocalization whatsoever. This is also clear from the analysis of single liposome line profiles (Fig 3.13 **B**). For the positive sample, there is overlapping of the membrane peaks in both channels. On the contrary, for the negative controls there are no significant peaks in the eGFP channel. The enrichment of liposomes is also very obvious from the distribution of average rim intensities. The rim intensity median, for the positive sample (about 500

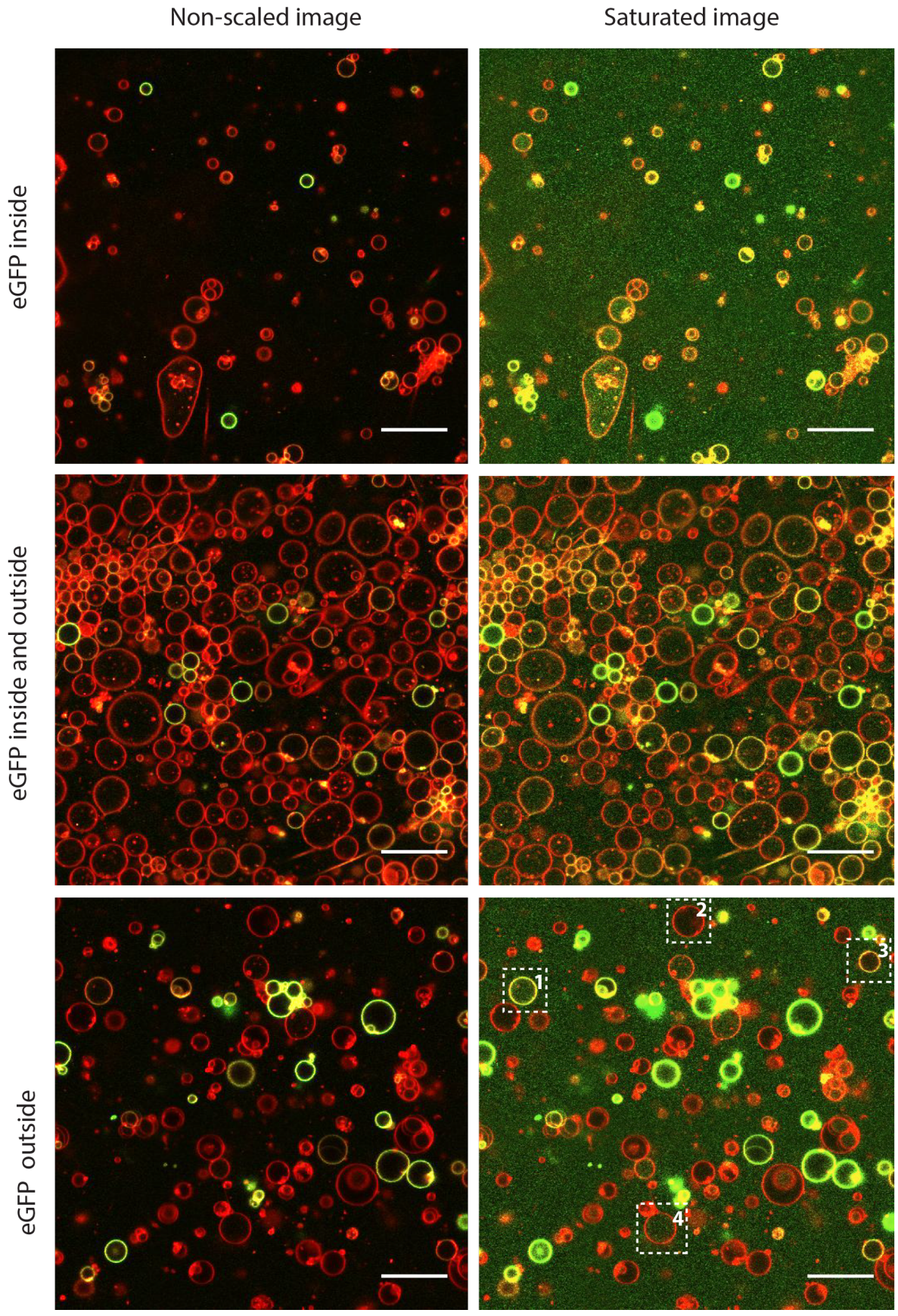


Figure 3.12: Possible permeability of liposomes to the eGFP-lactC2 protein probe. Micrographs of liposomes encapsulating PURE system, 1nM linearized pGEMM 7.0 and lipid precursors incubated with eGFP-lactC2 in different locations: swelling solution (inside), both swelling and feeding solution (inside and outside) and only feeding solution (outside). Scale bar is 20 μm.

a.u.), is higher than the enrichment threshold defined (T=422 a.u.) in Section 1.4.1, and much higher than the median for the negative controls (150 a.u.). The percentage of enriched liposomes is on average 55%, considering three repeats of the experiment, and above 40% for all of them.

In summary, PS synthesized inside liposomes is incorporated in the membrane and detected with the eGFPlactC2 probe. In comparison with the results obtained for experiments with expression everywhere, the number of enriched liposomes is lower, about 55% vs. 75%. This is in correspondence with the bulk experiments with LC-MS (section 1.4.3), where total PS accumulated was higher without confinement.

Even though smaller when compared to past experiments, the fraction of PS-enriched liposomes is surprisingly high given the amount of components necessary to be encapsulated. PURE system, by itself, has over 80 different molecules, plus the DNA template and lipid precursors added (o-CoA and G3P), as CTP and L-serine are already present. All of these should not only be present but also in concentrations high enough that lead to the production detectable amount of PS. However, since GUV were used in this project, calculations support that the large volumes of the vesicles allowed for encapsulation of all the components. Moreover, encapsulation restricts the diffusion of the components, potentially enabling a more efficient PS production.

