Université Sorbonne
Paris Cité

Thèse de doctorat

de l'Université Sorbonne Paris Cité

Préparée à l'Université Paris Diderot

BioSPC ED562

Laboratoire de Biologie Physico-Chimique / UMR7099

Molecular basis of membrane protein production and intracellular membranes proliferation in *E. coli*

Par Federica Angius

Thèse de doctorat de Biologie Moleculaire

Dirigée par: Bruno Miroux

Présentée et soutenue pubiquement à Paris le 13 Octobre 2017

Président du jury: Pebay-Peyroula Eva / Professeur / Université Joseph Fourier

Rapporteurs: Jean-Michel Jault / Directeur de Recherche / Institut de Biologie et Chimie des Protéines; Elisabeth Darrouzet / Chargée de recherche / Institut de Biosciences et Biotechnologies de Grenoble BIG

Examinateurs: Dirk Schüler / Professeur / Université de Bayreuth;

Elisabetta Bergantino / Professeur / Université de Padoue

Directeur de thèse: Bruno Miroux / Directeur de Recherche / Institut de Biologie Physico-Chimique

Base moléculaire de la production des protéines membranaires et de la formation des membranes intracellulaire dans *Escherichia coli*

Résumé: Le système d'expression le plus utilisé pour la production des protéines membranaires, est le système basé sur l'ARN polymérase T7 (ARNpol T7) (Hattab et al., 2015). L'inconvénient de ce système est néanmoins que la vitesse de transcription de l'ARNpol T7 est dix fois plus rapide que celle de l'enzyme bactérienne. Depuis l'isolement de mutants spontanés, notamment C41(DE3) et C43(DE3) (Miroux et Walker, 1996) et l'identification de leurs mutations dans le génome, il apparaît clairement que la toxicité provoquée par la surproduction des protéines membranaires est liée à la quantité trop élevée d'ARNpol T7 dans la cellule (Wagner et al., 2008; Kwon et al., 2015). Les protéines membranaires ont besoin d'une vitesse de transcription/traduction plus basse pour se replier correctement dans la membrane de la bactérie. Le premier objectif de ma thèse était d'étendre l'amplitude du promoteur du système T7 sur laquelle est basée l'expression des protéines. Pour cela, nous avons isolé et caractérisé de nouvelles souches bactériennes dans lesquelles le niveau d'ARNpol T7 était efficacement régulé par un mécanisme non transcriptionnel très favorable à l'expression des protéines membranaires (Angius et al., 2016). Le deuxième objectif était de comprendre la prolifération des membranes intracellulaires chez *E. coli* suite à la surexpression de la protéine AtpF, une sous unité membranaire du complexe de l'ATP synthétase (Arechaga et al., 2000). Pour mieux comprendre les voies métaboliques impliquées dans la biogenèse, la prolifération et l'organisation des membranes, nous avons utilisé une approche de séquençage d'ARN à haut débit à différents temps après induction de la surexpression de la sous-unité AtpF dans la souche C43(DE3). Ensuite, et en collaboration avec Gerardo Carranza and Ignacio Arechaga (Université de Cantabria, Espagne), nous avons construit et étudié des mutants de C43(DE3) déficients pour les trois gènes codants pour des enzymes de la biosynthèse des cardiolipides afin d'évaluer leur participation dans la biogénèse des membranes intracellulaires.

Mots clefs: *E. coli*, production des protéines membranaires, ARN T7polymérase, membranes intracellulaires, cardiolipides, interactions protéine-lipide.

Molecular basis of membrane protein production and intracellular membranes proliferation in *Escherichia coli*

Abstract: The most successful expression system used to produce membrane proteins for structural studies is the one based on the T7 RNA polymerase (T7 RNAP) (Hattab et al., 2015). However, the major drawback of this system is the over-transcription of the target gene due to the T7 RNAP transcription activity that is over ten times faster than the *E. coli* enzyme. Since the isolation of spontaneous mutants, namely C41(DE3) and C43(DE3) (Miroux and Walker, 1996) and the identification of their mutation in the genome, it becomes clear that reducing the amount of the T7 RNAP level removes the toxicity associated with the expression of some membrane proteins (Wagner et al., 2008; Kwon et al., 2015). Also, some membrane proteins require a very low rate of transcription to be correctly folded at the *E. coli* membrane. The first objective of my PhD was to extend the promoter strength coverage of the T7 based expression system. We used genetic and genomic approaches to isolate and characterize new bacterial strains (Angius et al., 2016) in which the level of T7 RNAP is differently regulated than in existing hosts.

A second objective was to understand intracellular membrane proliferation in *E. coli*. Indeed it has been shown that over-expression of membrane proteins, like over-expression of AtpF of *E. coli* F_1F_o ATP synthase is accompanied by the proliferation of intracellular membranes enriched in cardiolipids (Arechaga et al., 2000). To understand metabolic pathways involved in membrane biogenesis, proliferation and organization, we used a RNA sequencing approach at several time point upon over-expression of the F-ATPase b subunit in C43(DE3) host. On the other hand, in collaboration with Gerardo Carranza and Ignacio Arechaga (University of Cantabria, Spain) we studied C43(DE3) *cls* mutants, in which the cardiolipids genes A, B and C are deleted, to test how they participate to intracellular membranes structuration.

Keywords: *E. coli*, production of membrane proteins, T7 RNA polymerase, intracellular membranes, cadiolipin, protein-lipid interactions.

"We are spaniards, africans, phoenicians,
carthaginians, romans, arabs,
pisans, byzantines, piedmontese.
We are the golden yellow brooms
that droop on rocky paths,
as large lighted lamps.
We are the wild solitude,
the immense and profound silence,
the splendor of the sky,
the white flower of cistus.
We are the uninterrupted reign of mastic,
of waves that flow on ancient granites,
of wild roses, of the wind,
of the immensity of the sea.
We are an ancient land of long silences,
of large and pure horizons, of gloomy trees
of mountains burned by sunshine and revenge.
We are Sardinians."
(Grazia Deledda)

"Siamo spagnoli, africani, fenici,
cartaginesi, romani, arabi,
pisani, bizantini, piemontesi.
Siamo le ginestre d'oro giallo
che piovono sui sentieri rocciosi,
come grandi lampade accese.
Siamo la solitudine selvaggia,
il silenzio immense e profondo,
lo splendore del cielo,
il bianco fiore del cisto.
Siamo il regno ininterroto del lentisco,
delle onde che ruscellano i graniti antichi,
della rosa canina, del vento,
dell'immensità del mare.
Siamo una terra antica di lunghi silenzi,
di orizzonti ampi e puri, di piante fosche,
di montagne bruciate dal sole e dalla vendetta.
Noi siamo sardi."
(Grazia Deledda)

Acknowledgements

First and foremost I would like to thank my supervisor **Bruno**. Thanks so much for these three and a half years and for everything you have taught me. For all discussions, all the agreeing and disagreeing. Not many supervisors would have granted their Ph.D. students the amount of freedom you did. During these years, I really appreciate the positive attitude in front of all kind of dark moments (quite a number :P). Thank you for your support!

I would like to extend my gratitude to the jury members that have accepted to read and to evaluate this manuscript: **Elisabeth Darrouzet, Jean-Michel Jault, Eva Pebay-Peyroula, Dirk Schüler, Elisabetta Bergantino.**

Thanks to our collaborators in Spain, **Ignacio Arechaga** and **Gerardo Carranza** for working so hard to publish the paper.

All the present and past members of the entire unit UMR7099. It is hard to find the right words for how wonderful it was to get to know you. **Edith**, you deserve a medal for putting up with all the small and big problems in the unit. Thank you! **Marc**, you truly are an encyclopedia in microbiology, thank you to help me a lot with RNA and transduction, for your patience and suggestions. I will miss our conversations. **Hager**, I'm going to miss you so much, the best office mate ever and my "lab mother", thank you to be my reference point for lab things as well as for life problems. **Francesca**, thank you for the suggestions and your support during all these years, you were my little Italy in the lab together with **Silvia**, **Marina**, **Chiara** and **Ottavia**. **Oana**, thank you for your technical support during a big part of the thesis. **Martin**, thank you for your advice ON the manuscript (:P).
All the wonderful and fun "young team": **Amira**, **Lindsay**, **Quentin**, **Mathilde**, **Dhenesh**, **Yvette**, **Nathalie**, **Anaïs**, **Jordi** (Do you see? You are still young :P) and **Hager** (again :D). I have no words to thank you, you made my last year full of laughs and talks. Thank you all for making our unit a nice environment to work in!
I want also to thank all the other unit members that somehow crossed my years here.

I also got to know many great people outside of work, who made it very easy for me to feel at home in Paris. **Giulia**, **Ferran**, **Daniel** and **Pedro** (you're becoming a biologist too :D): thank you

to share with me this unique experience, for being there when no one understood French, when experiments did not work and for our talks in Spanish/Italian/English/French. I really hope that we'll stay in touch. I want to come to all your weddings :D

My Sardinian family in Paris: **Francesca**, **Mari**, **Charlotte** (the most Sardinian of us) and **Fabio**. I will miss all our dinners, Fabio and Charlotte's adventures, Sardinian cooking classes. I regret to have met you, guys, only so later, but I'm looking forward to see you all the next summers in Sardinia or wherever life gets you.

My present roommates: **Francesca**, thank you for putting up with me all these years, can't have been easy with all my mood swings -.-'; **Valentina**, thank you to understand me with only a look, I would like to stay in Paris only to support you across your Ph.D.

My run partner **Laura** for sharing with me joys and sorrows of runs.

My Sardinian friends: **Giovanna, Barbara, Erica** to make my holidays never boring :D.

Last but of cours not least **Viola**, because you were always there, even if I am still mad at you about your decision to leave Paris, you did everything to shorten distances. I really love you, friend!

Mattia…if it's true that little things are the most important, thank you for filling my life with a lot of little things. You are my strength.

I would like to thank **my extended family** as well as **Mattia's family** (including dogs, cat and chickens) to be my fan club.
A special thank you to my sponsors, **my parents**, this Ph.D. is yours too. Thank you for always encouraging me to go away and see the world outside even if you would want to be there with you. I love you!

Ringraziamenti

La prima e la più importante persona che vorrei ringraziare è il mio direttore di tesi, **Bruno**. Grazie per questi tre anni e mezzo e per tutto quello che mi hai insegnato. Per tutte le discussioni, in cui siamo stati d'accordo e anche quelle in cui non lo siamo stati. Non tutti i direttori di tesi avrebbero concesso ai loro dottorandi di essere così indipendenti e liberi, come tu hai fatto. Durante questi anni, ho veramente apprezzato la tua positività davanti ai momenti difficili (e non sono stati pochi :P). Grazie per avermi sempre supportato!

Vorrei estendere la mia gratitudine a tutti i membri della giuria che hanno accettato di leggere e valutare la tesi: **Elisabeth Darrouzet, Jean-Michel Jault, Eva Pebay-Peyroula, Dirk Schüler, Elisabetta Bergantino.**

Grazie ai nostri collaborator in Spagna, **Ignacio Arechaga** e **Gerardo Carranza,** per aver lavorato duramente nella pubblicazione dell'articolo.

Tutti i membri dell'unità UMR7099, anche quelli che stanno continuando la carriera altrove. E' difficile trovare le giuste parole per descrivere quanto sia stato bello conoscervi. **Edith**, dovresti avere una medaglia per tutti i grandi e piccoli problemi che riesci a risolvere all'interno dell'unità. Grazie per fare in modo che tutto vada avanti nel retroscena. **Marc**, tu sei un'enciclopedia vivente di microbiologia, grazie per avermi aiutato enormemente con gli RNA e le trasduzioni, per la pazienza che hai messo nello spiegarmi ogni dettaglio e per i tuoi suggerimenti. Le nostre chiacchierate mi mancheranno. **Hager**, mi mancherai tantissimo, la migliore collega d'ufficio che avrei potuto avere e la mia "mamma di laboratorio", grazie per essere stata il mio punto di riferimento per quanto riguarda sia il laboratorio sia la vita privata. **Francesca**, grazie per i suggerimenti e il tuo supporto durante questi anni, hai rappresentato la mia piccola Italia all'interno del laboratorio insieme a **Silvia**, **Marina**, **Chiara** e **Ottavia**. **Oana**, grazie per il tuo supporto tecnico durante una buona parte della tesi. **Martin**, grazie per aver corretto gli innumerevoli errori grammaticali della tesi (:P).
A tutto il bellissimo e divertente "team dei giovani": **Amira**, **Lindsay**, **Quentin**, **Mathilde**, **Dhenesh**, **Yvette**, **Nathalie**, **Anaïs**, **Jordi** (Hai visto?! Sei ancora giovane :P) e **Hager** (di nuovo :D). Non ho parole per ringraziarvi, avete reso il mio ultimo anno pieno di chiacchiere e risate. Grazie per aver reso il laboratorio un bell'ambiente dove lavorare.

Ugualmente, vorrei ringraziare tutte le persone dell'unità che ho incrociato durante questi anni.

Nel mio soggiorno a Parigi ho conosciuto anche tante persone al di fuori del lavoro grazie a cui mi sono potuta sentire a casa. **Giulia**, **Ferran**, **Daniel** e **Pedro** (ormai sei diventato un biologo anche tu :P) grazie per aver condiviso con me quest'esperienza unica, per essere stati lì quando nessuno di noi capiva una parola di francese, quando gli esperimenti non funzionavano e per le nostre chiacchierate in Spagnolo/Italiano/Francese/Inglese. Spero proprio che ci terremo in contatto. Voglio venire a tutti i vostri matrimoni :D.

La mia famiglia sarda a Parigi: **Francesca**, **Mari**, **Charlotte** (la più sarda di tutti) e **Fabio**. Mi mancheranno tutte le nostre cene, le avventure di Fabio e Charlotte e i corsi di cucina sarda. Mi dispiace di avervi conosciuto così tardi, ma non vedo l'ora di vedervi tutte le estati, in Sardegna o in ogni posto ovunque voi siate.

Le mie attuli coinquiline: **Francesca**, grazie per avermi sopportato tutti questi anni, non dev'essere stato facile visti i miei sbalzi d'umore -.-'; **Valentina**, grazie per capirmi solo con uno sguardo, vorrei stare a Parigi solo per supportati durante i tuoi anni di dottorato.

La mia compagna di corse **Laura** per aver condiviso con me gioie e dolori della corsa.

Le mie amiche sarde di sempre: **Giovanna, Barbara, Erica** per aver fatto sì che le mie estati non fossero mai noiose :D.

Ultima ma non per importanza **Viola**, perchè sei sempre stata lì e anche se l'ho ancora con te per aver lasciato Parigi, hai fatto di tutto per accorciare le distanze. Ti voglio bene, amica!

Mattia…se è vero che le piccole cose fanno la differenza, grazie per aver riempito la mia vita di piccole cose. Sei la mia forza.

Infine, vorrei ringraziare tutta la **mia famiglia estesa** e la **famiglia di Mattia** (inclusi cani, gatto e galline) per essere stati il mio fan club.
Un grazie speciale ai miei sponsors, i miei genitori, questo dottorato è anche vostro. Grazie per avermi sempre incoraggiato a uscire e vedere il mondo anche se mi avreste voluto lì con voi. Vi amo!

Preface

This thesis represents a research project from a three-year long Ph.D. carried out under the supervision of Bruno Miroux at the Institute of Physical-Chemical Biology (IBPC) (UMR7099) in Paris. The Labex Dynamo founded my PhD fellowship.

The manuscript consists of five main chapters: introduction, materials & methods, results, discussion and conclusions & perspectives. The sections of material and methods, results and discussions have been further divided in two sub-chapters each, in order to clarify the reading. The sub-chapters represent two different parts I worked on during the thesis.

The first part concerns the isolation of new mutant strains to further ameliorate membrane proteins production in *E. coli*. A manuscript is ready to be submitted.

In the second part we have investigated, through biochemistry and "omics" experiments, the phenomenon of intracellular membranes formation in *E. coli*. A BBA – Biomembranes article is published showing the importance of cardiolipids in intracellular membranes formation. The paper is the results of the collaboration with Ignacio Arechaga and Gerardo Carranza (IBBTEC, Santander, Spain).

During the Ph.D., I was an author of a chapter published in the second edition of the book "Heterologous Expression of Membrane Proteins" edited by Isabelle Mus-Veteau in Methods and Molecular Biology.

In the annexes section are available all mentioned documents.

Table of Contents

List of Figures

List of tables

Abbreviations

ACN, ACetoNitrile

BAE, Bacteria Adaptive Response

BCA, Branched-Chain Acid

BFP, Blue Fluorescent Protein

cDNA, complementary DNA

CFA, Cyclopropanated Fatty Acid

CL, CardioLipid

CoQ_1, Coenzyme Q_1

DDM, n-Dodecyl-D-Maltoside

DTT, DiThioThreitol

E. coli, Escherichia coli

ErmC, Erythromycin resistance protein

FACS,Fluorescence-Activated Cell Sorting

FC, Fold Change

FSC, Forward Scatter Channel

gDNA, genomic DNA

GFP, Green Fluorescent Protein

GPCR, G protein-coupled receptor

sfGFP, superfolder Green Fluorescent Protein

h-SQR, human sulphide-quinone oxydo-reductase

ICM, IntraCellular Membranes

IGV, Integrative Genomics Viewer

IM, inner membrane

IPTG,Isopropyl-β-D-thiogalactopyranoside

kDa, kiloDalton

LB, Luria-Bertani

LC-MS, Liquid Chromatography- Mass Spectrometry

LIC, Ligation-Independent Cloning

LPE, lyso-phosphatidyl-ethanolamine

MP, Membrane Protein

mRNA, messenger RNA

MS, Mass Spectrometry

NAO, 10-N-Nonyl Acridine Orange

ncRNA, non-coding RNA

OD, Optical Density

OM, Outer Membrane

OMP, Outer Membrane Protein

ORF, Open Reading Frame

PBS, Phosphate Buffered Saline

PCR, Polymerase Chain Reaction

PDB, Protein Data Bank

PG, PhosphatidylGlycerol

PMF, Proton-Motive Force

PMSF, PhenylMethylSulfonyl Fluoride

PSM, Peptide Spectrum Matches

PTS,phosphoenolpyruvate: phosphotransferase system

RBS, Ribosomal Binding Site

RCS, Regulator of Colanic acid Synthesis

RCSA, Relative Colony Surface Area

RFI, Relative Fluorescence Intensity

RT-qPCR, Real time quantitative PCR

SA, Surface Area

SCE, Stored Curvature Elastic

SEM, Standard Error of the Mean

SFA, Saturated Fatty Acid

SHX, Serine HydroXymate

sRNA, small RNA

SSC, Side Scatter Channel

T7 RNAP, T7 RNA Polymerase

TCS, two-component system

TEM, Transmission Electron Microscopy

TEV, Tobacco Etch Virus

TFA, TriFluoroacetic Acid

TM, TransMembrane

UFA, Unsaturated Fatty Acid

Introduction

Cellular proteins of different shapes and sizes are involved in many important cellular functions. To study their structure and function, they have to be produced, most of them, in heterologous expression systems. The protein production host can be either prokaryotic, like *Escherichia coli* (*E. coli*), or eukaryotic, like yeast, insect and mammalian cells. The Gram-negative bacterium *E. coli* is the most common prokaryotic model organism because of both its relative simplicity and the low cost at which it can be propagated and studied in the laboratory. The genome of *E. coli*, for example, consists of approximately 4.6 million base pairs and encodes about 4,000 different proteins. The small size of the *E. coli* genome provides obvious advantages for genetic analysis especially because the sequence of its entire genome can be determined for less than hundred euros.

Gram-negative bacteria, as *E. coli*, are composed by two kinds of proteins: membrane proteins (MPs) and soluble proteins. Soluble proteins are expressed in the cytoplasm and in the periplasm, while MPs are inserted into biological membranes. The latter account for more than 50% of the total cell membranes (Engelman, 2005); it is therefore not surprising that they represent the majority of drug targets (Terstappen and Reggiani, 2001; Overington et al., 2005) and are involved in important processes such as nutrient import, export of toxic compounds and other types of communication with the surrounding environment (Almén et al., 2009). Despite this, there is an extremely low number of unique MP structures available in the protein data bank (PDB) (Hattab et al., 2015). For this reason, there is the need to develop improved strategies for the production of MPs.

Complex soluble proteins, especially eukaryotic ones, can also be difficult to produce, mainly because the rate of protein translation is higher than the rate of protein folding when using microorganisms (Esposito and Chatterjee, 2006). Similar obstacles often are common for both soluble and membrane proteins and therefore benefits from the same optimization strategies.

Furthermore, contrary to what is believed, proteins are not expressed only in the cytoplasmic membranes or accumulate in inclusion bodies. A dozen of MPs have been reported to trigger massive membranes proliferation in the host cell when produced at very high level. The biogenesis of such natural storage compartments is

poorly understood and is one of the major questions I have asked during my Ph.D. project.

In the follow introduction sections, I will give a general overview about biological membranes in Gram-negative bacteria, the MPs biogenesis, the cell mechanisms to protect proteins homeostasis in bacteria and the strategies to optimize recombinant protein production. Subsequently, I will describe the lipid-protein interactions, to focus on membranes curvature and intracellular membranes (ICMs) formation in *E. coli*.

1.1 Biological membranes in Gram-negative bacteria

During the three years of my Ph.D., I have used the Gram-negative bacterium *E. coli*, in particular BL21(DE3) and BL21(DE3) derivative strains. For this reason I will focus on Gram-negative membranes composition.

A general description of a biological membrane is given as "a defined boundary between the cellular components and extracellular medium, which protects the cell from dangerous fluctuations in order to keep homeostasis in the cell" (Bowen, 2006; Keren, 2011). Membranes are involved in both structural and biological functions. They not only define a simple boundary but also create a domain in which certain cellular activities can be carried out and where they fulfil biological functions, such as the exchange of solutes with the exterior and generation of energy (Arechaga, 2013).

The cytoplasmic membrane, also known as inner membrane (IM), separates the cytoplasm, which contains soluble proteins and the chromosome, from the periplasm that houses the peptidoglycan cell wall (Grabowicz and Silhavy, 2017). The outer membrane (OM) isolates the inner of the cell from the outside environment. Both membranes consist of lipids and proteins. A hydrophilic head-group and a hydrophobic tail compose membrane lipids. Due to the amphipathic nature of phospholipids, membranes are structured as a lipid bilayer, with the hydrophobic tails facing inwards and the hydrophilic head-groups facing outwards (Tanford, 1978). The lipids are mainly represented by glycerol-phospholipids (phosphatidyl-ethanolamine (PE), phosphatidyl-glycerol (PG) and cardiolipin (CL)). In the outer leaflet of the OM is present another kind of lipid, the glycolipid (lipopolysaccharide (LPS)) (Silhavy et al., 2010). The main function of LPS is to serve as a permeability barrier towards external dangers such as antibiotics.

The different lipids composition is the cause of the asymmetry of the OM compared to the general symmetric nature of the IM.

The MPs of the IM are either integral or peripheral MPs (Papanastasiou et al., 2013). The integral MPs are either embedded in the membranes or covalently linked through a lipid anchor. All embedded integral MPs consist of one or multiple transmembrane domains (TMs), which have mainly a α-helical conformation and are connected to each other by hydrophilic loop-regions extending out from the membrane either on the periplasmic or the cytoplasmic side (Fairman et al., 2011). The peripheral monotopic MPs maintain contacts with one of the leaflets of the membrane by exploiting non-covalent bonds to the lipids and/or to integral MPs.

In contrast, the proteins of the OM consist of β-barrel proteins that function as porins, through which ions and polar compounds can diffuse, making the OM more permeable than the IM (Hancock, 1987).

Since the initial proposal of a fluid-mosaic model by Singer and Nicholson in 1972, the view of biological membranes has changed. Even if this model is still relevant, nowadays is evident that the lipid bilayer plays an important role in the regulation of MP function. Numerous studies on the relationship between MPs and their lipid environment have been published over the last years (Nicolson, 2014).

2 Folding, targeting, insertion and degradation of proteins in *E. coli*

In bacteria, all proteins are synthetized in the cytoplasm. Once the polypeptide chain of a protein has been synthesized, it is addressed to a specific biogenesis pathway depending on the final destination of the protein. All the proteins, with the exception of those designed to remain in the cytoplasm, are equipped with specific target sequences, such as labels that direct them to the correct location. These proteins have to go through membrane barriers either to be secreted, to stay in the periplasm or to fold in one of the membrane bilayers of the cell (Fekkes and Driessen, 1999).

The RNA polymerase (RNAP) transcribes the genes encoding proteins. It binds in a region located upstream of the gene on the DNA, called promoter-region, as a result of its binding to the σ-factor (Travers and Burgess, 1969). σ factors are important

regulators of gene expression as they determine to which promoters RNAPs are recruited (Chandrangsu and Helmann, 2014). More than one RNAP molecule can be present at different positions on the same gene simultaneously, generating more than one mRNA transcript at a time (Epshtein and Nudler, 2003). The strength of the promoter and the activity of the RNAP will affect the number of mRNA transcripts that are produced, which subsequently change the kinetics of protein synthesis, a fundamental concept for recombinant protein production (Goldbeck et al., 2013).

The mRNA transcript is translated in amino acids by the ribosome. In particular, the bacterial ribosome (70S) consists of the large (50S) and the small ribosomal (30S) subunits (Steitz, 2008). The insertion of the correct amino acid at the correct position in the polypeptide chain is dictated by another kind of RNA, the transfer RNA (tRNA). tRNAs recognise a certain codon on the mRNA thanks to an anti-codon sequence, which allow to insert a specific amino acid for that codon. The translation ends when the tRNA recognizes one of the stop codons, which have been classified as *amber* (UAG), *ochre* (UAA) and *opal* (UGA) codons. In some cases, a stop codon may be recognised as a coding codon and be replaced from an amino acid instead of determining the end of translation. This phenomenon is called nonsense suppression or ribosome read-through (Figure 1). It can be caused by either mutations in the translation machinery or by specific signals on the messenger RNA (Swart et al., 2016).

The efficiency of ribosome binding to the mRNA transcript largely depends on the ribosomal binding site (RBS) or Shine-Dalgarno sequence of the mRNA (Vellanoweth and Rabinowitz, 1992). The distance between the RBS and the start codon affects the efficiency of translation initiation. It has been shown that a distance of approximately eight nucleotides leads to the highest efficiency of translation initiation (Ringquist et al., 1992). Not surprisingly, the ribosome density, which depends on the sequence at and around the RBS, on an mRNA transcript influences protein accumulation kinetics (Steitz, 2008).

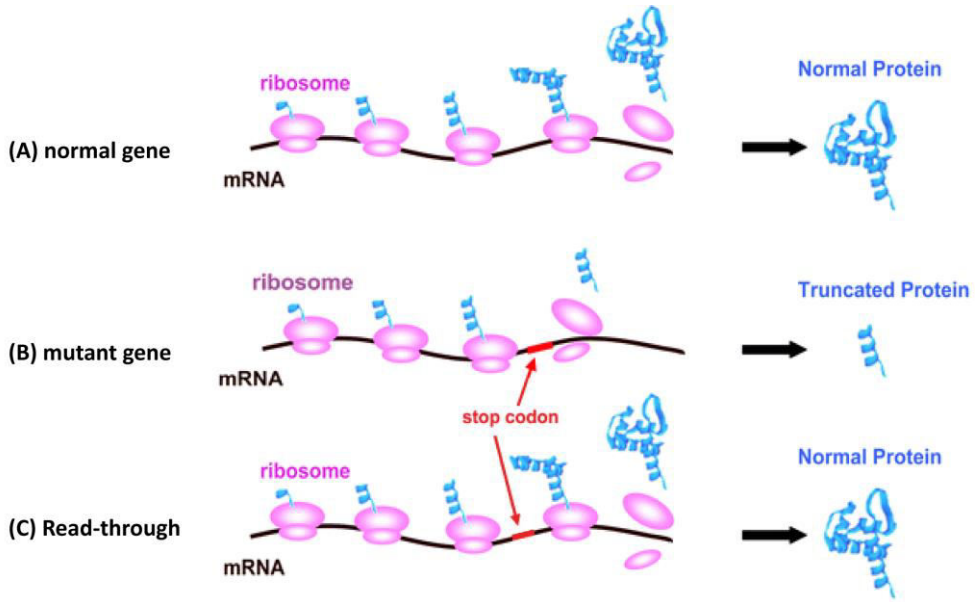


Figure 1. Ribosomal read-through mechanism. (A) Ribosome translation of a normal gene gives a normal protein. **(B)** Ribosome translation of a mutant gene gives a truncated protein. **(C)** Ribosome read-through of a mutant gene gives a normal protein. Figure adapted from (Hainrichson et al., 2008).

2.1 Proteins folding

The MP folding is a critical stage because of the high probability of proteins misfolding, indeed MPs are more sensitive to aggregation in the cytoplasm than soluble proteins (Jahn and Radford, 2005). To avoid the fold of MPs in non-functional state, the bacteria possesses a network of chaperones and targeting factors that assist the protein as it emerges from the ribosome (Diaz-Villanueva et al., 2015). The majority of known chaperones are proteins, but it has been shown that lipids and RNAs can help the folding of some proteins (Kim et al., 2010). In *E. coli* the best-known chaperone systems are the Trigger Factor (TF), DnaKJ/GrpE and GroEL/ES.

TF is an ATP-independent chaperone that is highly abundant in the *E. coli* cytoplasm (Merz et al., 2006). Ribosome-bound TF likely interacts with different kind of polypeptides but it has a preference for hydrophobic polypeptide. Although most TF substrates are cytoplasmic proteins, it plays an important role in directing secretory proteins through the post-translational translocation pathway (Lin et al., 2012) (see pathway III and IV in Figure 2). In many cases, the release of TF is followed by further co- and post-translational folding assistance by the chaperones DnaKJ/GrpE and GroEL/ES not associated to ribosome.

The ATP-dependent chaperone DnaK has a central role in the cytoplasmic chaperone network (Calloni et al., 2012). It receives *de novo* substrates from TF and gives many of them to the chaperone GroEL/ES (see pathway I in Figure 2). It is worth noting that although the majority of substrates are cytoplasmic proteins, DnaK appears also to have an important role in the targeting and biogenesis of some MP (Castanié-Cornet et al., 2014). DnaJ and GrpE support DnaK, regulating the substrate binding cycle and ensuring proper ATP hydrolysis, which control the opening and closing of the substrate-binding domain of DnaK (Kim et al., 2013). DnaK has a key role under stress conditions induced by protein aggregation (Deuerling et al., 1999; Teter et al., 1999).

GroEL function is dependent on ATP hydrolysis and on its co-chaperone GroES (Xu et al., 1997). GroEL own a hydrophobic pocket allowing the substrate to bind. Subsequently, ATP binding causes a large conformational change, making the interior hydrophilic (Skjærven et al., 2015). It has been shown that GroEL also actively catalyses the folding of the substrate (Georgescauld et al., 2014).

It has been shown that the activity of GroEL/ES retains only about 30% of refolding activity at 12°C, compared to its activity at 30°C (Schein, 2004). To overcome this obstacle, cells that co-express the cold-adapted chaperonins (ArcticExpress) Cpn10 and Cpn60 from the psychrophilic bacterium *Oleispira antarctica* have been engineered. They have 74% and 54% amino acid identity to the *E. coli* GroEL and GroES chaperonins, respectively, and show high protein refolding activities at temperatures of 4-12°C (Schein, 2004)

2.2 Protein targeting to the membrane

Once the protein is synthesised it is integrated to the IM through two main insertion/translocation machineries: the Sec-translocon (Figure 2) and the Twin-Arginine Translocation (TAT)-translocon.

The TAT-translocon machinery targets the proteins in a post-translational way, indeed, in contrast to the Sec-dependent pathway, proteins have to be folded prior to translocation *via* the TAT-translocase (Kudva et al., 2013). Most TAT-dependent proteins identified so far contain co-factors that need to be acquired during the folding process in the cytoplasm (Fröbel et al., 2012).

The targeting across the Sec-translocon may take place while translation is still on-going or after the complete polypeptide has been synthesized, defining two different pathways: the co-translational targeting pathway (see pathway II in Figure 2) and the post-translational pathway (see pathway III and IV in Figure 2), respectively. In order to choose between targeting pathways, proteins are equipped with a signal sequence, cleavable for the secretory proteins and non-cleavable for integral MPs. In the latter case, the signal sequence, also called "signal anchor sequence", will constitute the first TM (Lee and Bernstein, 2001; Luirink et al., 2012; Kudva et al., 2013). Many factors, however, can regulate the targeting pathway, for example the translation rate: a fast translation rate increases the risk of protein misfolding and the need of chaperone-assistance in the early stage of folding (Huber et al., 2005; Wagner et al., 2008; Castanié-Cornet et al., 2014).

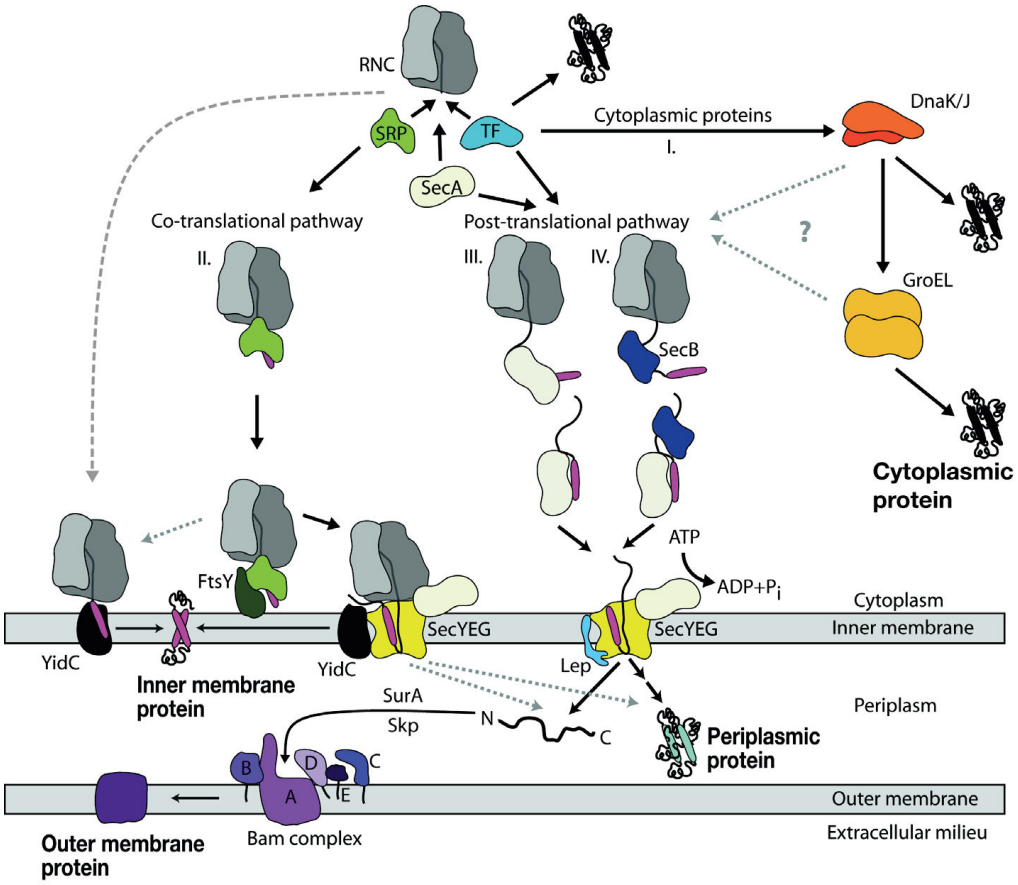


Figure 2. Biogenesis pathways of cytoplasmic, inner membrane and secretory proteins. (I) TF interacts strongly with ribosomes synthesizing cytoplasmic proteins. The protein thus generated, moves to DnaK/J and GroEL/ES, which constitute the cytoplasmic chaperone network. It may also play a role in the biogenesis of some membrane proteins. **(II)** Highly hydrophobic proteins are inserted through SRP binding in the IM by using the co-translational pathway. **(III & IV)** Soluble SecA targets the polypeptide chain to the Sec-translocon, either in the SecB independent **(III)** or SecB dependent **(IV)** way. Once the protein reaches the IM, Lep cuts off the signal sequence from the secretory protein, the other polypeptides are either folded into its native conformation or it can be targeted to the OM by SurA/Skp. The SecYEG complex plays a role in Sec-mediated IM insertion and translocation. YidC can insert inner MPs independently on the Sec-translocon. (Luirink et al., 2012).

2.3 Proteins insertion into the IM

In most of the cases, MPs insertion in the IM is achieved in a co-translational manner. Both insertion and translocation are mediated by the SecYEG-translocon, with or without assistance of SecA, YidC and other auxiliary components.

2.3.1 The Sec-translocon

The Sec-translocon is a hetero-trimer complex (Ohta et al., 2016) which exhibits the distinct ability to translocate substrates across a membrane or to insert them into the membrane. The bacterial translocon is composed of three integral MPs: SecY, the largest component of the translocon, essential for viability and translocation; SecE, a small integral MP that stabilizes SecY, is required for a full translocon efficiency; SecG, a 12 kDa protein that facilitates the binding and insertion of SecA into the translocon (du Plessis et al., 2011).

In the case of MP insertion into the IM, interactions between the highly hydrophobic signal anchor sequence and the Sec-translocon promote the opening of the lateral gate (du Plessis et al., 2011). The force of polypeptide elongation, at the expense of GTP hydrolysis, partially drives the co-translational insertion of the α-helices into the lipid bilayer (Natale et al., 2008). In addition, TMs interactions with lipids and interactions between other TMs also promote the insertion. However, efficient translocation of large proteins requires SecA (Collinson, 2017).

Others proteins supporting the SecYEG complex are YidC and SecDFYajC.

YidC is a highly conserved integral MP chaperone consisting of 6 TMs (Samuelson et al., 2000). Generally, YidC receives TMs from SecYEG and assists their insertion into the IM (Sachelaru et al., 2013; Scotti et al., 2000). Interestingly, YidC can also itself promote membrane protein insertion, constituting a second type of proteins insertion (Facey et al., 2007; Dalbey et al., 2014) even if the mechanism by which it regulates the insertion is still elusive. This protein is widely used, including by us in this study, as a model protein to assess MP production in different bacterial hosts (Schlegel et al., 2012; Zhang et al., 2015; Baumgarten et al., 2017).

The SecDFYajC complex works in association with the SecYEG complex in different structural arrangements: either with a SecA-bound SecYEG, for post-translational translocation or with a ribosome-bound SecYEG for co-translational

transport. The presence of this complex means that the transport of proteins is more dependent on the proton-motive force (PMF). SecDFYajC has been suggested to regulate the insertion/translocation process according to the properties of the substrate protein and to modulate protein composition of the membrane during different growth phases or stress conditions (Schulze et al., 2014).

2.3.2 Insertion is dependent on MPs topology

Integral MPs can assume different topologies depending on the C-terminal and N-terminal position, on the cytoplasmic side or on the periplasmic side of the IM. Furthermore, the TM-loop can be on the same IM side tilted compared to the lipid bilayer (Viklund et al., 2006). Its topology can be influenced by numerous factors: hydrophobicity and length of the TM, distribution of positively charged residues of the cytoplasmic loops connecting the TMs (Heijne, 1986) and the lipid composition of the membrane. In addition, the topology can be decided during insertion or flipped after insertion into the IM (Wahlberg and Spiess, 1997; Cymer et al., 2015).

According to Cymer and von Heijne (Cymer and von Heijne, 2013), tertiary folding of membrane proteins is initiated already at the immediate proximity of the Sec-translocon, in a co-translational way. As previously mentioned, YidC is believed to receive TMs as they emerge from the lateral gate of the translocon and help their insertion and folding. To support this hypothesis, it has been shown that YidC deletion leads to cell envelope stress response due to membrane protein misfolding, further pointing out a significant role of YidC as a molecular chaperone (Nagamori et al., 2004; Ruiz and Silhavy, 2005).