To compute this value the most unfavorable scenario can be considered, using both the lowest concentration component (1 nM of DNA) and a lower than average vesicle radius (1 μm). The liposome volume corresponding to this radius equals to 4.2 fL, assuming complete sphericity. The average number of molecules of DNA expected to be encapsulated is 2.5, for the referred concentration and volume. Considering that a Poisson distribution (Equation 3.2) describes molecule partitioning inside vesicles, the probability of having at least one molecule encapsulated is 0.918. The second least abundant component has a concentration of 15nM and an average number of molecules of 37.5. The computed probability of finding at least one molecule is ≈ 1, applying the Poisson distribution centered around 37.5. Since the remaining components have a higher concentration, the probability of having at least one encapsulated molecule is 1 for all of them. Assuming independent encapsulation of each component, Equation 3.3 can be applied, where j is the number of events (in this case j is about 80 encapsulation events). As a result the final probability of encapsulating at least one molecule of each of the approximately 80 components is 0.918.

$$f(k;\lambda) = \frac{e^{-\lambda}\lambda^{k}}{k!} \tag{3.2}$$

$$P = \prod_{1}^{j} p_j \tag{3.3}$$

The values observed were lower than predicted as slightly more than half of liposomes were enriched, on average. We hypothesize this could be due to several reasons. Partitioning of some molecules in liposomes has shown not to follow necessarily a Poisson distribution. For instance, the aggregates formed by some of the molecules leads to liposomes with very high concentration of ribosomes or proteins [35]. These interactions between components also challenges the assumption of independent encapsulation. It was verified that the complexes formed between DNA and other macromolecules can hinder transcription, and concluded that possibly only ca. 10% of the bulk DNA was transcriptionally active [38].

On another perspective, to allow gene expression, lipid synthesis and detectable lipid incorporation more than only one molecule of each is needed. It can be possible that some liposomes encapsulate all components but the necessary threshold concentration is not achieved or the relative concentrations between components do not allow reactions. Aditionally, a stochastic behaviour was observed in our system, since a higher concentration of DNA does not lead necessarily to higher concentration of protein [38].

Moreover, the liposome-to-liposome variability can also be explained by differences in permeability. On one hand, if the probe does not access the inner leaflet and flip-flop does not occur, the liposome would not be stained even with incorporation of PS. On the other hand, if the liposome is permeable, DNAse could also enter and degrade intra-vesicular pGEMM 7.0.

3.5 Kinetic of PS incorporation in the membrane of liposomes

Investigating the kinetic nature of the incorporation could elucidate on the time scale of gene expression, lipid synthesis and incorporation. To analyze this, liposomes of the standard composition were made, encapsulating PURE system, 1nM of linearized pGEMM 7.0 and the precursors except o-CoA, which was present in the feeding solution. The difference was that, for this particular experiment, the incubation was performed in the microscope, to allow imaging from the beginning of gene expression.

After 15 hours of incubation, the end-point micrographs show liposomes enriched in PS Fig. 3.14 **A**, as in previous equivalent experiments (Section 1.4.5). Micrographs of one enriched liposome through time show the increase in eGFP signal colocalized with the membrane (Fig. 3.14 **B**). It is clear that the intensity of fluorescence increases through each of the close-ups, eventually achieving a plateau indicated by the similar intensity for six and fifteen hours. On a side note, it is important to clarify that the close-up taken at 15h presents two smaller liposomes attached since these aggregated while imaging. This phenomena of liposome aggregation was observed often in kinetic experiments. Moreover, the radius of this liposome varied slightly during imaging. However, slight liposome movement during acquisition was often observed in both the positive samples and negative controls. This was likely due to fluctuations on the imaged z-plane. It was also observed that liposomes move in and out the imaged field of view. For this reason, it is not possible to conclude on possible growth due to lipid incorporation.

In an attempt to quantitatively describe the fluorescence increase in the membrane, the average rim intensity of individual liposomes that showed PS enrichment was extracted, for each time point. The curves of eGFP fluorescence over time, for three independent repeats, are shown in Fig. 3.15. A negative control without o-CoA in the feeding solution was also performed and imaged at similar time-points. The liposomes of the positive sample showed a steep increase of fluorescence over time, followed by a plateau. The negative control displayed an approximately constant fluorescence, about one order of magnitude lower than the final intensity of the positive sample.

A sigmoid function was fitted to the average rim intensity of enriched liposomes, for each experimental repeat. The expression is shown Eq. 3.4, where B is the growth rate, k is the upper assyntote and t_f is the time corresponding to the inflection point. The parameters obtained for each repeat are presented in Table 3.2.

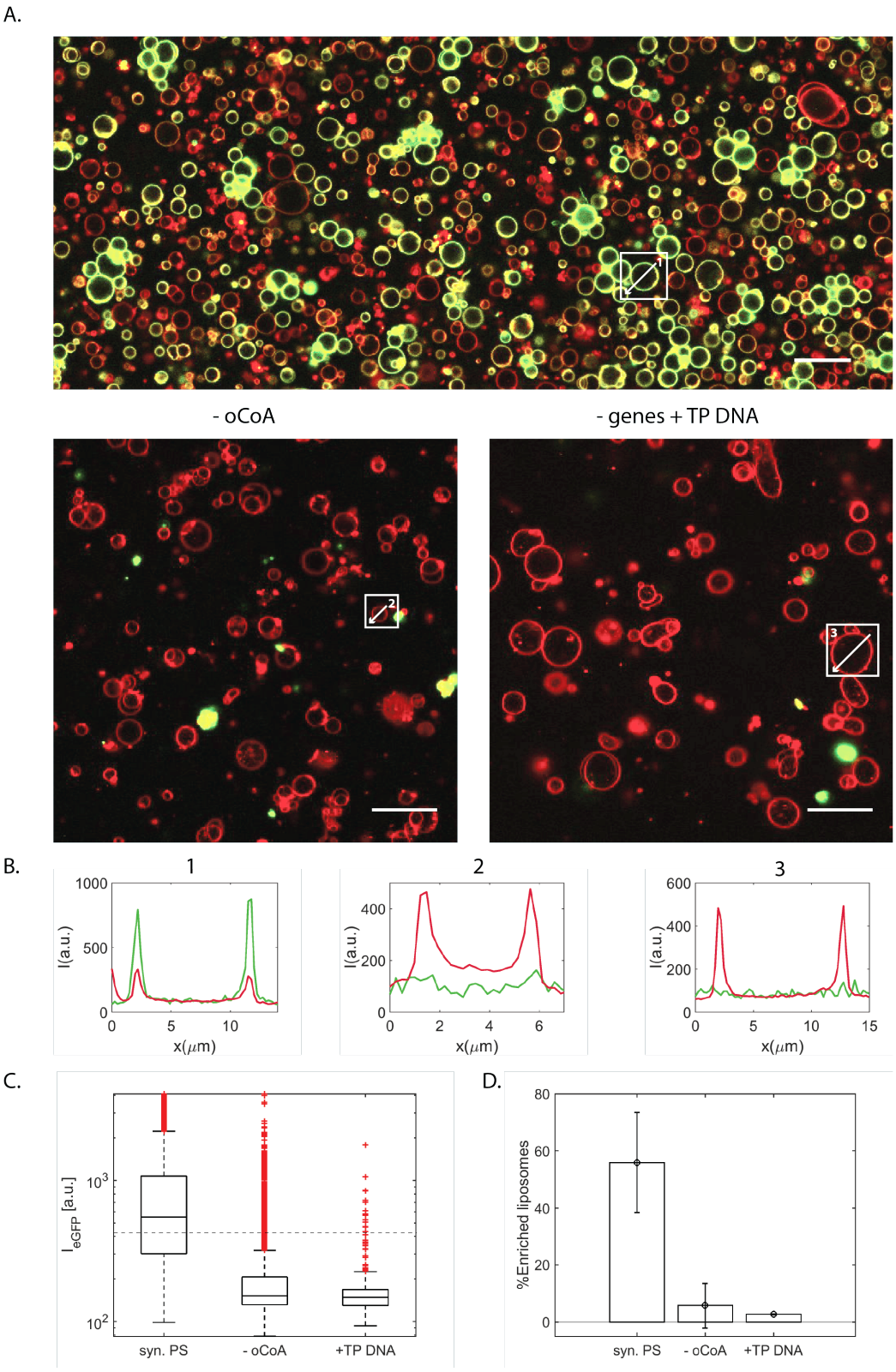


Figure 3.13: Synthesized PS in liposomes with confined gene expression incorporates in the membrane of liposomes.A. Confocal micrographs of liposomes (labelled in red) encapsulating linearized pGEMM 7.0, PURE system, lipid precursors (G3P, CTP and L-serine) and incubated with 150 nM of eGFP-lacC2 (labelled in green), 100μM of o-coA and DNAse outside. Two negative controls are presented: (left) no o-CoA was fed to the system and (right) TP DNA was encapsulated instead of linearized pGEMM 7.0. Scale bar is 20 μm. **B.** Line intensity profile for three liposomes (framed in **A.**), given that each of the colors represent the corresponding channel. **C.** Distribution of average intensities across the rim of liposomes in the eGFP channel. The dashed line indicates the threshold used to consider a liposome enriched in PS. **D.** Percentage of PS-enriched liposomes for each sample. Four repeats of each sample were made, from which several liposomes (n= [4048, 3642, 569], respectively for the samples in the presented order) were taken into account. The error bars represent the inter-sample standard deviation.

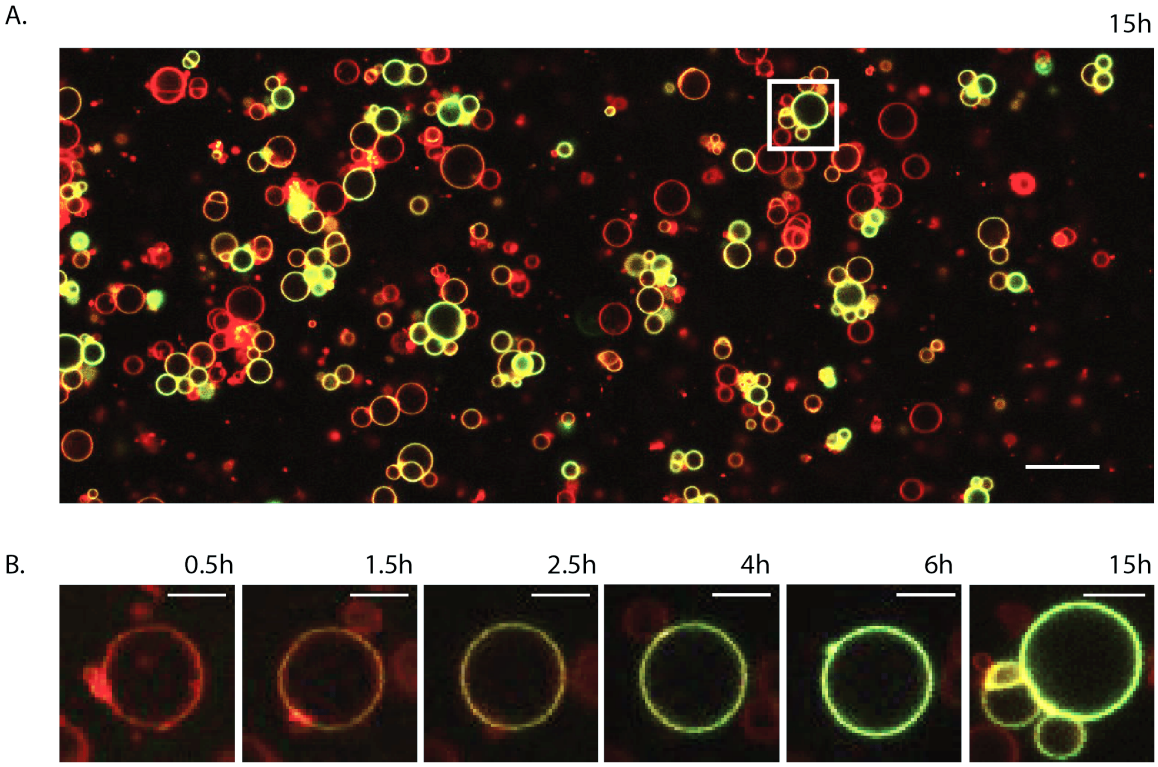


Figure 3.14: Time-lapse of synthesized PS incorporation in liposomes.A. End-point confocal micrographs of liposomes (labelled in red) encapsulating linearized pGEMM 7.0, PURE system, L-serine and CTP, incubated with 150 nM of eGFP-lactC2 (labelled in green), 100μM of o-coA and 0.07U/μL DNAse outside. Scale bar is 20μm. **B.** Single liposome (highlighted in **A.**) followed over time. Scale bar is 10μm.