2.4 Proteins degradation

The proteins level in a cell is determined not only by the rates of synthesis, but also by rates of degradation. The half-lives of proteins vary widely, from minutes to several days, and differential rates of proteins degradation play an important role in maintaining the protein homeostasis within cells. The rapid turnover of some proteins, such as transcription factors, allows to quickly changing their level in response to

external stimuli. Furthermore, the apparatus of proteases and peptidases are involved in degradation of proteins incorrectly synthesized or aggregated into the cytoplasm.

There are five major ATP-dependent proteolytic systems in the *E. coli* cytoplasm: ClpXP, ClpAP, HslUV, Lon and FtsH. All of them containe a peptidase domain and an ATPase domain. The peptidase domain is located in the inner cavity of a β-barrel structure, which can be accessed through a narrow pore, while the ATPase domain is located at the top of the β-barrel and unfolds protein substrates prior to their entry into the barrel. Once the protein is inside the barrel, it is exposed to the catalytic site and fragmented into smaller peptides, subsequently released for further degradation by peptidases (Bochtler et al., 2000; Sousa et al., 2000; Park et al., 2006; Suno et al., 2006; Wang et al., 2009; Vieux et al., 2013).

The membrane protein protease FtsH is the only one to be essential for *E. coli* viability (Katz and Ron, 2008). This protein has the ability to extract integral MPs from the IM to degrade them. It has been suggested that FtsH works in close collaboration with YidC and the inhibitory proteins HflK and HflC forming a quality control network (Kihara et al., 1996). Interestingly, it was shown that GFP equipped with a signal, promoting its degradation by FtsH, or a signal recognized by both FtsH and ClpXP, cannot be degraded by FtsH as long as the GFP moiety is folded. In contrast, ClpXP readily degraded folded GFP equipped with the respective degradation signal. This suggests that FtsH has only a rather weak unfolding activity and will not be able to act on targets that contain a stably folded domain (Herman et al., 2003).

The protease Lon of *E. coli* is involved in the acquisition of tolerance to UV, tolerance to high pressure, regulation of cell division and regulation of MP synthesis quality control (Sugiyama et al., 2013; Kwon et al., 2015). Its task is to clear the cell from misfolded protein or partially translated that hamper the ribosome during co-translational membrane insertion process. The strain BL21(DE3) is Lon-deficient due to IS4 insertion in the promoter region of *lon* (saiSree et al., 2001). Its phenotype is restored in C43(DE3), as I will explain in following sections.

Generally speaking, it has been shown that re-folding a protein may be much more favourable than degrading and re-synthesizing it (Sharma et al., 2010). Perhaps this is the rationale behind the development of bacterial stress responses.

3 Bacterial stress responses

In order to react efficiently to different kinds of damage conditions, *E. coli* has developed various types of stress-responses: heat-shock response, acid stress response, starvation response, envelope stress response and many others. In general, a common regulatory mechanism is based on σ-factors. As mentioned before, these small proteins direct the RNAP to promoters, thus controlling the expression of certain sets of genes. However, many proteins that are part of a stress response are also produced under non-stressed conditions but at a lower level (Chung et al., 2006). Under stress conditions, alternative σ-factors are induced by σ^{70}; in particular σ^{S} (RpoS) is the master regulator of the general stress response. Under non-stress conditions, RpoS is degraded by the ClpXP protease, while it is synthetized at the beginning of the stationary phase and also in responses to stresses such as weak acids, starvation, high osmolarity and high or low temperature (Lange and Hengge-Aronis, 1994).

During my Ph.D., the most investigated stress responses were those involved in the heat-shock response and the envelope stress responses. The latter were studied in the context of intracellular membranes formation.

3.1 The heat-shock response

The heat-shock response in bacteria is a protective mechanism regulated by σ^{32} and induced by elevated temperatures or misfolded proteins (Nonaka et al., 2006; Zhao et al., 2007). Some of the proteins that are regulated during the heat-shock response are: DnaKJ, GroEL, FtsH, Lon and HtpX, proteases and chaperones that have already been described. Additional proteins are the protease ClpB and the inclusion body binding proteins IbpA and IbpB. These proteins are believed to assist DnaK in resolving protein aggregates and assisting re-folding of misfolded proteins (Mogk et al., 2003). Furthermore, the small heat-shock proteins IbpA and IbpB are shown to be over-expressed during proteomic experiments in presence of inclusion bodies (Wagner et al., 2007). Interestingly, approximately 25% of the genes of the σ^{32} regulon encode MPs (Nonaka et al., 2006).

The levels of σ^{32} are strictly regulated to avoid a stress response under normal conditions. The gene encoding σ^{32} factor, *rpoH*, is constitutively transcribed and its

regulation appears to occur primarily after transcription. It has been shown that the σ^{32} mRNA factor form secondary structures, which inhibit association with the ribosome. Upon elevated temperatures, these secondary structures are resolved, which allow the synthesis of σ^{32} (Morita et al., 1999).

At a post-translational level, DnaKJ/GrpE and GroEL/ES have been shown to regulate σ^{32} activity by degrading the factor and reducing its level (Zhao et al., 2005).

3.2 The envelope stress response networks

Under most conditions, a functional cell envelope is essential for Gram-negative bacteria, especially because of antimicrobials that target its components. Polymyxin B is an effective antimicrobial peptide that specifically targets the outer membrane of bacteria (Bader and Teuber, 1973), lysozyme and beta-lactams such as penicillin target the peptidoglycan wall (Lehrer et al., 1989), defensins create pores in the bacterial membrane (Ganz and Lehrer, 1998) and the accumulation of aggregated material in the periplasm can be lethal to the bacteria (Snyder and Silhavy, 1995).

To efficiently protect the envelope integrity, *E. coli* possesses at least five different mechanisms: σ^E, Cpx, BAE, RCS, PSP (Ruiz and Silhavy, 2005; Flores-Kim and Darwin, 2016). Acting together, the network of stress response pathways responds to different types of envelope damage and is capable of maintaining intact the entire envelope.

3.2.1 The σ^E pathway

The σ^E pathway is activated as a result of outer membrane protein (OMP) assembly defects. The alternate σ factor (σ^E, RpoE), enables RNAP to increase the transcription of genes involved in cell envelope biogenesis (Rhodius et al., 2006).

Under non-stress conditions, the IM anti-σ-factor RseA holds it to the plasma membrane (Ades et al., 1999). The presence of misfolded OMP in the *E. coli* periplasm triggers a cascade of proteolytic events (Grabowicz and Silhavy, 2017) (**Figure 3**, A). First, the IM serine protease DegS removes a periplasmic domain of RseA. RseA is then subjected to a second cleavage by the zinc metalloprotease, RseP. At this stage, RseA is still being held from the cytoplasmic domain. ClpXP proteases degrade this last

fragment releasing the σ^E factor (**Figure 3**, A). The cleavage site on RseA is normally occluded by the periplasmic protein, RseB (Barchinger and Ades, 2013). It has been shown that in addition to unfolded OMP (to activate DegS), the activation of the σ^E factor requires sensing folding defects in periplasmic LPS (to displace RseB from RseA) (Lima et al., 2013).

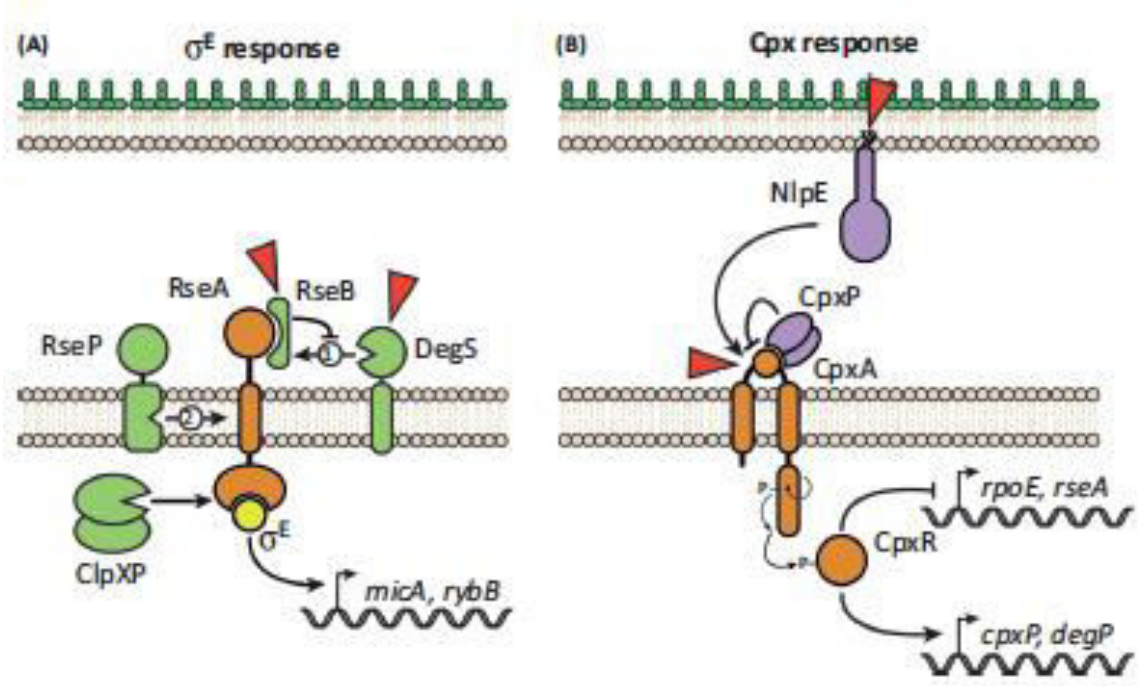


Figure 3. The σ^E and Cpx envelope stress responses in *E. coli*. (A) The **σ^E** response is activated by proteolysis steps that release **σ^E** into the cytosol, where it can activate the expression of target genes. RseB is displaced from RseA in response to LPS binding, allowing DegS to cleave the RseA periplasmic domain and initiate the response. Subsequently, RseP further cleaves RseA. The last cytoplasmic portion of RseA is then degraded by the ClpXP protease to release σ^E. **(B)** The Cpx TCS responds to numerous signals detected by the periplasmic sensory domain of CpxA. CpxP inhibits CpxA activation, likely by interacting with the sensory domain directly. NlpE can transduce cell adhesion signals across the envelope to activate the Cpx response. Stimulated CpxA histidine kinase autophosphorylates and transfers the phosphate to the CpxR response regulator, which alters gene expression (Grabowicz and Silhavy, 2017).

3.2.2 The Cpx pathway

Unlike the σ^E stress response, which specifically responds to misfolded OMP, the Cpx pathway is activated by a broad set of conditions: elevated pH, high salt concentrations and alterations in IM lipid composition (Raivio, 2014). In addition, recent works suggest a role of Cpx in sensing phosphatidylglycerol (PG) wall defect (Delhaye et al., 2016). Sensing is nonetheless specific and misfolded OMP fail to activate Cpx. In contrast, Cpx senses all misfolded IM, periplasmic proteins and defects in protein translocation across the IM. Thus, it appears that the primary task of Cpx is defending IM integrity (Raivio, 2014).

Cpx is a two-component system (TCS) composed by the sensor histidine kinase, CpxA, which is inhibited by the periplasmic protein CpxP (Figure 3, B). When misfolded proteins accumulate in the periplasm, CpxP binds to them and can no longer inhibit CpxA. CpxA activation leads to autophosphorylation of its cytoplasmic domain and can transfer the phosphate to the response regulator CpxR (Raivio and Silhavy, 1997).

NlpE is an auxiliary protein that modulates positively CpxA. It is an OM lipoprotein that activates Cpx upon *E. coli* adhesion to abiotic surfaces (Otto and Silhavy, 2002). It has been proposed that unfolded NlpE directly contact CpxA from the OM to induce signalling (Hirano et al., 2007).

Cpx enhances the transcription of numerous genes, among them is *cpxP*, which promotes a negative-feedback loop by inhibiting CpxA kinase activation (Price and Raivio, 2009). Furthermore, Cpx can up-regulate the periplasmic protease DegP, the periplasmic disulphide isomerase DsbA and the heat-shock factor σ^{32} (Price and Raivio, 2009).

Interestingly, Cpx activation results in down-regulation of several genes, including the σ^{E} pathway genes *rseABC* and *rpoE* that encode the σ^{E} factor (De Wulf et al., 2002).

3.2.3 The phage shock response (Psp) pathway

The Psp response is an important envelope stress response pathway widely conserved, indeed, some Psp-like proteins occur in both Gram-negative and Gram-positive bacteria and even in archaea and plant chloroplasts (Darwin, 2005; Huvet et al., 2011; Joly et al., 2010).

It was initially described as a response to phage protein pIV production in *E. coli*. While, nowadays, it is studied as a stress response that monitor the integrity of the IM (Flores-Kim and Darwin, 2016). The protein PspA is the key player of the response and is encoded by the *pspABCDE* operon. It is regulated by the protein PspF, which also controls the expression of the gene *pspG*, of unknown function (Figure 4, a) (Engl et al., 2009).

It has been shown that *pspA* operon expression is induced by numerous envelope damages, including mislocalized secretins, extreme temperatures (superior at 50°C) or high osmolarity changes, the disruption of fatty acid biosynthesis and by

proton ionophores (Flores-Kim and Darwin, 2016). Psp-inducing events are sensed by PspA recruitment either in proximity to PspBC or directly to the IM. The release PspF can then activate the *psp* operon and *pspG* (**Figure 4**, a).

Recently, it has been demonstrated *in vitro* that *E. coli* PspA binds preferentially to phospholipid vesicles which have higher stored curvature elastic (SCE) stress (McDonald et al., 2015). SCE stress occurs in lipid bilayer when external forces generate deformation of lipids and force them to pack unfavorably. This increased curvature occurs during osmotic stress or upon changes in membrane fluidity due to extreme temperature variations (Nielsen et al., 1998). It seems that PspA oligomers can stabilize membranes in compromised area of the IM (**Figure 4**, b). Furthermore, this function is shared with the PspA-homolog Vipp1 in plant chloroplasts and cyanobacteria, which has been linked to thylakoid membrane formation and integrity (Westphal et al., 2001; McDonald et al., 2015).

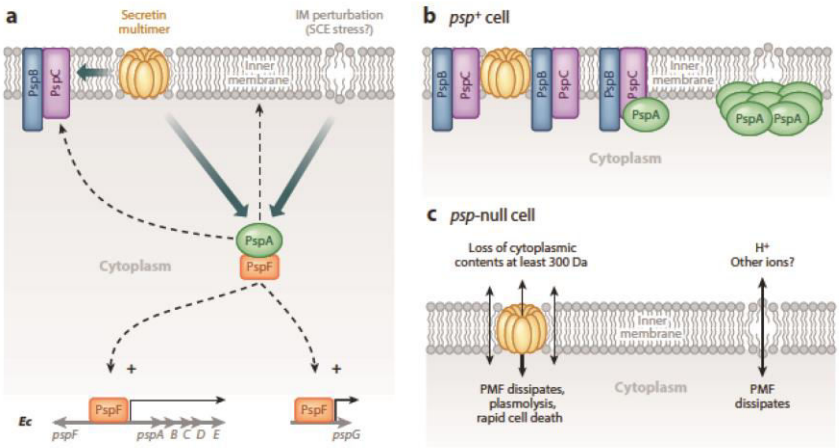


Figure 4. The phage shock response in *E. coli*. (a) A stress is caused by a mislocalized multimeric secretin, by other aberrant proteins or by environmental extremes that perturb the membrane. PspA senses the stress and interacts directly with PspBC or the IM. Sequestration of PspA releases PspF, which induces expression of the *pspA* operon and *pspG*. **(b)** In cells containing WT PspA protein, its oligomers stabilize compromised areas of the IM. **(c)** In cells with non-functional Psp response, the absence of PspA is associated with dissipation of the proton-motive force (PMF). If PspB and PspC are absent, in addition to PMF dissipation, mislocalized multimeric secretins also cause more severe IM permeability, resulting in loss of cytoplasmic contents and rapid cell death. Figure adapted from (Flores-Kim and Darwin, 2016).

3.2.4 The Rcs and Bae Pathway

Two other envelope stress responses are the Rcs and Bae pathways, here described only in general terms because not essential for the results section of the manuscript.

The regulator of colanic acid synthesis (Rcs) pathway is activated by osmotic stresses and damages to the peptidoglycan wall (Laubacher and Ades, 2008). Furthermore, Rcs regulon is activated by cell interactions with surfaces and plays a role in the proper development of biofilms on surfaces. The Rcs senses the stress signals across the IM *via* a TCS that results in activation of the response regulator RcsA. RcsA is degraded by the cytoplasmic protease, Lon, which serves to repress the system when the cell is not experiencing a stress (Majdalani and Gottesman, 2005).

The BaeSR TCS mediates the bacterial adaptive response (Bae) pathway. BaeS is the histidine kinase sensor and BaeR the response regulator that receives the phosphate. Bae stress response is induced by exposure to toxic compounds such as indole and ethanol. It regulates the expression of genes encoding multidrug transporters such as the *mdtABCD* operon, *acrD*, and *emrK*, in addition to the OM porin TolC and to the periplasmic chaperone Spy (Raffa and Raivio, 2002) along with few other conserved proteins (Sikdar et al., 2013).

4 Recombinant MP production in *E. coli*

Since the development of recombinant technologies in the late 1970s, the use of recombinant proteins in life science and industrial biotechnology has revolutionized academic research and opened new opportunities for the pharmaceutical industry. The possibility to produce proteins had a considerable importance in the therapeutics field. For example in 1980s, the recombinant human insulin from *E. coli* became the first recombinant pharmaceutical product to enter the market.

By May 2015, the number of protein-based recombinant pharmaceuticals licensed by the U.S. Food and Drug Administration and European Medicines Agency reached 400 (Sanchez-Garcia et al., 2016). Also, it is now possible to engineer proteins with enhanced functions and to equip them with accessories to favour their production, purification and detection.

Despite the huge potential and importance of recombinant protein production, MPs are still particularly difficult to produce. MPs are essentials for several biological processes such as cell signalling, metabolite transport, energy conversion and ion homeostasis (Villalobo, 1990; Cornelius, 1991; Tate, 2012). Approximately 30% of genes

encode MPs and they cause a growing number of human diseases. Unsurprisingly, about 50% of all existing drugs act on MPs (Lundstrom, 2007). It is known that one of the main bottlenecks in MPs production is the insufficient capacity of the Sec-translocon (Wagner et al., 2007). The overloading of the Sec-translocon leads to accumulation of misfolded proteins in the cytoplasm, which results to cell toxicity and consequently to low cell densities.

The low amount of proteins and the difficulty to produce them in a functional state, limit the structural and biochemical investigations of MPs. Although refolding procedures sometimes provide pure and homogeneous samples for structural analysis (Casiraghi et al., 2016) a current challenge is to produce them in biological membranes (Lyons et al., 2016).

Over the last twenty years, eukaryotic and prokaryotic expression systems have evolved in parallel and, together with developments in MP crystallization and improvement of crystal diffraction (more focused and more brilliant X-ray beams), have boosted the number of MP structures (Freigassner et al., 2009). For example, more than 30 unique G protein-coupled receptors (GPCR) structures have been obtained after overproduction either in insect (Sf9) and human (HEK) cells (Tate, 2012). Despite an increment in the use of eukaryotic systems, about half of the unique MP structures present in the Protein Data Bank (PDB) (Hattab et al., 2015) have been produced in *E. coli*, suggesting that it is still a powerful vehicle for recombinant MP production. Unfortunately, not all proteins of eukaryotic origin can be successfully produced in bacteria. Among the differences, the lack of glycosylation machinery in *E. coli* and the different transcription and translation rates are the most important. In such cases, yeast, insect, fungal, mammalian or plant cells are maybe a better choice. However, if the aim is to produce high amounts of protein, the use of bacteria is usually more relevant. Bacteria grow faster, they are cheaper and easier to maintain and they are physiologically and genetically less complex.

A first aim of my Ph.D. has been to isolate new bacterial hosts from BL21(DE3) with improved recombinant MPs expression ability. In the following sections I will describe the T7 expression system and its historical BL21(DE3) host widely used for MP production. In addition, I will briefly describe various strategies to isolate new bacterial hosts.

4.1 The protein production strain BL21(DE3) and the T7 RNAP expression system

E. coli BL21 is well known for the purpose of protein production as it lacks two major proteases, the outer membrane protease OmpT and the cytoplasmic protease Lon. Both proteases have been shown to interfere with protein production and isolation (Shineberg and Zipser, 1973; Baneyx and Georgiou, 1990).

BL21(DE3) was developed by Studier and co-workers for the production of soluble proteins (Studier and Moffatt, 1986). The nomenclature "DE3" denotes a prophage, which has been integrated into the genome of BL21. The prophage sequence contains the gene of the T7 RNAP controlled by the *lacUV5* promoter. The gene encoding the protein of interest is expressed on a plasmid under the control of a T7 promoter. The addition of the inducer isopropyl-β-D-thiogalactoside (IPTG) initiates the transcription of the T7 RNAP gene. Subsequently, the T7 RNAP transcribes the target gene inserted in the expression plasmid (Figure 5).

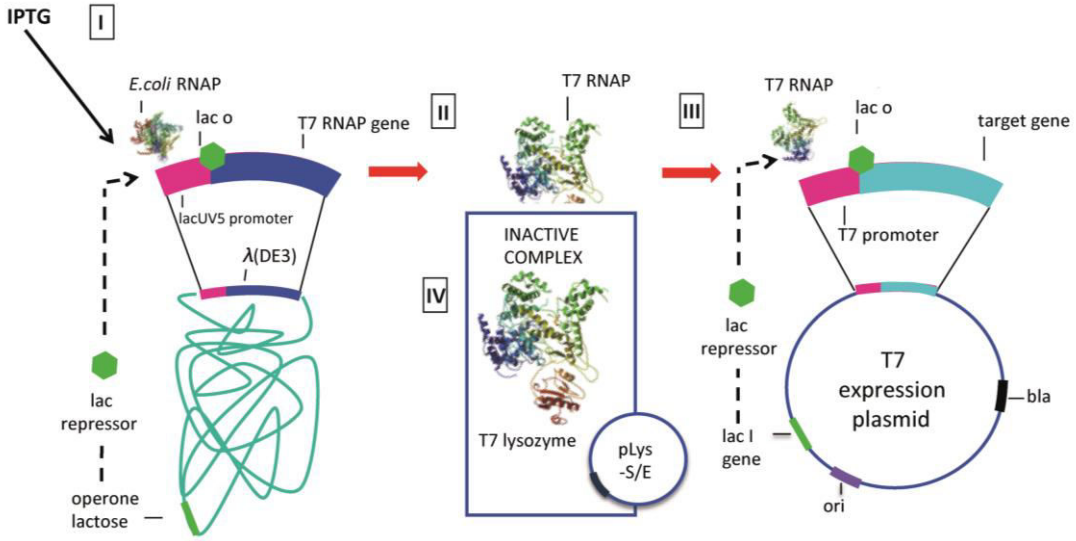


Figure 5. Global view of the T7-based expression system in BL21(DE3). The T7 RNAP gene is lysogenic in the genome of BL21(DE3) and its expression is under the control of the IPTG-inducible *lacUV5* promoter. **(I) & (II)** After addition of IPTG, the T7RNAP is produced. **(III)** The RNA polymerase will then specifically transcribe the target gene expressed in the T7 expression plasmid. **(IV)** The activity of the T7 RNAP can be lowered by expressing in BL21(DE3) a plasmid producing lysozyme, the natural inhibitor of the T7 RNAP. Figure adapted from (Angius et al., 2016).

The *lacUV5* promoter is derived from the *lac* promoter; it consists of the regions -35 and -10, three operator sites and a binding site for the cAMP-receptor-protein CRP.

In the absence of inducer, the *lac* inhibitor (LacI) binds to, at least, two operator sites (out of three) and prevents RNAP binding (Ippen et al., 1968; Lewis, 2005). The addition of IPTG releases LacI from the operator sites (*lacO*), which allows RNAP to transcribe the genes placed downstream of the *lac* promoter.

The *lacUV5* promoter exhibits four single point mutations as compared to the *lac* promoter. These are located in the promoter (-10) and operator (-35) as well as in the CRP-binding site. The mutations seem to increase efficiency of RNAP binding and to make the promoter less sensitive to glucose catabolic repression (Arditti et al., 1968; Wanner et al., 1977). This makes *lacUV5* promoter stronger in presence of glucose than the wild-type *lac* promoter.

The combination of *lacUV5* and T7 RNAP, which transcribes eight times faster compared to *E. coli* RNAP (Iost et al., 1992) and targets only the target gene promoter, makes BL21(DE3) a powerful tool for proteins production. However, this system is often leaky due to the partial repression of the *lacUV5* promoter. Since even small amounts of T7 RNAP can be lethal for the cell when the target gene is toxic, additional levels of control are often provided. For example, the natural inhibitor, T7 lysozyme, decreases the background levels of T7 RNAP (Figure 5) (Studier, 1991).

According to the global analysis carried out in our laboratory (Hattab et al., 2015), the six promoters currently used for protein production are: T7, T5, arabinose, lactose, rhamnose and tetracycline. The T7 and arabinose promoters cover 80% of all unique MP structures (63% and 17%, respectively). In the next section, I will discuss the other promoter systems used for recombinant proteins production.

4.1.1 Other promoters used for recombinant proteins production

In the T5 promoter system the *E. coli* RNA polymerase recognizes the T5 promoter upstream of the target gene and induction is accomplished by IPTG addition, which displaces LacI from the two operator sites (Bujard et al., 1987).

The arabinose inducible promoter, *araBAD*, is a relatively weak promoter that is susceptible to catabolic repression (Guzman et al., 1995). The *araBAD* promoter-based gene expression system depends on the regulator protein AraC, which acts both as an activator and an inhibitor. AraC represses transcription in the absence of arabinose and activates transcription in its presence. Arabinose can be transported into the cell *via*

two transporters: AraE (low affinity) and AraFGH (high affinity) (Schleif, 2010). This promoter was successfully used in our laboratory to obtain structural insights about the heme-haemopexin acquisition system Hxu from *Haemophilus influenzae* (Zambolin et al., 2016).

The rhamnose promoter *rhaBAD* is derived from the operon responsible for utilization of the sugar L-rhamnose (Giacalone et al., 2006). In the presence of L-rhamnose, the transcriptional regulator RhaR activates transcription of RhaS and RhaR, RhaS binds to the *rhaBAD* and *rhaT* promoters, controlling expression of genes involved in rhamnose metabolism and uptake, respectively (Wickstrum et al., 2005). Just as the *lac* operon, the *rha* operon is susceptible to glucose catabolic repression. One of the main advantages of this system is that the addition of increasing amounts of inducer leads to a gradual increase in the level of transcription, whereas *lac* promoters are difficult to titrate (Giacalone et al., 2006).

The tetracycline promoter (*tetA*) confers resistance to tetracycline. The same *tet* operon also encodes the *tetA* repressor protein TetR, which regulates the expression of *tetA*. The Tet promoter is a strong and non-leaky promoter and it is not sensitive to catabolic repression (Skerra, 1994; Muthukrishnan et al., 2012). Even if its features seems to be highly promising to MPs production, the Tet promoter contributions to structural studies are to date very low (Hattab et al., 2015).

4.2 Isolation of *E. coli* strains for MP production

Over the last ten years, *E. coli* strains with improved MP production abilities have been developed to overcome proteins toxicity during production processes. The best-known example is the isolation of C41(DE3) and C43(DE3) strains by Miroux and Walker. Understanding how these strains have been isolated and the mechanisms by which they allowed a successful improvement of MP production was a major task of my Ph.D. work. I will discuss in detail their isolation and the isolation of other *E. coli* strains from BL21(DE3) or by using engineering strategies.

4.2.1 The BL21(DE3) derivatives C41(DE3) and C43(DE3)

The insertion of a target gene in BL21(DE3) is often associated with cellular toxicity. In this case the plasmid containing the target gene was difficult cloned, the bacteria stopped growing immediately after induction of expression and cells growth was impaired on plates containing IPTG.

This intrinsic toxicity of the T7 system was exploited by Miroux and Walker in a genetic screen to isolate two mutant strains adapted for the production of soluble and membrane proteins (Miroux and Walker, 1996). *E. coli* C41(DE3) was isolated from BL21(DE3) upon over-expression of the oxoglutarate mitochondrial carrier protein (OGCP) gene. After addition of IPTG to the culture, BL21(DE3) cells harbouring the expression plasmid died. Screening on IPTG containing agar plates for surviving cells that still produced OGCP, led to the isolation of C41(DE3). In a similar manner, C43(DE3) was screened for its ability to produce the *E. coli* integral MP F_oF_1ATPase b subunit. These strains have proved useful for a large set of difficult proteins (Dumon-Seignovert et al., 2004; Miroux and Walker, 1996). Authors concluded that improving the fitness of the cell by fine-tuning the T7 promoter activation, allowed higher production levels of the target MP in a long-term induction experiment. Both mutant hosts have had a great impact on the structural biology of MP, especially for the most fragile α-helical MP (Hattab et al., 2015).

Subsequently, it has been shown that three single nucleotide polymorphisms (SNPs) are responsible for the improved production in C41(DE3) and C43(DE3), which are located in the sequence of the *lacUV5* promoter (Wagner et al., 2008). Consequently, the levels of T7 RNAP, upon addition with IPTG, are much lower than in BL21(DE3) (Wagner et al., 2008). This results in a lower expression rate of the gene encoding the target protein and consequently, the overproduced proteins do not saturate the protein biogenesis machinery.

C43(DE3) strain harbours the same mutation as C41(DE3), but in addition a fourth mutation was found in the *lac*-repressor gene, *lacI* (Kwon et al., 2015). It has been proposed that this mutation causes a higher affinity of LacI for the *lac* operator-binding site. This fits with a decrease of T7 RNAP accumulation levels in C43(DE3) compared to C41(DE3).

Both C41(DE3) and C43(DE3) acquired additional mutations during their isolation. The majority of mutations identified involved IS-elements (Kwon et al., 2015; Schlegel et al., 2015). IS-elements are small mobile genetic elements identified in most organisms. Generally, they carry only the information necessary for their own transposition (Siguier et al., 2014). The number of these elements change across the organisms. In BL21(DE3) there are 62 IS-elements, whereas the genome of MG1655 only harbours 54 Is-elements (Kim et al., 2017). IS-elements induce mutations simply by moving within the genome, since they are able to jump out of one position and into another or to move *via* replication intermediates (Mahillon et al., 1999).

Given that the activity of IS-elements has been linked to stressful conditions, it has been proposed that at least a part of the observed mutations might be a consequence of the isolation procedures used (Schlegel et al., 2017). For example, in both C41(DE3) and C43(DE3) strains, *rbsD* gene function was restored by IS-element excision. In *E. coli* B strain REL606 this mutation mediated by IS-elements caused difficulties for ribose uptake and catabolism (Cooper et al., 2001). In a similar manner, the expression of *lon* is restored in C43(DE3) strain (Kwon et al., 2015). As mentioned before, the Lon protease interferes quite often with the production of proteins in the cytoplasm.

4.2.2 Other examples of improved *E. coli* strains

Other mutants better suited for the production of membrane and secretory proteins have been isolated. In some cases, the spontaneous occurrence of mutations was coupled with additional mutagenic events.

Four different *E. coli* strains named NG2, NG3, NG5 and NG6 were generated upon overproduction of different MPs (Gul et al., 2014). In this selection GFP was used as a folding reporter and the erythromycin resistance protein (ErmC) to select improved MPs production by culturing cells in the presence of increasing amounts of antibiotic. The so-obtained strains carried different mutations in the gene encoding the protein HNS. It is plausible that this gene is implied in the function of the membrane protein biogenesis machinery (Gul et al., 2014).

In the so-called EXP strains, the accumulation of mutations was facilitated by chemical mutagenesis and by using a strain proofreading-deficient of DNAP III (Massey-

Gendel et al., 2009). The gene encoding the target protein was simultaneously expressed from two separate plasmids. On one plasmid the target gene was fused to a kanamycin resistance gene and on the other plasmid, the same gene was fused to a trimethoprim resistance gene. The clones were selected for the ability to grow in presence of both antibiotics. The screen yielded several mutants, even if the nature of mutations remains unknown. The authors proposed that the increased protein production level was linked to a reduced load on the biogenesis machinery (Massey-Gendel et al., 2009).

The transposon-mediated mutagenesis was used with the aim to improve the production of certain GPCR (Skretas and Georgiou, 2009). Also in this case, the GPCR was fused to GFP enabling the isolation of clones with improved characteristics by fluorescence-activated cell sorting (FACS). The isolated strain harboured a transposon in the gene encoding the DnaK co-chaperone DnaJ. The proposed scenario to explain the increased production yield of the host was that DnaJ competes with the targeting factor SRP of the GPCR or that DnaJ is involved in the degradation of this particular protein (Skretas and Georgiou, 2009).

In two different publications, the same research group found first that co-expressing the *ybaB* gene encoding a putative DNA-binding protein, led to a prominent increase in the production levels of several MPs (Skretas and Georgiou, 2010). More recently they screened for improved GPCR production a library of approximately 10^7 different *E. coli* chromosomal fragments (Skretas et al., 2012). The co-expression of three different gene fragments resulted in improved GPCR production: *nagD* encoding the ribonucleotide phosphatase NagD*,* a fragment of *nlpD* encoding a truncation of the predicted lipoprotein NlpD and the three genes cluster *ptsN-yhbJ-npr* encoding three proteins of the nitrogen phosphotransferase system.

A screening method for the generation of scFv antibody fragments was developed by using the periplasmic expression with cytometric screening (PECS) (Chen et al., 2001). *E. coli* cells producing scFvs in the periplasm are incubated with a fluorescent ligand, which can traverse the OM and bind to the scFv. Only ligands that are bound to its corresponding fragments are retained in the periplasm. Subsequently, the cells exhibiting the highest fluorescence are selected using FACS. This method was

used to isolate strains producing high amounts of full-length antibody (IgG) (Makino et al., 2011).

4.2.2.1 Developing E. coli strains through engineering approaches

When factors improving protein production yields are known, the use of targeted engineering is an attractive alternative approach to engineer new bacterial hosts. The characteristics of the strains C41(DE3) and C43(DE3) were used to develop the Lemo21(DE3) strain (Wagner et al., 2008). In Lemo21(DE3) the activity of the T7 RNAP can be tuned precisely by inserting, in a companion expression plasmid, the T7 lysozyme, a natural inhibitor of the T7 RNAP, under the control of the *rhaBAD* promoter. As mentioned before, the concentration of rhamnose provides a tight control of the T7 lysozyme levels and consequently, an optimal amount of active T7 RNAP can be found for each protein target.

Another example is the deletion of genes encoding transcription factors, which had beneficial effects on the recombinant protein production of some MPs (Nannenga and Baneyx, 2011). Another attempt was the deletion of the phosphoenolpyruvate:phosphotransferase system (PTS) (Bäcklund et al., 2011). These mutants had a slower growth rate, accumulated lower levels of acetate and produced higher levels of recombinant proteins.

Furthermore, the co-production of biogenesis factors have been linked to improved MPs levels in *E. coli*. For example, the levels of membrane-integrated CorA, an *E. coli* magnesium transporter, increased markedly when DnaK and DnaJ were co-produced (Chen et al., 2003). CorA could be produced at high level even in the absence of additional chaperones, but the vast majority of the protein accumulated as inclusion bodies. The authors argue that, given the large N-terminal extension of CorA prior to the first TM it is likely that this transporter is targeted to the membrane in a post- rather than a co-translational fashion.

To produce disulphide bond-containing proteins, *E. coli* strains that allow disulphide bond formation in the cytoplasm were engineered. One example of this kind, is the production of DsbC in the cytoplasm, which facilitates the biogenesis of proteins containing multiple disulphide bonds (Hatahet et al., 2014).

5 Summary on MP production

The amount of produced protein in *E. coli* depends on different factors: strain background, cultivation conditions, amount of inducer, and co-production of other biogenesis factors. In this introduction I have described some factors that may affect recombinant protein production levels as the *E. coli* strain choice or the promoter controlling the expression of the target protein. However, the target protein sequence also influences expression level of the protein and must be optimized by adapting the sequence of the gene to the *E. coli* machinery or by adding N- or C-terminal tags.

There are numerous factors that can affect protein yields, this makes difficult to design *a priori* a strategy that guarantees the highest amount of a given target protein. Consequently, the screening of culture conditions is a mandatory step to optimize protein production yields of the target gene.

Hattab and co-workers (Hattab et al., 2015) have identified several optimization rules and have shown that IPTG concentration and growth temperature are important parameters complementary to the choice of a bacterial host. Authors have also revealed that high copy number plasmids are preferentially used with C41(DE3) consistent with the fact that this strain was isolated using a high copy number plasmid (pMW7 encoding the oxoglutarate mitochondrial carrier).

6 Consequences of MP over-expression in *E. coli*

In the previous sections I focused on Gram-negative proteins and on how optimize their production. We have seen that a lot of research group, including ours, aim to express the target protein in native membranes rather than as inclusion bodies to obtain a folded and active protein. In some cases, over-expression of MPs results, not only in its expression at *E. coli* membranes, but also in intracellular membranes (ICM) in which the protein is stored. The morphology of the ICMs seems to be based on lipid-protein interactions.

In the second part of my Ph.D., I used molecular biology and biochemistry tools to understand ICMs proliferation upon over-expression of the F-ATPase b subunit. In

the following next sections I will pursuit this introduction focusing on membrane lipids and lipid-protein interactions, which can generate membrane curvatures.

7 Membrane lipids

For years, lipids were believed to act as a rather passive solvent for membrane proteins. Recent studies, however, indicate that they have significant roles in cellular processes (van Meer et al., 2008). For example, the water-lipid interface is a rough surface where the lipid headgroups create an active interface that change the concentrations of charged molecules or ions close to the membrane surface (Lee, 2004).

Lipids have distinct features that make them important in the construction of biological membranes. First of all their oxidation releases a great amount of energy allowing the cell to store its energy in lipids and use it later for energy requiring processes. Another interesting feature arises in their chemical structure. Membrane lipids are amphiphilic molecules that have hydrophilic headgroups and hydrophobic acyl chains. These properties enable them to self-assemble in an aqueous environment and thus segregate the cell from the outer environment.

Lipids play an important role in signal transduction and molecular recognition processes (van Meer et al., 2008), in cellular processes such as energy transduction, cellular trafficking, endocytosis/exocytosis, translocation of membrane proteins across membranes and even cellular defense mechanisms (Williamson et al., 1984; Pohnert, 2002; van Meer et al., 2008; Santos and Riezman, 2012).

E. coli IM is composed of ~40% (wt/wt) phospholipids and ~60% (wt/wt) proteins. The phospholipid composition is ~75-80% phosphatidylethanolamine (PE), ~15-20% phosphatidylglycerol (PG) and ~2-5% cardiolipin (CL) (Cronan, 2003) (Figure 6). The shapes of these phospholipids have a deep impact on the physicochemical characteristics of the bilayer, which in turn affect the functions of membrane proteins (van den Brink-van der Laan et al., 2004; Wikström et al., 2009).