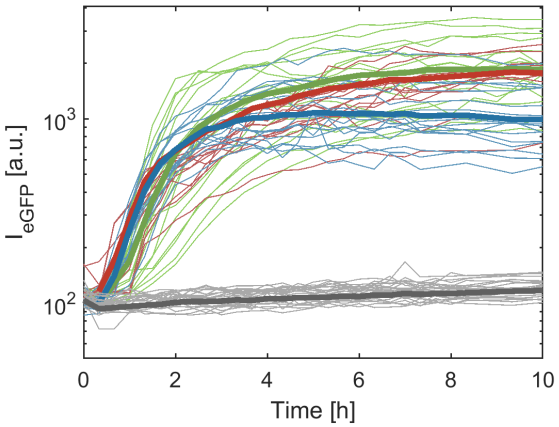


Figure 3.15: Sigmoidal behaviour of fluorescence increase in enriched liposomes. The average rim intensity of liposomes was extracted over time. Each color is a different repeat (green, blue and red represent repeats of the positive sample) and grey represents a negative control without o-CoA. The thin lines corresponds to individual liposomes and the thick line corresponds to the average for the liposomes of the corresponding repeat. Several liposomes were analyzed : n=18 (green), n=12 (red), n=17 (blue), n=28 (grey).

$$I = \frac{k}{1 + e^{-B(x - t_f)}} \tag{3.4}$$

The third repeat differs significantly from the other two. This repeat was done with a different stock of protein (produced with Rosetta 2 instead of Rosetta ER2566 cells) unlike all previous experiments. Possibly, a different fraction of protein was active and influenced the binding and fluorescence signal. This reflects a drawback of the method: the concentration of protein can affect the upper assyntote of the sigmoid (likely since not all PS is saturated at 150nM) and the time to achieve that plateau. In any case, a rough estimate of the time the system takes to produce and incorporate lipids can be extracted. The time to achieve the plateau can be estimated by doubling the t_f: between 3.1 and 5.7 hours.

Table 3.2: Parameters of the sigmoid function fitted for each of the repeats.

	B (h^{-1})	k (a.u.)	$t_f(h)$
Repeat 1	0.90	1850	2.83
Repeat 2	0.73	1722	2.85
Repeat 3	1.84	1004	1.58

The lipid incorporation process comprises several steps: transcription and translation, incorporation or association of the enzymes with the membrane, lipid synthesis from precursors and incorporation of the phospholipid in the bilayer. In previous projects of our lab, the kinetic of the PURE system transcription and translation was characterized by expressing and measuring YFP fluorescence intensity inside single liposomes (unpublished work by Duco Blanken). It was concluded that gene expression achieved a plateau after about 5h for the PUREfrex 2.0 system.

Even though it might not seem clear why gene expression takes longer to reach a plateau than lipid synthesis and incorporation, since for the latter there are more processes involved, these results are not necessarily incompatible. Firstly, all the sequential processes from transcription to phospholipid incorporation occur in parallel inside vesicles. That is, from the moment DNA is

transcribed to mRNA (and a specific minimum threshold is achieved), translation can also start and so on. Combining this notion with the characteristics of our detection method studied in this project, particularly the low concentration of PS needed to achieve saturation (Section 3.3.3), it is conceivable that the plateau for the signal of PS detection is lower than the plateau achieved for gene expression.

There was previous kinetic analysis of lipids synthesized by the enzymes expressed with the PURE system [12]. These experiments were done by the expression of the GPAT and LPAAT enzymes in the presence of liposomes and using p-CoA as a precursor. This means the gene expression occurred in the extravesicular medium and the liposomes served as a membrane support for the lipids. After the bulk lipid fraction was extracted and analyzed with MS (for each time-point). It was verified that the concentration plateau of PA was reached in 12h.

Comparatively, the results obtained in the present work suggest a much faster process. Since the past experiments were done in bulk and in the expression occurred in the extra-vesicular solution, the contribution of the dilution of the components and the diffusion restraints can explain the comparative delay. The plateau of production can be achieved for a lower concentration of end-product when the experiment is done *in vesiculo*. Furthermore, as seen from the third repeat of our experiment small changes in protocol can lead to heterogeneous results.

The work by Scott *et al.* was done using PUREfrex 1.0 (and not 2.0). Recently, it has been observed in our lab that PUREfrex 2.0 has increased protein yield. With a higher concentration of enzymes produced in the same time-frame, the concentration threshold necessary to trigger lipid synthesis might be achieved earlier. This theory is confirmed by the lag-phasepresent in the previous kinetic experiments (about 3 hours) and that is not observed in the present work.

Our work detected PS presence, which corresponds to two more enzyme catalyzed steps when compared to the production of PA. However, each step does not correspond to a linear increase in the time needed to reach the plateau. Not only because the enzymatic steps occur in parallel (as mentioned previously) but also because it was observed that the enzymes may work in tandem when incorporated simultaneously in the membrane, leading to higher plateau concentrations [12].

This work gave further insight on the incorporation of lipids synthesized by enzymes expressed with PURE system, confined to the lumen of GUV. The incorporation was imaged in real-time and promising kinetic data extracted.

Chapter 4

Conclusions and Outlook

In this work we detect for the first time incorporation of phospholipids synthesized *in vesiculo*. Previous work done by our lab, in the scope of the lipid synthesis in the minimal cell, detected production of phospholipids in bulk. Herein, we demonstrate the lipids synthesized incorporate in the membrane of liposomes, in a single-vesicle approach.

To reach this breakthrough we followed a list of milestones. Firstly, we verified that the lactC2 bound specifically to the membranes containing PS. However, it was also observed that for high concentrations of the probe, non specific binding can occur. Throughout the work we used a lower concentration that allowed a binary response regarding the presence of PS. The presence of non-specific binding did not allow to verify if the working concentration was in the plateau of fluorescence, i.e. if all the PS L-serine moieties were saturated by the probe. Moreover, after testing the binding response to the relative quantity of PS in the membrane, we observed a non-linear behaviour. For these reasons, it is not possible to use this method to quantify how much PS actually incorporates in the membrane and therefore what is the relative liposome growth.

Secondly, we developed a experimental setup to confine gene expression inside liposomes by incubating with DNAse after swelling. Afterwards, different probe localizations with respect to the liposome (intra- and/or extravesicular) were tested and surprisingly, the rim intensity was similar in all the configurations. We also observed eGFP fluorescence in the liposome lumen, in the cases of extravesicular supply of probe, which led us to conclude on possible diffusion of probe through the membrane. Even though the observations suggested spontaneous lipid flip-flop, the possibility that the membrane is permeable to the probe does not allow to conclude in this regard. Since the flip-flop of synthesized lipids is essential for the liposome growth, further work should be done to clarify its occurrence in our system.

Subsequently, we proceeded to image the incorporation of PS synthesized by enzymes expressed with the PURE system and confined to liposomes. From the micrographs and posterior image analysis, we observed clear enrichment of 55% of the liposomes, on average. This was the main result of this work as it demonstrates the incorporation of the lipids in the membrane. Furthermore, this work set the ground for an imaging technique to track enriched liposomes with different purposes.

In an over-time imaging experiment, the kinetic of lipid synthesis and incorporation was analyzed for enriched liposomes. As a rough estimate, the process takes up to 5.7 hours. Nevertheless, we observed a high variability, within sample (liposome-to-liposome) and within repeats (day-to-day). The complexity of the system does not allow for complete control of the effect of all external variables.

The current and past efforts of our lab to characterize and model the separate parts (for instance the dependence of the PURE system expression levels on external factors or the correlation between DNA and protein concentration) will allow for higher control and reproducibility in the future. Regarding the intra-sample variability, we reasoned macromolecule aggregation and intrinsic stochasticity of encapsulation could be the cause.

In the future, the data generated in this project can contribute to elucidate the time-frame needed for growth-based division. Fluorescence-based methods pose as an interesting strategy for both growth and flip-flop detection. Regarding growth, there are fluorescence techniques used to detect membrane fusion that can be adapted. These could be based on fluorophores that self-quench and, with growth, start having a fluorescent signal or, on the contrary, liposomes that contain both fluorescent energy transfer donor and acceptor lipid probes (e.g. rhodamine-PE and NBD-PE) and with the expansion, fluorescence intensity would decrease. Techniques with similar principles could be applied for flip-flop detection, for instance using a lipid bound to a fluorophore and incubating the liposomes with a molecular quencher in the feeding solution (for which the liposomes would be impermeable). Both of the approaches could be combined, achieving detection of incorporation, growth and flip-flop. To allow deformation and division by increasing the surface-to-area ratio, high quantities of lipid have to be produced. Even though quantification was not established with this work, in previous experiments it was verified that the bulk amount of lipids increased about 5%. One of the ways currently being researched to lead to higher lipid yield is to expand the pathway to allow feeding the system with simpler precursors that are more water soluble and allow for higher initial concentrations. Better understanding and modelling of the PURE system dynamics will also reveal the most relevant parameters to manipulate in order for lipid yield improvement.

In conclusion, we believe this work demonstrated the membrane presence of the synthesized phospholipids and poses as an advance for future studies regarding liposome growth and division.

Bibliography

1. Dowhan, W. Molecular basis for membrane phospholipid diversity: Why are there so many lipids? *Annual Review of Biochemistry* **66,** 199–232 (1997).
2. Baker, C. D., Basu Ball, W., Pryce, E. N. & Gohil, V. M. Specific requirements of nonbilayer phospholipids in mitochondrial respiratory chain function and formation. *Molecular Biology of the Cell* **27,** 2161–2171 (2016).
3. Sohlenkamp, C. & Geiger, O. Bacterial membrane lipids: Diversity in structures and pathways. *FEMS Microbiology Reviews* **40,** 133–159 (2015).
4. Gurtovenko, A. A. & Vattulainen, I. Molecular mechanism for lipid flip-flops. *Journal of Physical Chemistry B* **111,** 13554–13559 (2007).
5. Contreras, F. X., Sánchez-Magraner, L., Alonso, A. & Goñi, F. M. Transbilayer (flip-flop) lipid motion and lipid scrambling in membranes. *FEBS Letters* **584,** 1779–1786 (2010).
6. Epand, R. M. *The Biophysics of Cell Membranes* (2017).
7. Gidden, J., Denson, J., Liyanage, R., Ivey, D. M. & Lay, J. O. Lipid compositions in Escherichia coli and Bacillus subtilis during growth as determined by MALDI-TOF and TOF/TOF mass spectrometry. *International Journal of Mass Spectrometry* **283,** 178–184 (2009).
8. E.P. Kennedy and S.B. Weiss. The Function of Cytidine Coenzymes in the Biosynthesis of Phospholipids. *J Biol Chem* **222,** 193–214 (1956).
9. 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).
10. Yao, J. & Rock, C. O. Phosphatidic acid synthesis in bacteria. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1831,** 495–502 (2013).
11. Dowhan, W. A retrospective: Use of Escherichia coli as a vehicle to study phospholipid synthesis and function. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids* **1831,** 471–494 (2013).
12. Scott, A. *et al.* Cell-free phospholipid biosynthesis by gene-encoded enzymes reconstituted in liposomes. *PLoS ONE* **11,** 1–23 (2016).
13. Exterkate, M., Caforio, A., Stuart, M. C. & Driessen, A. J. Growing Membranes in Vitro by Continuous Phospholipid Biosynthesis from Free Fatty Acids. *ACS Synthetic Biology* **7,** 153–165 (2018).
14. Schmidli, P. K., Schurtenberger, P. & Luisi, P. L. Liposome-Mediated Enzymatic Synthesis of Phosphatidylcholine as an Approach to Self-Replicating Liposomes. *Journal of the American Chemical Society* **113,** 8127–8130 (1991).
15. Kuruma, Y., Stano, P., Ueda, T. & Luisi, P. L. A synthetic biology approach to the construction of membrane proteins in semi-synthetic minimal cells. *Biochimica et Biophysica Acta - Biomembranes* **1788,** 567–574 (2009).
16. Fuchs, B., Süß, R., Teuber, K., Eibisch, M. & Schiller, J. Lipid analysis by thin-layer chromatography-A review of the current state. *Journal of Chromatography A* **1218,** 2754–2774 (2011).