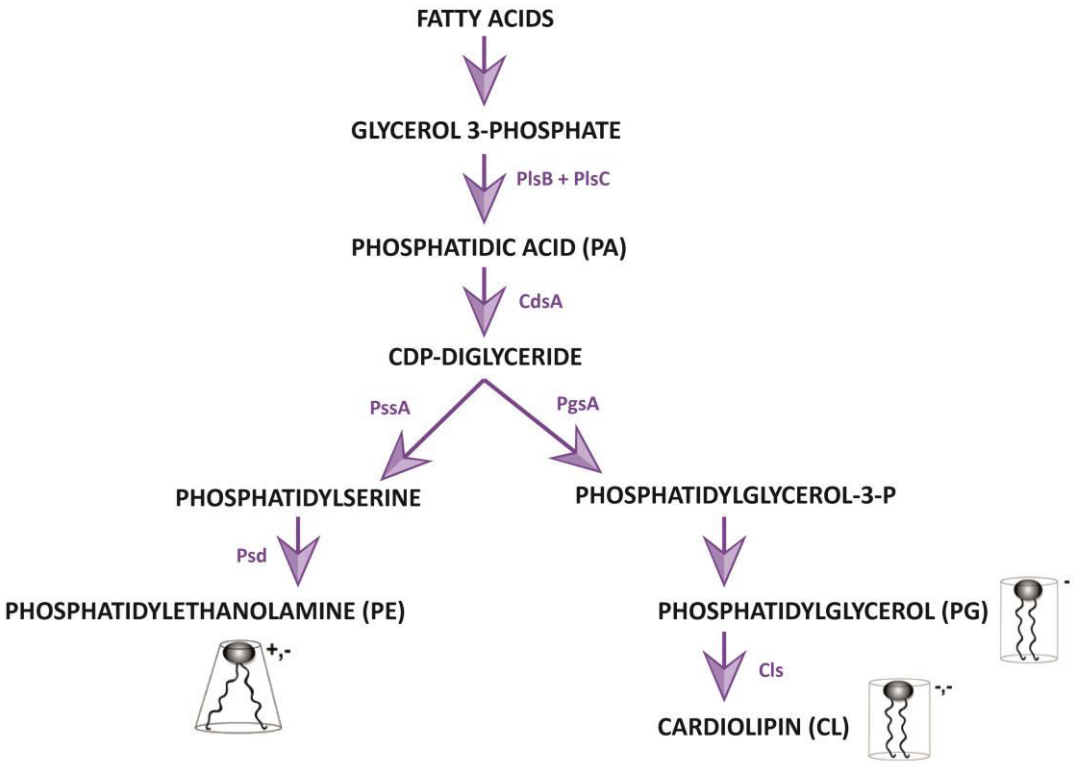


Figure 6. Phospholipid biosynthesis pathways in *E. coli*. Phospholipids are constituted by fatty acids and glycerol backbones. Both building blocks are synthetized from acetyl-coA, which is derived from amino acids in growth medium. Figure adapted from (Cronan, 2003).

Structural differences in lipids are known to affect their physicochemical properties and the broad lipid diversity is a cell-strategy for environmental changes adaptation. Acyl chain and headgroup modifications are the most common strategies to diversify the lipid species in a biological membrane. Bacteria can regulate the synthesis of new lipids but also modify the existing species in order to maintain the required membrane properties (Zhang and Rock, 2008).

7.1 Acyl chain modifications

Fatty acids or acyl chains determine the viscosity of a biological membrane and influence its permeability. Four types of fatty acids exist in bacterial membranes: saturated (SFA), unsaturated (UFA), branched-chain (BCA) and cyclopropanated (CFA). Saturated fatty acids, such as palmitic acid (16:0), are linear and pack tightly to form a bilayer with high temperature phase transition and low permeability. Bacteria can introduce insaturations, i.e. double bonds into growing fatty acids *via* FabB or FabZ enzymes, which results in a pronounced angle in the chain (Zhang and Rock, 2008). If a

double bond is inserted among the chain, they cannot pack tightly and the order of the bilayer is disrupted.

Membranes with UFAs have lower transition temperatures and higher permeability. Cis-UFAs (in which the carbon moieties lie on the same side of the double bond) build membranes with lower phase-transition temperatures, increased fluidity and high permeability to solutes compared to trans-UFAs. However, few bacteria have evolved to convert cis-UFAs to trans-UFAs and generate membranes with higher phase-transition temperatures, increased rigidity and decreased permeability to solutes.

Besides unsaturation, branching of the acyl chains is another common modification in bacterial membranes. Branched-chain fatty acids create an effect similar to UFAs, which leads to altered membrane order. Since the methyl group addition affects the alignment of acyl chains, the position is important to determine the effects on the bilayer. Anteiso-fatty acids have their methyl branch close to the mid-point of the acyl chain and promote a more fluid membrane compared to the bilayers with iso fatty acids (the methyl group is in the end of the acyl chain). Bacteria are known to adjust the iso:anteiso ratio and modulate membrane properties in response to the fluctuations in the environment (pH, temperature, pressure etc.) (Kaneda, 1991).

Another type of modification is called cyclopropanation and this is usually observed in bacteria entering into the stationary phase. Cyclopropanation is achieved by the conversion of pre-existing cis-UFAs to cyclic forms of the methylated fatty acids. The membrane-associated enzyme cyclopropane fatty acid synthase (CFA) performs this conversion (Zhang and Rock, 2008). Transcription of the *cfa* gene is controlled by the σ^S transcription factor, which is activated when the cells enter in stationary phase. In the membrane, CFA interacts with anionic lipids and the surface charge is proposed to be important for the up-regulation of *cfa* transcription. The increase in CL content during stationary phase seems to support this phenomenon (Raetz, 1978; Grogan and Cronan, 1997; Tan et al., 2012). The advantage of having cyclopropanated species arises from the high stability of the cyclopropane bond compared to SFAs and UFAs. Some pathogenic *E. coli* strains with higher levels of cyclopropanated fatty acids were observed to be more resistant to acid stress, antibiotics and temperature shifts (Zhang and Rock, 2008).

The complexity of biological membranes does not originate only from fatty acid variations but also from the diversity in the headgroups and lipid backbones.

7.2 Headgroup variations

In bacteria, phosphatidic acid is the main precursor of all glycerophospholipids (containing glycerol as the backbone) and slight modifications on the headgroup generate a diverse collection of phospholipids with different physicochemical properties. The most important feature is the surface charge brought by the lipid headgroups. A diverse collection of lipid headgroups, is generated by bacteria to balance an optimum surface charge in the membrane (Dowhan and Bogdanov, 2012; Parsons and Rock, 2013). Maintenance of an optimum surface charge is important, since many enzymatic reactions are regulated by electrostatic interactions (Dowhan and Bogdanov, 2012; Zhang and Rock, 2008). Some membrane-associated enzymes, like PssA, are regulated by surface charge. For PssA, the decrease of negativity at the surface of the membrane causes the inactivation of the enzyme and thus zwitterionic (+,-) (Figure 6) PE levels decrease, leading to increase of anionic lipid levels. An increase of anionic lipid content contributes to the negative surface charge and re-attracts PssA to the membrane causing re-increase of PE levels (Linde et al., 2004).

7.3 Backbone composition

Some bacteria use alternative backbones to glycerol-3-phosphate when phosphate levels are low in growth medium. The common substitution is the replacement of glycerol backbone with amino acids such as ornithine (most common), lysine, glycine, glutamine and serine-glycine (Parsons and Rock, 2013).

The second common backbone type observed in lipids is the sphingoid backbone. A sphingoid backbone is a long aliphatic amino alcohol N-linked to a fatty acid and O-linked to a charged group (ethanolamine, serine or choline). Sphingolipids are a class of lipids that contain sphingoid backbone and are classified as mechanically stable and chemically resistant. They help bacterial cells to be more tolerant against oxidative stress and heat shock. In eukaryotic cells, sphingolipids form micro domains in

the membrane and play significant roles in signal transduction and cell recognition (Parsons and Rock, 2013).

7.4 Lipid/protein ratio

An important physical parameter of a biological membrane is membrane viscosity. It can be changed by modifying fatty acids and headgroups, and also by changing lipid/protein ratio. Proteins are more rigid than lipid molecules and increasing protein content in membranes leads to more rigid membranes with low permeability and mechanical instability. In contrast, increasing lipid content will decrease the order and form a flexible stable membrane with high permeability. In eukaryotes, increasing the content of cholesterol, a lipid with rigid structure, controls this equilibrium. Higher cholesterol content is known to increase viscosity and thus to create a more rigid membrane with lower permeability to solutes (Bastiaanse et al., 1997).

7.5 Bilayer-prone/non-bilayer prone ratio

Attractive and repulsive forces between the headgroups and fatty acids of two lipids with different shapes usually create a lateral pressure in the membrane, which affects membrane curvature.

Membrane phospholipids usually have a cylindrical shape if the cross-sectional area of the headgroup is similar to the cross-sectional area of the acyl chains. Such lipids are called bilayer prone lipids and they aggregate to form a liquid-crystalline lamellar (Lα) phase, a phase that is usually seen in biological membranes (bilayer shape, zero curvature) (Figure 7).

If the cross-sectional areas of headgroup and acyl chains are different from each other, lipids are referred as non-bilayer prone lipids. There are two types of non-bilayer prone lipids in biological membranes. The first type has a larger headgroup compared to its acyl chains and forms positive curvature (van den Brink-van der Laan et al., 2004; Brown, 2012). Lipids with positive curvature form micelles in water and tend to have normal hexagonal HI phase (Brown, 2012).

The second type of non-bilayer prone lipids has relatively small headgroups compared to their acyl chains (Figure 7). These lipids generate negative curvature and form an inverted conical molecular shape. They are capable of forming reverse hexagonal (HII) phase in the bilayer (Brown, 2012).

Adjusting the bilayer prone/non-bilayer prone ratio is extremely important since the overall contribution of their molecular forces affects the mean curvature of biological membranes. In 1977, Israelachvili and co-workers (Israelachvili et al., 1977) had already presented a general theory of lipid self-assembly in which the mechanism involves the interplay of thermodynamics, interaction forces and molecular geometry.

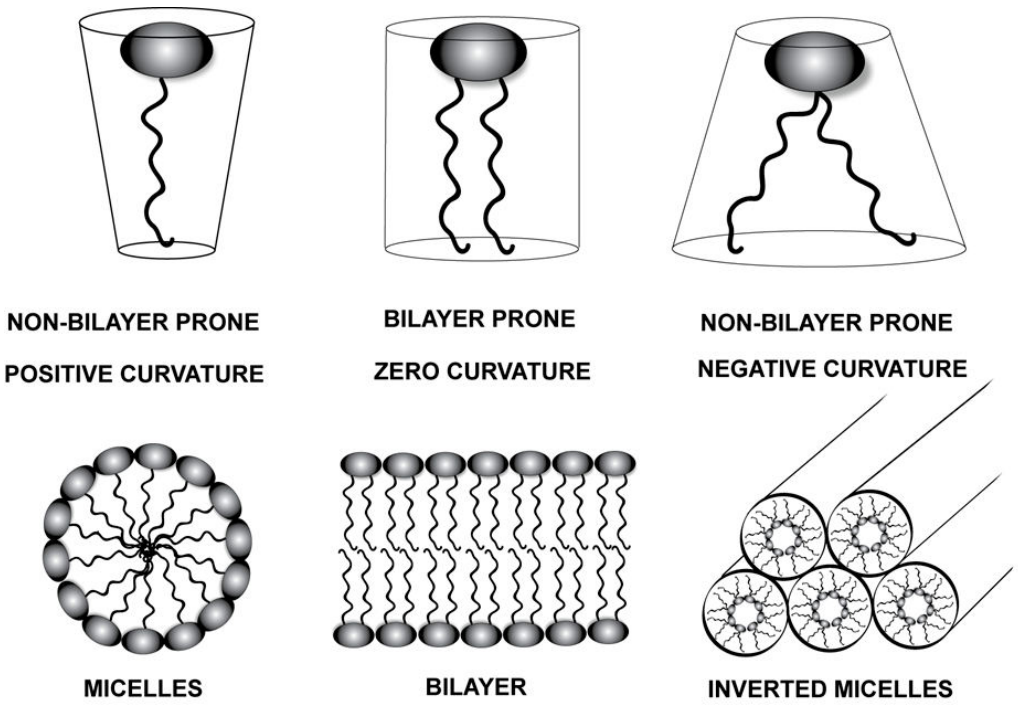


Figure 7. Different kind of lipids assembling. Lipids causing positive and negative curvature are grouped as non-bilayer prone lipids and tend to aggregate as micelles and inverted micelles, respectively. Bilayer-prone lipids do not induce curvature creating a lipid bilayer. Adapted from (van den Brink-van der Laan et al., 2004).

8 Lipid-protein interactions

8.1 Types of interactions

Lipids are often described by their proximity to the membrane proteins. Depending on the position where these lipids are found they are termed differently and the type of interaction differs with the orientation of lipids with respect to the mid-plane of the bilayer. According to their proximity, lipids can be classified as 1st shell lipids, 2nd shell lipids and bulk lipids (Lee, 2011).

8.1.1 1st shell lipids (annular lipids)

The 1st shell lipid is considered as an annular lipid because it forms a structure ring-like around the membrane protein. The length of the interaction among lipid and membrane protein depends on the amino acid affinity of the residue for the lipid molecule of interest. If this interaction is strong, then the lipid molecule has slow off-rate compared to bulk lipids and usually the interaction point is considered as a hot spot for lipid binding.

Annular lipids have regulatory effects on the functions of membrane proteins. A well known example of protein-annular lipids interactions is the large-conductance mechanosensitive channel (MscL) (Marius et al., 2005; Powl et al., 2007). MscL from Mycobacterium tuberculosis is a pentameric channel in which a cluster of three positively charged residues Arg98, Lys99 and Lys100 form a hot spot where anionic lipids are retained on the protein surface and released upon deprotonation (Marius et al., 2005; Powl et al., 2008a, 2008b).

8.1.2 2nd shell lipids (non-annular lipids)

Another class of lipids found in the membrane protein structures is non-annular lipids. They are similarly attached to the membrane protein surface as annular lipids but they are likely trapped in small pockets formed by amino acid residues in unusual positions (headgroup below the membrane plane and/or nonperpendicular to the bilayer) and have much lower off-rates compared to annular lipids (Lee, 2004; Palsdottir and Hunte, 2004; Marius et al., 2005). Some non-annular lipids have major roles in the regulation of membrane protein function by contributing to their folding and self-assembly in the membrane (Palsdottir and Hunte, 2004). One of the best-known examples is phosphatidylinositol (PI), which is observed in the structure of the cytochrome bc1 complex. PI resides in an interhelical (an unusual) position where side chains of cytochrome bc1 stabilize the headgroup of PI through the formation of several hydrogen bonds with the inositol headgroup. Lipid binding dissipates the torsion forces generated by the fast movement of the extrinsic domain of the Rieske protein. In this position, PI acts as a regulatory molecule for the cytochrome bc1 complex and could also be important for the self-assembly of the complex because it is located in a

position where four TM subunits of the monomer get in contact with each other (Palsdottir and Hunte, 2004).

8.1.3 Bulk lipids

Bulk lipids contribute to the macroscopic properties of the membrane, including viscosity (fluidity), internal pressure (lateral stress) and spontaneous membrane curvature (curvature elasticity) (Lee, 2004).

Membrane curvature influences a diverse variety of processes such as membrane fusion and fission, molecular transport, recruitment of proteins into the cytoplasmic surface, regulation of the activities of membrane proteins, molecular trafficking etc. (Mim and Unger, 2012). Biological membranes often contain lipids that adopt the hexagonal (HII) or non-bilayer phase. By modulating the proportion of bilayer and non-bilayer lipids, it is possible to modulate membrane curvature.

8.2 Membrane curvature

Membrane curvature defines the ability of a membrane to curve its mid-plane towards the interior or exterior of the bilayer itself. According to Helfrich model, this movement requires energy (Jarsch et al., 2016). The energy needed is created in the bilayer by the free-energy released arising from the spatial arrangements of lipid headgroups and chains.

As mentioned before, membranes can be positively curved (towards the cytoplasm) or negatively curved (away from the cytoplasm), and their deformability varies depending on the tension in the membrane (McMahon and Gallop, 2005).

8.2.1 Lipid-asymmetry-depending membrane curvature

There are two different mechanisms to generate membrane asymmetry in a flat bilayer. One way of altering the membrane curvature is by modifying the local lipid composition in terms of either total amount of lipid molecules or diversifying lipid species (modifications of headgroups or fatty acids), or both (Graham and Kozlov, 2010).

The second mechanism relies on the intrinsic shape of some membrane proteins, which creates an imbalance in the bilayer, and introduces asymmetry. In addition, the insertion of hydrophobic protein domains can trigger asymmetric insertion of hairpins or amphipathic helices (Jarsch et al., 2016). It seems that the molecular mechanism depends on surface charge fluctuations and on the proton-motor force (PMF) (Bakowski et al., 2010).

In mitochondria cristae membranes, calculations of the electrostatic field strength demonstrate that regions of high membrane curvature display a significant increase in charge density. It has been suggested that cardiolipin (CL) segregation to membrane compartments of high curvature may further contribute to the proton sink and enhance ATP synthase function (Bogdanov et al., 2008).

8.2.2 Modulating membrane curvature by protein scaffolds

In addition to membrane-intrinsic mechanisms, various cytoplasmic protein machineries modify membrane shape. In this case, curved peripheral binding proteins act as monomeric scaffolds or as homo- or heterodimers. For example, the Bin-amphiphysin-Rvs (BAR) domain can be curved to various degrees, generating either positive or negative curvature (Zimmerberg and Kozlov, 2006).

It is worth noting that the shape of the protein is not the only factor that can generate curvature. Protein crowding on the membrane surface may influence the local surface density of the proteins. The accumulation of proteins disrupts the homogeneous distribution of both lipids and membrane proteins and leads to an asymmetry in the membrane causing curved protrusions (Zimmerberg and Kozlov, 2006).

In the result section, we have illustrated the role of cardiolipin interactions with the F-ATPase b subunit in the intracellular membrane morphology. The nature of the interactions has not been elucidated but some of the mechanisms described above are likely to be involved.

9 Intracellular membranes in bacteria

Membrane curvature in prokaryotes is often linked to a photosynthetic apparatus. Photosynthetic bacteria, like *Rhodobacter sphaeroides* (Sener et al., 2007) and *Synechocystis* (Nevo et al., 2007) have an intricate system of membrane invaginations to accommodate the photosystem complexes, called thylakoids. In cyanobacteria, as *Synechocystis,* the thylakoid membrane maintenance and biogenesis are assigned to the Vesicle-inducing protein 1 (Vipp1), a peripheral membrane protein broadly conserved in cyanobacteria and plants. Vipp1 interacts with thylakoid membranes *via* its N-terminal amphipathic domain (Otters et al., 2013). In addition, Vipp1 can form oligomeric structures, which are capable of binding to lipid vesicles *in vitro* (McDonald et al., 2015).

Recently, two bacterial membrane remodelling systems, discussed in the following sections, have received a growing attention. These are the magnetosome biogenesis in the Alphaproteobacterium *Magnetospirillum gryphiswaldense* and the unconventional cytoplasmic organization of Planctomycetes.

9.1 Magnetosome biogenesis

Magnetosomes are sophisticated systems for detection of magnetic fields developed by magnetotactic bacteria. Electron tomography established that magnetosomes derived from the inner cell membrane invaginations (Greene and Komeili, 2012). The structuration of these membranes is under control of a high iron amount which is transported inside the invagination and contributes to the creation of well-ordered cuboctahedral crystals of magnetite (Raschdorf et al., 2016).

Invaginations are supported by genes organized in a large genomic island called the Magnetosome Island (MAI) (Grünberg et al., 2001; Raschdorf et al., 2016). Interestingly, among the genes that constitute the MAI island are the actin homologue MamK (Komeili et al., 2006) and MamY, which encodes a putative BAR domain. MamY was shown to tabulate liposomes *in vitro* by direct interaction with phospholipids (Tanaka et al., 2010; Cornejo et al., 2016).

Additionally, membrane invaginations during magnetosome biogenesis depend on six genes *mamB, mamM, mamE, mamO, mamQ* and *mamL* (Raschdorf et al., 2016).

MamL contains a positively charged C-terminal helix, which is predicted to interact with the inner cytoplasmic membrane and trigger membrane curvature during magnetosome biogenesis (Komeili, 2012). MamB and MamM proteins are believed to transport iron into magnetosomes. Furthermore, recent studies have ascribed to MamB the most important role in magnetosomes formation (Raschdorf et al., 2016).

9.2 Cytoplasmic compartmentalization in Planctomycetes

For many years, Planctomycetes were believed to be evolutionary linked to eukaryotes (Devos and Reynaud, 2010; Forterre and Gribaldo, 2010; Fuerst and Sagulenko, 2011). This hypothesis was supported by the presence of a proteinaceous wall rather than a peptidoglycan cell wall (König et al., 1984) and a compartmentalized cytosol that is separated *via* an intracytoplasmic membrane into a paryphoplasm and a pirellulosome (Lindsay et al., 1997).

Phylogenetically, the Planctomycetes are Gram-negative bacteria (Wagner and Horn, 2006). Their cell plan differs from a Gram-negative bacteria by the presence of interconnected invaginations leading to an enlarged periplasmic space (Boedeker et al., 2017). Interestingly, results suggest an inner membrane localization of the F_1F_o ATPase in the Planctomycete *G.obscuriglobus.* This observation is consistent with the phenomenon of membranes cristae formation in mitochondria, as mentioned before and with the observation that the membrane curvature is linked to a different PMF corresponding to membrane curvature regions.

9.3 MP-induced intracellular formation

Although *E. coli* do not generate intracellular membranes, the formation of additional invaginations from the inner membrane, tubules, sacks and vesicles have been reported when native or foreign membrane proteins are over-expressed. Some examples are the fumarate reductase (Weiner et al., 1984a), the ATP synthase complex (Von Meyenburg et al., 1984), the glycerol acyl transferase PlsB (Wilkison et al., 1986a), the manitol permease MtlA (Weeghel et al., 1990), the FtsY-receptor-ribosome complex

(Herskovits et al., 2002), the chemotaxis receptor Tsr (Weis et al., 2003a; Lefman et al., 2004).

In addition, intracellular membranes have also been observed upon heterologous expression in *E. coli* of the bacteriophage PM2 sps6.6 gene (Armour and Brewer, 1990), the foot and mouth disease 3A gene product (Weber et al., 1996), the alkane hydrolase AlkB from *Pseudomonas oleovorans* (Nieboer et al., 1996) and the lipid glycosyltransferases alMGS and alDGS from *Acholeplasma laidawii* (Eriksson et al., 2009).

During the second part of my Ph.D., I started some experiments to understand the molecular basis of intracellular formation in *E. coli* C43(DE3) upon over-expression of the F-ATPase b subunit. In the following section I will further explore the MP feature and the intracellular membrane characteristics.

9.3.1 F-ATPase b subunit-induced intracellular membranes (ICMs)

In 2000, Arechaga and co-workers (Arechaga et al., 2000) reported that the abundant over-production of the *atpF* gene encoding the b subunit of *E. coli* F_1F_o ATPase in the mutant *E. coli* strain C43(DE3) was followed by massive ICMs formation (**Figure 8**). The electron micrographs of C43(DE3) cell sections showed tubular-like membranes.

In addition the F-ATPase b subunit over-expression affected the lipid composition of cells. It has been observed a CL level increment in proliferating ICM up to 14% as compared to 2-4% upon physiological conditions. A high CL level is consistent with the observation that the over-expression of a protein on the membrane surface may influence the local surface density. The accumulation of proteins, indeed, may change the distribution of both lipids and membrane proteins in an attempt to maintain a constant lipid/protein ratio. Thus, the disruption of conventional ratios may lead to membrane asymmetry causing curved protrusions, as discussed above.

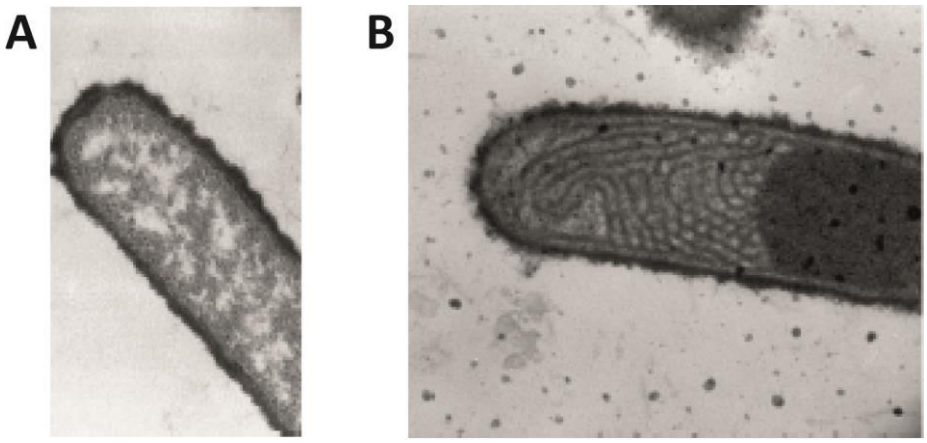


Figure 8. Thin sections of *E. coli* C43(DE3). (A) Cells without plasmid. **(B)** Cells over-expressing the F-ATPase b subunit plasmid. Figure adapted from (Arechaga et al., 2000).

9.3.1.1 F-ATPase b subunit topology

The *E. coli* b subunit protein is part of the F_1F_o ATP synthase, a multi-subunit enzyme complex responsible for intracellular ATP production. Structurally and functionally the complex can be divided into two major domains: the membrane intrinsic F_o domain and the peripheral F_1 domain. The two copies of the b subunit are necessary for the stable binding of F_1 to the membrane and for the correct assembly of the complex. The *E. coli* b subunit is a molecule of 156 residues with a hydrophobic N-terminal membrane-spanning α-helix and a highly polar cytoplasmic domain (Arechaga, 2013).

The protein is divided in four functional domains: a membrane domain (residues 1-24), a tether domain (residues 24-53), a dimerization or coiled-coil domain (residues 53-122) and the δ- binding domain (residues 122-156) (Revington et al., 1999) (Figure 9).

Figure 9. F-ATPase b subunit domain structure. Adapted from (Revington et al., 1999).

The N-terminal 24-residues sequence is highly hydrophobic and is embedded in the membrane with the other F_o subunits. The sequence form residues 25 to 52 is called tether domain because it joins the membrane region with the dimerization

region. In this region, loss of enzyme activity was observed following substitution of Arg36 in the tether domain (Welch et al., 2008). Revington and co-workers demonstrated that the b subunit dimers depend on residues from 53 to 122 and that the amino acid 105 appears to be proximal to its counterpart in the other b subunit.

It should be noted that the fragment from residues 24 to 156 tends to form dimers at 20°C when its concentration is higher than about 1 mg/ml, while at temperatures of 40°C the dimer interactions are not stable and the protein changes into the monomeric state (Revington et al., 1999).

10 Summary on ICM formation

In contrast to what was believed, membranes are dynamic structures responding to a large variety of different stimuli; they are not simple and inert mechanical envelopes separating the cell from the extracellular environment. Membrane proteins and lipids interact between them and communicate with the interior of the cell.

The possibility to induce membrane proliferation in a broadly used organism such as *E. coli* gives the advantage of both: to facilitate the study of these interactions and to provide a great biotechnology platform for the overproduction of membrane proteins. A better understanding of how ICMs in *E. coli* are produced would help to find a strategy to trigger their proliferation upon foreign MP overproduction. Such of storage organelle would be extremely useful not only in structural biology of MP but also for the storage of lipophilic compounds (carotenoids or fatty acids) that are usually difficult to accumulate in large amount in bacteria.

Materials and Methods

Chapter 1: Isolation and characterization of C44(DE3) and C45(DE3) E. coli strains

1 Bacterial strains

E. coli strain BL21(DE3) (Studier and Moffatt, 1986) was used to isolate C44(DE3) and C45(DE3) strains. C41(DE3) and C43(DE3) were described previously (Miroux and Walker, 1996) and were helpful to new strains characterization.

2 Vectors

The heat shock transformation was used to insert the vectors into the competent strains.

The high copy number T7 expression plasmid pHis17 derives from pMW7 (Way et al., 1990); it was constructed by Mike Runswick (MRC, Cambridge, UK) by adding three 6*His-Tag immediately in frame after BamHI, HindIII and EcoRI cloning restriction sites.

The pWALDO vector (Waldo et al., 1999) is an engineered pET28(a+) plasmid that contains the *BglII/XhoI* fragment from pET21(a+). In addition, the *BamHI/EcoRI* site was replaced with the DNA fragment GGATCCGCTGGCTCCGCTGCTGGTTCTGGCGAATTC coding for amino acid linker GSAGSAAGSGEF. This vector contains the GFPmut1 variant which consists of the red-shift S65T mutation and the folding F64L mutation (Cormack et al., 1996).

The pWALDO derivative pGFPe has an extended multiple cloning site (5' XhoI and 3' EcoRI/KpnI/BamHI), carries a TEV protease site in the linker sequence, and has a 8*His-tag placed in the GFP C-terminal (Rapp et al., 2004).

The plasmids pCT2Phe and pCT2His were used to suppress the T7RNAP amber stop codon in C44(DE3) strain. Both vectors are chloramphenicol resistant and were made by inserting HincII - HindIII fragment within tet of pACYC184 replaced by a 189 bp PvuII-HindIII fragment from pGIFIBI carrying synthetic suppressor genes.

3 Proteins

The pMW7-GFP-Xa construction (made by Mike Runswick) includes the soluble GFP with a C-terminal cleavage factor Xa.

The gene encoding the superfolder-GFP (sfGFP) protein was amplified from vector pDHL1029 (Ke et al., 2016) and subcloned between BamHI and HindIII into pHis17. The PCR reaction was carried out using the forward primer 5'-CAGGGATCCAGTAAAGGTGAAGAACTG-3' and the reverse primer 5'-CAGAAGCTTTTATTTGTAGAGTTCATC-3'.

All membrane proteins (MPs) used to characterize the strains are listed in Table 1.

Gene / Function	Origin	Molecular weight / TM	Plasmid	UniprotKB number
yidC - Membrane protein insertase	*Escherichia coli*	60 kDa 6 TM	pWaldo	NP_418161
pheP - Phenylalanine-specific permease	*Escherichia coli*	50 kDa 12 TM	pGFPe	NP_415108
yijD - Inner membrane protein	*Escherichia coli*	13 kDa 4 TM	pGFPe	P_418399
gltP - Glutamate-aspartate symport protein	*Escherichia coli*	47 kDa 10 TM	pWaldo	ACI73820
yfbF - Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase	*Escherichia coli*	36 kDa 2 TM	pGFPe	YP_002408357
ygfU - Uric acid transporter	*Escherichia coli*	50 kDa 13 TM	pGFPe	KLX82054
yqcE - Inner membrane protein	*Escherichia coli*	46 kDa 12 TM	pGFPe	ANK03046
tehA - Potassium-tellurite ethidium and proflavin transporter	*Escherichia coli*	36 kDa 8 TM	pGFPe	AAA19563
Ct6 - Mot A proton channel	*Chlorbium tepidum*	46 kDa 4 TM	pGFPe	Q8KC40
Dr10 - MscS channel	*Deinococcus tepidum*	40 kDa 5 TM	pGFPe	Q9RXU5

Dr35 - Multidrug efflux transporter	*Deinococcus tepidum*	44 kDa 12 TM	pGFPe	Q9RZF0
Dv1 - Na/Ca exchanger family	*Desulfovibrio vulgaris*	44 kDa 12 TM	pGFPe	Q723816
Ph1 - CIC channel	*Pyrococcus horikoshi*	40 kDa 11 TM	pGFPe	Q8U0I6
Dv12 - Na/H antiporter	*Desulfovibrio vulgaris*	45 kDa 13 TM	pGFPe	Q72AL4
Dv3 - MFS transporter	*Desulfovibrio vulgaris*	40 kDa 11TM	pGFPe	Q72WJ3
Sp17 - Formate/nitrate family transporter	*Silicibacter pomeroy*	26 kDa 10TM	pGFPe	Q5LW76
Oo3 - MscS channel	*Oenococcus onei*	32 kDa 3 TM	pGFPe	Q04DD4
Tt14 - transporter	*Thermus thermophilus*	42 kDa 10 TM	pGFPe	Q72HL7
Gs21 - Ms channel superfamily	*Geobacter sulfurreducens*	60 kDa 6 TM	pGFPe	Q74CF1
Ss21 - amino acid transporter-like iron(III) ABC transporter, permease	*Sulfolobus solfataricus*	49 kDa 12 TM	pGFPe	NP_343113.1
sqr - Sulphide:quinone oxidoreductase	*human*	50kDa monotopic	pHis17	Q9Y6N5

Table 1. Membrane proteins used in this work.

E. coli MPs (Zhang et al., 2015) were given to us by Dr. Hjelm and Dr. De Gier (Stockholm University, Sweden).

MPs from others prokaryotic organism (Hammon et al., 2009) were provided by Dr. Minor (UCSF, USA). They were cloned using a ligation-independent cloning (LIC) of PCR products containing LIC overhangs and inserted in pGFPe vector containing a LIC cassette.

A synthetic optimized version of the human sulphide-quinone oxydo-reductase (h-SQR) (Jackson et al., 2012) was obtained from Eurofins genomics and subcloned between the BamHI and EcoRI sites of the pHis17 plasmid.

4 BL21(DE3) derivatives selection

The *E. coli* strains C44(DE3) and C45(DE3) were isolated from BL21(DE3) in two independent selection experiments according to (Angius et al., 2016; Miroux and Walker, 1996). *E. coli* BL21(DE3) harbouring the pMW7-GFP-Xa was inoculated in 2xYT medium supplemented with ampicillin (100µg/ml). At OD_{600} between 0.4-0.6, 0.7 mM of Isopropyl-β-D-thiogalactopyranoside (IPTG) was added. One, two and three hours after induction, bacteria (1 ml of culture) were collected by centrifugation (300 xg for 2 minutes), and gently re-suspended in 1 ml of sterile water. In order to obtain separated colonies, serial dilutions were performed and µlplated on 2xYT-agar plates containing ampicillin and IPTG

overnight , plates were screened for low green fluorescence under far UV light (300 nm UV lamp to avoid further mutations) and subcloned onto 2xYT ampicillin plates. Isolated bacterial mutants were further analysed and cured from the expression plasmid by growing the cells in 2xYT without antibiotic addition. Serial dilutions were performed daily and 100 µl of 10^{-8} and 10^{-7} dilutions were plated on 2xYT agar plates.

A normal size colony showing no green fluorescence was isolated, transformed with the pMW7-GFPxa and plated on 2xYT plate with ampicillin and with or without 0.7 mM of IPTG to verify the green fluorescence in presence of inducer.

5 Culture conditions

E. coli C41(DE3), C43(DE3), C44(DE3) and C45(DE3) were grown overnight in 2xYT (16 g/l tryptone, 10 g/l yeast extract, 5 g/l NaCl) broth supplemented with ampicillin (100 µg/ml) when harbouring the pHis17 or pMW7 plasmids or with kanamycin (50 µg/ml) for all the other plasmids.

Overnight pre-cultures were diluted 1:100 and cultures were grown at 37°C. At OD_{600} of 0.4 proteins production was induced by the addition of 0.7 mM IPTG or different IPTG concentrations (from 0 to 0.7 mM) for titration experiment. The BL21(DE3) host has been grown at 30°C in Luria-Bertani (LB) (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl) broth according to Zhang and co-workers (Zhang et al., 2015) to optimize proteins production.

The h-SQR has been grown at 30°C in BL21(DE3), C44(DE3) and C45(DE3) strains, and at 25°C in C43(DE3) preserving the same induction conditions.

6 Flow-cytometry

The BD Accuri™ C6 Flow Cytometer was used to measure and to analyse single cells physical features. It can measure particle's relative size through the forward scatter channel (FSC), internal complexity through the side scatter channel (SSC), and relative fluorescence intensity by using 4 different detectors:

- FL1 (533/30nm), which detects green fluorescence;
- FL2 (585/40nm), which detects orange-red fluorescence;
- FL3 (>670nm), which detects red fluorescence;
- FL4 (675/25), which detects ultra-red fluorescence.

GFP fluorescence was monitored after overnight IPTG induction. 1 ml of culture were washed with Phosphate Buffered Saline (PBS) and diluted up to an OD_{600} of 0.01. Cells were analysed and counted during one minute by using a threshold of 10,000 on FSC and 3,000 on SSC to remove electronic noise, small contaminants and bacterial intrinsic fluorescence. Fluorescence was analysed with FL1 detector by gating on the population to further enhance the analysis precision.

All fluorescence cytometry experiments were based on a minimum of three biological replicate and row data were analysed and processed with FlowJo10.1 software.

Statistical analysis were performed by using Prism 6 taking into account the standard error of the mean (SEM).

7 Genomic analysis

Genomic DNA of BL21(DE3), C41(DE3), C43(DE3), C44(DE3) and C45(DE3) were extracted from an overnight culture with the Wizard® Genomic DNA Purification Kit (Promega, U.S.A.). DNA quality was checked on 0.7% agarose gel.

The reference genome for this analysis was *E. coli* BL21(DE3) (GenBank ID AM946981.2). The sequences were examined comparing them with the reference

genome through the software Integrative Genomics Viewer (IGV). The mutations identified were further confirmed by PCR using the primers listed in Table 2.

Name	Sequence	Bacterial mutant
Gene 1 for	GGCCCTTGAGCATGAGTCTT	C44(DE3)
Gene 1 rev	GAGACTCGTGCAACTGGTCA	C44(DE3)
rbsD for	TGATATTTCATCGGTGATCTCCC	C44(DE3)
rbsD rev	CGAATTTCAATGGTATTTCCCTG	C44(DE3)
Gene 1 for (PCR)	GGCCCTTGAGCATGAGTCTT	C45(DE3)
Gene 1 rev (PCR)	GAGACTCGTGCAACTGGTCA	C45(DE3)
Gene 1 sequencing primer	TCTGGCTTGCCTAACCAGTG	C45(DE3)
rbsD for	TGATATTTCATCGGTGATCTCCC	C45(DE3)
rbsD rev	CGAATTTCAATGGTATTTCCCTG	C45(DE3)
gltL for	CGGTGCAGCAAGGTGAAATC	C45(DE3)
gltL rev	CTGCGCCGGAAACTTATTGG	C45(DE3)
YcdX for	AGTGGTTGATGGGGTAGGGA	C45(DE3)
YcdX rev	TCGCTTGTGTATTGGTCGCT	C45(DE3)
B-lom for	GCGCGAATATGCCGGTTATC	C45(DE3)
B-lom rev	GCCACCTCTTCCACCATCAG	C45(DE3)

Table 2. PCR primers used to confirm mutations in C44(DE3) and C45(DE3).

8 Human-SQR purification and activity assay

Overnight cultures were harvested by centrifugation at 6000 xg for 10 minutes and resuspended in 30 ml of 50 mM Tris-Cl buffer, pH 7.4 supplemented with 1x protease inhibitor cocktail. The cells were disrupted with the CellD (Constant Systems LTD) using a pressure of 2 kbar. After cell debris removal, the presence of possible inclusion bodies was verified with a 20 minutes, 10,000 xg centrifugation.

Membranes were isolated after a one-hour centrifugation at 100,000 xg of the low speed supernatant. Based on the colorimetric Bradford assay, membranes were diluted at 2 mg/ml proteins in 50 mM Tris-HCl buffer (pH 7.4), 50 mM NaCl, 10% glycerol, 1x protease inhibitor cocktail (Thermo scientific) and solubilised in presence of 2% DDM for one hour at 4°C.

Insolubilized materials were removed by centrifugation at 100,000 xg at 4 °C for 1 hour. The high speed supernatant was collected and diluted in 100 ml final volume of 100 mM Tris-HCl buffer (pH 8.0) containing 150 mM NaCl, 10% glycerol and 5 mM Imidazole. Sample was loaded onto a His Pur^{TM} Ni-NTA Chromatography column (Thermo Scientific) previously equilibrated with a 40 mM Tris-HCl pH 8, 150 mM NaCl, 5 mM Imidazole and 0.05 % DDM buffer. h-SQR was eluted with a buffer containing 40 mM Tris pH 8, 150 mM NaCl, 0.05% DDM, 10% glycerol and 200 mM Imidazole.

Activity assays were conducted under anaerobic conditions at room temperature using a screw-cap cuvette equipped with a Teflon-silicon septum. h-SQR activity was measured by monitoring the reduction of DMSO-solution Coenzyme Q_1 (CoQ_1) at 278 nm. A 50 mM sodium sulphide (Na_2S) stock solution was prepared in argon flushed assay buffer. Cuvettes containing assay buffer, 100 mM potassium phosphate, 0.5 mM EDTA pH 8 with 0.05% DDM and different CoQ_1 concentrations were flushed with Argon for 10 minutes before recording. The concentration of CoQ_1 used was determined on the basis of its absorbance at 278 nm (ε = 12 000 M^{-1} cm^{-1}). Prior to initiation of reaction, 400 µM Na_2S were added and the reaction were started in response to 0.2 µg of purified h-SQR. All reagents were added using gastight Hamilton syringes. The CoQ_1 reduction rates were determined after subtraction of the initial slope obtained before h-SQR addition. The CoQ_1 absorption at 278 nm was corrected at 700 nm.