17. Bele, A. An overview on thin layer chromatography. *International Journal of Pharmaceutical Sciences and Research* **2,** 256–267 (2011).

18. Bierhanzl, V. M. *et al.* Direct injection mass spectrometry, thin layer chromatography, and gas chromatography of Bacillus subtilis phospholipids. *Monatshefte für Chemie - Chemical Monthly* **147,** 1385–1391 (2016).

19. Hoffmann, E. D. & Stroobant, V. *Mass Spectrometry - Priniples and Applications.* **6,** 945–61. arXiv: NIHMS150003 (2007).

20. Cajka, T. & Fiehn, O. Comprehensive analysis of lipids in biological systems by liquid chromatography-mass spectrometry. *Trends in Analytical Chemistry,* 192–206 (2015).

21. Maekawa, M. & Fairn, G. D. Molecular probes to visualize the location, organization and dynamics of lipids. *Journal of Cell Science* **127,** 4801–4812 (2014).

22. Daemen, S., van Zandvoort, M. A., Parekh, S. H. & Hesselink, M. K. Microscopy tools for the investigation of intracellular lipid storage and dynamics. *Molecular Metabolism* **5,** 153–163 (2016).

23. Zhao, C., Du, G., Skowronek, K., Frohman, M. A. & Bar-Sagi, D. Phospholipase D2-generated phosphatidic acid couples EGFR stimulation to Ras activation by Sos. *Nature Cell Biology* **9,** 707–712 (2007).

24. Bohdanowicz, M. *et al.* Phosphatidic acid is required for the constitutive ruffling and macropinocytosis of phagocytes. *Molecular Biology of the Cell* **24,** 1700–1712 (2013).

25. Iwamoto, K. *et al.* Curvature-dependent recognition of ethanolamine phospholipids by duramycin and cinnamycin. *Biophysical Journal* **93,** 1608–1619 (2007).

26. Waehrens, L. N., Heegaard, C. W., Gilbert, G. E. & Rasmussen, J. T. Bovine lactadherin as a calcium-independent imaging agent of phosphatidylserine expressed on the surface of apoptotic HeLa cells. *Journal of Histochemistry and Cytochemistry* **57,** 907–914 (2009).

27. Otzen, D. E., Blans, K., Wang, H., Gilbert, G. E. & Rasmussen, J. T. Lactadherin binds to phosphatidylserine-containing vesicles in a two-step mechanism sensitive to vesicle size and composition. *Biochimica et Biophysica Acta - Biomembranes* **1818,** 1019–1027 (2012).

28. Shao, C., Novakovic, V. A., Head, J. F., Seaton, B. A. & Gilbert, G. E. Crystal structure of lactadherin C2 domain at 1.7 A° resolution with mutational and computational analyses of its membrane-binding motif. *Journal of Biological Chemistry* **283,** 7230–7241 (2008).

29. Shi, J., Pipe, S. W., Rasmussen, J. T., Heegaard, C. W. & Gilbert, G. E. Lactadherin blocks thrombosis and hemostasis in vivo: Correlation with platelet phosphatidylserine exposure. *Journal of Thrombosis and Haemostasis* **6,** 1167–1174 (2008).

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

31. Van Nies, P., Canton, A. S., Nourian, Z. & Danelon, C. *Monitoring mRNA and protein levels in bulk and in model vesicle-based artificial cells* 1st ed., 187–214 (Elsevier Inc., 2015).

32. Wingfield, P. T. *Overview of the Purification of Recombinant Proteins* 1–50 (2016).

33. Goldfine, H. Bacterial membranes and lipid packing theory. *Journal of lipid research* **25,** 1501–7 (1984).

34. Bravo, A. & Salas, M. Initiation of bacteriophage 29 DNA replication in vivo: Assembly of a membrane-associated multiprotein complex. *Journal of Molecular Biology* **269,** 102–112 (1997).

35. De Souza, T. P., Fahr, A., Luisi, P. L. & Stano, P. Spontaneous Encapsulation and Concentration of Biological Macromolecules in Liposomes: An Intriguing Phenomenon and Its Relevance in Origins of Life. *Journal of Molecular Evolution* **79,** 179–192 (2014).

36. Sakuma, Y. & Imai, M. From Vesicles to Protocells: The Roles of Amphiphilic Molecules. *Life* **5,** 651–675 (2015).

37. Bennett, W. F. D. & Tieleman, D. P. The Importance of Membrane Defects—Lessons from Simulations. *Accounts of Chemical Research* **47,** 2244–2251 (2014).

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

Appendix A

Primers list

Table A.1: Primers used in this study.

Primer	Sequence 5'-3'	Description
ChD 25	GATGCTGTAGGCATAGGCTTGG	Forward primer - Colony PCR / Sanger sequencing primer
ChD 288	CGATGCGTCCGGC	Sanger sequencing primer
ChD 310	GGATCTCGACGCTCTCCCTTATG	Reverse primer - Colony PCR
ChD 471	CCGCTGAGCAATAACTAGC	Reverse primer - Amplification of pET-11a backbone
ChD 848	TATACATATGGGCAGCAGCCATCATCATCATCATCACAGCAG CGGCCTGGTGCCGCGCGGCAGCCATATGGTGAGCAAGGGCG AGG	Forward primer - Amplification of pET-11a backbone
ChD 849	AACTCAGCTTCCTTTCGGGCTTTGCTAACAGCCCAGCAGCTC C	Forward primer - Amplification of eGFP-lactC2 genes
ChD 850	GATGATGGCTGCTGCCCATATGTATATCTCCTTCTTAAAGTT AAAC	Reverse primer - Amplification of eGFP-lactC2 genes

Appendix B

Mass-spectrometry detection settings

Table B.1: Mass spectrometry settings for detection of the phospholipids analyzed in this work. Asterisk (*) indicates species incorporating 13C-G3P, resulting in a 3 Da mass shift with respect to the regular species (6 Da for PG). Table adapted from unblished work by Duco Blanken and Dr. David Foschepoth.