9 Fluorescence-detection size-exclusion chromatography (FSEC)

A 200 ml culture of cells producing the different MPs were harvested by centrifugation at 5,000 xg for 10 min and bacterial pellet was suspended in 20 mM Tris-HCl pH 8.0, 150mM NaCl buffer (TBS) supplemented with 1% (v/v) PMSF. All steps were carried out on ice or at 4°C.

Isolated cells were broken twice using a CellD (Constant Systems) at 2.4 kBar. After unbroken cells removal, the supernatant was centrifuged for one hour at 100,000xg and membranes were suspended in 5 ml solubilisation buffer (20mM Tris-

Hcl, pH 8, 150 mM NaCl and 2% DDM). After one-hour solubilisation at 4°C, insoluble material was discarded by 30 minutes of ultracentrifugation at 100,000 xg.

An alternative method for solubilisation was also employed to avoid different protein/detergent ratio between different strains. After suspension of the membrane pellet in TBS, protein concentration was measured (Pierce BCA assay) and 18 mg (2 mg/ml) of MPs were solubilized with 1% DDM in 20 mM Tris-Hcl, pH 8.0, 150 mM NaCl buffer for one hour at 4°C. Solubilised MPs (100 µl) were loaded on a Superose 10/300GL column (GE-Healthcare) or onto a Superdex 5/150GL column (GE-Healthcare). Size-exclusion chromatography was performed in presence of TBS supplemented with 0.03% DDM. Intelligent fluorescence detector (FP-2020, Jasco) was employed in order to measure GFP fluorescence with an excitation wavelength of 488 nm and emission wavelength of 512 nm. Data were analysed with Prism 6.

10 Real time quantitative PCR (RT-qPCR) gene expression analysis

Total RNAs from BL21(DE3), C44(DE3) and C45(DE3) were extracted from cells using a phenol-chloroform protocol (Angius et al., 2016).

Briefly, 700 µl of cell culture were added to a lysis solution (35.5 µl of 20 % SDS, 7 µl of 200 mM Na-EDTA and 500 µl of water-saturated phenol) maintained at 65°C. After a few cycles of intermittent mixing, the solution was cooled on ice and centrifuged for 2 minutes at maximum speed to separate aqueous phase from organic phase. The aqueous phase was then transferred in new Eppendorf tubes. The RNA was extracted in the same way two more times by adding an equal volume of phenol or phenol/chloroform 1:1 solution (chloroform was added with isoamyl alcohol prior to use in a proportion v/v 24:1), respectively. RNA was precipitated with 1/10 of 3M Na-acetate, pH 5 and 2.5 volume ethanol and left at −20 °C for 2 hours. The tubes were then centrifuged and the pellet was washed with 1 ml of 70 % ethanol. Once the pellet was dried the samples were suspended in 400 µl of sterile water and treated with RNase-free DNase I (Qiagen).

Next, a second phenol/chloroform extraction and a second RNA precipitation were performed. RNA concentration was assessed with the spectrophotometer and

purity calculated by estimating the ratio 260/280 nm. RNA quality was checked by depositing 10 µg of each sample on 0.7 % agarose gel.

cDNA synthesis was performed using the QuantiTect reverse transcription Kit (Qiagen). According to manufacturer's protocol, in order to avoid genomic DNA contamination, 1 µg of RNA was treated with 2 µl of gDNA Wipeout buffer at 42°C for 2 minutes and the entire solution was transcribed in cDNA with 1 µl of Quantiscript Reverse Transcriptase at 42°C for 15 minutes.

The so-obtained cDNA was diluted 5 times and 1 µl was used in the reaction mix along with 1 µl of the reverse and forward primer of interests (Table 3) and 5 µl of SsoAdvanced™ Universal SYBR® Green Supermix (Biorad). The 96-well plate was analysed with the CFX96 Touch™ Real-Time PCR Detection System (Biorad). All the samples were made in triplicate and for all the primer pairs a melting curve was performed. The Integration Host Factor (IHF) mRNA was used as internal control for standardization. Statistical analyses were made using Prism 6 software.

Name	**Sequence**
sfGFP f	ACCCGGATCACATGAAGCAG
sfGFP r	CAGGATGTTGCCGTCCTCTT
ihf f	CAAGACGGTTGAAGATGCAGT
ihf r	GCAAAGAGAAACTGCCGAAA
T7 RNA P f	AGTCAAGCTGGGCACTAAGG
T7 RNA P r	CACTGCGAGTAACACCGTGA

Table 3. Primers used to perform RT-qPCR analysis.

Chapter 2: E. coli intracellular membranes proliferation analyses

1 Bacterial strains

In order to study intracellular membranes proliferation, the *E. coli* strains C43(DE3) and C44(DE3) were used.

2 Vectors

E. coli atpF gene, which encodes for the F-ATPase b subunit, was cloned (Arechaga et al., 2000) into the high copy number plasmid pHis17 (characteristics described above). The same vector was used empty as control (Figure 10).

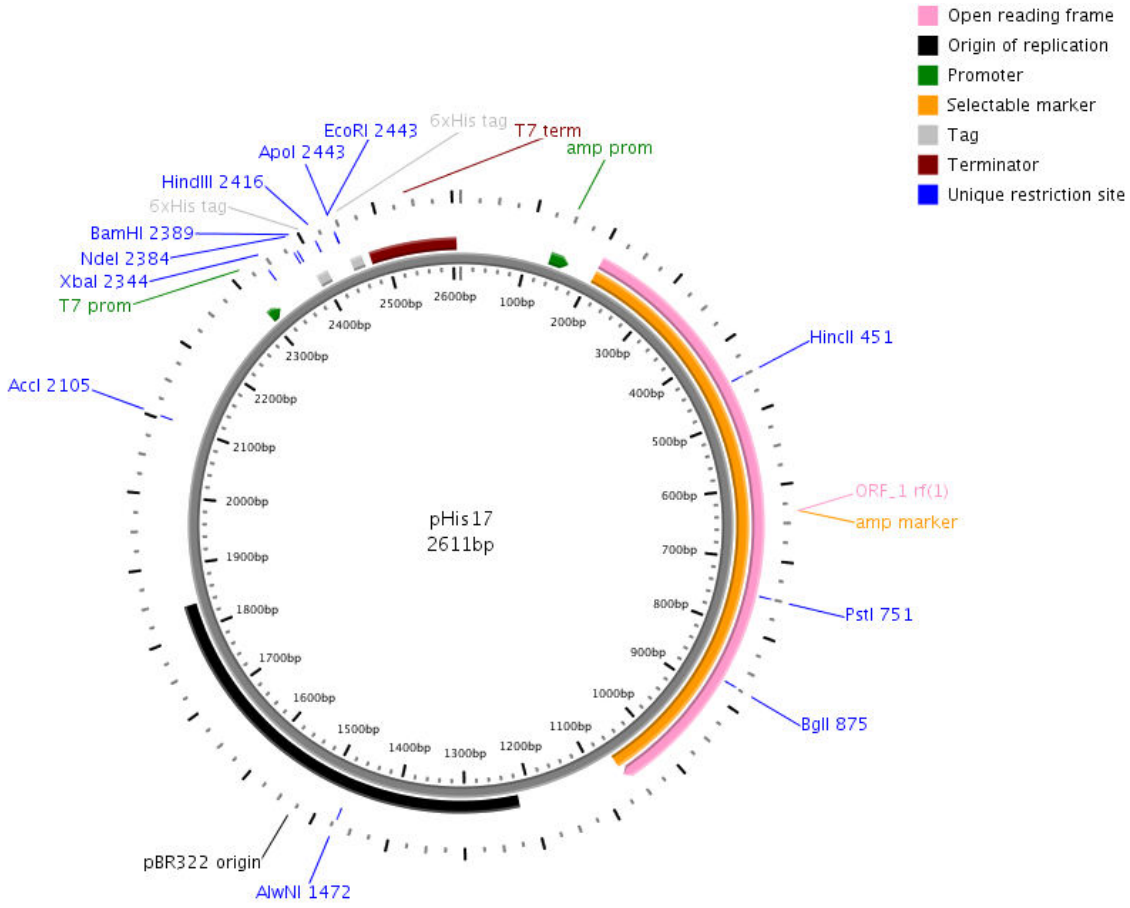


Figure 10. pHis17 graphic map.

TrwB and TrwC genes were cloned into pET3a plasmid using conjugative R388 plasmid as template (Carranza et al., 2017).

The mCherry and the blue fluorescent protein (BFP) tags were inserted in co-expression or fusion into F-ATPase b subunit expression vector. After amplification, tags were cloned between SalI and KpnI restriction sites to yield co-expression vectors. Instead, fusion vectors were obtained by insertion of tags between HindIII and KpnI restriction sites, in frame with a stop codon-deleted b subunit. Primers used are listed in Table 4.

Colony PCR was used to screen for appropriate colonies and the PCR products sequenced by Sanger method (Eurofins MWG Operon).

Name	Sequence	Restriction site
phis sub forw	5'-TACAGTCATATGAATCTTAACGCAACAATCCTC-3'	NdeI
phis sub revers	5'-TGCATGAAGCTTCAGTTCAGCGACAAGTTTATCCAC-3'	HindIII
Fus b mche f	5'-CATATAAAGCTTGTGAGCAAGGGCGAGGAGGAT-3'	HindIII
Fus b mche r	5'-TGCATGGGTACCTTACTTGTACAGCTCGTCCATGCC-3'	KpnI
bfp fusion f	5'-ATATACAAGCTTGTGAGCAAGGGCGAGGAGCT -3'	HindIII
bfp fusion r	5'-TCAATGGGTACCTTACTTGTACAGCTCGTCCATGCC -3'	KpnI
mCherrycoex f	5'-TATCGTGTCGACATGGTGAGCAAGGGCGAGGAGGAT-3'	SalI
mCherrycoex r	5'-TGCATGGGTACCTTACTTGTACAGCTCGTCCATGCC-3'	KpnI
Bfpcoexp f	5'-ATCATAGTCGACATGGTGAGCAAGGGCGAGGAGCT -3'	SalI
Bfpcoexp r	5'-TCAATGGGTACCTTACTTGTACAGCTCGTCCATGCC-3'	KpnI

Table 4. Primers used to clone tags in co-expression or fusion with F-ATPase b subunit.

3 Culture conditions

For fluorescence and electron microscopy analysis bacteria were grown in LB liquid medium at 37°C to an optical density of 0.6 at 600 nm. IPTG was added to a final concentration of 0.7 mM. The cells were grown for 3 hours at 37°C with appropriate antibiotics (100 µg/ml of ampicillin and/or 25 µg/ml of kanamycin).

For all the other experiments, cells were grown at 37°C in 1 L 2xYT liquid medium to an OD_{600} of 0.4 and induced with 0.7mM of IPTG. At that moment the temperature was lowered to 25°C until the next morning.

4 Construction of *cls* deletion mutants

Keio collection (Baba et al., 2006) was used to develop C43(DE3) *cls* deletion mutants. Keio mutants harbour a kanamycin cassette flanked by two FLP recombine sites that allow the cassette excision. P1vir transduction and calcium availability were used to regulate infectivity and to transfer the cassettes to C43(DE3) strain. Kanamycin cassette excision was carried out using the ampicillin-plasmid pCP20 that shows temperature-sensitive replication and thermal induction of FLP synthesis (Yu et al., 2000). Mutants for pCP20 were selected on ampicillin plates at 30°C. Colonies were purified at 42°C without antibiotics to cure the pCP20 plasmid.

C43(DE3) mutants were replicated on plates with or without antibiotics. The colonies that did not grow on plates with antibiotics were verified using PCR and DNA sequencing. The derived strains were subjected to another round of P1vir transduction and elimination of the kanamycin marker to introduce additional null mutations (Carranza et al., 2017).

5 Intracellular membranes (ICMs) isolation

Bacteria were grown as explained in section 3 of this chapter. According to (Arechaga et al., 2000) overnight cultures were harvested by centrifugation at 5,000 xg for 10 minutes and the pellet suspended in TEP buffer (10 mM Tris-HCl pH 8, 1 mM EDTA and 1% (v/v) PMSF) added with 5 µl of Benzonase (250 units/ µl, Sigma Aldrich). All steps were carried out on ice or a 4°C. A French Press (1,000 PSI) has been used twice to break the cells. After application of a low-speed centrifugation (2,500 xg, 10 minutes) the bacterial membranes (BMs), in the supernatant, were separated from the ICMs, in the pellet. The supernatant containing BMs was centrifuged at 100,000 xg for 1 hour to allow bacterial membranes sedimentation. While, to further purify the ICMs, the pellet was washed with TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA), homogenised and subjected to another low-speed centrifugation*. This last supernatant containing ICMs was centrifuged at 100,000 xg for 1 hour to allow membranes

* It is highly unusual to collect membranes at 2,500 xg. It might be due to the high number of cells (1L culture at OD_{600nm} = 8 concentrated in 25 ml), or the high density of the membranes and their association with DNA and cell debris. However, after washing the pellet at 2,500 xg, membranes will not anymore pellet at 2,500 xg but as expected, at 100,000 xg.

sedimentation. Procedures are illustrated in **Figure 11**. The Pierce BCA Protein Assay Kit was used to estimate ICMs proteins concentration. All samples were analysed by loading them onto an SDS-PAGE gel 14%.

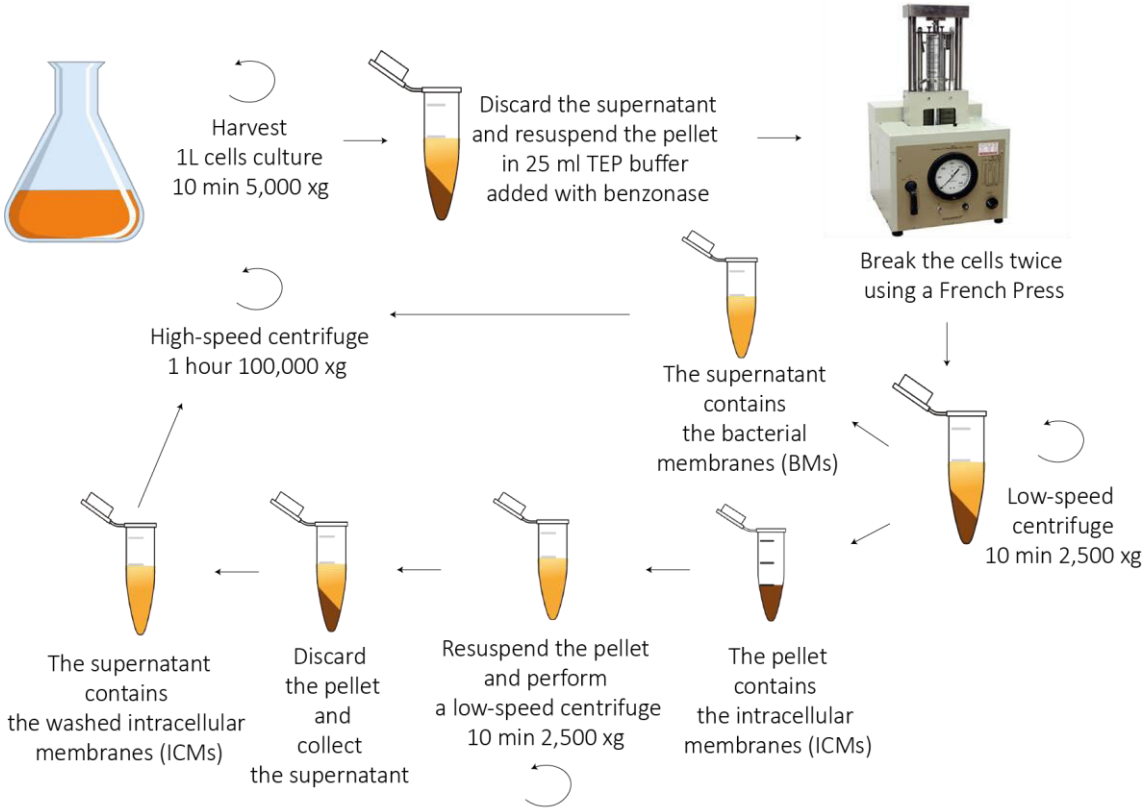


Figure 11. ICMs isolation protocol.

6 Differential centrifugation on sucrose gradient

To obtain continuous gradients, two sucrose solutions (5% and 50% w/v) in 10 mM Tris-HCl pH 8.0 buffer were prepared. The solutions were independently poured into each column of a gradient maker. The mixing was made down to the bottom into Beckman Coulter Ultra-Clear™ 12 ml centrifuge tubes thanks to a peristaltic pump. The tube was filled up with 1 ml of 2 mg/ml protein sample. Centrifugation was made at 100,000 xg for 18 hours at 4°C. The day after, 1 ml fractions were collected in Eppendorf tubes from top to the bottom and 20 µl of each fraction were loaded on SDS-PAGE gel.

7 Flow-cytometry

Cells were stained with the acridine orange derivative 10-N-nonyl acridine orange (NAO, Sigma-Aldrich). NAO contains a NH group in position 10 of acridine orange resulting in an increase hydrophobicity of the reagent and the inability to form hydrogen bonds with DNA or RNA. Therefore, NAO only binds to anionic phospholipids of the cells owing to an interaction between its quaternary amine and the phosphate residue of phospholipids and an intercalation of the hydrophobic acridine moiety into the membrane bilayer (Mileykovskaya and Dowhan, 2009).

NAO staining (2 µM) was performed 22 hours after induction with IPTG in 2xYT medium and incubated for 1 hour at 37°C. Excess NAO was removed by washing the cells with PBS (10 mM Phosphate, 150 mM NaCl, pH 7.4). The analysis was made as described in section 6 of chapter 1, through the FL1 detector.

8 Fluorescence microscopy

Cells were immobilized on microscope slides covered with a thin film of 1.2% agarose in water. Membranes were displayed with Nile-Red (Sigma-Aldrich) (1.5 µM) and NAO (Sigma-Aldrich) (2 µM). The fluorescent dyes were excited using a mercury lamp. Red fluorescence (with excitation at 550 nm and emission at 640 nm) from Nile Red was detected by using a Standard Rhodamine filter unit (Ex. 546/12–Em. 608/65).

Green fluorescence (with excitation at 495 nm and emission at 525 nm) from NAO was detected by using a standard GFP filter unit (Ex. 470/40–Em. 525/50). Standard fluorescence microscopy was carried out using Zeiss Axio Imager M1 upright fluorescence microscope (Zeiss Plan-Neofluar_100/1.30 NAOil Ph3objective) equipped with a 12 bits B&W camera (AxioCam MRm). The image analysis was carried out using ImageJ v.1.38 (National Institute of Health) (Carranza et al., 2017).

9 Electron microscopy

C43(DE3) and C43(DE3)Δcls derivatives were harvested and resuspended in PBS. Pelleted-cells were fixed with 3% glutaraldehyde and then with 2% osmium tetroxide. Fixed cells were dehydrated in acetone and embedded in araldite (Durcupan, Fluka, Switzerland). Cross sections were obtained by cutting embedded cells using a Leica

ultracut UCT ultramicrotome. Ultrathin sections were stained with 1% uranyl acetate in the dark and examined with a JEOL 2010 electron microscope (Santander, Spain).

C44(DE3) cells were subjected to a standard fixation in 2% glutaraldehyde/0.1 M cacodylate pH 7.2 (Sigma Aldrich) and to a post-fixation with 1% Osmium Tetraoxide (OsO_4, EMS), 1,5% potassium ferrocyanide in 100 mM cacodylate buffer. Cells dehydration was performed first by using different ethanol concentration and the second in acetone and embedded in EPON resin (TAAB). A Leica ultracut UCT ultramicotome was used to obtain embedded cells-cross sections. Ultrathin sections were picked up on formvar-coated grids, stained with 1% uranyl acetate and lead citrate to be examined with a FEI Tecnai Spirit 120 kV (Institute Curie, Paris, France).

10 Lipid analysis

Bacteria were grown in 2xYT medium as described in section 3 of this chapter. ICMs from C43(DE3) bacterial host and C43(DE3)Δcls derivatives were isolated as described in section 5 of this chapter. In control cells harbouring pHis17 empty plasmid, ICMs are not present and therefore total membranes were isolated by 100,000 xg ultracentrifugation for 1 hour at 4°C.

A Folch method (Folch et al., 1957) was used to extract lipids in triplicate from about 150 mg of isolated membranes. After added the latter to H_2O, membranes were broken by 30 minutes sonication (1 minute sonication at 40 W and 1 minute pause) on ice. Each solution was transferred into a separating funnel along with 100 ml chloroform/methanol (2:1 v/v) solution and 20 ml of H_2O/100 mM NaCl. Following decantation, the organic phase containing lipids was separated from the aqueous phase and evaporated on a rotovap. A vacuum pump was used to evaporate the remaining chloroform drops.

Dry lipids were analysed by mass spectrometry at the SAMM mass spectrometry facility following the protocol described in (Moulin et al., 2015). The phospholipids analysis was performed by coupling a liquid chromatography RSLC Dionex-U3000 (ThermoFisher Scientific) equipped with two quaternary pumps and a charged-aerosol detector Corona-CAD Ultra (ThermoFisher Scientific), and a LTQ-Orbitrap Velos Pro (ThermoFisher Scientific) mass spectrometer.

11 RNA-sequencing

E. coli C43(DE3) harbouring the F-ATPase b subunit and empty pHis17 plasmid were grown in 2xYT at 37°C to an OD_{600} of 0.4 and then induced with 0.7 mM IPTG. After induction the temperature was lowered to 25°C. Total RNA was extracted after 5 minutes, 3 hours and 22 hours after induction as described in section 10 of chapter 1.

RNA sequencing was performed thanks to the Transcriptome and EpiGenome Platform at Institut Pasteur (Paris). RNA quality and quantity was confirmed on a Bioanalyzer (Agilent). Enriched mRNA was obtained from 10 µg of total RNA using the rRNA capture hybridization approach from the RiboZero kit (Epicentre, Singapore), according to the manufacturer's instructions. For high-throughput sequencing, non-directional cDNA libraries were prepared from enriched fragmented mRNA using the TruSeq Stranded RNA LT sample prep kit, set A et B (Illumina). cDNA fragments of about 150 bp, ligated with Illumina adapters and amplified per PCR, were purified from each library. Sequencing of 51 bases was performed in single-end mode, using an Illumina HiSeq2000 instrument (Illumina).

The Bowtie program was used to align the reads to BL21(DE3) (NC_012892.2) genome. Statistical analysis were performed with R software and Bioconductor packages (Gentleman et al., 2004). Normalization and differential analysis were carried out according DESeq2 model and package (version 1.8.2) (Love et al., 2014). A Benjamini and Hochberg (BH) p-value adjustment was performed to take into account multiple testing and control the false positive rate to a chosen level of 0.05.

Gene expression values of the genes differentially expressed in at least one induction condition are represented in the form of heatmap by using the following script on R software:

```
>if (!require("gplots")) {
   install.packages("gplots", dependencies = TRUE)
   library(gplots)
   }
>if (!require("RColorBrewer")) {
   install.packages("RColorBrewer", dependencies = TRUE)
   library(RColorBrewer)
```

```
   }
>rna <- read.csv("~/Desktop/rna.csv", sep=";")
> rnames<-rna[,1]
> mat_rna <- data.matrix(rna[,2:ncol(rna)])
> rownames(mat_rna) <- rnames
> my_palette <- colorRampPalette(c("red", "yellow", "green"))(n
= 299)
>png("~/Desktop/rna.png",width = 5*100,height = 6*600,res =
300,pointsize = 6)
>heatmap.2(mat_rna,
main="Correlation",notecol="black",density.info="none",trace="no
ne",
margins =c(12,6), col=my_palette, dendrogram="none", Colv=FALSE,
Rowv=FALSE)
> dev.off()
```

12 ppGpp quantification

Measurement of ppGpp accumulation was made on cell cultures grown in 1x MOPS minimal medium added with 1x micronutrients and 0.4 mM phosphate concentration (2mM for overnight culture), low enough to achieve ^{32}P specific activities enabling nucleotide detection but high enough to partially repress *pho* regulon expression (Mechold et al., 2002).

Overnight cultures were diluted 1:100 into 4 ml of fresh medium and growth was monitored. At OD_{600} of about 0.05, carrier-free ^{32}P was added (final specific activity 150 µCi/ml) to 200 µl of culture and other eppendorf tubes were prepared with 200 µl of culture to monitor cells growth. Samples were grown until OD_{600} of about 0.4 and induced with IPTG. At different time point before induction and after induction, 15 µl of samples were mixed with an equal volume of 13M formic acid to allow cells to break. After freeze cycles, 5 µl of samples were applied to polyethyleneimine (PEI) cellulose thin layer chromatography plates, resolved with 1.5 M KH_2PO_4 (pH 3.4) and quantified by densitometry by using the Typhoon FLA9000 (GE Healthcare). Estimates of ppGpp were calculated by dividing the counts in GTP to the counts in ppGpp.

13 Proteomic analysis

After loading intracellular isolated membranes of C43(DE3), harbouring empty pHis17 plasmid or F-ATPase b subunit vector, on sucrose gradient, 20 µl of each fractions were applied on SDS-PAGE gel. In order to concentrate all membrane proteins in one band, sucrose gradient fractions 9 and 10 were run only until the sample entered in the running gel. Each band was cut off by separating it in two pieces (A and B) and submitted to trypsin digestion according to the following protocol already establish in our laboratory:

Coomassie-blue bands destaining
1. Place the band in an Eppendorf tube and wash with 200 µl of H_2O, 10 minutes under agitation.
2. Remove the supernatant and cover with 100 µl ammonium bicarbonate (NH_4CO_3) 25 mM/ Acetonitrile (ACN) 50%, 10 minutes under agitation.
3. Remove the supernatant and cover with 100 µl ACN 100%, 10 minutes under agitation.
4. Remove the supernatant and dry 5 minutes at 30°C using a SpeedVac™ (ThermoScientific).
Reduction & Alkylation
5. Put 50 µl of Dithiothreitol (DTT) 10 mM/ NH_4CO_3 25 mM, 1 hour at 56°C.
6. Replace it with 50 µl of iodoacetamide 55 mM/ NH_4CO_3 25 mM, 45 minutes at room temperature protected from light.
7. Remove the supernatant and add 100 µl NH_4CO_3 25 mM, 10 minutes under agitation.
8. Replace it with 100 µl NH_4CO_3 25 mM/ ACN 50%, 10 minutes under agitation.
9. Replace it with 100 µl ACN 100%, 10 minutes under agitation.
10. Remove the supernatant, dry 5 minutes at 30°C using a SpeedVac™ (ThermoScientific)
Trypsin digestion
11. Prepare a stock solution: • 2/3 µl trypsin stock 1 µg/µl (Promega) • 50 µl NH_4CO_3 50 mM

• 44.9 µl H_2O • 4.1 µl $CaCl_2$
12. Rehydrate the samples on ice with 10 µl stock solution freshly prepared. The bands should be completely covered.
13. Remove the supernatant and add about 30 µl of NH_4CO_3 25 mM.
14. Incubate for 4 or 12 hours at 37°C.
Peptides extraction
15. Transfer the supernatant in another Eppendorf tube.
16. Extract the peptides with 35 µl ACN 50% / trifluoroacetic acid (TFA) 1%, vortex 5 minutes and sonicate 10 minutes.
17. Transfer the supernatant into the tube.
18. Add 35 µl ACN 100% into the tube with the bands, vortex 5 minutes and sonicate 10 minutes.
19. Transfer the supernatant into the tube.
20. Dry the tube A mix using a SpeedVac™ (ThermoScientific) for about a hour and freeze the samples.

Once peptides were extracted, the analysis was performed thanks to Marion Hamon (IBPC Proteomic Platform). Peptides were diluted into 15 µl ACN 3% / formic acid 0.05%. Then 5 µl of tube A were mixed with 5 µl of tube B. 5 µl samples were analysed in four replicates by nano LC-MS/MS using an easy 1000 nano LC coupled to a Q-Exactive plus mass spectrometer (ThermoScientific). The sequence was made in a way to avoid the fluctuation due to the instrument during time by injecting samples in order of replicates. Peptides were desalted by a C18 pre-column (Acclaim PepMap®, 75 µm x 2 cm nano viper ThermoScientific) and separated on a 50 cm C18 column (PepMap® RSLC, 2 µm, 100 Å, ThermoScientific) using the gradient in Table 5 with an equilibration time of 30 min before each run.

Time (min)	Duration	%B
0	0	0
1	1	0
71	70	20
91	20	32
96	5	85
106	10	85
111	5	0
126	15	0

Table 5. Separation gradient set up.

The MS analysis was performed on a Q-Exactive with a top 10 method with 60 seconds of dynamic exclusion. Full MS was performed at 70,000 resolution with a max AGC target of 3.106 within a scan range of 400-1800 m/z and MS/MS were performed at 17500 resolution with a max AGC target of 1.105 with an isolation window of 2 m/z. The maximum trap filling was 100 ms. Peptides were identified using Mascot software and protein data bank UniProt.

Results

Chapter 1: BL21(DE3) derivatives isolation and characterization

As mentioned in the introduction sections, since the bacterial strains C41(DE3) and C43(DE3) were isolated in 1996 (Miroux and Walker, 1996), it becomes clear that reducing the amount of T7 RNA polymerase (T7 RNAP) was a trump card to improve membrane proteins production. To further ameliorate the T7 RNAP expression system by extending its promoter strength coverage, we decided to apply the same isolation protocol previously used to get C41(DE3) and C43(DE3).

Instead of using a membrane protein (MP), like the OGCP or the F-ATPase b subunit, the jellyfish protein green fluorescent protein (GFP) was chosen. The rationale behind this choice is that GFP is a very stable protein and can easily be monitored on agar plates, in whole cells flow cytometry and can be detected in standard SDS-PAGE (Drew et al., 2008). The main purpose was to create a whole set of production strains able to support the expression of a larger number of MPs.

In Chapter 1, I will describe the isolation and characterization of the new strains, named C44(DE3) and C45(DE3). They have been isolated, by Annabelle Suisse (C45(DE3)) and Oana Ilioaia (C44(DE3)) in our laboratory. My task was to characterize them and to identify the mechanisms of regulation responsible for improved MP production.

1 C44(DE3) and C45(DE3) hosts isolation

Since GFP production in the bacterial host BL21(DE3) is toxic, we selected BL21(DE3) derivatives on IPTG containing plates and made a second round of selection on plates without IPTG to keep mutants exhibiting no fluorescence in absence of IPTG. BL21(DE3) was transformed with a pMW7-GFP-xa vector in two independent experiments. The pMW7-GFP-xa construction includes the soluble GFP with a C-terminal cleavage factor “Xa”.

The transformation mixture was plated on 2xYT-agar plates containing 100µg/ml of ampicillin and 0.7 mM IPTG. Starting from a 2xYT culture, one, two and three hours

after induction, serial ten-times dilutions were performed and plated on 2xYT-agar plates containing ampicillin with or without IPTG (Figure 12). After an overnight incubation at 37°C, the screen yielded quite a number of IPTG-resistant colonies. About 1/100,000 cells survived the induction of GFP expression by IPTG as described in (Miroux and Walker, 1996).

To verify GFP fluorescence, plates were exposed to a natural light (Figure 12, A) and to UV light (Figure 12, B). The plates contained either fluorescent or non-fluorescent colonies; several clones of various fluorescence intensities were isolated on 2xYT plates. Those with the lowest fluorescence intensity in the absence of IPTG were further analysed. A long time culture in 2xYT medium was required to cure C44(DE3) and C45(DE3) from the plasmid. Once the strain was isolated, it was re-transformed with the vector pWM7-GFP-xa to validate that the phenotype was stable and that the mutation was in the bacterial genome.

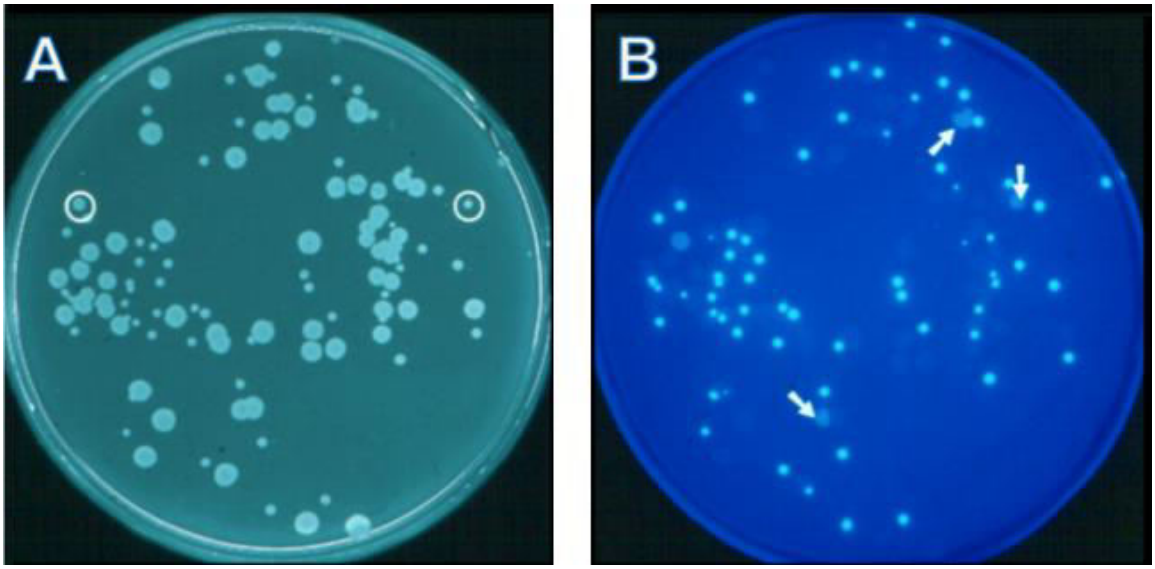


Figure 12. Selection of bacterial hosts using GFP-xa as gene reporter. Serial dilutions of BL21(DE3) over-expressing GFP-xa were plated on 2xYT-agar plates containing ampicillin and 0.7 mM IPTG. The next day plates were exposed to a natural light **(A)** and to UV light **(B)**. Arrows show low fluorescence colonies. Figure adapted from (Walker and Miroux, 1999).

2 C44(DE3) and C45(DE3) characterization

The superfolder-GFP (sfGFP) fluorescence was used to compare the T7 system regulation in BL21(DE3) derivatives. The rational behind is that sfGFP is a faster fold variant, resistant to random mutations and it is more soluble than conventional GFP (Pédelacq et al., 2006)

The relative fluorescence intensity (RFI) of cells harbouring the pHis17-sfGFP plasmid was measured by flow cytometry in different culture conditions (**Figure 13**). A constant amount of 0.7 mM of IPTG was initially monitored during a time course experiment. One, two and three hours after IPTG addition, sfGFP RFI increased proportionally to reach, after three hours, 4 and 7-fold increased of RFI in C44(DE3) and C45(DE3), respectively. Surprisingly, a subsequent RFI increased of 22-fold in C44(DE3) and 18-fold in C45(DE3) between 3 and 22 hours after induction were observed (**Figure 13**, A). This result shows that recombinant protein accumulation (which in the case of sfGFP is well reflected by the measured RFI) continues to accumulate in the stationary phase.

Next, different IPTG concentrations were compared on overnight cultures. In contrast to BL21(DE3), which is difficult to titrate with IPTG (**Figure 14**), both C44(DE3) and C45(DE3) can be titrated over an unprecedented range of IPTG concentration starting from any addition of IPTG and up to 700 µM (**Figure 13**, B).

IPTG titration is a way to understand if the amount of target protein produced in the cell is proportional to the concentration of the promoter regulator. Such titration may also provide information about the mechanism of regulation. Indeed, some research group have demonstrated that the best conditions of proteins production for the BL21(DE3) strain are due to the addition of low amount of IPTG (Alfasi et al., 2011) or omitting the inducer thanks to leaky production (Zhang et al., 2015).

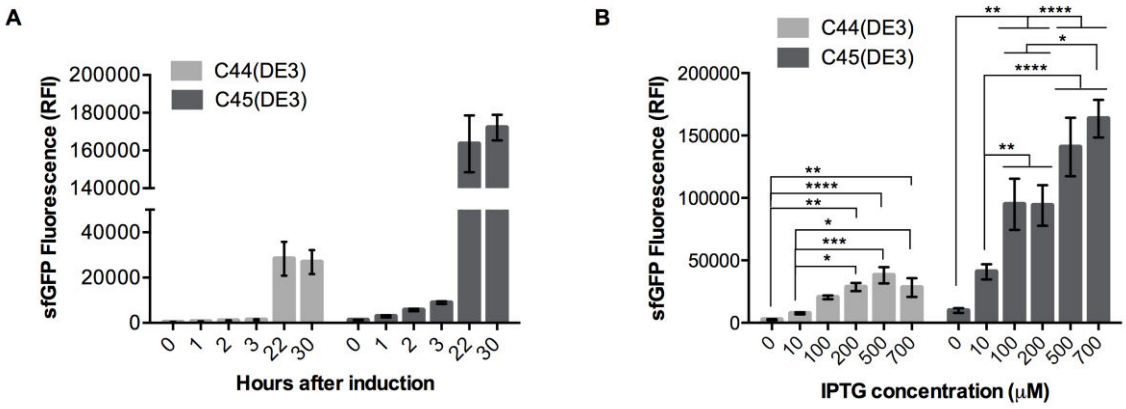


Figure 13. Regulation of the T7 expression system in C44(DE3) and C45(DE3) bacterial hosts. C44(DE3) and C45(DE3) bacterial hosts were transformed with pHis17-sfGFP expression vector. Two hundred thousand cells were counted and analysed for sfGFP relative fluorescence intensity (RFI). **(A)** Time course experiment of sfGFP RFI accumulation after addition of 700 µM of IPTG at $OD_{600\ nm}$=0.4 during the time. **(B)** IPTG titration experiment. Cells were induced at $OD_{600\ nm}$=0.4 with increasing concentration of IPTG (from 0 up to 700 µM). The sfGFP RFI was recorded after overnight induction. Each plot shows a mean value of six independent experiments. Standard error of the mean is indicated. Statistical significance was analysed using One-way ANOVA test: * p value<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

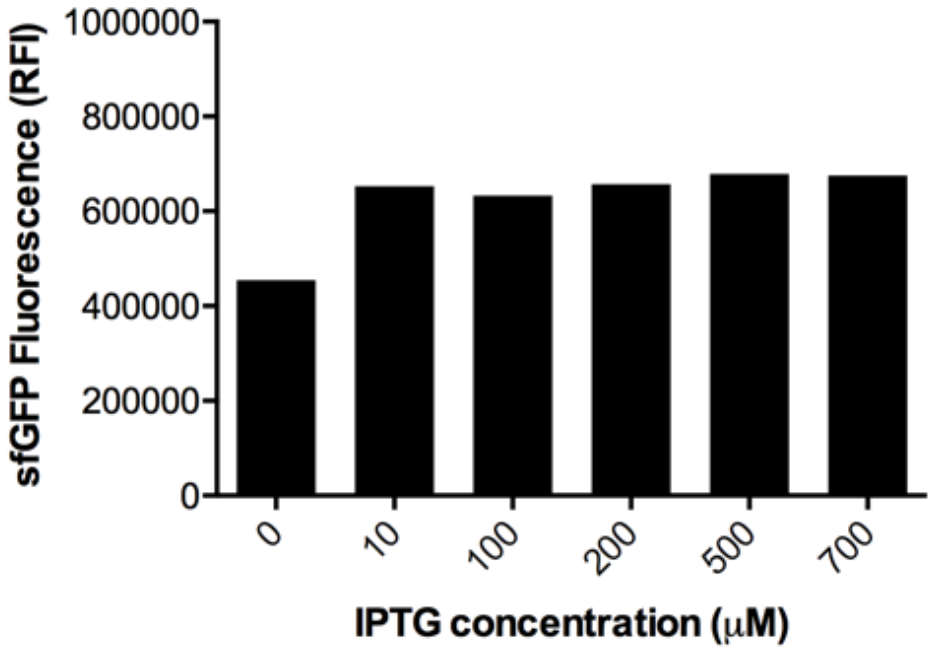


Figure 14. Regulation of the T7 expression system in BL21(DE3). BL21(DE3) was transformed with pHis17-sfGFP expression vector. About two hundred thousand cells were counted and analysed for sfGFP relative fluorescence intensity (RFI). Cells were induced at $OD_{600\ nm}$=0.4 with increasing concentration of IPTG (from 0 up to 700 μM). The sfGFP RFI was recorded the next day after overnight induction.

2.1 Comparison of sfGFP fluorescence in BL21(DE3) derivatives

The production of sfGFP protein was analysed by flow cytometry, in all existing BL21(DE3) derivatives (Figure 15). In BL21(DE3), the highest RFI was achieved without any addition of IPTG (Figure 15, A) as previously reported (Zhang et al., 2015). The IPTG toxicity was reflected by a large number of cells (80% of the bacterial population) that had lost the sfGFP fluorescence (Table 6 and Figure 15, A solid line).

Figure 16 shows that a better regulation of sfGFP production was already achievable by employing the strain BL21(DE3)pLysS compared to BL21(DE3). The use of lysozyme inhibits the activity of the T7 RNAP providing a means to tune the T7 regulation system upon induction (Moffatt and Studier, 1987; Wagner et al., 2008). However, in the host mutant C41(DE3) the basal level of sfGFP fluorescence was decreased 3.5-fold compared to BL21(DE3) and addition of IPTG led to an homogeneous population of cells (93%) exhibiting the highest level of sfGFP RFI (Table 6 and Figure 15, B solid line).

The sfGFP RFI mean of C43(DE3) strain in the absence of IPTG amounted to 36,615 +/- 346 (Table 6) represented by two bacterial cell populations. Addition of IPTG

had the beneficial effect to induce a single fluorescent population showing that all cells kept the ability to produce sfGFP (**Figure 15**, C).

C44(DE3) and C45(DE3) most remarkable feature was their extremely low levels of sfGFP RFI in the absence of IPTG (**Figure 15**, D and E). 89% of C44(DE3) cell population exhibited a sfGFP RFI of 359, which is only about twice the bacterial intrinsic fluorescence (**Table 6**). Consequently, the induction fold change of sfGFP RFI estimated by flow cytometry increased up to 33 times in C45(DE3) and 48 times in C44(DE3) after overnight induction with IPTG (**Figure 16**).

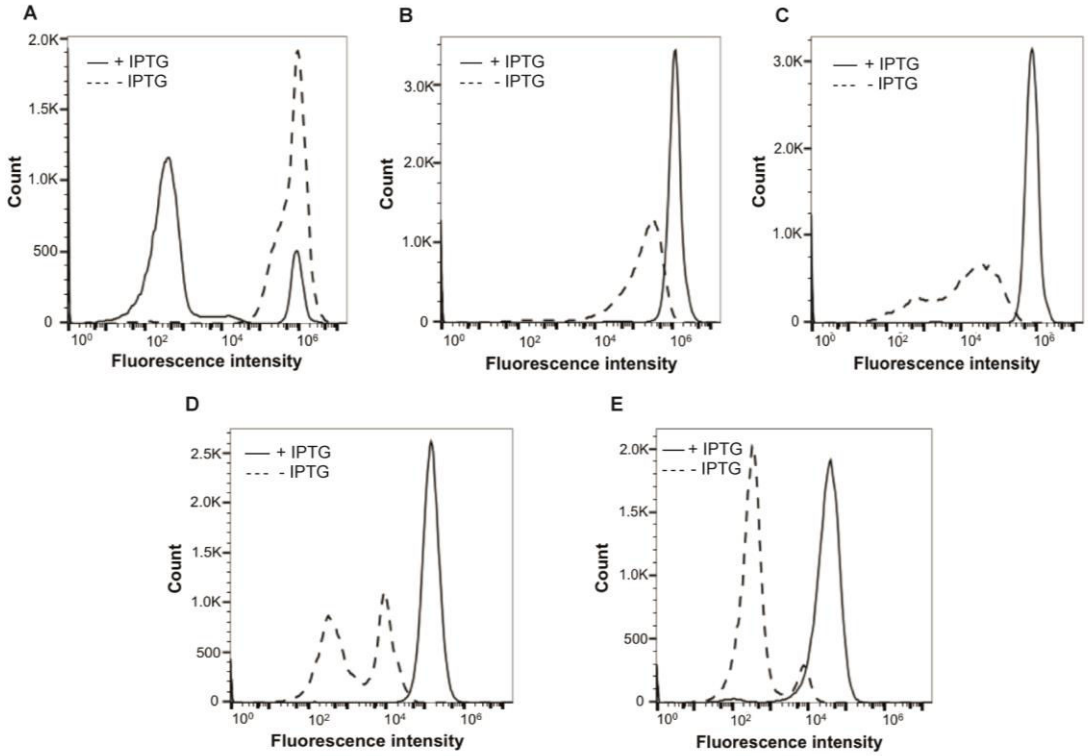


Figure 15. Flow cytometry analysis of an overnight culture of BL21(DE3) **(A)** C41(DE3) **(B)** C43(DE3) **(C)** C44(DE3) **(D)** and C45(DE3) **(E)** transformed with pHis17-sfGFP vector. Dashed line indicates RFI of non-induced cells while continuous line shows the RFI of cells induced at $OD_{600\,nm}$=0.4 with 700 μM of IPTG.

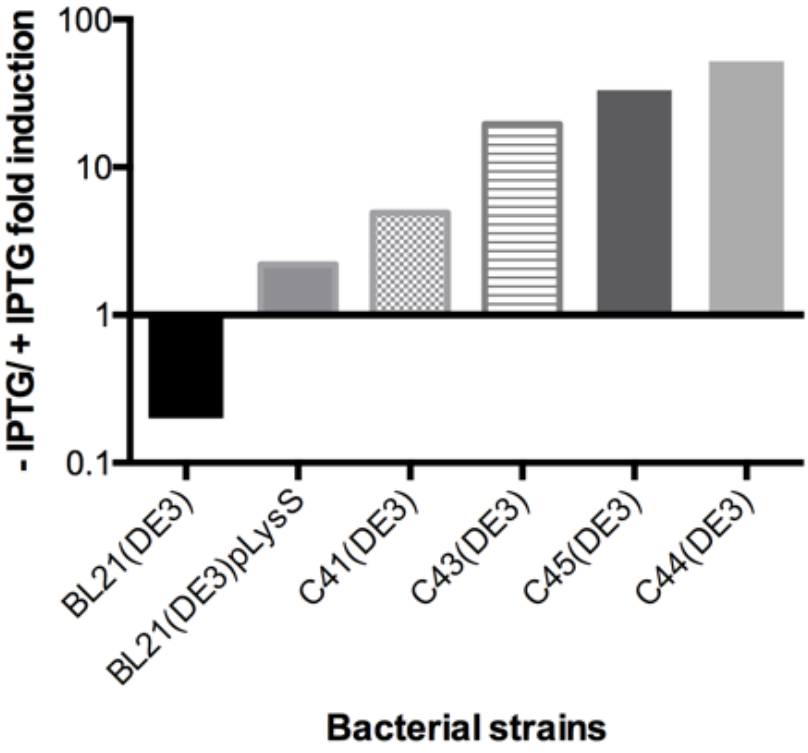


Figure 16. RFI mean value ratio of induced over non-induced cells.

Bacterial host	Induction	Total mean RFI (SEM[1])	RFI mean (% population)	
			Low[2]	High[3]
BL21(DE3)	+IPTG	184,728 (29,173)	869 (81)	965,333 (19)
BL21(DE3)	-IPTG	788,656 (35,374)	368 (5)	832,333 (95)
C41(DE3)	+IPTG	1,102,447 (24,083)	5,153 (7)	1,186,667 (93)
C41(DE3)	-IPTG	223,500 (10,838)	434 (6)	238,935 (94)
C43(DE3)	+IPTG	713,016 (27,837)	980 (6)	754,666 (94)
C43(DE3)	-IPTG	36,615 (346)	792 (32)	53,394 (68)
C44(DE3)	+IPTG	60,988 (23,242)	130 (3)	73,979 (97)
C44(DE3)	-IPTG	1,268 (52)	359 (89)	7,688 (11)
C45(DE3)	+IPTG	157,115 (5,239)	262 (2)	159,361 (98)
C45(DE3)	-IPTG	4,717 (385)	468 (60)	11,082 (40)

Table 6. Analysis of the sfGFP fluorescent intensity in expression hosts. [1] SEM, standard error of the mean of three independent experiments. [2] Low RFI population. [3] High RFI population.

3 Analyses of MPs production in C44(DE3) and C45(DE3)

C44(DE3) and C45(DE3) mutant hosts were challenged with two collections of expression plasmids encoding MP fused with GFP from *E. coli* and other bacterial species (Materials and Methods, Table 1). First, the toxicity associated with the overproduction of the MP-GFP fusion was assessed on solid medium. For each construct, cells were diluted 10^7-fold in order to have less than 100 colonies on each plate and were plated on 2xYT medium with or without IPTG.

The surface area (SA) of colonies was measured from photographic pictures using ImageJ software. The relative colony surface area (RCSA) upon induction was calculated by using the following formula SA(+IPTG)/SA(-IPTG)*100 and data can be found in Table 7. In the BL21(DE3) host, all expression vectors, including the soluble protein sfGFP, prevented colony formation on 2xYT plates supplemented with 0.7 mM IPTG (RCSA=0, Table 2). In contrast, almost none of the MP-GFP fusions prevented colony formation on IPTG containing plates in either C44(DE3) or C45(DE3) with the exception of Ph1-GFP and GS21-GFP (Table 7).

It should be noted that in the case of PheP-GFP, YqcE-GFP, GltP-GFP and YfbF-GFP fusions, which facilitate phenylalanine, glutamate/aspartate and Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transport, addition of IPTG increased the size of colonies in C45(DE3). This suggests that production of these permeases increased nutrient uptake, which could explain the positive effect on cell growth. No positive effect on colony SA was observed with non *E. coli* MP-GFP fusions cloned in the pGFPe expression vector. The RCSA values were generally lower in C45(DE3) for most of the MP-GFP fusions as well as for the control plasmid pHis17-sfGFP encoding sfGFP.

Bacterial host (growth medium)	**BL21(DE3) (LB)**	**C44(DE3) (2xYT)**	**C45(DE3) (2xYT)**
Expression plasmids	**RCSA=SA(+IPTG)/SA(-IPTG) (%)**		
sfGFP	0	116	40
tehA - GFP	0	67	127
yijD - GFP	0	82	53
pheP - GFP	0	102	200

ygfU - GFP	0	105	57
yqcE - GFP	0	62	156
yidC - GFP	0	77	43
gltP - GFP	0	90	143
yfbF - GFP	0	143	200
Ct6 - GFP	0	81	30
Dr10 - GFP	0	81	70
Dr35 - GFP	0	71	32
Dv1 - GFP	0	41	40
Ph1 - GFP	0	32	0
Dv12 - GFP	0	85	67
Dv3 - GFP	0	49	42
Sp17 - GFP	0	40	38
Tt14 - GFP	0	57	34
Oo3 - GFP	0	49	45
Ss21 - GFP	0	58	11
Gs21 - GFP	0	16	0
Bacterial host	**C43(DE3)**	**C44(DE3)**	**C45(DE3)**
h-SQR	120	126	59

Table 7. Toxicity analysis of the expression plasmids.

3.1 *E. coli* MPs production in C44(DE3) and C45(DE3)

GFP fluorescence was then measured upon induction of the T7 RNAP in liquid cultures. *E. coli* MPs were chosen based on an already published paper (Zhang et al., 2015). In order to confirm the observations already made by Zhang and co-workers, all the proteins were over-expressed in BL21(DE3) either in LB medium without addition of IPTG at 30°C or in 2xYT with addition of IPTG at 37°C, condition frequently used in the laboratory (Figure 17). Given that half of the proteins were better expressed in the first condition and half in the second one, we decided to keep the conditions previously identified for optimal production in BL21(DE3) (Zhang et al., 2015).

While, the strains C44(DE3) and C45(DE3) were grown in 2xYT medium at 37°C with addition of IPTG, following the best growth condition identified for C41(DE3) and C43(DE3) hosts (Walker and Miroux, 1999) and used with sfGFP.

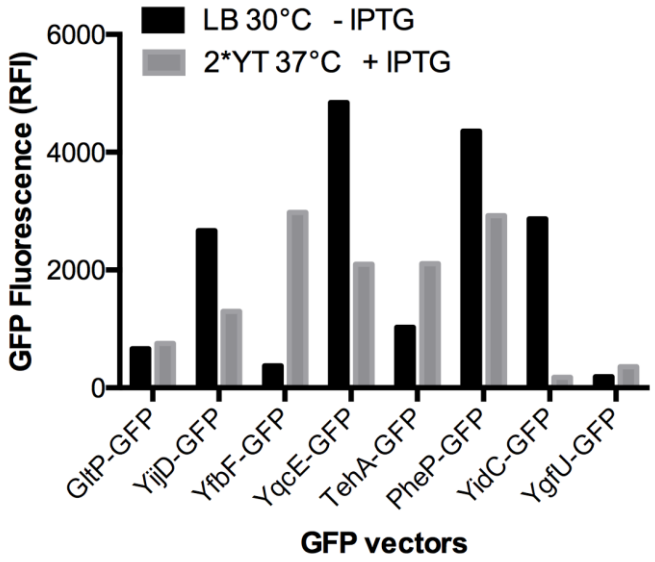


Figure 17. Different culture conditions of BL21(DE3) over-expressing *E. coli* MPs. BL21(DE3) growth was tested on LB at 30°C without addition of IPTG and 2xYT at 37°C after IPTG induction.

The IPTG concentration was the last condition evaluated (Figure 18). Indeed, even if Zhang and co-workers assessed 0.4 mM of IPTG as the optimal induction condition, we observed that 0.7 mM of IPTG is actually the better one for C44(DE3).

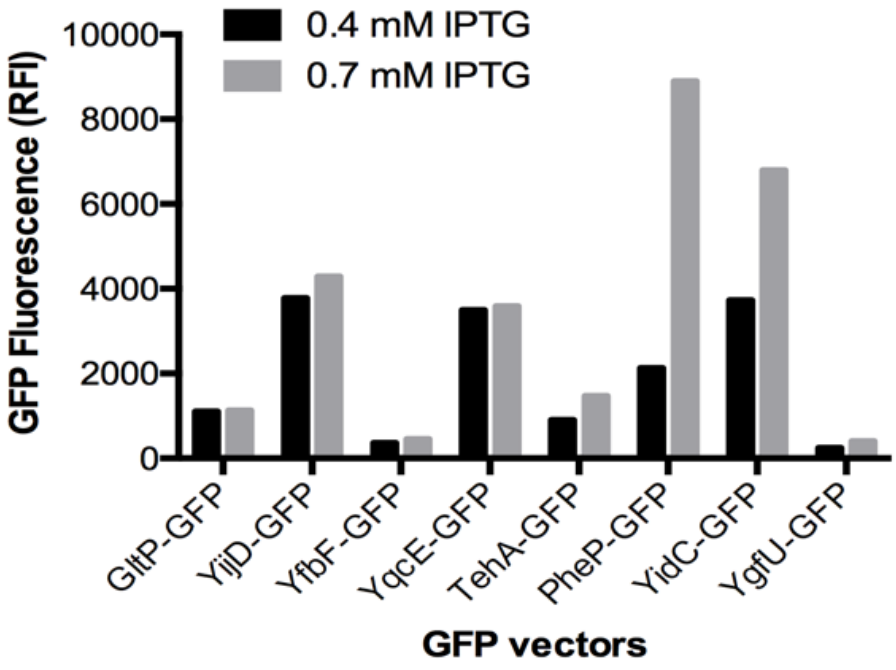


Figure 18. Overnight cultures of C44(DE3) harbouring the eight *E. coli* MP under consideration. The cultures were induced using either 0.4 mM of IPTG or 0.7 mM of IPTG.

By employing the conditions thus determined, eight *E. coli* MPs were expressed in BL21(DE3) and in its derivatives. As shown in Figure 19, four proteins YijD, PheP, YidC and GltP were produced at significantly higher levels in either C44(DE3) or C45(DE3) bacterial hosts, while three constructs (YgfU, YfbF and YqcE) remained better produced in the BL21(DE3) host. Interestingly, the protein YidC, a inner MP required for the insertion and folding of integral membrane proteins was 32-fold better expressed in C45(DE3) compared to both BL21(DE3) and C44(DE3).

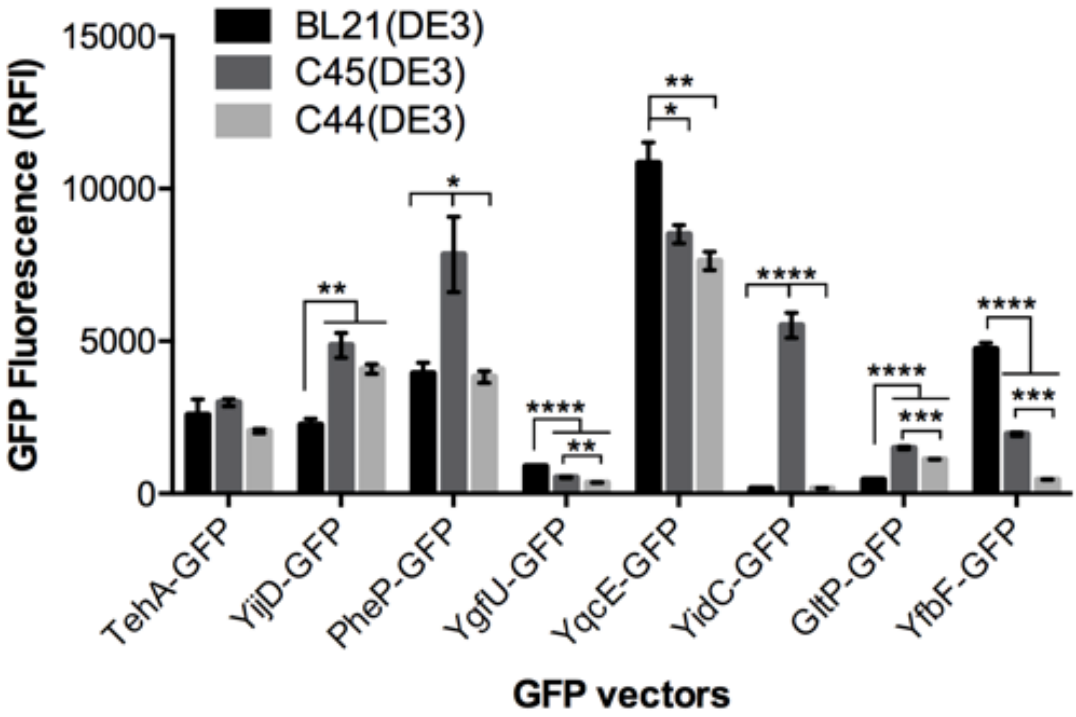


Figure 19. Triplicate cultures producing *E. coli* MP-GFP fusions. BL21(DE3) cells were grown overnight in LB medium at 30°C. All other BL21(DE3) derivatives were grown in in 2xYT medium, induced at OD_{600} nm=0.4 with 0.7 mM IPTG and further grown overnight at 37°C. GFP RFI was recorded by flow cytometry.

3.2 Non-*E. coli* MP production in C44(DE3) and C45(DE3)

To strengthen our results, twelve non-*E. coli* prokaryotic MPs fused to GFP (Hammon et al., 2009), were produced in BL21(DE3) and its derivatives, including C41(DE3) and C43(DE3) (Figure 20). We used the same culture conditions previously employed for the *E. coli* MPs.

Almost all constructs were produced at significantly higher levels in either C44(DE3) or C45(DE3) bacterial hosts. The best production was obtained for the protein

Dv12 in C44(DE3) which expression was about 6-fold higher than in BL21(DE3), C41(DE3) and C43(DE3). Among non-*E. coli* proteins, the only example where BL21(DE3) turned out to be an optimal host is in the case of the protein Ss21.

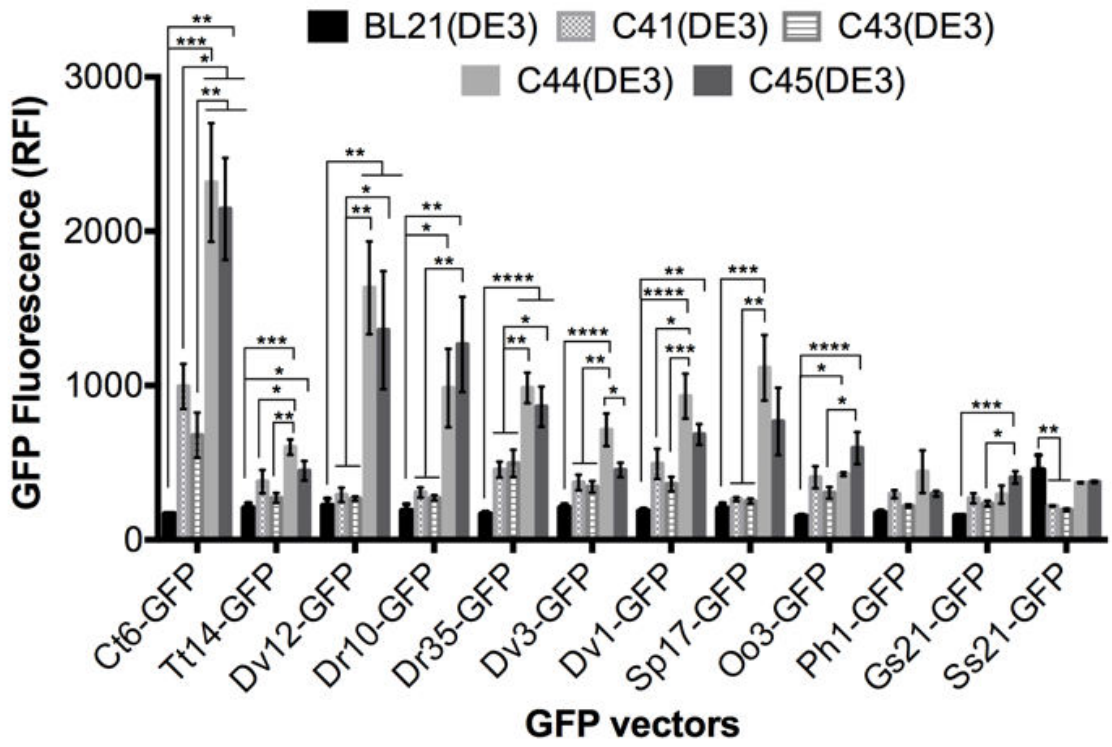


Figure 20. Six-time replica cultures producing non-*E. coli* MPs fused with GFP. BL21(DE3) cells were grown overnight in LB medium at 30°C. All other BL21(DE3) derivatives were grown in in 2xYT medium, induced at $OD_{600\ nm}$=0.4 with 0.7 mM IPTG and further grown overnight at 37°C. GFP RFI was recorded as described by flow cytometry.

3.3 Fluorescence-detection size-exclusion chromatography (F-SEC)

It has been demonstrated that GFP fluorescence analyses can be misleading due to GFP cleavage from the MP. On the whole cell GFP RFI does not adequately reflect the exact fraction of MP extracted from bacterial (García-Fruitós et al., 2005) or yeast (Sarramegna et al., 2002) membranes. To quantify precisely the amount of proteins correctly folded into the bacterial membranes, we used the F-SEC method developed by Eric Gouaux (Kawate and Gouaux, 2006) (Figure 21).

The MPs YijD and YidC were produced in C44(DE3) and C45(DE3), respectively. Membranes were prepared from 200 ml of culture and solubilized in presence of DDM,

a mild detergent widely used in membrane protein chemistry. 100 µl of solubilised material were subjected to F-SEC analysis as described in Material and Methods.

F-SEC analysis was normalized either per volume of culture (**Figure 21**, A, B, D, E) or per milligram of MP (**Figure 21**, C, F). **Figure 21** (A-C) show the total absorbance of proteins at 280 nm, whereas GFP RFIs are displayed in **Figure 21** (D-F). YijD fluorescence was 34-fold increased compared to BL21(DE3) when extracted from 200 ml of C44(DE3) culture (**Figure 21**, D). While, YidC fluorescence increased 105-fold in C45(DE3) solubilised membranes (**Figure 21**, E) compared to BL21(DE3).

In order to validate our analysis, MPs from both bacterial hosts were solubilised with the same detergent-protein ratio of 5 (w/w), we used a fixed amount of MP (18 mg with 1% DDM in 9 mL final volume). The F-SEC analysis showed that the fusion protein level increased 20-fold in C45(DE3) bacterial membranes (**Figure 21**, F). Since the total amount of MP obtained from a one-litter culture of C45(DE3) cells was four times higher than in BL21(DE3) (126 mg in C45(DE3) versus 32 mg in BL21(DE3)), the yield per litter of YidC-GFP culture increased 70-fold, a fold change close to the one measured per volume of culture in panel E.

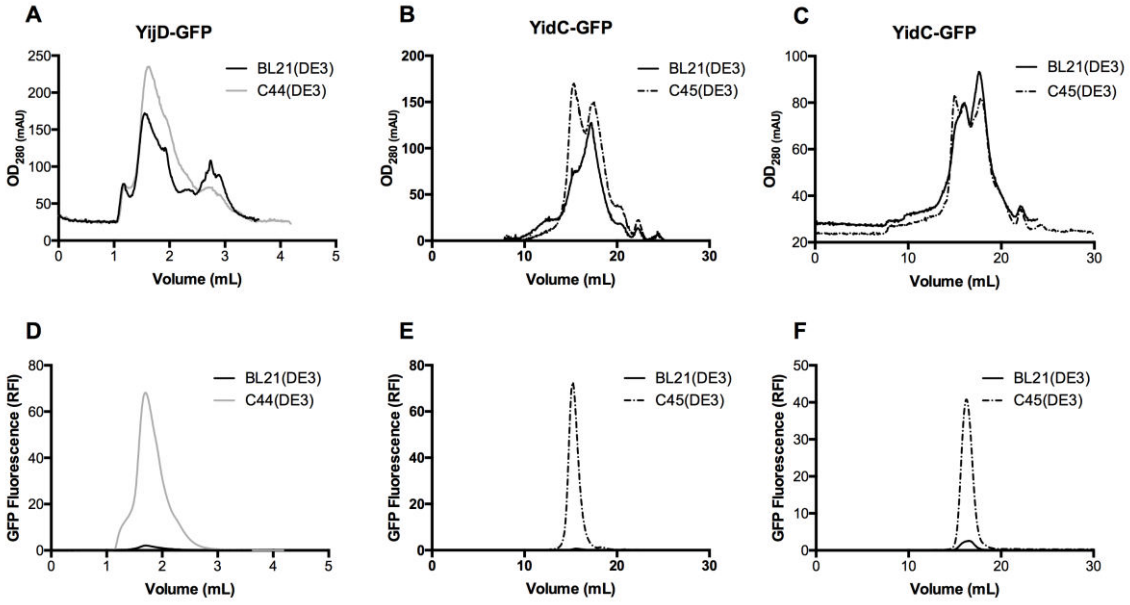


Figure 21. F-SEC analysis of YijD and YidC proteins. BL21(DE3), C44(DE3) and C45(DE3) host strains were transformed either with yijD-GFP **(A, D)** or yidC-GFP **(B and C, E and F)** expression vectors. Cell pellets from 200 ml culture were solubilized with DDM (1% final concentration) and 100 µl of solubilised material were injected on a Superdex-200 5/150GL column for GFP-yijD or on Superose-6 10/300GL column for GFP-yidC. 18 mg yidC-GFP were characterised by F-SEC after solubilisation with 1% DDM at a 2 mg/mL protein concentration **(C, F)**. Absorbance at 280nm **(A-C)** was recorded while GFP fluorescence was analysed using a Jasco detector **(D-F)**.

3.4 Human-SQR production in C44(DE3) and C45(DE3)

Following the positive results obtained after MPs production in C44(DE3) and C45(DE3) mutants, we wondered if these strains could be helpful to express eukaryotic proteins. To corroborate this hypothesis, we tested the human sulphide quinone oxydo-reductase (h-SQR). h-SQR is a monotopic MP that has been previously produced in *E. coli* at moderate levels (250 µg/l of culture) by employing the companion plasmid pCPN10/60, which produces cold-adapted chaperon proteins from *Oleispira antarctica* (Jackson et al., 2012).

The same synthetic gene used by Jackson and co-workers encoding for the h-SQR was introduced in pHis17 and transformed either in BL21(DE3) or in BL21(DE3) derivatives. The recombinant protein was produced as inclusion bodies in BL21(DE3), C41(DE3) and C43(DE3) but was readily targeted to the membrane of C44(DE3) and C45(DE3) (Figure 22, A).

To confirm that the protein was correctly folded in the bacterial membranes, h-SQR was purified (Figure 22, B) from a one-litter culture (600 µg of purified protein) and the activity was tested using Na2S as a substrate. The resulting activity curves (Figure 22, C) were similar to those previously published (Jackson et al., 2012), suggesting that the protein is correctly folded, can bind to its cofactor FAD and is able to oxidize Na_2S.

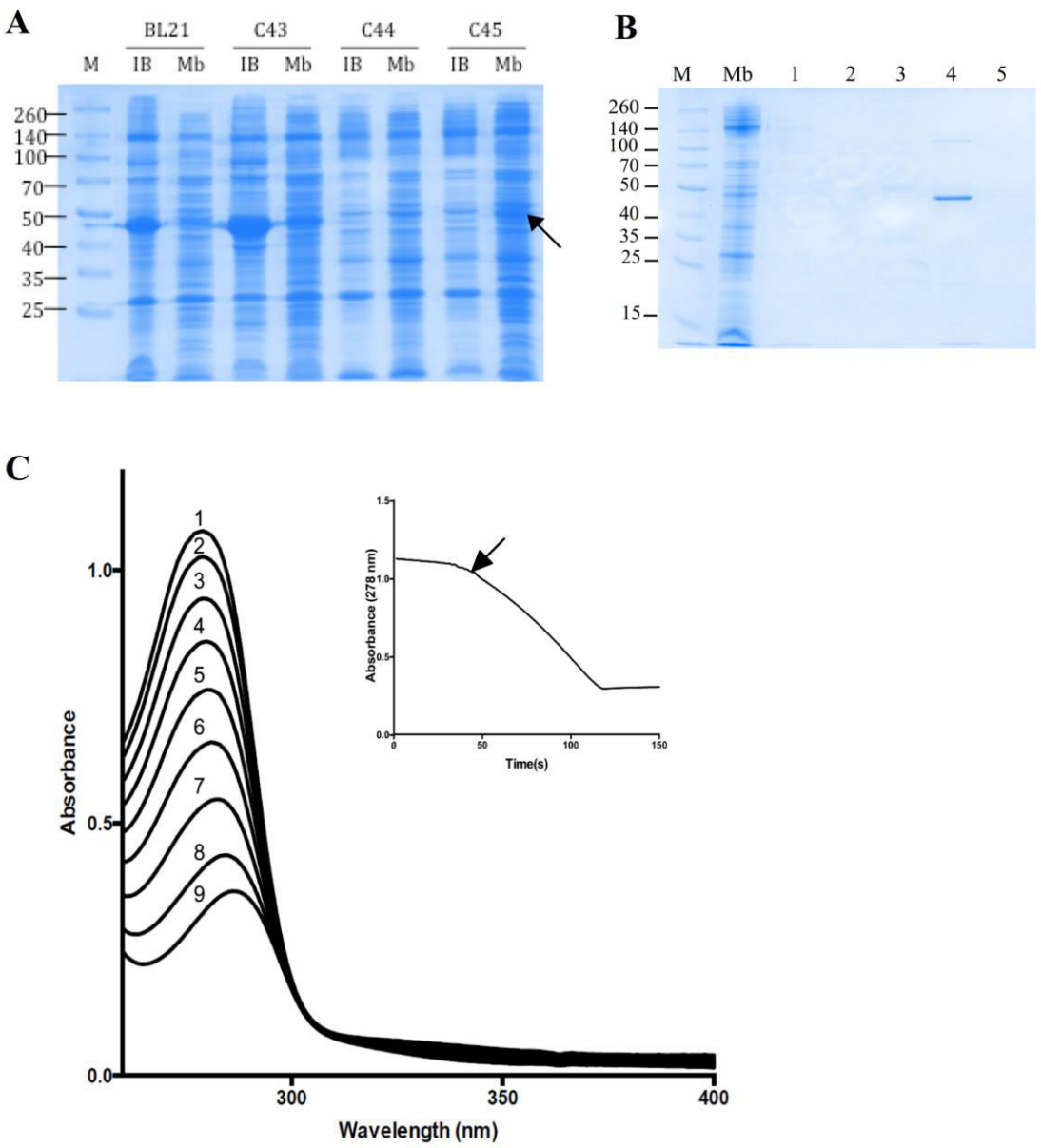


Figure 22. Production of h-SQR in C44(DE3) and C45(DE3) host strains. (A) SDS-PAGE analysis of h-SQR inclusion bodies formation in BL21(DE3) and BL21(DE3) derivatives. 30 µg of inclusion bodies (IB) or membrane fraction (Mb) proteins were loaded in a 12% acrylamide gel. The gel was strained with Coomassie Blue solution. Lane M: molecular weight marker. **(B)** SDS-PAGE analysis of h-SQR purification. Membrane fraction (Mb), column washes with 20 mM (1) and 40 mM (2-3) Imidazole, column elution with 200 mM (4) and 500 mM (5) Imidazole. **(C)** Spectral course of SQR catalytic assays. Reactions were performed in degased 100 mM potassium phosphate, 0.5 mM EDTA pH 8 buffer containing 0.05 % DDM. The reaction mixture contained 100 µM CoQ1 and 400 µM of Na2S. The activity of 0.2 µg h-SQR was recorded during 600 seconds at 287 nm. Curves 1 to 9 were recorded every 10 seconds after h-SQR addition. The inset shows the time course of absorbance changes at 287 nm, black arrow shows h-SQR addition.

4 Mechanistic insights of C44(DE3) and C45(DE3) regulation

4.1 Genomes analyses

In order to understand the mechanism underlying the regulation of C44(DE3) and C45(DE3), the genomes of BL21(DE3) and its derivatives were entirely sequenced. The genomic DNA was extracted by using Wizard® Genomic DNA Purification Kit (Promega, U.S.A.) and DNA quality was checked on 0.7% agarose gel (Figure 23).

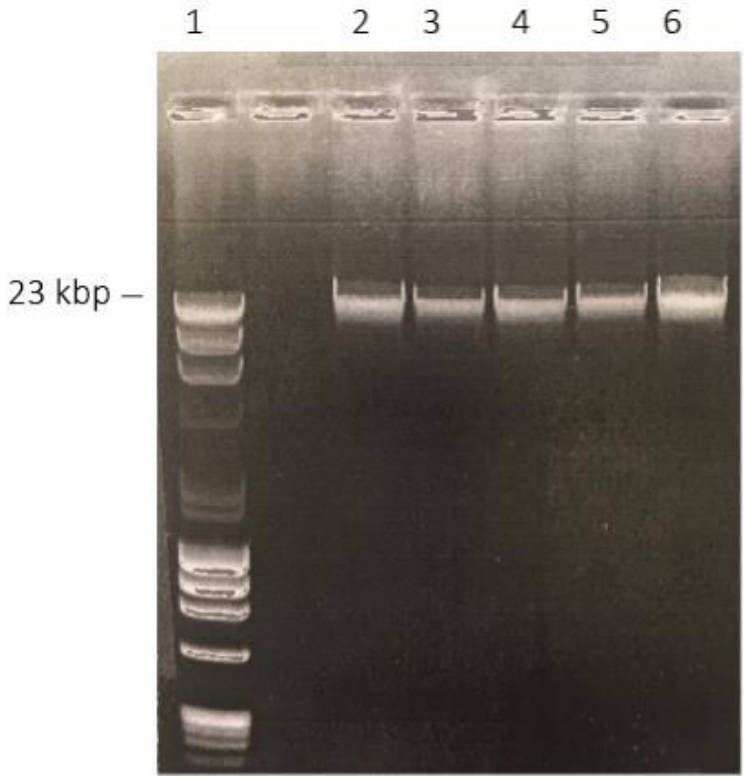


Figure 23. 0.7% agarose gel of BL21(DE3) and its derivatives genomes. Lane 1: marker; lane 2: BL21(DE3); lane 3: C41(DE3); lane 4: C43(DE3); lane 5: C44(DE3); lane 6: C45(DE3). For each lane 100 ng/µl of genome were loaded.

The genome analyses of C41(DE3) and C43(DE3) confirmed the data previously published (Kwon et al., 2015; Schlegel et al., 2015). We found the same genomic mutations in the T7 promoter that were not present in BL21(DE3) control strain (Table 8).

Gene	Function	Mutation	Position from nucleotide 1 of coding sequence	Region	Effect
Present only in C41(DE3)					
proY	Predicted cryptic proline transporter	T → A*	438	coding	Synonymous
melB	Melibiose:sodium symporter	C → T*	653	coding	Ala → Val
ycgO	Potassium/proton antiporter	C → A*	1103	coding	Gly → Val
yhhA	Hypothetical protein	C → T*	290	coding	Pro → Leu
ydcD	Hypothetical protein	G → A	71	coding	Gly → Glu
zwf	glucose-6-phosphate 1-dehydrogenase	C → A	815	coding	Ala → Asp
rpoC	DNA-directed RNA polymerase subunit beta	G → T	3023	coding	Gly → Val
Present only in C43(DE3)					
dcuS	Sensor of fumarate two component system; frame-shifted in BL21(DE3)	GCGCC deletion*	866–870	Coding	Frameshift
fur	Ferric uptake regulator	GTC insertion*	392	Coding	Val insertion
lacI	*lac* repressor	G → T*	574	Coding	Val → Phe
lon	DNA-binding ATP-dependent	IS*4* excision*	− 156 from *lon*	Intergenic	Activation of Lon protease

	protease La				
yibJ	Predicted Rhs-family protein	C → A*	90	Coding	Synonymous
yjcO	Hypothetical protein	A → G*	665	Coding	Glu → Gly
cydA	Cytochrome d terminal oxidase, subunit I	IS*1* insertion*	−262 from *cydA*	Promoter	N.D.
*ccmF ~ ompC**		IS*1* deletion*	21-kb large deletion	N.D.	
*yjiV ~ yjjN**		IS*1* deletion*	37-kb large deletion	N.D.	
Present in both C41(DE3) and C43(DE3)					
1	T7 DNA-directed RNA polymerase	AA − > GT, G − > A	P*lac*UV5 − 10 P*lac*UV5 CAP binding site	Promoter	
rbsD	Predicted cytoplasmic sugar binding protein; disrupted by insertion of IS in BL21(DE3)	IS*3* excision**	217 - 1586	Coding	N.D.
yehU	Predicted sensory kinase in two component system	T − > G	679	Coding	Phe → Val

Table 8. Genetic mutations in C41(DE3) and C43(DE3) bacterial hosts. *Mutation identified in the publication of Kwon and co-workers; **change also identified in C44(DE3) and C45(DE3).

In *E. coli* C44(DE3) and C45(DE3) as in C43(DE3) and C41(DE3) was found a mutation in the *rbsD* gene (Table 8 and Table 9). In BL21(DE3), the *rbsD* gene encoding D-ribose pyranase which converts β-D-ribopyranose to β-D-ribofuranose, is interrupted by an IS element, IS150. According to Schlegel and co-workers, the presence of an IS element in *rbsD* prevents efficient growth of BL21(DE3) on ribose as its sole carbon and

energy source. Deletion of IS150 insertion restores the wild-type *rbsD* gene in all BL21(DE3) derivatives, enabling efficient growth on ribose (Schlegel et al., 2015).