Compound class	Compound name	Precursor Ion (Da)	Product Ion (Da)	Fragmentor (V)	Collision energy (eV)	Polarity
Matrix	DOPC	786.4	184.0	135	50	Positive
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
C16:0-product	DPPS*	737.5	255.1	180	41	Negative
C18:1-product	DOPS*	789.5	281.2	130	40	Negative
C18:1-product	DOPE*	745.5	281.2	140	25	Negative
Mixed product	POPS*	763.5	281.2	130	40	Negative
			255.1	180	41	Negative

Appendix C

Average rim intensity

C.1 Rationale

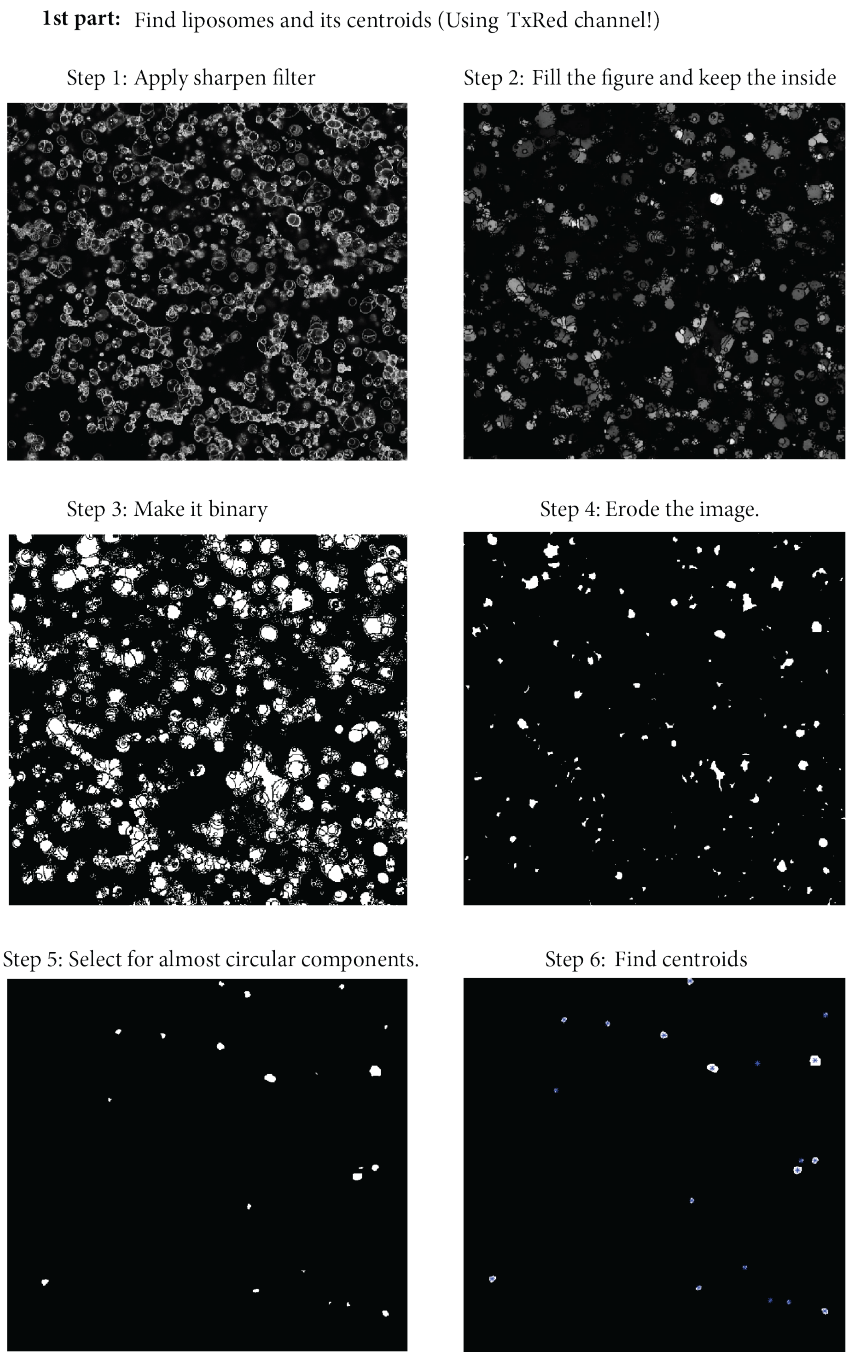


Figure C.1: Steps for centroid detection in a confocal imaging micrograph.

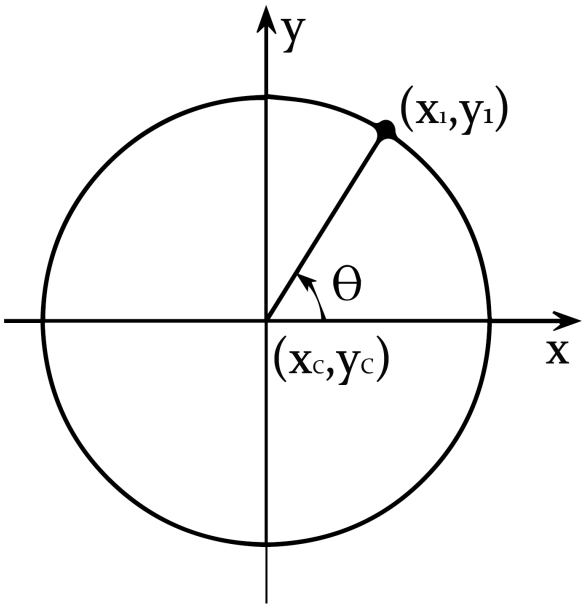


Figure C.2: Average rim intensity extraction for a single liposome.