The only mutation that distinguishes C44(DE3) from the other derivatives strains consists of an amber stop codon at the position Glu359 of the T7 RNAP. Interestingly, we found that C45(DE3) also carries an ochre stop codon at position Gln656 of the T7 RNAP, which highlights a common genetic feature between both C44(DE3) and C45(DE3) mutants. However, the pattern of genetic changes in C45(DE3) bacterial mutant is more complex because of three additional deletions in *gltL*, *B-lom* and *ycdX* loci (Table 9). We can assume that the insertion in *gltL* gene can affect efficient growth of BL21(DE3) cells on glutamate/aspartate source and lost of insertion can restore the wild-type phenotype, as in the case of *rbsD*.

E. coli C45(DE3) harbours a 16 kb deletion causing the deletion of 18 genes in the prophage sequence DE3 integrated into the strain genome (Table 10).

Gene	**Function**	**Mutation**	**Gene position**	**Effect**
Present only in C44(DE3)				
1	T7 DNA-directed RNA polymerase	C → T	752,380	Gln_{656}→ TAG-*amber* codon
Present only in C45(DE3)				
1	T7 DNA-directed RNA polymerase	T→ G	751,489	Glu_{359}→ TAA *ochre* codon
gltL	Glutamate/aspartate ATP binding protein subunit disrupted by insertion in BL21(DE3)	Deletion of insJK2 insertion	642,461 -> 643,903	Restore Leu_{118} in frame with Lys_{119}
B-lom	DE3 □ (B, C, D, E, Fi, Fii, Z, U, V, G, T, H, M, L, K, I, J, lom)	Deletion	771,453 -> 788,063	from B(Leu_{49}) to lom(Phe_{206})
ycdX	Hypothetical protein	Deletion	1,101,730 -> 1,101,744	Frameshift from Phe_{100}
Present in both C44(DE3) and C45(DE3)				

rbsD	D-ribose pyranase disrupted by insertion in BL21(DE3)	Deletion of IS150 insertion	3,822,152 -> 3,824,017	Restore Val_{64} in frame with Asp_{65}

Table 9. Genetic mutations in C44(DE3) and C45(DE3) bacterial hosts.

Gene	Function
B	Capsid component
C	Capsid assembly protease
D	Head-DNA stabilization protein
E	Capsid component
Fi	DNA packaging protein
Fii	Head-tail joining protein
Z	Tail component
U	Tail component
V	Tail component
G	Tail component
T	Tail component
H	Tail component
M	Tail component
L	Tail component
K	Tail component
I	Tail component
J	Tail: host specificity protein
lom	Outer host membrane

Table 10. List of genes deleted in C45(DE3) strain with relative function.

All mutations were further confirmed by PCR analysis (Figure 24) by using primers listed in Material and Methods, Table 2.

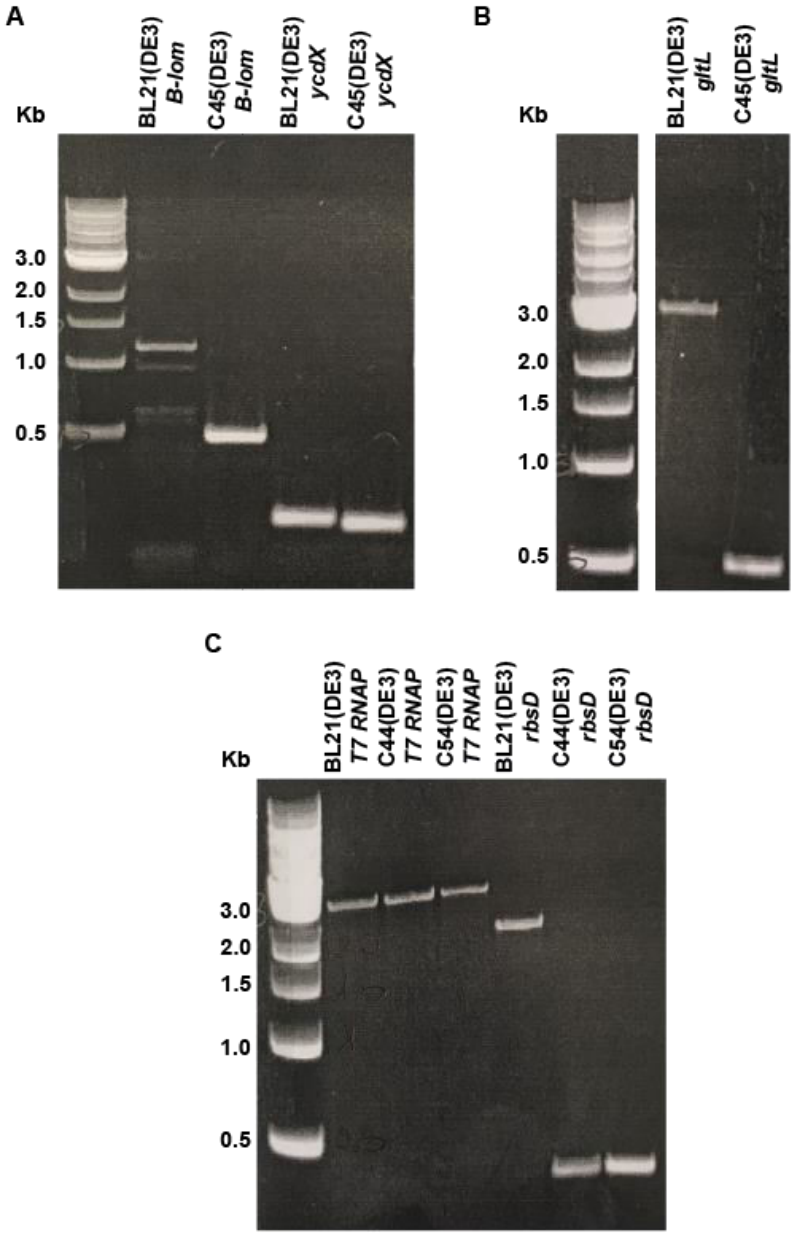


Figure 24. Confirmation of mutations in C44(DE3) and C45(DE3) bacterial hosts by PCR amplification, sequencing and agarose gel. (A) B-lom and ycdX C45(DE3) mutations compared to BL21(DE3). **(B)** gltL C45(DE3) mutation compared to BL21(DE3). **(C)** T7 RNAP and rbsD mutation in C44(DE3) and C45(DE3) compared to BL21(DE3).

4.2 RT-qPCR gene expression analysis

We hypothesized that the observed phenotypes in both mutants were due to the stop codon positioned within the T7 RNAP gene. To demonstrate this hypothesis, total RNA was extracted from BL21(DE3), C44(DE3) and C45(DE3) before induction as well as 30, 60 and 120 minutes after induction with 0.7 mM of IPTG. T7 RNAP and sfGFP

mRNA levels were then measured (Figure 25) by RT-qPCR experiments using primers listed in Material and Methods, Table 3. Results were compared with IHF gene to normalize values.

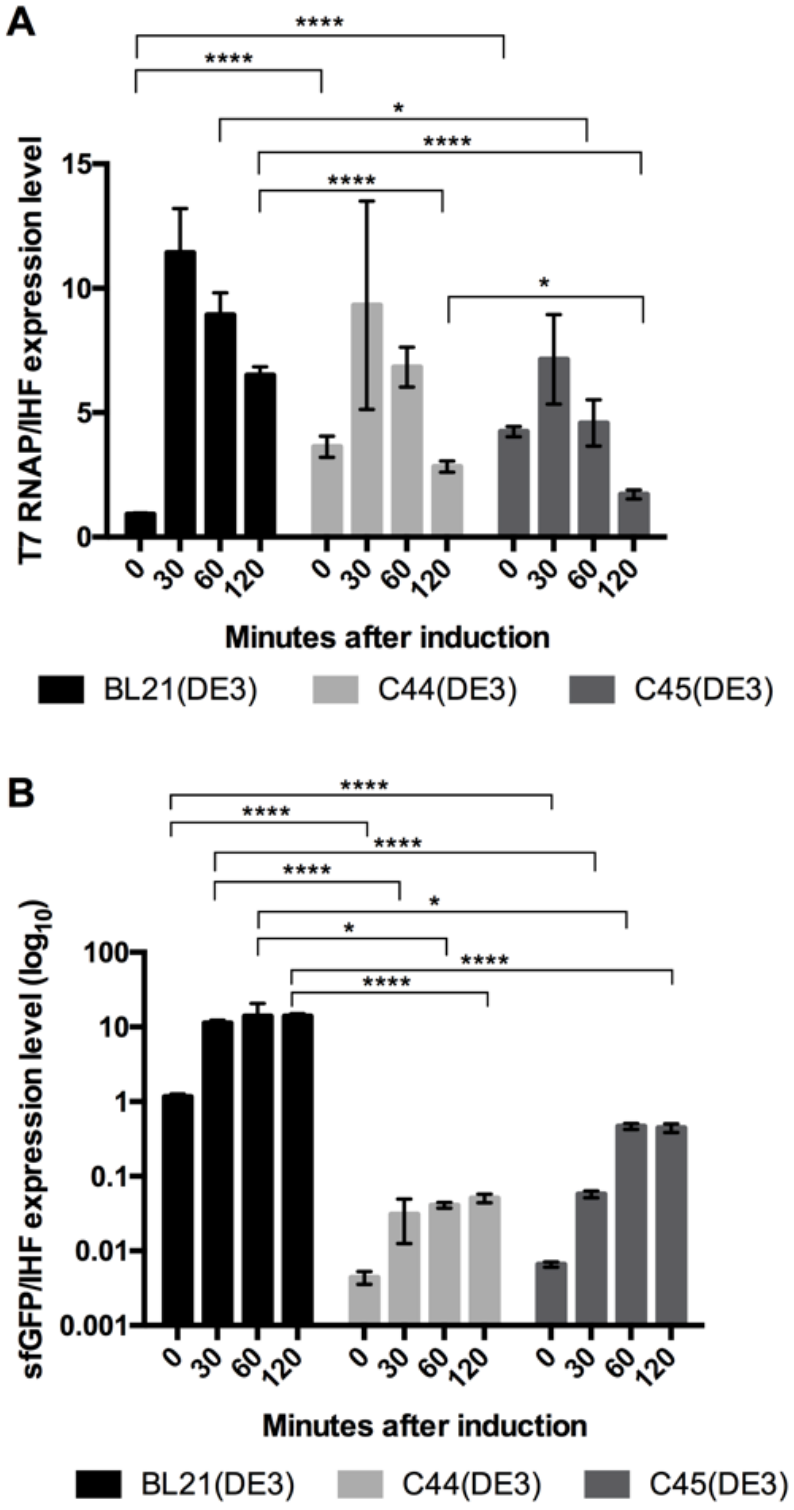


Figure 25. Quantitative PCR analysis of T7 RNA Polymerase and sfGFP mRNA levels. IHF was chosen as reference gene. **(A)** T7 RNAP/IHF expression levels **(B)** sfGFP/IHF expression levels. Three independent experiments were performed with technical triplicate. Standard error of the mean is indicated on each plot. Statistical significance was analysed using One-way ANOVA test: * p value<0.05; **p<0.01; ***p<0.001; ****p<0.0001

Figure 25 (A) shows only moderate reductions in T7 RNAP mRNA levels in C44(DE3) and C45(DE3) compared to BL21(DE3) host. For all three strains we can observe an expression peak 30 minutes after induction and a subsequent decrease of the mRNA expression level.

In contrast, the target sfGFP mRNA levels, controlled by the T7 RNAP activity, were dramatically decreased in both mutant hosts (Figure 25, B) compared to BL21(DE3) host. First, the basal levels of sfGFP mRNA, before addition of IPTG, showed a 200–fold decrease in C45(DE3) and a 300-fold decrease in C44(DE3), when compared to BL21(DE3). Second, the maximal levels of sfGFP mRNA showed a 270-fold and 30-fold decrease in C44(DE3) and C45(DE3), respectively, when compared to BL21(DE3) after addition of IPTG.

Western-blot analysis of the T7 RNAP enzyme shows that protein levels fall below the antibody detection levels (Figure 26). All together our results suggest that regulation of the T7 RNAP production does not occur at the transcriptional level but rather at the translational level.

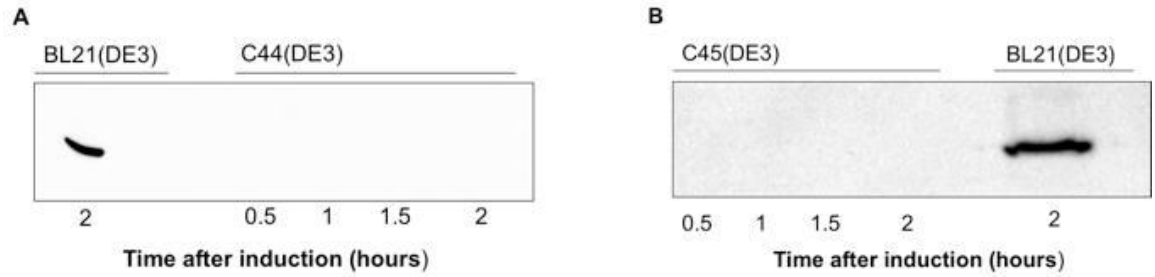


Figure 26. T7 RNA polymerase immunodetection. Cells were grown in 2xYT medium and induced with 0.7 mM IPTG. Total cell extract proteins (30 µg) were loaded on SDS-PAGE 12% gel and transferred to nitrocellulose membrane. Anti-T7 RNAP antibody (Novagen) was used at 1/3000 dilution. Peroxydase conjugated anti-mouse antibody (Sigma) was used at 1/10000 dilution and chemo-luminescence was recorded on a Chemi-Doc imaging system (Biorad). Each panel shows the time course of induction of the T7 RNAP in C44(DE3) **(A)** or C45(DE3) **(B)** host. BL21(DE3) extracted 2 hours after addition of IPTG was used as control.

4.3 C44(DE3) stop codon suppression

To demonstrate bonafide that *amber* stop codon mutation at Gln_{656} position of the T7 RNAP was responsible for the C44(DE3) phenotype, we expressed known *amber* suppressors in C44(DE3) and tested whether sfGFP expression phenotypes could be reversed to that of the parental strain, BL21(DE3). We used the pCT2Phe and pCT2His *amber* suppressor vectors (Lenfant et al., 1990), which mediate the insertion of a phenylalanine or a histidine at the *amber* stop codon, respectively. C44(DE3) cells

transformed with both sfGFP and pCT2Phe or pCT2His plasmids were unable to form colonies on IPTG containing plates (data not shown).

Flow cytometry analysis showed that in C44(DE3) cells expressing amber suppressor, sfGFP RFI reached the level of the BL21(DE3) parental host in both overnight induced and non-induced cultures (Figure 27). Both observations confirm that the C44(DE3) sfGFP expression phenotypes were caused by the amber stop codon mutation.

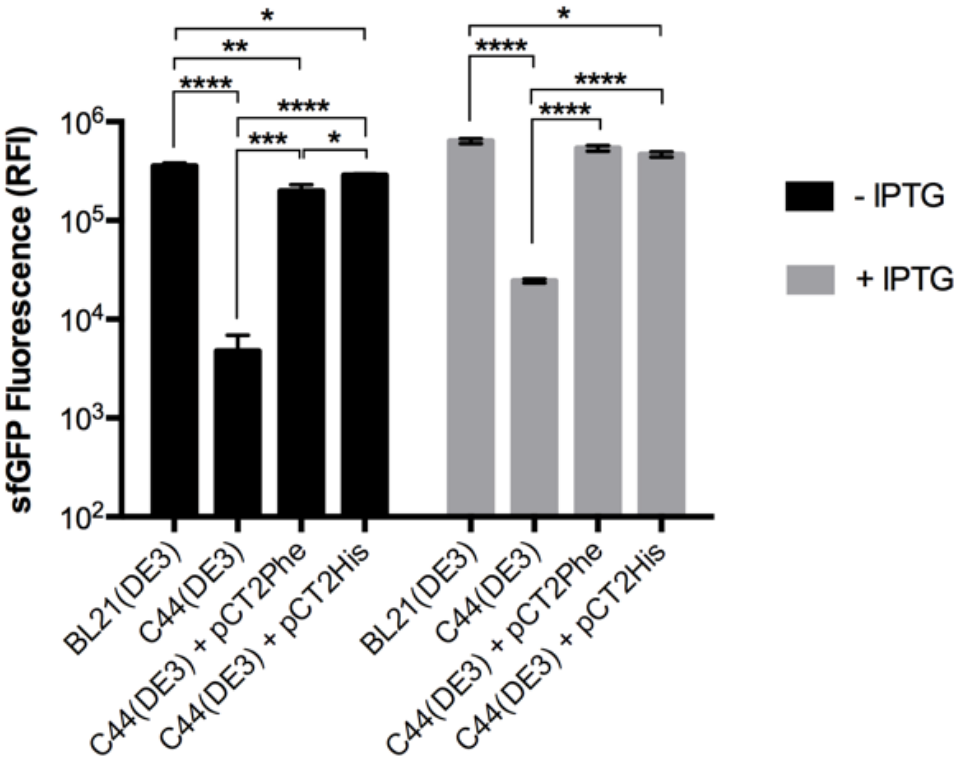


Figure 27. Amber stop codon suppression in C44(DE3). Cells were transformed either with sfGFP as control or with sfGFP and pCT2Phe or pCT2His plasmid. The graph represents sfGFP RFI analysed by flow cytometry after overnight cultures with (grey columns) or without addition of IPTG (black columns). Three independent experiments were performed. Standard error of the mean is indicated on each plot. Statistical significance was analysed using One-way ANOVA test: *p-value<0.05; **p-value<0.01; ***p-value<0.001; ****p-value<0.0001.

Chapter 2: Intracellular membranes in cells overproducing F-ATPase b subunit

In 2000 F-ATPase b subunit overexpression in the *E. coli* strain C43(DE3) was demonstrated to promote ICMs formation (Arechaga et al., 2000). The morphology of these membranes looks like those found in mitochondrial inner cristae membranes, indeed, *E. coli* ICMs are characterized by an increased content in cardiolipin (CL) (Weiner et al., 1984b; Arechaga et al., 2000).

In Chapter 2, I will describe the experiments made to study the ICMs properties from a structural level and a more "omic" point of view. The majority of results set out in the first part are published (Carranza et al., 2017) (Annexe). In this article, together with our collaborator, we revealed the importance of CL in membranes structuration by deletion of the three CL synthase genes.

In the second part, in order to understand the mechanisms that trigger ICMs formation, transcriptomic and proteomic experiments were started.

1 TEM cross-section of ICMs

Thin-layer micrographs of cells embedded into EPON resin were obtained from an overnight culture at 25°C of C43(DE3) cells overexpressing F-ATPase b subunit to confirm the presence of ICMs (Figure 28, A). Horizontal and vertical cross-sections confirmed the ICMs formation in C43(DE3) and their tubular-like structuration.

The same membranous structures were present also in C44(DE3) strain upon F-ATPase b subunit overexpression (Figure 28, B). Here, the ICMs were present in less amount that in C43(DE3) but the tubular network was comparable to the one already observed in 2000 (Arechaga et al., 2000).

In contrast, the expression of this protein in *E. coli* C45(DE3) was toxic after induction with IPTG (data not shown).

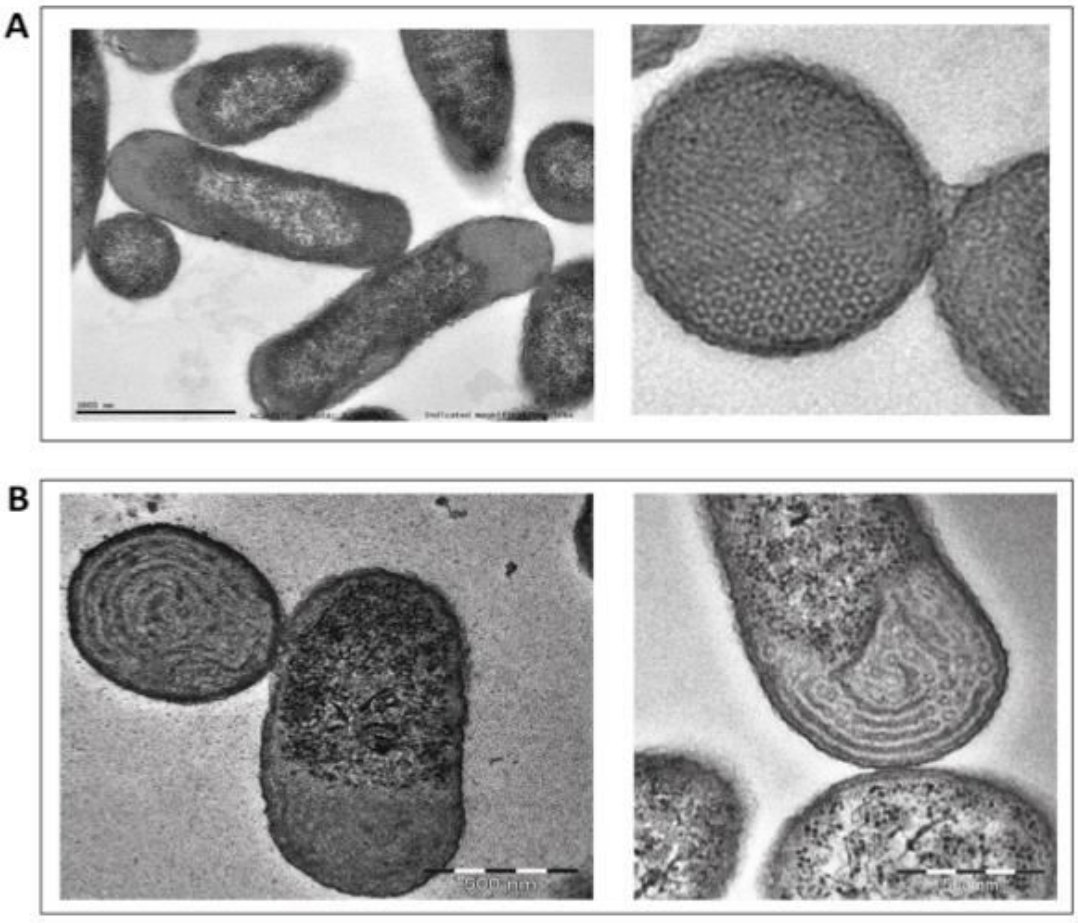


Figure 28. Cross-section of *E. coli* bacteria producing F-ATPase b subunit viewed under the transmission electron microscope (TEM). Cells were grown overnight at 25°C and embedded into EPON resin. A Leica ultracut UCT ultramicrotome was used to obtain embedded cells-cross sections. **(A)** TEM examination of C43(DE3) ICMs. Scale bar 1000 nm. **(B)** TEM examination of C44(DE3) ICMs. Scale bar 500 nm.

1.1 ICMs isolation

C43(DE3) and C44(DE3) ICMs isolation from bacterial membranes (BMs) was possible, after cells breakdown at French press, thanks to a low-speed centrifuge at 2,500 xg. Unbroken cells were removed by a washing step and the ICMs containing F-ATPase b subunit were purified with a 100,000 xg centrifuge. After purification, ICMs were detectable as one big brown and translucent pellet that was not present in control cells not expressing the F-ATPase b subunit (Figure 29, A).

The purified ICMs contained a large amount of F-ATPase b subunit compared to the outer membranes (OM) and to the control (C) (Figure 29, B).

Arechaga et al. demonstrated that membranes containing F-ATPase b subunit had the appearance of tubes or ribbons linking large vesicles. They studied the buoyant

densities on sucrose gradient finding that vesicles were denser than tubes and precipitated in the heavier fractions of gradient. We confirmed these results (Figure 30, A) and observed a comparable behaviour for the C44(DE3) ICMs (Figure 30, B), even if the amount of F-ATPase b subunit was less than in C43(DE3) light fractions.

Analysis of mutant hosts by sfGFP in flow cytometry experiments showed that C44(DE3) maximum RFI was only 5% of C43(DE3) maximum RFI; it is therefore not surprising that F-ATPase b subunit is produced at lower levels in C44(DE3).

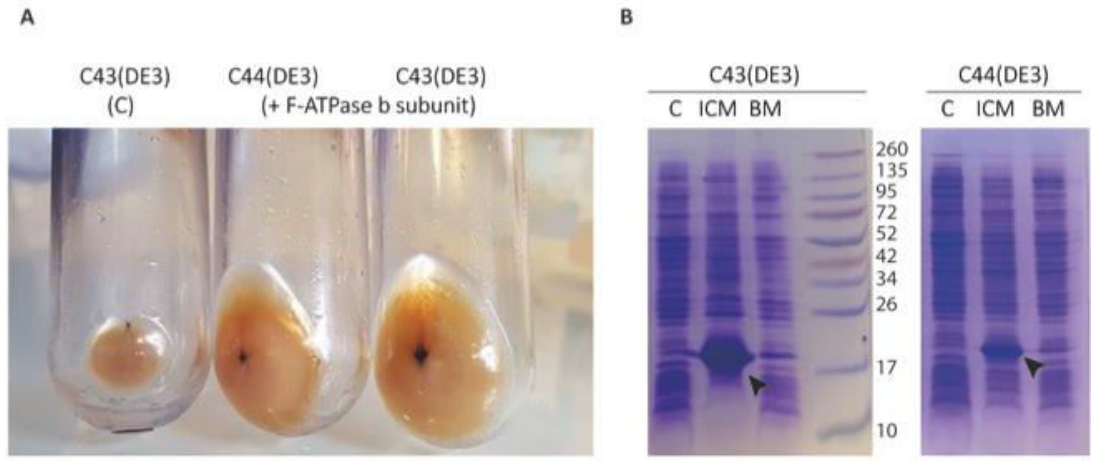


Figure 29. C43(DE3) and C44(DE3) ICMs isolation. **(A)** ICMs in C43(DE3) over-expressing the empty plasmid pHis17 (C) and in C44(DE3) and C43(DE3) over-expressing the F-ATPase b subunit (indicated by the arrow). The pellet was obtained by breaking the cells at French Press and by subjecting them to a centrifuge series. The picture was taken after the last centrifuge at 100,000 xg. **(B)** SDS-PAGE gel stained with Coomassie blue. The F-ATPase b subunit is concentrated in the ICMs membranes rather than in bacterial membranes (BM) in both C43(DE3) and C44(DE3) hosts compared to the control (C).

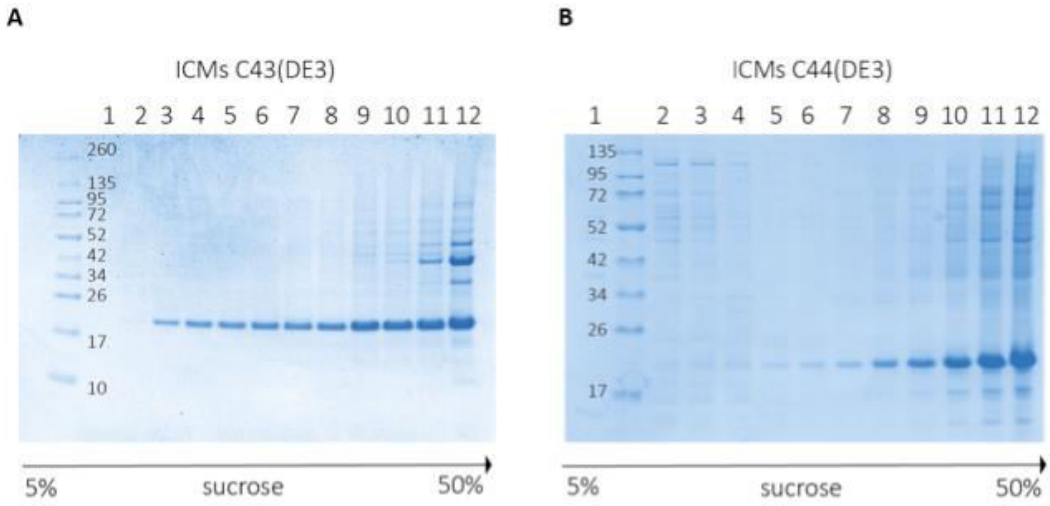


Figure 30. C43(DE3) and C44(DE3) ICMs sucrose gradients. 2 mg/mL of ICMs were deposited onto a continuous sucrose gradient and centrifuged at 100,000 xg for 18 hours. 1 mL fractions were charged on SDS-PAGE gel 14% and stained with Coomassie blue. **(A)** 14% SDS-PAGE gel of C43(DE3) ICMs sucrose gradient fractions. **(B)** 14% SDS-PAGE gel of C44(DE3) ICMs sucrose gradient fractions.

1.2 Dynamic analyses of ICMs

Although ICMs monitoring by TEM cross-sections of fixed cells has proven to be useful to study ICMs formation (Von Meyenburg et al., 1984; Weiner et al., 1984b; Wilkison et al., 1986b; Arechaga et al., 2000; Weis et al., 2003b), this procedure is not applicable for studying their dynamic properties. To achieve this goal, we generated chimeric fusion proteins of F-ATPase b subunit with blue fluorescent protein (BFP) or mCherry, already used in literature for this purpose (Kremers et al., 2007; Li and Young, 2012). However, over-expression of these fusion proteins resulted in undetectable ICMs on bacterial thin section and in accumulation of the fusion protein in the last fraction of sucrose gradient (Figure 31, A); both observations suggest that ICMs formation is strongly impaired. The same was demonstrated in Figure 31, B from the cross-section of cells over-expressing the fusion protein F-ATPase b subunit-BFP.

These results confirm that the membranes network structure is dependent on the soluble part of F-ATPase b subunit (Arechaga et al., 2000).

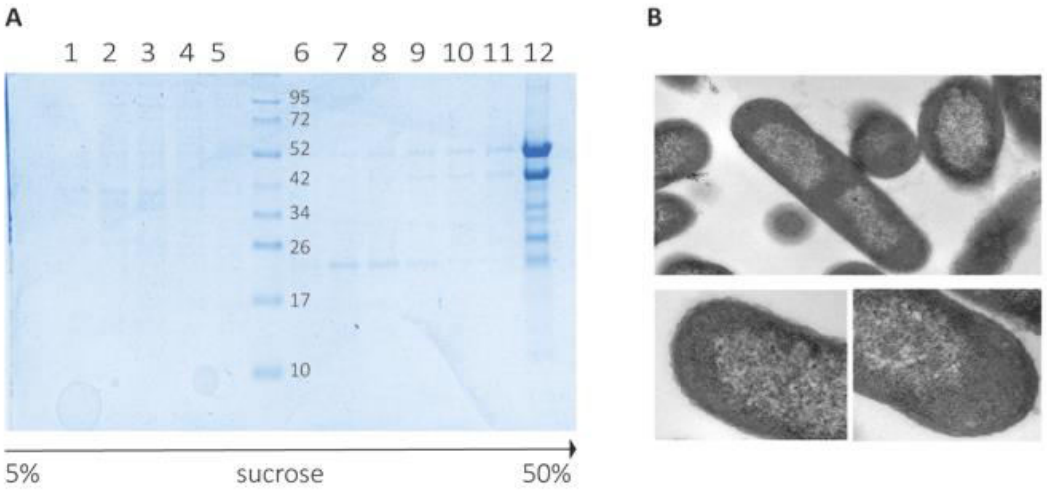


Figure 31. Chimeric fusion proteins over-expression. C43(DE3) cells were transformed with F-ATPase b subunit in fusion with BFP or mCherry. ICMs formation was verified by sucrose gradient **(A)** after membranes isolation and by TEM cross- section **(B)** of entire cells.

To overcome this limitation two fluorescent lipid-probes widely used for membranes and organelles staining were employed: 10-N-nonyl acridine orange (NAO) and Nile Red. Preliminary tests were made on cells over-expressing the F-ATPase b subunit or harbouring the control pHis17 empty plasmid by flow cytometry after NAO staining for one hour at 37°C. This probe stains preferentially anionic phospholipids (Oliver et al., 2014) and it has been used to measure cardiolipid (CL) levels in

mitochondria (PETIT et al., 1992; Gallet et al., 1995; Garcia Fernandez et al., 2004; Kaewsuya et al., 2007) and bacteria (Mileykovskaya and Dowhan, 2000, 2009).

After overnight induction 99% of the control cells showed a RFI of 4733 (dark grey, Figure 32, A). In contrast, the RFI of cells over-producing F-ATPase b subunit increased 2.5 times (light grey, Figure 32, A). Cell population analysis showed that, among the cells over-producing the F-ATPase b subunit, 52.6% were NAO negative while 47.4% of cells exhibited a RFI of 21,324 (Figure 32, A). In addition, for the latter, the size of both NAO negative and positive cells was compared. Figure 32, B shows that NAO positive cells (continuous line) are 40% bigger that NAO negative cells (dashed line).

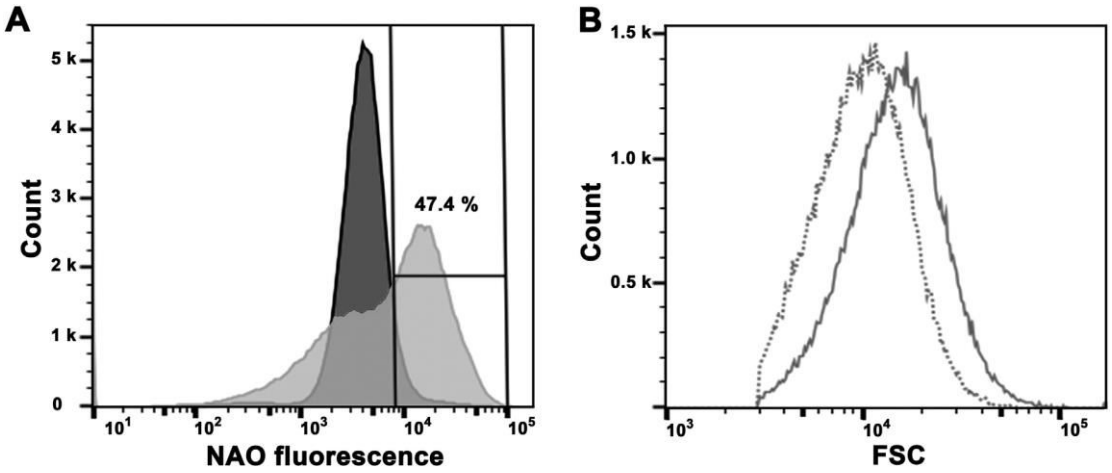


Figure 32. Flow cytometry analysis of C43(DE3) cells upon ICMs formation. Cells were grown at 37 °C and left overnight at 25 °C after addition of IPTG. After NAO staining (2mM, OD_{600} nm=0.1), cells were analysed by flow cytometry. **(A)** RFI of NAO stained cells harbouring a control empty plasmid (dark grey) or cells overproducing F-ATPase b subunit (light grey). **(B)** FSC analysis of both NAO negative (dashed line) and positive populations (continuous line) of C43(DE3) cells overproducing F-ATPase subunit b.

Next, cells were analysed by fluorescence microscopy with NAO and Nile Red which is a lipophilic stain used as membranes label (Greenspan et al., 1985; Mukherjee et al., 2007). *E. coli* C43(DE3) cells over-expressing F-ATPase b subunit stained with 2 µM of Nile Red revealed the presence of multiple internal and cytoplasmic foci, which were absent in control C43(DE3) cells harbouring pHis17 plasmid (Figure 33, A). To confirm the results obtained with Nile Red staining, C43(DE3) cells over-producing F-ATPase b subunit were also stained with 2 µM NAO. As Figure 33, B shows, NAO stain co-localizes with Nile Red, demonstrating the presence of anionic phospholipids (CL and PG) in the ICMs.

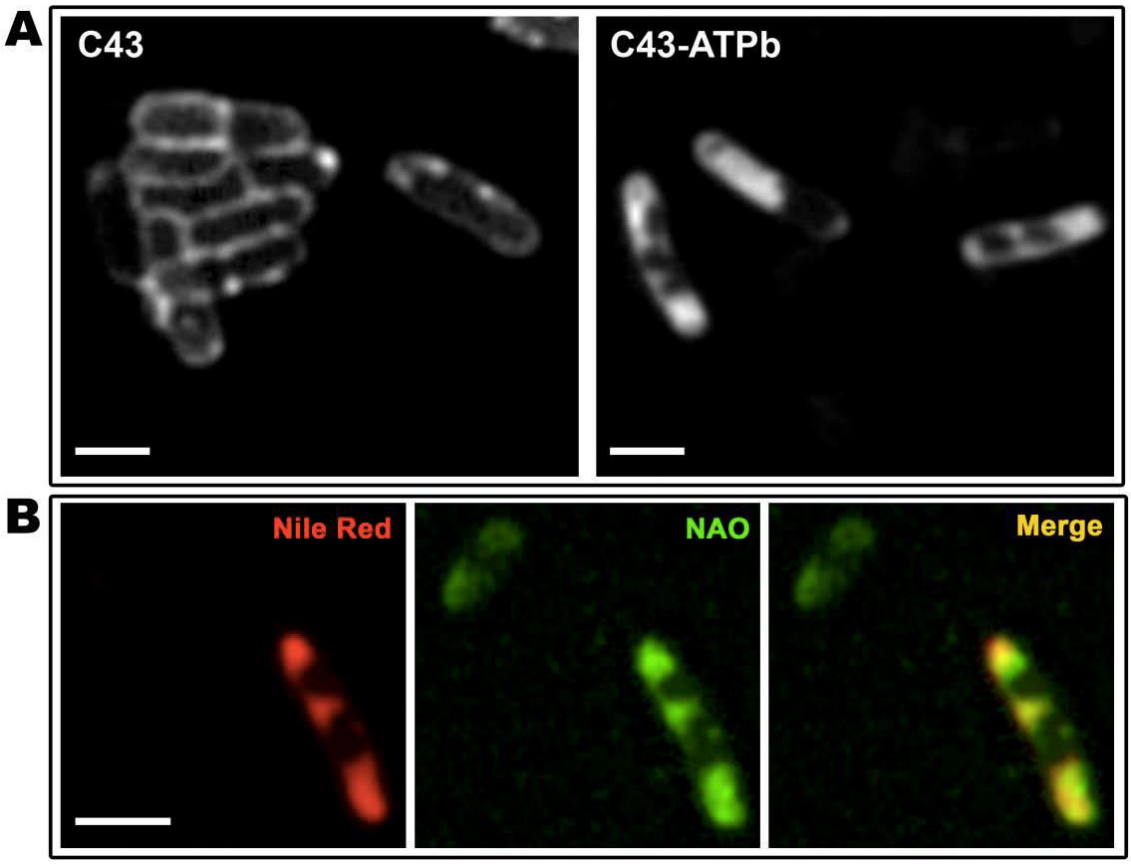


Figure 33. Membrane staining of *E. coli* C43(DE3) over-expressing F-ATPase b subunit. (A) C43(DE3) cells were stained with Nile Red (2 µM) and examined by fluorescence microscopy. Wild-type C43(DE3) cells harbouring either an empty plasmid (left panel) or producing F-ATPase b subunit (right panel). **(B)** Co-localization of Nile Red and NAO staining on C43(DE3) cells producing F-ATPase b subunit. Scale bar, 1 µm.

2 Deletion of cardiolipin synthase genes *clsABC*

In order to assess the role of cardiolipin in ICMs formation, we deleted sequentially from C43(DE3) cells the three *cls* genes responsible of CL biosynthesis in *E. coli* (Tan et al., 2012) by using the *E. coli* Keio Knockout Collection. This is a set of single-gene, knockout mutants for all nonessential genes in *E. coli* K-12 designed to create in-frame deletions upon excision of a kanamycin resistance cassette.

2.1 Lipid composition in C43(DE3) cls deletion mutants

Lipid composition in total cell membranes from each C43(DE3) mutant strain was analysed by mass-spectrometry in cells harbouring an empty pHis17 plasmid or over-producing the F-ATPase b subunit. Deletion of *clsA* gene hardly affects phospholipid composition of cells harvested at stationary phase compared to wild-type

C43(DE3) (Table 11). However, by deleting both *clsA* and *clsC* genes we observed a reduction of PE and CL levels (69% and 2%, respectively) and concomitant increase of LPE that reaches 5% of total lipid amount in cells. Triple deletion of *cls* genes results in CL absence and in PG increase (from 21% to 27% compare to wild-type C43(DE3) cells).

Strain	PE	PG	CL	LPE
C43(DE3)	75	21	4	*n.d.*
C43(DE3) *ΔclsA*	73	25	2	*n.d.*
C43(DE3) *ΔclsAC*	69	24	2	5
C43(DE3) *ΔclsABC*	68	27	*n.d.*	4

Table 11. Lipids composition in C43(DE3) control cells. Control cells harbouring the empty pHis17 plasmid were harvested and total membranes fractions were purified without separating ICMs from BMs. PE, phosphatidylethanolamine; PG, phosphatydilglycerol; Cl, cardiolipin; LPE, lysophosphatidylethanolamine. Amount of phospholipids are determined by mass-spectrometry and are given as molar percentage. n.d. not detectable.

Then, F-ATPase b subunit was over-produced in all three C43(DE3) *cls* mutants and membranes were isolated after overnight induction at 25°C (Figure 34). SDS-PAGE gel shows that the F-ATPase b subunit protein was produced in all three mutant strains compared to cells over-producing pHis17 plasmid and that F-ATPase b subunit was localised in ICMs.

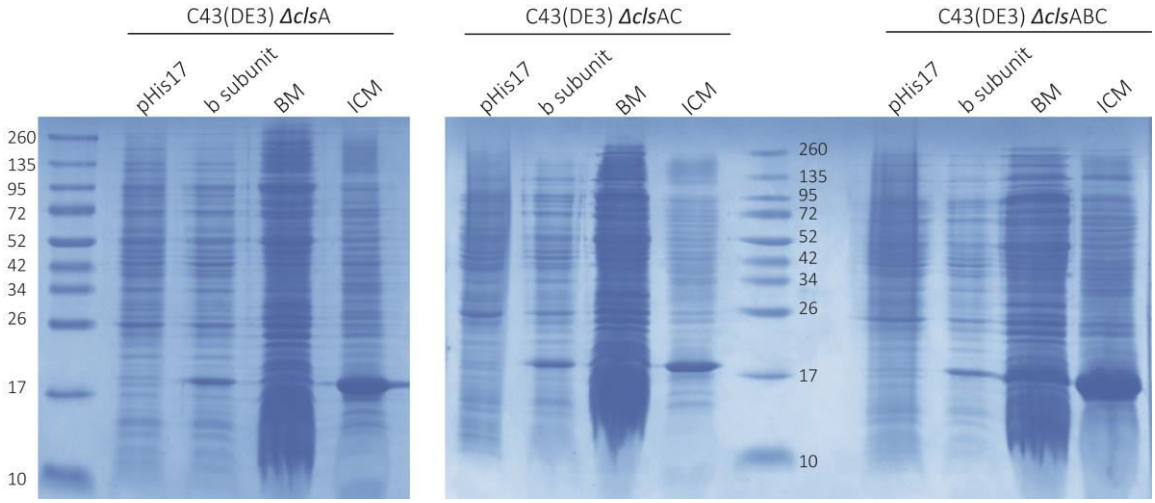


Figure 34. Over-production of F-ATPase b subunit in C43(DE3) cls mutants. Cell cultures of C43(DE3) ΔclsA, C43(DE3) ΔclsAC and C43(DE3) ΔclsABC were grown at 37°C until OD_{600}=0.4 and temperature was switched at 25°C after IPTG addition. For each mutant strain SDS-PAGE gel shows total extract of cells over-expressing an empty pHis17 plasmid and over-expressing F-ATPase b subunit. In addition, we loaded on gel 20 µl of bacterial membrane extract (BM) and 20 µl of intracellular membrane fractions (ICM).

Phospholipid composition was determined by extracting lipids from about 150mg of ICMs. In all three *cls* deletion mutants CL levels were below the detection

limit, which was associated with a large increase in PG (+307%) and reduction (-140%) of PE levels (Table 12). In contrast, ICMs isolated from C43(DE3) cells over-producing F-ATPase b subunit showed an increment of 5% of CL compared to control.

Strain	PE	PG	CL	LPE
C43(DE3)	77 ± 4	14 ± 6	9 ± 0.5	*n.d.*
C43(DE3) *ΔclsA*	50	50	*n.d.*	*n.d.*
C43(DE3) *ΔclsAC*	56	44	*n.d.*	*n.d.*
C43(DE3) *ΔclsABC*	55 ± 3	43 ± 4	*n.d.*	2

Table 12. Phospholipid composition in ICMs from cells expressing F-ATPase b subunit. C43(DE3) cells and C43(DE3) cls deletion mutants containing F-ATPase b subunit were harvested after overnight induction at 25°C. Membranes were fractioned and processed for lipid analysis. Amount of phospholipids are determined by mass-spectometry and are given as molar percentage. n.d., not detectable.

The fatty acids composition of CL in control cells and in cells over-producing F-ATPase b subunit was analysed by LC-MS (Figure 35 and Table 13). We observed that the F-ATPase b subunit over-expression did not significantly affect the fatty acids composition of CL, being palmitic (16:0) and vaccenic (18:1) acids, the most abundant fatty acids, which were similar to previously reported findings (Garrett et al., 2012).

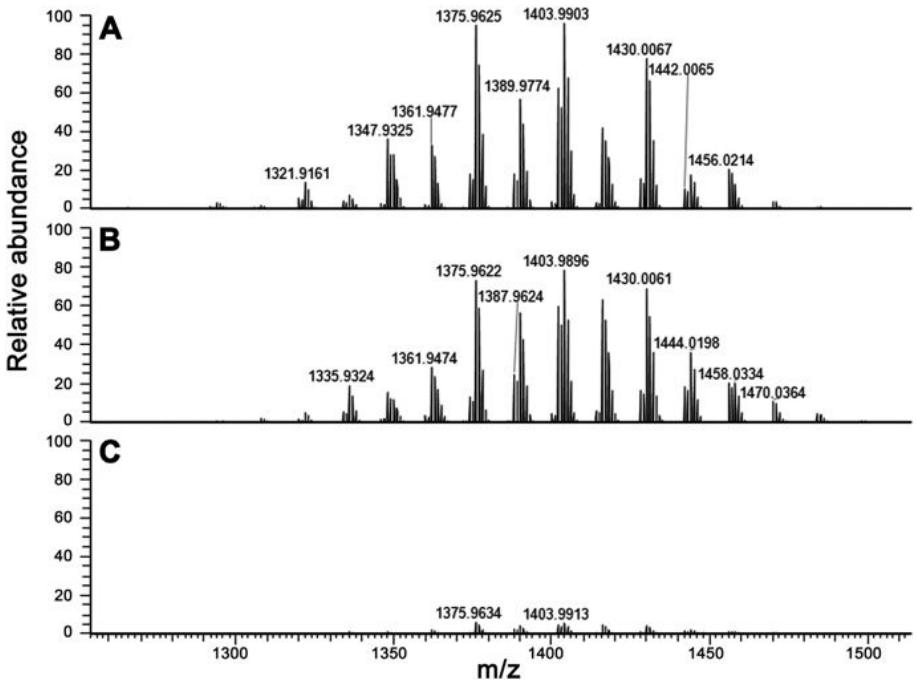


Figure 35. Identification of cardiolipin molecular species by mass spectrometry. After lipid extraction, phospholipid analysis was performed by LC-CAD-ESI/MS. Cardiolipin molecular species identified in: **(A)** C43(DE3) cells over-producing F-ATPase b subunit; **(B)** C43(DE3) control cells harbouring the empty pHis17 plasmid; **(C)** C43 (DE3) ΔclsABC triple mutant cells overexpressing F-ATPase b subunit. The identification of the fatty acids in CL compatible with these molecular masses is shown in Table 8.

Experimental mass $[M–H]^-$	Exact mass $[M–H]^-$	CL molecular species	Molecular formula $[M–H]^-$
1321.9161	1321.9180	62:1	$C_{71}H_{135}O_{17}P_2^-$
1333.9179	1333.9179	63:2	$C_{72}H_{135}O_{17}P_2^-$
1335.9327	1335.9336	63:1	$C_{71}H_{137}O_{17}P_2^-$
1347.9325	1347.9336	64:2	$C_{73}H_{137}O_{17}P_2^-$
1349.9342	1349.9493	64:1	$C_{73}H_{139}O_{17}P_2^-$
1361.9477	1361.9493	65:2	$C_{74}H_{139}O_{17}P_2^-$
1373.9683	1373.9505	66:3	$C_{75}H_{139}O_{17}P_2^-$
1375.9625	1375.9649	66:2	$C_{75}H_{141}O_{17}P_2^-$
1387.9630	1387.9649	67:3	$C_{76}H_{141}O_{17}P_2^-$
1389.9774	1389.9806	67:2	$C_{76}H_{143}O_{17}P_2^-$
1401.9775	1401.9806	68:3	$C_{77}H_{143}O_{17}P_2^-$
1403.9903	1403.9962	68:2	$C_{77}H_{145}O_{17}P_2^-$
1415.9920	1415.9962	69:3	$C_{78}H_{145}O_{17}P_2^-$
1427.9920	1427.9962	70:4	$C_{79}H_{145}O_{17}P_2^-$
1430.0067	1430.0119	70:3	$C_{79}H_{147}O_{17}P_2^-$
1442.0065	1442.0131	71:4	$C_{80}H_{147}O_{17}P_2^-$
1444.0205	1444.0275	71:3	$C_{80}H_{149}O_{17}P_2^-$
1456.0214	1456.0275	72:4	$C_{81}H_{149}O_{17}P_2^-$

Table 13. Cardiolipin molecular species identified in *E. coli* C43(DE3) upon over-production of F-ATPase b subunit.

2.2 Analyses of C43(DE3) deletion mutants ICMs

C43(DE3) *cls* deletion mutants were stained with Nile Red and analysed by optical microscopy (Figure 36, A). Staining comparison between ICMs produced in wild-type C43(DE3), single *ΔclsA* mutant and *ΔclsABC* triple mutant cells upon over-expression of F-ATPase b subunit revealed the absence of intracellular *foci* in the triple mutant. Interestingly, in the case of C43(DE3) *ΔclsA* mutant some inner foci could be observed, although in much lower amount than in wild-type. Co-localization experiment

with both NAO and Nile Red showed that PG accumulates at the poles of C43(DE3) *ΔclsABC* triple mutant cells (Figure 36, B).

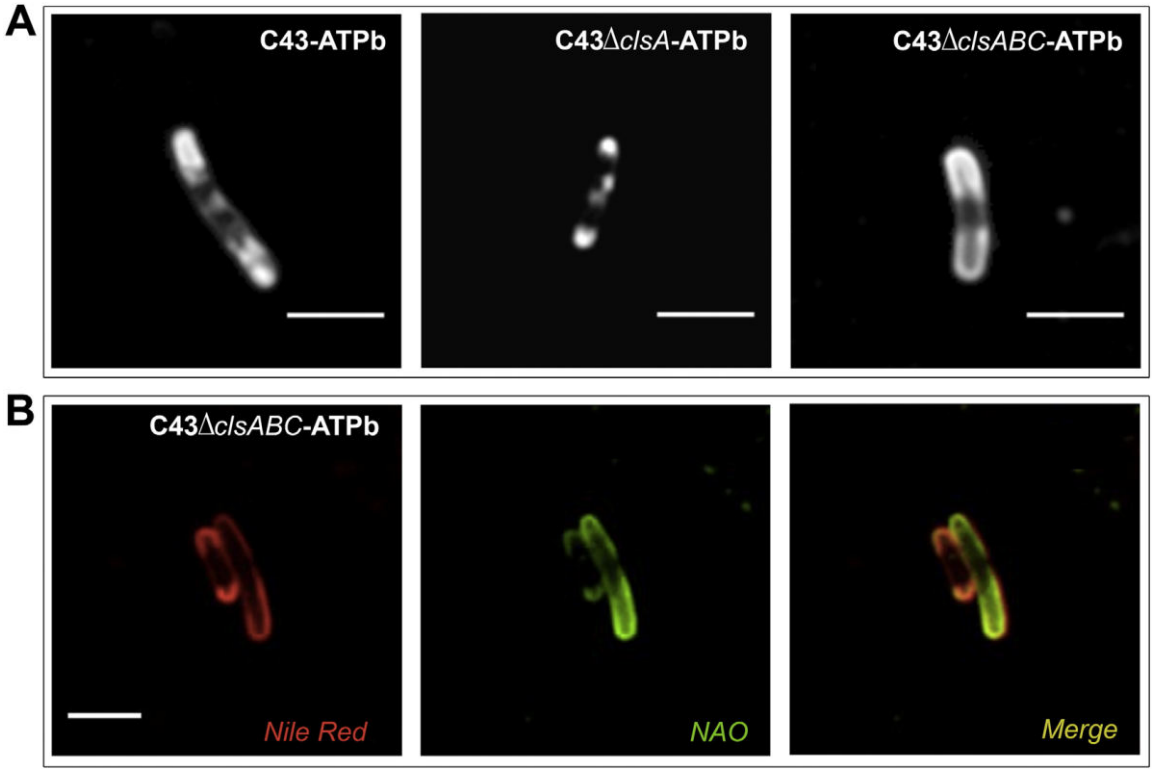


Figure 36. Membrane staining of C43(DE3) cls mutants. (A) Nile Red staining of cells over-producing F-ATPase b subunit. From left to right: wild-type C43(DE3), C43(DE3) ΔclsA and C43(DE3) ΔclsABC, respectively. **(B)** Co-localization of Nile Red and NAO staining in C43(DE3)ΔclsABC cells over-producing F-ATPase subunit b.

In order to visualize in more detail the membranes, cross sections of cells embedded in EPON resin were analysed by electron microscopy. C43(DE3) cells over-producing F-ATPase b subunit (Figure 37, A) showed the characteristic ICMs morphology already observed in 2000 (Arechaga et al., 2000).

In contrast, as observed in fluorescence experiments, the *cls* mutants ICMs morphology was affected. Analysis of C43(DE3) *ΔclsA* cells over-producing F-ATPase b subunit revealed a multi-lamellar structure within the cytoplasm of the cell (Figure 37, B). In the case of C43(DE3) *ΔclsABC* cells, we observed a similar lamellar phenotype but more severe compared to the single C43(DE3) mutant (Figure 37, C). Analysis of the double *cls* mutant C43(DE3) *ΔclsAC* revealed a morphology similar to that found in the triple mutant. Control C43(DE3) cells do not exhibit any ICMs formation (Arechaga et al., 2000).

The phenotype observed in *clsABC* mutant (**Figure 37**, C) is even more different from the one observed upon over-expression of co-expression construction between F-ATPase b subunit and BFP (**Figure 31**). Indeed, we demonstrated that CL depletion does not affect ICMs formation but they are not structured in the presence of CL. This strongly suggests that the soluble part of the F-ATPase b subunit is important for the interaction with CL.

However, internal foci disappear in optic microscopy. This could be due to the fact that F-ATPase b subunit concentration in the pole of the cell rather than in other cell part.

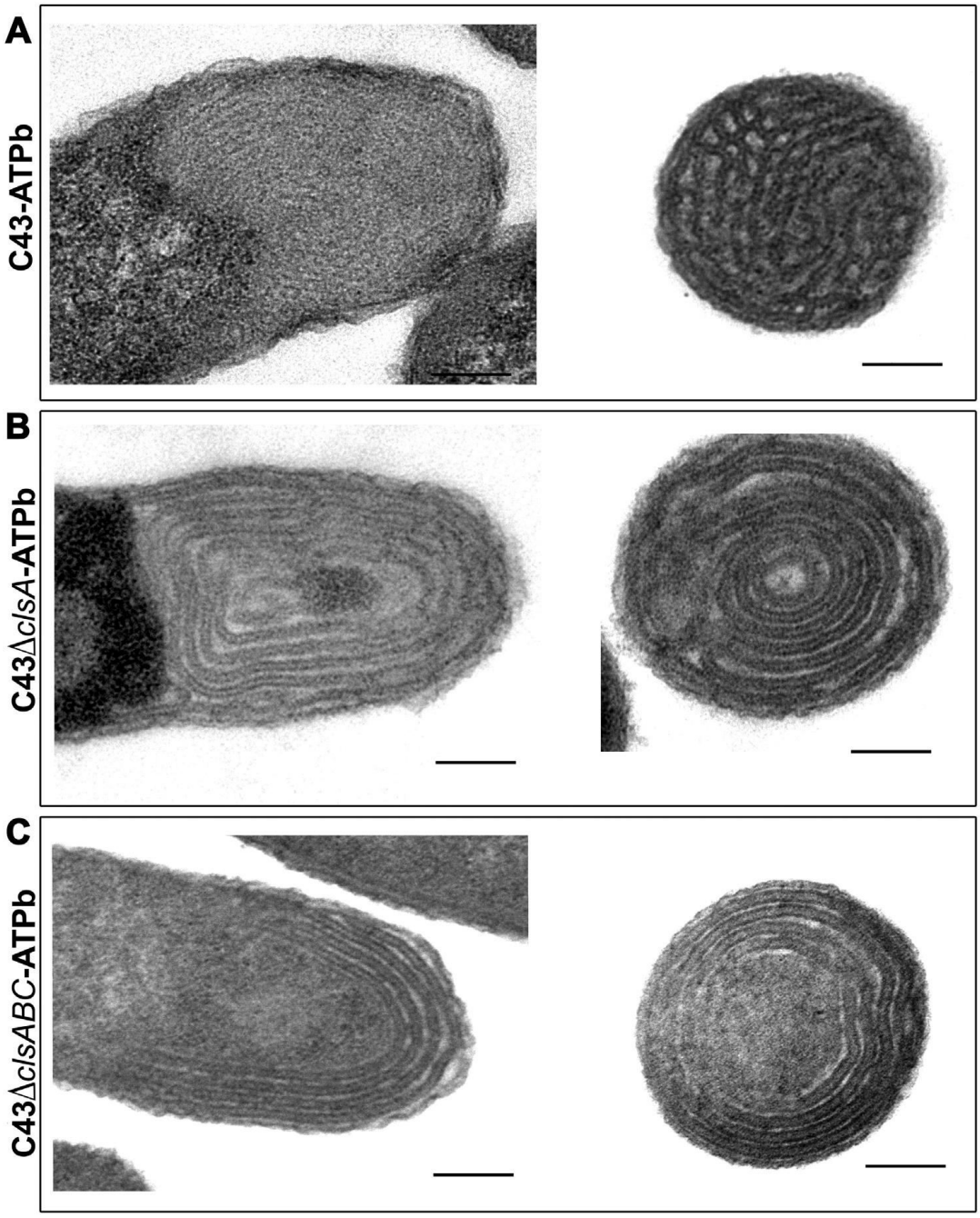


Figure 37. Electron microscopy of C43(DE3) cls mutants. (A) Electron micrographs of thin sections of C43(DE3), **(B)** C43(DE3)ΔclsA and **(C)** C43(DE3)ΔclsABC cells over-producing F-ATPase b subunit. Scale bar, 0.2 μm.

3 C43(DE3) cells RNA-sequencing analysis

To characterize the expression profile changes during ICMs formation, RNA was isolated from control C43(DE3) cells harbouring the empty pHis17 plasmid and C43(DE3) cells over-producing F-ATPase b subunit. Cells were grown aerobically on 2xYT medium at 37°C, at an $OD_{600\ nm}$=0.4 cultures were induced with 0.7 mM of IPTG and temperature was lowered to 25°C. Protein production, from the moment of induction

until 22 hours after induction, was monitored by loading each-hour total cell extract on 14% SDS-PAGE gel (Figure 38).

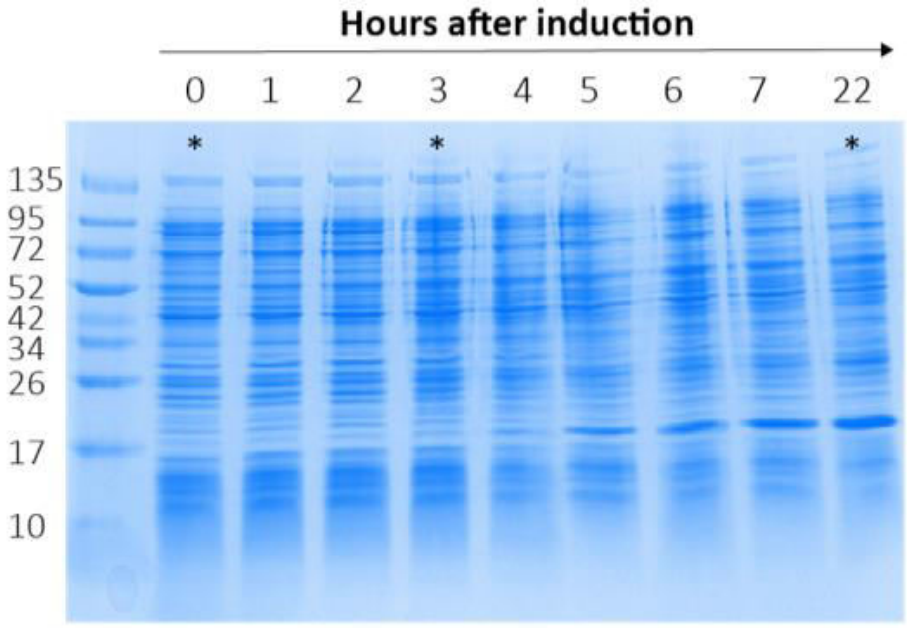


Figure 38. SDS-PAGE gel of C43(DE3) cells over-expressing F-ATPase b subunit. 1 ml of samples from time zero (moment of induction) until 22 hours after induction was taken and diluted to $OD_{600\ nm}$ equal to 0.05. Gel was stained with Coomassie blue. Asterisks indicate time point of RNA extraction.

For expression profiling, samples were taken in triplicate 5 minutes after induction, as well as 3 hours and 22 hours after induction from both C43(DE3) cells either harbouring the control pHis17 plasmid or over-expressing the F-ATPase b subunit. Cells were broken adding 700 µl of each sample to a lysis solution kept at 65°C composed by phenol, SDS and EDTA. Subsequently, RNA was extracted with repeated extraction of phenol and phenol/chloroform. To check RNA integrity, samples were charged on 0.7% agarose gel (Figure 39) and RNA integrity number (RIN) was calculated by using a Bioanalyzer (Agilent) (Table 14).

Enriched mRNA was converted to non-directional cDNA, ligated with Illumina adapters and sequenced with the Illumina HiSeq2000. The reads of C43(DE3) cells in triplicate carrying either the empty pHis17 plasmid or the pHis17-F-ATPase b subunit plasmid, were aligned to the reference BL21(DE3) genome with the Bowtie program.

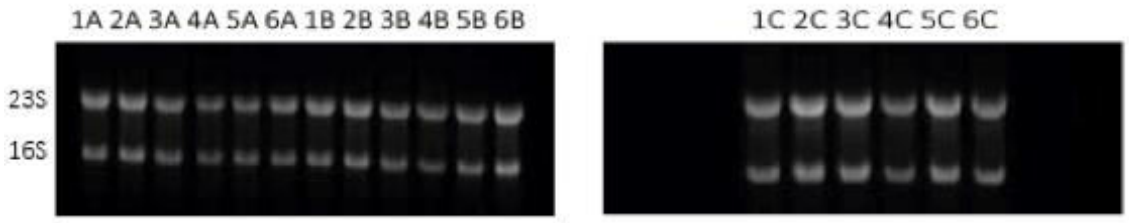


Figure 39. C43(DE3) cells RNA agarose gel. 10 µg of RNA extracted in triplicate from C43(DE3) cells harbouring either an empty plasmid (1-3) or the F-ATPase b subunit (4-6). RNA was extracted 5 minutes (1A-6A), 3 hours (1B-6B) and 22 hours (1C-6C) after IPTG induction.

Name	Plasmid	Time after IPTG induction	Concentration (µg/ul)	RIN
1A	empty pHis17	5 minutes	3,4	8,5
2A	empty pHis17	5 minutes	2,8	9,3
3A	empty pHis17	5 minutes	3,2	9,1
4A	F-ATPase b subunit	5 minutes	3,4	9,6
5A	F-ATPase b subunit	5 minutes	3,3	9,3
6A	F-ATPase b subunit	5 minutes	2,9	9,6
1B	empty pHis17	3 hours	8,7	9,6
2B	empty pHis17	3 hours	7	9,4
3B	empty pHis17	3 hours	6,2	9,3
4B	F-ATPase b subunit	3 hours	7,1	9,5
5B	F-ATPase b subunit	3 hours	6,6	9,4
6B	F-ATPase b subunit	3 hours	7,3	9,5
1C	empty pHis17	18 hours	2,4	9,1
2C	empty pHis17	18 hours	4,1	8,1
3C	empty pHis17	18 hours	4,5	8,7
4C	F-ATPase b subunit	18 hours	2,5	8,7
5C	F-ATPase b subunit	18 hours	3,3	8,8
6C	F-ATPase b subunit	18 hours	3	8,8

Table 14. RNA integratity number (RIN) values for each sample. The algorithm assigns 1 to 10 RIN score: level 3 corresponds at a RNA strongly degraded, level 5 partial degraded and level 10 intact RNA.

3.1 Expression profile overview

An initial screen of about 4000 identified genes was based on p-value. Only genes with a p-value above 0.05 were taken into account.

Fold change (FC) was the second parameter evaluated. FC describes a ratio between an initial value and a final value. In our case it is the ratio between the numbers of reads determined for a gene in the control and the number of reads for the same gene upon over-production of the protein of interest. In genomics, log ratios are often used for analysis and visualization of FC and the $\log_2$ FC is the most common used (Love et al., 2014).

Among these, genes with a $\log_2$ FC ratio > 1 were considered as over-expressed and genes with a $\log_2$ FC > -1 were evaluated as down-expressed. **Figure 40** shows that after 5 minutes of induction the main changes take place, 15% of genes were up regulated and 10% were down regulated. Meanwhile, 3 and 22 hours after induction the total of genes differentially regulated decreased at 15% and 3%, respectively. To note that, while 5 minutes after induction the up-regulated/down-regulated ratio was positive and the number of over-expressed genes were higher than under-expressed genes; 3 hours after induction the ratio is reversed and the number of genes down-regulated was double than the over-expressed genes.

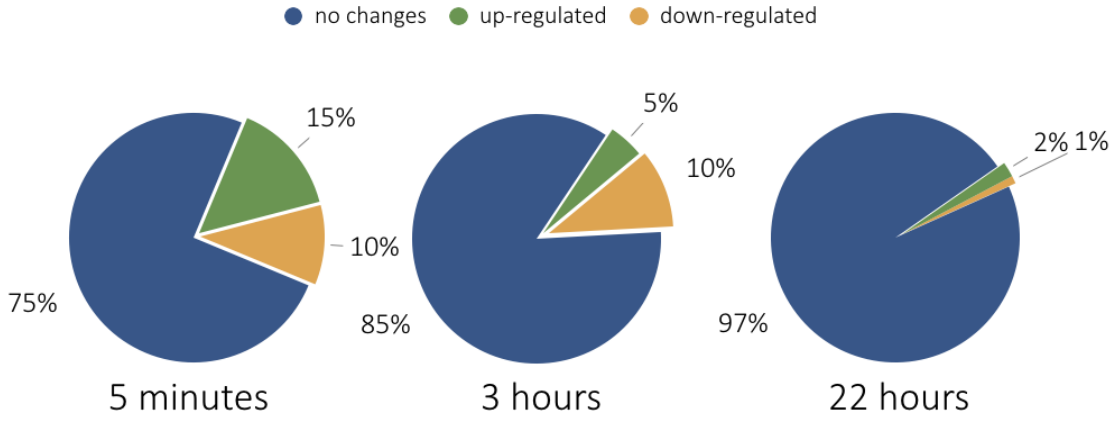
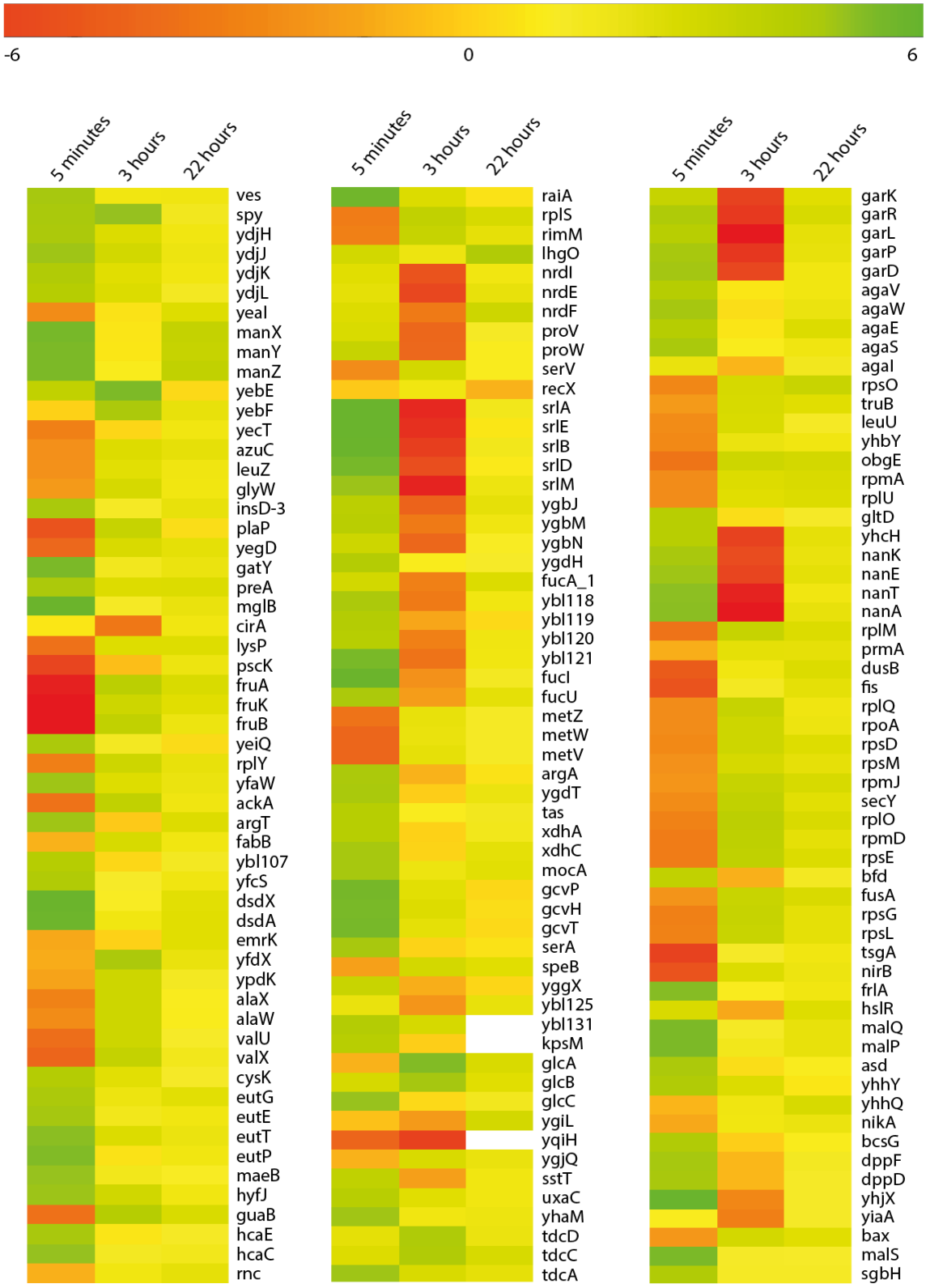


Figure 40. Graphic representation of differentially expressed genes distribution in C43(DE3). Different organization of up-regulated (green) and down-regulated (orange) genes after 5 minutes, 3 hours or 22 hours of IPTG induction. The $\log_2$ FC was used to discriminate between up- or down-regulated genes expressed as a percentage. The percentage of genes that expression do not change compared to the control are shown in blue.

Finally, 458 genes significantly up- or down-regulated in at least one condition were identified and gathered in the heatmap in **Figure 41**. Classification is based on $\log_2$ FC values, from a minimum of -6 (red) up to a maximum of 6 (green).


-6
0
6
5 minutes
3 hours
22 hours
satP
yaaI
dnaJ
nhaA
ybl3_1
rluA
sgrT
setA
ppdD
pdhR
aceE
aceF
yadV
rpsB
tsf
pyrH
fadE
yagP
paoB
paoA
yagU
ykgO
ykgM
yahE
yahN
prpR
prpB
prpD
lacA
lacY
mhpC
mhpF
mhpE
tauA
tauB
malZ
htpG
ybbC
rhsH
allA
argU
nmpC_1
borD
hokE
entD
fepA
fes
entF
fepE
entC
cstA
ybdD
rihA
glnX
glnV
metU
leuW
metT
asnB
chiQ
kdpA
ybgI
ybgJ
ybgK
ybgL
gltA
sucA
sucB
sucC
sucD
mngA
cydA
lysY
lysQ
ybgS
ybhJ
gam
kil
cIII
cII_2
S
R
D
U
G
M
K
ybiC
mcbA
dps
yliF
bssR
ylil
ybl31
artJ
poxB
serW
lrp
ycaC
ycaD
ycaO
rpsA
fabA
yccT
hyaA
hyaC
gfcA
yccE
agp
ymdF
efeU
efeO
efeB
phoH
csgC
yceK
rluC
yceD
ndh
ycfJ
ldtC
ycfZ
potB
ycgX
ycgF
ymgC
ymgF
ymgH
ymgE
tyrV
puuP
puuB
pspA
pspB
pspC
ycjM
ycjN
ycjO
ycjQ
ydaF
ydaW
rzpR
rzoR
uspF
pfo
ydcI
yncJ
ydcX
yddH
nhoA
narU
ddpB
ddpX
lsrF
dgcZ
rem
ybl81
speG
ynfE
asr
fumC
rsxA
sodB
ynhF
ydiB
aroD
ydiF
pps
ppsR
ydiE
rpmI
pfkB
ydjO


-6
0
6
5 minutes
3 hours
22 hours
ves
spy
ydjH
ydjJ
ydjK
ydjL
yeaI
manX
manY
manZ
yebE
yebF
yecT
azuC
leuZ
glyW
insD-3
plaP
yegD
gatY
preA
mglB
cirA
lysP
pscK
fruA
fruK
fruB
yeiQ
rplY
yfaW
ackA
argT
fabB
ybl107
yfcS
dsdX
dsdA
emrK
yfdX
ypdK
alaX
alaW
valU
valX
cysK
eutG
eutE
eutT
eutP
maeB
hyfJ
guaB
hcaE
hcaC
rnc
5 minutes
3 hours
22 hours
raiA
rplS
rimM
lhgO
nrdI
nrdE
nrdF
proV
proW
serV
recX
srlA
srlE
srlB
srlD
srlM
ygbJ
ygbM
ygbN
ygdH
fucA_1
ybl118
ybl119
ybl120
ybl121
fucI
fucU
metZ
metW
metV
argA
ygdT
tas
xdhA
xdhC
mocA
gcvP
gcvH
gcvT
serA
speB
yggX
ybl125
ybl131
kpsM
glcA
glcB
glcC
ygiL
yqiH
ygjQ
sstT
uxaC
yhaM
tdcD
tdcC
tdcA
5 minutes
3 hours
22 hours
garK
garR
garL
garP
garD
agaV
agaW
agaE
agaS
agaI
rpsO
truB
leuU
yhbY
obgE
rpmA
rplU
gltD
yhcH
nanK
nanE
nanT
nanA
rplM
prmA
dusB
fis
rplQ
rpoA
rpsD
rpsM
rpmJ
secY
rplO
rpmD
rpsE
bfd
fusA
rpsG
rpsL
tsgA
nirB
frlA
hslR
malQ
malP
asd
yhhY
yhhQ
nikA
bcsG
dppF
dppD
yhjX
yiaA
bax
malS
sgbH

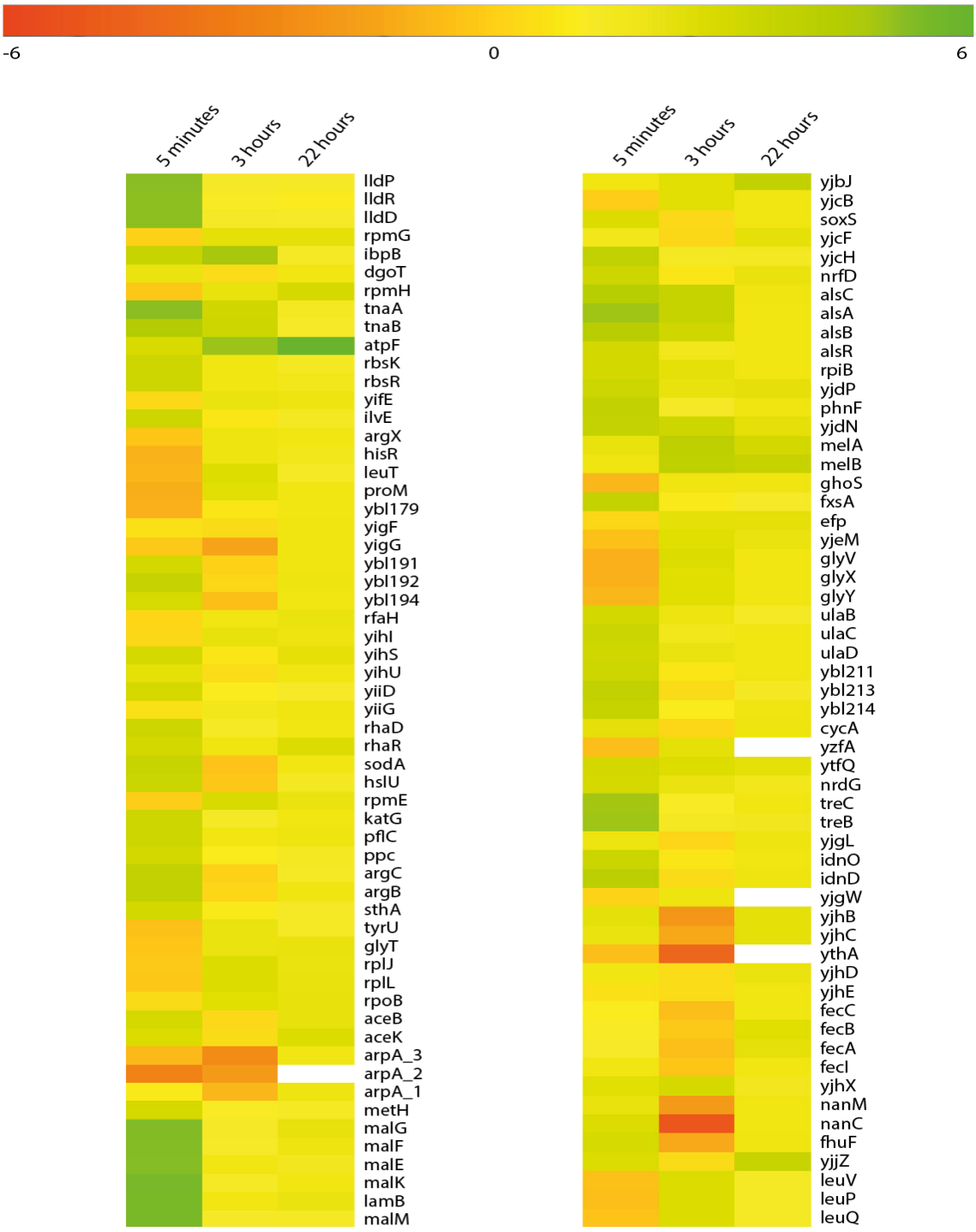


Figure 41. Heatmap comparing the expression profile through 5 minutes, 3 hours and 22 hours after induction. 458 set of genes up- (green) or down-regulated (red) significantly in at least one condition. Data based on $\log_2$ FC were gathered in a file.csv and elaborated with R to create a heatmap.

Next, differentially expressed genes involved in the same pathway were grouped in 18 categories according to KEGG annotations (Table 15). Even if genes with unknown functions were dramatically over-represented among up- and down-regulated genes

(the percentage increase over the different time point from about 20% up to 60% 22 hours after induction), genes were generally spread across functional categories throughout the experiment.