Firstly, the centroids of a population of liposomes within the micrograph are computed with the method presented in Fig. C.1. After having having the coordinates of the centroid ($x_c and y_c$) and the aproximate radius (R), the average rim intensity can be extracted (Fig. C.2). For this, two variables are used. The first, θ, varies from 0 to 2π, in steps 0.1 rad. For each of these angles, a line starting on the centroid and with a length of 1.5R is drawn. The pixel intensity of all pixels within this line is extracted. The maximum pixel intensity is computed and should correspond to the rim of the liposome (x_1, y_1). After going through all the 63 angles, the average is calculated. The MATLAB script to perform both steps is presented below. The input files of the rim_int function are the pixel intensity matrices of the images of each separate channel (I_TR and I_GFP).

C.2 MATLAB Script: Rim intensity

```
function [mean_TR, mean_GFP] = rim_int(I_TR,I_GFP)

% define sharpen filter
h = [-1 -1 -1; -1 12 -1; -1 -1 -1]./4;

%apply sharpen filter
I_sharp = imfilter(I_TR,h);

%fill the areas inside liposomes
I_filled = imfill(I_sharp);

%imshow(I_filled,[])

%stay only wih the filling
I_inside = I_filled-I_sharp;

%create a matrix the size of I_TR
fgm = zeros(size(I_TR));
```

```
19
20 %Define threshold and make it binary
21 fgm(I_inside<=100)= 0;
22 fgm(I_inside>100) =1;
23
24 %%Fill the holes
25 fgm = imfill(fgm,'holes');
26
27 %Erode 5 pixels
28 n_erosion = 5;
29 se = strel('disk',n_erosion);
30 fgm = imerode(fgm,se);
31
32 %bwconncomp takes information about the image ( number of components,
       image
33 %size and pixelidxlist
34
35 CC = bwconncomp(fgm);
36 pixels = CC.PixelIdxList;
37 props = regionprops(CC,'PixelIdxList','Perimeter','Area','Centroid');
38
39
40 % select for circular things
41 for j = 1:length(pixels)
42     %Place{j} = props(j).Centroid;
43     P(j) = props(j).Perimeter;
44     A(j) = props(j).Area;
45
46     C(j) = P(j)^2/(4*pi*A(j));
47
48     if C(j) >1.05
49         fgm(CC.PixelIdxList{j}) = 0;
50     elseif C(j) < 0.95
51         fgm(CC.PixelIdxList{j}) = 0;
52     end
53
54 end
55
56 CC = bwconncomp(fgm);
57
58 pixels = CC.PixelIdxList;
59
60 %Compute the centroids
61 props = regionprops('table',CC,'Centroid','MajorAxisLength','
       MinorAxisLength');
62
63 centroids = cat(1, props.Centroid);
64
65 %compute the radii of the liposomes
66 diameters = mean([props.MajorAxisLength props.MinorAxisLength],2);
67 radii = diameters/2 + n_erosion;
```

```
68
69 clear CC
70 clear pixels
71
72 %For each liposome found compute the rim intensity for 63 angles
73 and perform the average
74
75 theta_vector = 0:0.1:2 pi;
76 for n = 1:length(centroids)
77 d = 0:0.1:1.5 radii(n);
78 for i = 1:63
79 theta = theta_vector(i);
80 for j = 1:length(d)
81     try
82 TR_line{n}{i}(j) = double(I_TR(round(centroids(n,2)+sin(theta) d(j)),
    round(centroids(n,1)+cos(theta) d(j))));
83 GFP_line{n}{i}(j) = double(I_GFP(round(centroids(n,2)+sin(theta) d(j)),
    round(centroids(n,1)+cos(theta) d(j))));
84     catch
85     end
86 end
87 TR{n}(i) = double(max(TR_line{n}{i}));
88 GFP{n}(i) = double(max(GFP_line{n}{i}));
89 end
90 mean_TR(n) = mean(TR{n});
91 mean_GFP(n) = mean(GFP{n});
92 end
```

Appendix D

Boxplot interpretation

Box plots provide a graphical visualization for the summary statistics of the data. The box stretches from the 25^{th} to the 75^{th} percentile, and the distance between the top and bottom of the box is considered the interquartile range. The line dividing the box corresponds to the median of the sample (50^{th} percentile). Lines extend from the edges of the box, designated whiskers. Whiskers are drawn from either the top or bottom of the interquartile ranges to the furthest observations within the whisker length. The values beyond the whisker length are plotted as outliers. An example of a box plot is presented in Fig. D.1.

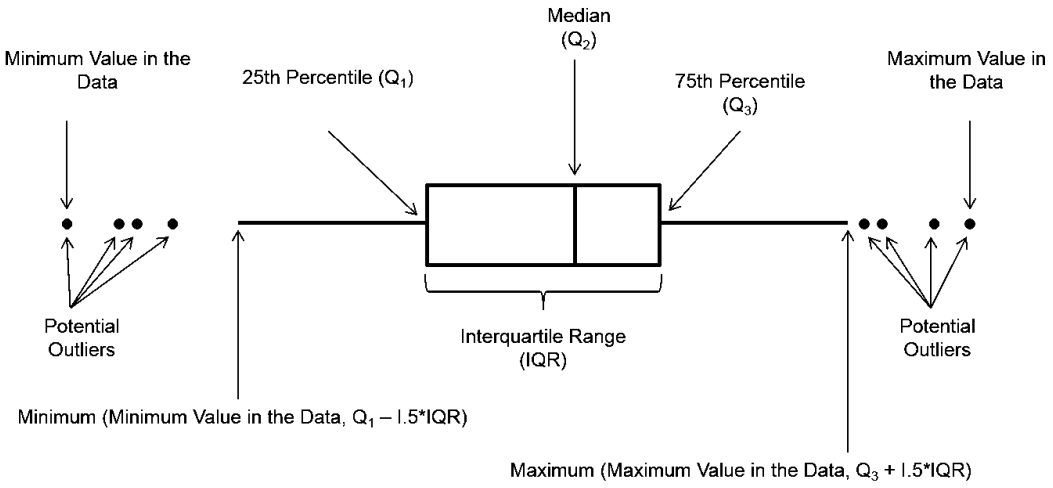


Figure D.1: Interpretation of a box plot.