Pathway involved*	**No. of genes**					
	5 minutes		**3 hours**		**22 hours**	
	Up-regulated	Down-regulated	Up-regulated	Down-regulated	Up-regulated	Down-regulated
Carbohydrate metabolism	63	6	3	19	1	0
Membrane transport	42	16	5	22	1	0
Energy metabolism	11	3	4	0	1	0
Amino acid metabolism	29	2	3	7	0	1
Transcription & Translation	8	72	0	3	0	0
Lipid metabolism	4	2	1	2	1	0
Metabolism of cofactors and vitamins	8	1	0	1	0	0
Phage lambda	8	1	4	1	0	0
Nucleotide metabolism	7	5	2	2	0	0
Cellular community	5	1	4	0	0	0
Signal transduction	1	2	4	1	0	0
Folding, sorting and degradation	3	3	1	4	0	0
Drug Resistance	3	1	1	0	0	0
Cell motility	3	0	0	2	0	0
Replication	2	0	0	0	0	1

and repair						
Biosynthesis of secondary metabolites	2	0	0	3	0	0
Xenobiotics biodegradation and metabolism	2	0	1	1	0	0
Unknown	34	30	14	36	5	0
Total**	196	138	44	97	8	2

Table 15. Functional groups of differentially expressed genes. * genes with putative function are classified as unknown. ** some genes are involved in more than one pathway.

3.1.1 Gene expression profile 5 minutes after induction

After 5 minutes of IPTG induction the most over-expressed pathway are those of carbohydrate metabolism (32%), membrane transport (21%) and amino acid metabolism (15%). To note that 5,6% of genes were shared between the carbohydrate metabolism and the membrane transport pathway.

We observed an impressive over-expression of the ABC transporter complexes *malKlamBmalM* and the divergent *MalEFGK* operon (Johansson et al., 1998) involved in maltose/maltodextrin transport (log_2 FC between 4.6 and 5.8). In addition, other operons involved in sugar transport were over-expressed (*srlABDER*, *lldDPR*, *manZYX*, etc…) as well as operons involved in the amino acid metabolism (*astCADBE*, *dsdAXC*, *tnaABC*, etc…).

Unexpectedly, the gene *crp* known to be a master gene regulator of transport and metabolism of carbon source (Shimada et al., 2011) was slightly over-expressed (log_2 FC of 0.8). Similarly, a weak over-expression of genes involved in cardiolipid synthesis was observed.

The most under-expressed genes were those involved in 50S and 30S ribosome production, in particular genes implicated in transcription and translation pathways represented 52% of down-regulated genes. At this time point, furthermore, the gene *fis* known to be an important metabolism regulator in *E. coli* (González-Gil et al., 1996) was severely down-regulated. This was quite surprisingly because it is one of the most abundant DNA binding proteins in *E. coli* under nutrient-rich growth conditions. It is

implicated in many reactions including phage lambda site-specific recombination, transcriptional activation of rRNA and tRNA operons, repression of its own synthesis and oriC-directed DNA replication (Finkel and Johnson, 1992).

3.1.2 Gene expression profile 3 hours after induction

Comparison between the time points at 5 minutes and 3 hours after induction shows that the number of genes differentially expressed decreased. We observed an about-turn of genes involved in carbohydrate metabolism and membrane transport that in this case represent 20% and 22%, respectively, of down-regulated genes.

In particular, the *fecIRABCDE* transport gene operon involved in citrate-dependent iron transport (Ochs et al., 1995) and other genes as the ferrienterobactin outer membrane transporter, *fepA* and the ferric iron-catecholate transporter, *cirA* were under-expressed 3 hours after induction.

A strong down-regulation was observed for the *nan* operon, *nanATEK* which $\log_2$ FC arrives to a value of -5.5. This operon is involved in the catabolism of sialic acids (Kalivoda et al., 2013) used by some bacteria as a nutrient source, as it contains both carbon and nitrogen and can be converted to fructose-6-phosphate. The *nanR* gene, that controls the operon expression, in our analysis retains a normal expression level compared to C43(DE3) strain harbouring an empty pHis17 plasmid.

As showed in Table 10, up-regulated genes 3 hours after induction are two times less than down-regulated genes. The genes differentially regulated are spread across different metabolic pathways as membrane transport, energy metabolism, cellular community, etc..

However, it is possible to observe the presence of some genes involved in stress response. This is the case of the *psp* operon, activated in the phage-shock stress response operon induced upon stress that dissipate the proton motive force (Model et al., 1997). In the same way, the periplasmic ATP-independent protein-refolding chaperone, *spy* that prevents protein aggregation and assists protein refolding is up regulated.

3.1.3 Gene expression profile 22 hours after induction

After overnight induction the cell machinery seems to be stopped. The function of 5 out of 8 significantly up-regulated proteins were unknown. The three remaining genes were the *atpF* gene, which encodes the protein of interest, the F-ATPase b subunit that was over-expressed 716-fold compared to C43(DE3) harbouring an empty plasmid. Similarly, the genes that belong to the melibiose operon, *melAB* coding for the α-galactosidase and the melibiose:H^+/Na^+/Li^+ symporter, respectively, were up regulated. *melA* mutants are unable to utilize melibiose as the sole source of carbon (Schmitt, 1968). *melB* mediates melibiose symport with sodium, lithium or proton as the coupling ion (Roepe and Kaback, 1990).

Only the spermidines N-acetyltransferase, *speG* and the DNA strand exchange and recombination protein inhibitor, *recX* were slightly down regulated with a $\log_2$ FC of about -1.

3.1.4 ncRNA expression

59 ncRNA were identified in BL21(DE3) in our laboratory by homologies with the *E. coli* strain K-12 MG1655 and inserted in the analysis. Recent BL21(DE3) upgraded annotation confirmed the expression of 45 out of 59 ncRNA (Kim et al., 2017). Among the ncRNA identified, 23 were represented with a p-value under 0.05 in at least one condition and are listed in Table 16.

ncRNA identified in BL21(DE3)	**Expression parameters**					
	5 minutes		**3 hours**		**22 hours**	
	$\log_2$ FC	p-value	$\log_2$ FC	p-value	$\log_2$ FC	p-value
sgrS	-0.412	0.0017	-1.107	0.0138	-0124	0.5189
tff	-1.451	0.0001	0.619	0.0003	0.024	0.9271
ffs	-0.001	0.9964	0.948	0.0018	-0.319	0.3021
mcaS	0.561	0.3964	-0.100	0.8347	-0.829	0.0068
mgrR	-0.272	0.4502	0.831	0.0182	-0.237	0.4412
sdsR	-0.912	0.0043	-0.427	0.3106	-0.760	0.0121
micL	0.418	0.0655	0.715	0.0007	-0.191	0.3940
sibB	0.993	0.0008	0.542	0.0408	0.432	0.1242

sibF	0.856	0.0038	0.454	0.0984	0.326	0.2333
sibG	1.055	2.5600	0.275	0.2087	0.501	0.0484
ryfA	-0.471	0.0008	0.519	2.0111	1.095	3.3488
glmY	-0.971	0.0001	0.882	0.0005	1.180	9.8559
gcvB	-0.695	0.0070	0.609	0.0182	0.610	0.0268
omrA	-1.302	0.0061	-0.283	0.5407	-0.002	0.9952
ssrS	-0.359	0.0520	1.109	1.9054	-0.262	0.2480
sibD	0.321	0.3397	-0.661	0.0511	0.39	0.1154
sibE	0.218	0.3735	0.021	0.9304	0.504	0.0368
rnpB	-0.363	0.0207	0.457	0.0036	-0.956	9.9604
gadY	0.167	0.8006	-0.138	0.7746	-0.967	0.0020
glmZ	-1.049	0	-0.008	0.9467	0.609	1.4393
csrC	0.477	0.0001	0.318	0.0096	0.913	3.3772
ryjA	0.607	0.2249	0.943	0.0154	0.382	0.1407
rydB	-0.857	0.0866	1.678	0.0021	0.545	0.0806

Table 16. Expression of ncRNA in C43(DE3) during F-ATPase over-expression.

We were able to observe only few ncRNA differentially expressed in the different time points. 5 minutes after the F-ATPase b subunit induction two ncRNA were up regulated, *sibB* known to be involved in the toxic protein IbsB regulation and *sibF* of unknown function.

Symmetrically, two ncRNA were down-regulated including *omrA,* an antisense regulator that affects the OM proteins expression and that is positively controlled by the two-component system OmpR/EnvZ (Brosse et al., 2016; Kim et al., 2017). Interestingly, Guillier and Gottesman (Guillier and Gottesman, 2006) observed the down-regulation of the OM cirA, fecA and fepA upon over-expression of *omrA* and *omrB*. In our analysis, the expression of these OM is strongly down-regulated 3 hours after induction (Figure 41) but this is not combined with the over-expression of *omrA* or *omrB.*

We observed an over-expression of the sRNA *rydB* 3 hours after induction. The $\log_2$ FC was equal to 1.68 at this time point (Table 16). *rydB* is involved in inhibition of the σ^S subunit of RNA polymerase (*rpoS)*, which is the master regulator of general stress response in *E. coli* (Hengge-Aronis, 2002). Indeed, *rpoS* mRNA expression does not

change during induction period compared to C43(DE3) cells harbouring an empty plasmid.

3.2 ppGpp quantification

The RNA sequencing profile that we observed 5 minutes after induction seemed to partially correspond to the transcription profile of stringent response identified by Durfee and coworkers (Durfee et al., 2008) in MG1655 *E. coli* cells upon serine hydroxymate (SHX) treatment. The serine hydroxymate is a structural analogue of L-serine that induces the stress stringent response by inhibiting charging of seryl-tRNA synthetase (Dalebroux and Swanson, 2012). This stress response is coordinated by the guanosine tetraphosphate (ppGpp) and the guanosine pentaphosphate (pppGpp), which rapidly inhibit synthesis of stable RNA, ribosomes and proteins, leading to growth arrest. The proteins known to control the cellular pool of ppGpp in the cell, RelA and SpoT (Traxler et al., 2008, 2011; Dalebroux and Swanson, 2012), in our transcriptomic study $\log_2$ FC values are 0.11 and -0.86, respectively.

Interestingly, after 30 minutes of SHX treatment Durfee and coworkers observed a significant decline in transcripts for tRNA coding genes and the activation of eight amino acids operon. Our data likewise, show that 5 minutes after induction 72 genes involved in transcription and translation pathways are down regulated and 29 genes that regulate amino acid biosynthesis are up regulated.

In addition the following observations matches our data, 5 minutes after SHX treatment they observed down-regulation of *fis*, which regulates positively rRNA and tRNA operons, as well as *stpA* gene. The *rmpA* gene encoding the RNase P subunit required for processing of both 4.5 S RNA and tRNAs was under-expressed. Similarly, genes involved in nucleotide biosynthesis such as *pyrF, purU* and *apt* were also down regulated. Equally, the anti-sigma factor *rsd* and the universal stress protein (Usp) were induced after SHX treatment (Durfee et al., 2008) and in our study.

In order to verify the production of the alarmone ppGpp in C43(DE3) strain harbouring F-ATPase b subunit, C43(DE3) cells upon production of either F-ATPase b subunit, sfGFP or the empty pHis17plasmid were cultured in MOPS minimal medium with 0.4 mM of phosphate and labelled with ^{32}P following the protocol described by

Cashel in 1993 (Svitil et al., 1993). In contrast to what we expected, thin-layer chromatography ppGpp quantification 5, 20, 50 minutes, 2 hours and 22 hours after induction showed that the ppGpp production levels were unaffected by the overproduction of F-ATPase b subunit (Figure 42). The quantity of ppGpp was calculated as a ratio between the GTP and ppGpp spot density (Figure 43).

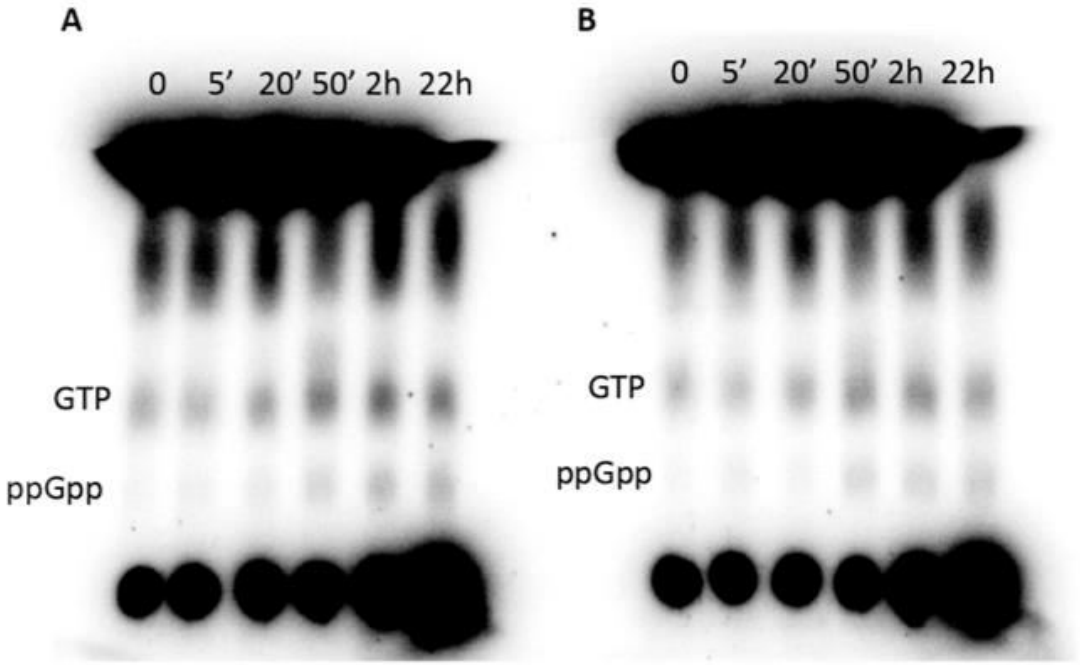


Figure 42. Thin-layer chromatography of C43(DE3) over-expressing F-ATPase b subunit with (A) or without addition of IPTG (B). 5 µl of samples were applied at the bottom of the membrane and the ascending chromatography was developed in 1.5 M phosphate buffer. Results were analysed with a Typhoon FLA9000 (GE Healthcare).

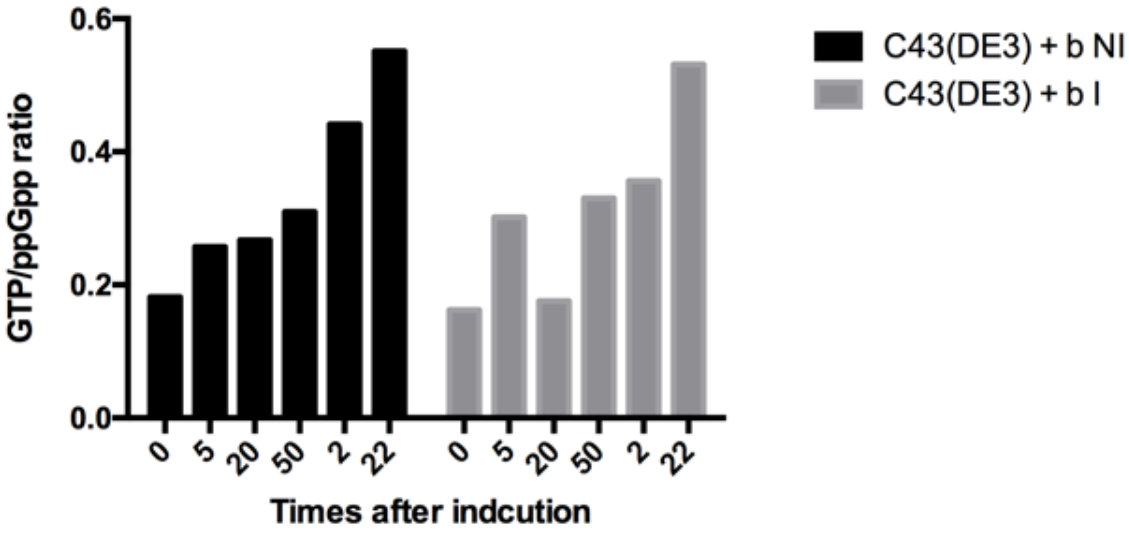


Figure 43. ppGpp quantification. The ppGpp amount in C43(DE3) cells harbouring F-ATPase b subunit with or without IPTG induction was calculated as a ratio between the quantity of GTP and ppGpp.

4 Proteomic qualitative analysis

To further analyse ICMs composition, membranes extracted from overnight C43(DE3) cultures expressing either the F-ATPase b subunit or the empty plasmid pHis17 were separated on sucrose gradient (Figure 44) and analysed by LC-MS/MS analysis.

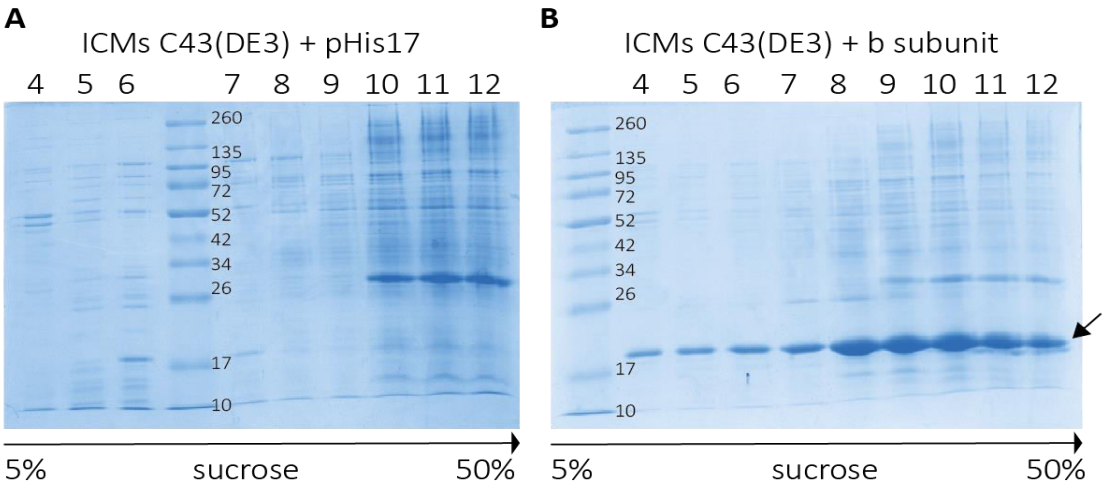


Figure 44. ICMs sucrose gradients of C43(DE3) cells. 2 mg/mL of ICMs were deposited onto a continuous sucrose gradient and centrifuged at 100,000 xg for 18 hours. A little amount of 1 mL fractions were charged on SDS-PAGE gel and stained with Coomassie blue. **(A)** SDS-PAGE gel of C43(DE3) ICMs upon expression of the empty plasmid pHis17. **(B)** SDS-PAGE gel of C43(DE3) ICMs upon expression of the F-ATPase b subunit. The arrow indicates the F-ATPase b subunit.

In order to analyse the entire content of fractions 9 and 10 of sucrose gradient (Figure 44), 20 µl of each fraction were applied on SDS-PAGE gel and run until the sample entered in the running gel.

The digested bands were submitted to mass-spectrometry analysis. Overnight cultures proteomic profile revealed an unexpected scenario, different from the expression of RNA sequencing. Even if the nature of the analysis was rather qualitative than quantitative, proteins with a fold change ratio peptide spectrum matches (PSM) ≥ to 2.5 between the control cells and cells harbouring the F-ATPase b subunit were selected. Furthermore, the PSM value indicates the total number of identified peptide spectra matched for protein. The PSM value may be higher than the number of peptides identified for high-scoring proteins because peptides may be identified repeatedly.

A large extent of proteins identified were presents only in sucrose gradient fractions of C43(DE3) cells harbouring the F-ATPase b subunit and surprisingly 41% of these were cytoplasmic proteins (Figure 45). 26% of inner membrane proteins and 9% of outer membrane and periplasmic proteins followed this data.

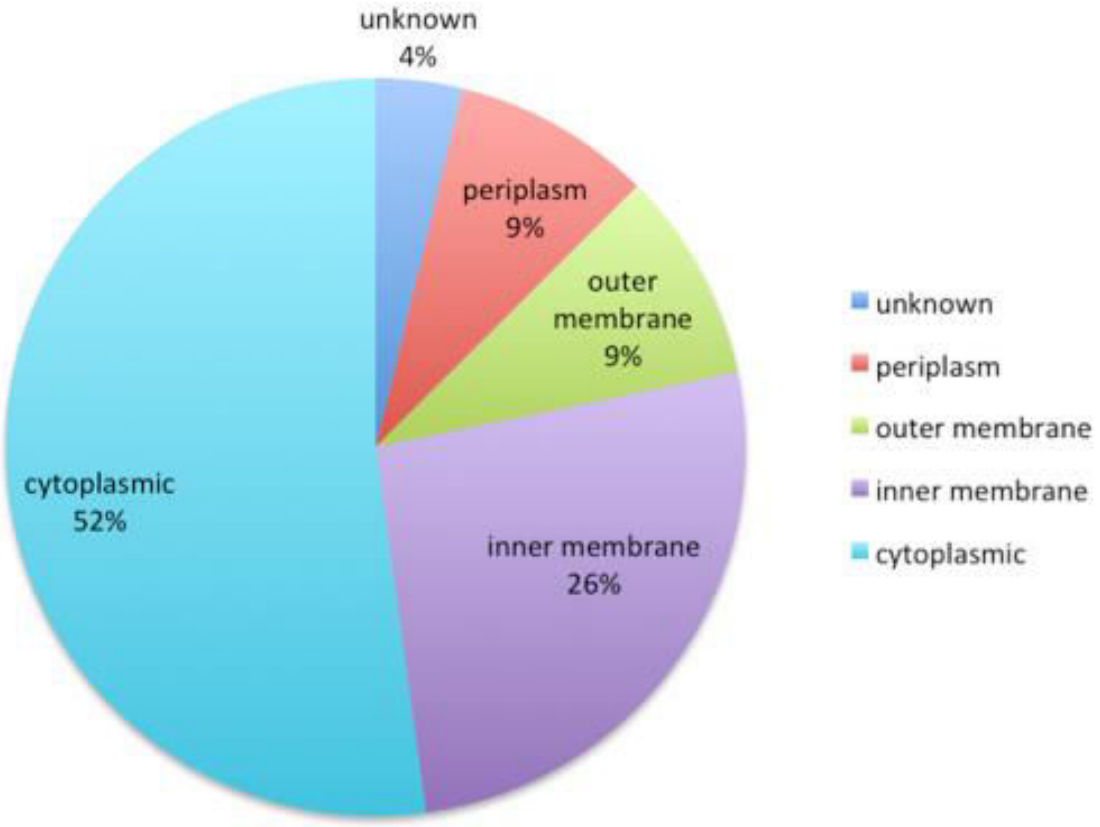


Figure 45. Pie chart of ICMs distribution in C43(DE3) cells. 184 proteins were identified as present in higher proportion in C43(DE3) cells harbouring F-ATPase b subunit and were categorized based on compartment occupied in the cells.

The mass spectrometry profile was quite different from the RNA-sequencing profile; however this initial proteomic experiment was not designed for precise quantification and it should be repeated after isotopic labelling. It is also difficult to compare a whole cell RNA sequencing experiment with a subcellular (ICM) proteomic analysis.

Among the proteins with a PSM ratio higher than 10 we observed protein as the outer membrane protein F, ompF, the phosphoglycerol transferase I, mdoB, the transcriptional regulator of maltose operon, malT that were not differentially expressed in RNA sequencing experiment.

Interestingly, the proteins of Psp operon and the maltose outer membrane porin, LamB were represented in both profiles. However, that data are different from the proteome expression observed by Wagner and co-workers (Wagner et al., 2007) where they pointed out the toxicity effect of membrane protein over-expression through Sec translocon overload. In both our analyses of RNA sequencing and proteomic seem that the cell behaviour is adapted to F-ATPase b subunit overproduction.

Discussion

Chapter 1: Isolation of new BL21(DE3) derivatives

1 MPs production improvement

The *E. coli* strains C41(DE3) and C43(DE3) have significantly impacted recombinant MPs production in bacteria (Miroux and Walker, 1996; Jia and Jeon, 2016). According to Hattab and co-workers (Hattab et al., 2015), they contributed to 28% of non *E.* coli MP structures produced in the T7 expression system.

C41(DE3) mutant host was isolated from BL21(DE3) by exploiting the toxicity of the mitochondrial oxoglutarate malate carrier protein (OGCP). Even though many proteins were produced more efficiently in C41(DE3) than in its ancestor BL21(DE3), the F-ATPase b subunit was one exception for two main reasons: (*i*) the expressed protein accumulated in the cells as inclusion bodies; (*ii*) the expression of the F-ATPase b subunit affected cell growth after addition of IPTG. The toxicity caused by the induction of this MP was used to isolate C43(DE3) from C41(DE3) bacterial host (Miroux and Walker, 1996).

Isolation of strains adapted to MP production by using genetic screen proven to be very successful (Dragosits and Mattanovich, 2013). Following in the footsteps of Miroux and Walker, some laboratories used this method to improve MPs production in *E. coli*, as described below.

Zhao and co-workers (Zhao et al., 2007) ameliorated single-chain antibodies production isolating BL21(DE3) derivatives based on a T7 expression system, namely BLR(DE3) and BLRM(DE3). The *recA*-mutant strain BLR(DE3) was used to select colonies of different sizes on LB agar-plates containing antibiotic and IPTG. The approach used is quite interesting because a *recA*-deficient allele inhibits random homologous recombination-mediated events, which occur quite often in this kind of experiments.

Alfasi and co-workers (Alfasi et al., 2011) used a GFP-based screening coupled with Fluorescence-activated cell sorting (FACS) to improve the production of the chemotaxis protein, CheY.

Hatahet and co-workers (Hatahet et al., 2014) isolated *E. coli* strains capable to express functionally the cytoplasmic alkaline phosphatase, PhoA. In normal conditions, this protein is inactive in the bacterial cytoplasm due to the absence of disulphide

bridge formation, required for its activity. This genetic screen was performed in tandem with an engineering approach. The disruption of the *gor* gene of the glutathione biosynthesis pathway was added to the mutation in the gene *trxB* (thioredoxin reductase) naturally arisen from the selection. This double mutant, called Origami(DE3), enhanced markedly the disulphide bond formation in the cytoplasm.

Recently, Baumgarten and co-workers (Baumgarten et al., 2017) exploited the fusion between the molecular chaperone YidC and the GFP to isolate viable BL21(DE3)-derivative strains upon IPTG induction. The strain Mt56(DE3) so-obtained improves MP production by means of a non-synonymous amino acid substitution in the *T7 RNAP* gene (Ala305Asp), decreasing the affinity of the polymerase for its promoter.

Here we performed a random selection of bacterial mutants using a cytoplasmic protein widely known screening-protein, GFP (Drew et al., 2001; Sarramegna et al., 2002; Drew et al., 2008; Hammon et al., 2009). The rationale behind this idea was to further explore the possibilities of improving the T7-dependent expression system, especially by decreasing the basal level of target gene expression. Others and we have indeed previously shown that, in the T7 expression system (Hammon et al., 2009), any protein can be used to isolate its regulatory mutants (Walker and Miroux, 1999; Schlegel et al., 2017).

Even if a soluble protein was used to perform the genetic screen, these strains enhanced MPs production. To track protein expression, GFP-MP fusion proteins with different TMs numbers were used as it is a common useful approach. Indeed, the fluorescence is visible only if the upstream target MP integrates into the membrane (Drew et al., 2005; Hammon et al., 2009). Moreover, GFP fluorescence is helpful to rapidly, accurately and easily measure protein expression both in liquid cultures and standard SDS-PAGE gels (Drew et al., 2006). Furthermore, once protein expression has been optimized, the fluorescence from GFP can considerably accelerate detergent screening and purification (Kawate and Gouaux, 2006; Drew et al., 2008).

Interestingly, in C45(DE3) host, the reduction of colony size on IPTG plates, linked to YidC production, was similar to that observed with the soluble protein sfGFP (**Table 7**). Solubilised YidC levels were 100 times increased in C45(DE3) membranes as compared to BL21(DE3) host (**Figure 21**). Flow cytometry experiments performed by Baumgarten and co-workers (Baumgarten et al., 2017) show that YidC production was

increased 5 times in Mt56(DE3) compared to BL21(DE3). This difference is likely due to different growth conditions used for the selection procedure. Indeed, Mt56(DE3) cells were grown on LB medium at 30°C whereas the strains C44(DE3) and C45(DE3) were grown on 2xYT medium at 37°C. LB medium is the most commonly used medium for *E. coli* cultures, but the growth stops at a relatively low density, because it contains low amounts of carbohydrates and divalent cations (Sezonov et al., 2007).

We have selected the bacterial hosts using 2xYT medium because, like other rich medium as terrific broth (TB) and super broth (SB), it has been shown to be superior to LB in terms of higher cell densities (Studier, 2005). Finally, it should be underlined that selecting a bacterial host that survives the expression of a given protein, does not necessarily means that the highest expression level for this protein is achieved. C44(DE3) is an example where expression levels of sfGFP are suboptimal. The highest expression level of sfGFP is actually obtained in C41(DE3). At the time of writing this manuscript Mut56(DE3) could not be compared directly to C45(DE3) but the data on YidC expression levels are in favour of C45(DE3).

In addition to growth medium, growth temperature also affects protein production and together with the concentration of inducer, they are essential parameters to get a successful recombinant protein expression (Hattab et al., 2015; Angius et al., 2016; Jia and Jeon, 2016). Often, cultures are shifted to temperatures well below 37°C upon prior to addition of inducer to increase protein yields (Gordon et al., 2008; Skretas et al., 2012) and this facilitated also the production of the h-SQR in *E. coli* C45(DE3). Indeed, a lower temperature during the production process may also improve the quality of membrane-integrated material rather than increase the total amount of produced protein (Chen et al., 2003). This observation is consistent with the idea that by decelerating transcription and/or translation the amount of misfolded MPs is reduced.

2 Genomic modifications in BL21(DE3) derivatives

Genome analysis of C41(DE3) revealed the conversion of the *lacUV5* promoter, upstream of the T7 RNAP gene in the genome of lambda DE3, to wild-type *lac* promoter probably through RecA recombination (Wagner et al., 2008; Schlegel et al., 2015).

Indeed, in BL21(DE3) derivative lacking RecA they were unable to identify hosts in which the promoter *lacUV5* had mutated in its wild-type variant.

Three single nucleotide polymorphisms (SNPs), that specifically change the -10 region and the operator site, are responsible for this conversion (Wagner et al., 2008) (Figure 46). However, the promoter in C41(DE3) appears to be even less strong than the wild-type *lac* promoter (Schlegel et al., 2015), as it cannot be triggered by the CRP-cyclic AMP complex (Ippen et al., 1968) because of the presence of a mutation in the CRP-binding sequence (Figure 46).

Figure 46. The lac promoter (P_{lacWT}) region controls the lac operon. The promoter variant lacUV5 (P_{lacUV5}) controls the expression of the gene encoding the T7 RNAP in BL21(DE3). This variant differs from the wild-type promoter in four positions (asterisk). In C41(DE3) the lacUV5 promoter has reverted to a weaker variant ($P_{lacWeak}$). This variant still harbours the altered CRP-cAMP binding site of the lacUV5 promoter region, but reverted to the wild-type promoter in the -10 and operator sites (Schlegel et al., 2017).

In both strains, C41(DE3) and C43(DE3) one SNP mutation in the coding sequence of *yehU* gene and the excision of an IS-element from the *rbsD* gene (described before) were observed (Kwon et al., 2015). *yehU* encodes a sensor histidine kinase, which, together with the response regulator YehT, positively regulates the expression of the gene encoding the putative peptide transporter YjiY. The amino acid substitution in *yehU* gene leads to high accumulation levels of YjiY (Schlegel et al., 2015).

In addition, three non-synonymous mutations in the coding regions of *melB*, *ycgO* and *yhhA* were observed only in C41(DE3)(Kwon et al., 2015). However, the function of these mutations has not yet been studied in detail.

In C43(DE3), in addition to mutations observed in *lacUV5* promoter of C41(DE3), a mutation in the *lac*-repressor gene, *lacI*, appears to be the key mutation explaining the improved protein production phenotype (Kwon et al., 2015). This mutation decreases the binding of the inducer (IPTG in our case) to the repressor producing a super repressor I^s phenotype (Markiewicz et al., 1994), therefore inhibiting transcription initiation of the *T7 RNAP* gene. Interestingly, Kwon and co-workers underline that the

mutation in *lacI* could only be selected in presence of the weak promoter observed in C41(DE3).

Besides, more diverse genomic variations were detected in C43(DE3) genome (Table 8). SNPs were observed in *yibJ* and *yjcO* genes. Whereas the *yibJ* mutation was synonymous, the *yjcO* mutation was non-synonymous, but the gene function is unknown. In addition, a duplication of the tri-nucleotide GTC was detected in *fur* gene, which function should be further investigated. Furthermore, one of the two CGCCG sequences duplicated in BL21(DE3) in the *dcuS* gene was deleted, thus shifting the gene back in frame (Kwon et al., 2015). According to Schlegel and co-workers (Schlegel et al., 2015) *dcuS* reversion is reported also in C41(DE3), suggesting that this mutation can be easily selected in response to stress conditions. *dcuS* encodes the sensor histidine kinase of the DcuS/DcuR two-component regulatory system. This system regulates the expression of genes involved in uptake and utilization of C_4-dicarboxylates. Authors state that, BL21(DE3) harbours *dcuS*, which contains an internal stop codon and cannot efficiently utilize C_4-dicarboxylates, such as succinate, as its sole carbon and energy source. The deletion in C41(DE3) restores a full-length ORF, which allows the efficient utilization of C_4-dicarboxylates. Moreover, an IS-element was inserted into the promoter region of the cytochrome d terminal oxidase subunit I, *cydA*. The introduction of this mutation in BL21(DE3) by homologous recombination resulted in cell death upon over-expression of OGCP in BL21(DE3) and the F-ATPase b subunit in C41(DE3). Cell death was observed also after introducing *lon* gene restored upon IS-element excision (see introduction section 4.2.1 for other information about *lon* restored mutation in C43(DE3)), except those possessing the mutation in *lacI*, suggesting a phenotype-dependence on *lacI* mutation of C43(DE3). Two large-scale genomic deletions seem to be associated also to IS-element: the *ccmF – ompC* deletion and *yjiV – yjjN* deletion. Concerning the first deletion, authors do not give insights about their impact on C43(DE3) phenotype, but an IS-element insertion in *ompC* gene was already known to be present in B strain (Schneider et al., 2002) and we identified it also in genomic data of BL21(DE3) genome. The insertion is associated with a deletion of the upstream promoter and first 114 bp of *ompC*. In addition, the deletion eliminates the sRNA *micF*, which negatively controls expression of *ompF*. The corresponding OM proteins, OmpC and OmpF, are the two major porins of *E.*coli strain K-12. Thus, it seems that *E. coli* B

expresses only OmpF and its possible implication in C43(DE3) behaviour upon F-ATPase b subunit 5 minutes after induction will be discussed in section 3 of chapter 2. However, the new mapping of BL21(DE3) published by Kim and co-workers (Kim et al., 2017) locates the OmpC porin in the BL21(DE3) genome replacing the coordinates of the OmpC fragment not disrupted by the IS-element and confirms the absence of the sRNA *micF* from the genome. While it seems that the *yjiV – yjjN* region is involved in foreign DNA restriction and the *mcrC – mrr* genes in this region are deleted also in commercial *E. coli* strains such as DH10B and TOP10, generally used to plasmid production.

Thus, it would not be surprising if some mutations identified in C43(DE3), like *dcuS*, *lon* or *rbsD* have an impact on up-regulation of carbohydrate metabolism and membrane transporters during ICMs formation upon F-ATPase b subunit over-expression. Also, we observed a massive production of OmpF in F-ATPase b subunit induced ICMs that could be explained by sRNA *micF* deletion.

Genome sequencing of C44(DE3) and C45(DE3) genomes revealed stop codons at position 656 and 359, respectively, in the T7 RNAP. In contrast to all previously identified T7 mutant, where the amount of T7 RNAP enzyme was regulated at a transcriptional level, stop codon mutations impact the T7 RNAP expression at post-transcriptional level.

The T7 RNAP enzyme contains three C-terminal domains (A, B and C) that are critical for the catalytic activity of nucleotide polymerases (Bonner et al., 1992; Delarue et al., 1990). In C45(DE3) host, the truncated T7 RNAP lacks all three domains and is therefore inactive. In C44(DE3) host, the T7 RNAP enzyme is truncated just after the second domain and the 228 amino-acid deletion removes the C-terminal domain. Although C-terminal deletions of this enzyme have not been systematically studied, the C motif is conserved in all nucleotide polymerases and the point mutation D812N in the C terminal domain completely abolishes its activity (Osumi-Davis et al., 1994). These observations strongly suggest that the T7RNAP truncated form in C44(DE3) is also inactive. The production of an active T7 RNAP enzyme could be possible thanks to a ribosomal stop codon read-through mechanism previously observed in BL21(DE3) (Figure 1) (Odoi et al., 2013). We could formally demonstrate that *amber* suppressor plasmids transformation in C44(DE3) fully restore the BL21(DE3) phenotype, but we did

not have the genetic tools, as *ochre* suppressor strains to transduce *ochre* suppressor in C45(DE3).

This unexpected regulation mechanism appears to be highly efficient to overcome promoter leakage in absence of inducer and to control the T7 RNAP expression during the bacterial growth. The knowledge earned, so far, about the expression of recombinant MPs in *E. coli* highlights the need to lower the accumulation rate of the target protein in the cell. However, recently Hjelm and co-workers (Hjelm et al., 2017) demonstrated that the ability of the rhamnose promoter protein production is actually due to rhamnose consumption rather than regulation based on production rates.

In addition, in both C44(DE3) and C45(DE3) strains, we observed the IS150 insertion deletion from the *rbsD* gene, as already observed in C41(DE3) and C43(DE3) bacterial hosts (Kwon et al., 2015). In C45(DE3), the same kind of mutation appears in the gene *gltL* that encodes the glutamate/aspartate ATP binding protein subunit. We can speculate that this mutation confers a better growth of the strain in presence of glutamate/aspartate, as observed for C41(DE3) and C43(DE3) hosts (Schlegel et al., 2015) in which the deletion of the insertion within *rbsD* gene allows the cells to better grow in media containing ribose.

E. coli C45(DE3) harbors, in addition, a large deletion of about 16 kb in the DE3 region, but the origins of this deletion is unclear. Furthermore, a deletion of 14 bp causes frameshift in *ycdX* gene, which seems to be implicated in swarming motility (Redelberger et al., 2011) defined as a rapid multicellular bacterial surface movement powered by rotating flagella.

As illustrated here, one major drawback of bacterial selection is that additional mutational events (insertions, deletion of insertions) almost systematically occur. These mutational events arise from a first stress upon the overproduction of a foreign protein and a second stress during the long period of growth to cure the expression plasmid. It would be interesting to reintroduce the principal mutations that regulate efficiently the T7 system in the original BL21(DE3) background, at least to restore the advantage of the Lon protease deficiency in BL21(DE3).

3 Toward a large-scale production of eukaryotic MP?

Three main features differentiate the new isolated hosts from the previous ones: (*i*) tight control of the target gene expression in rich medium at 37°C; (*ii*) the amount of target express protein correlates directly with increasing IPTG concentration (Zhang et al., 2015); (*iii*) slow and continuous target protein accumulation in both the exponential and stationary phases at 37°C.

This is achieved without the need to express in a plasmid the natural T7 RNAP inhibitor, T7 lysozyme under the control of the rhamnose promoter as in Lemo21(DE3) strain (Wagner et al., 2008; Schlegel et al., 2012). In addition, there is no need to tune the T7 RNAP mRNA levels as in C41(DE3) and C43(DE3) strains.

Auto-induction medium developed to produce a greater proportion of target proteins in the stationary phase (Studier, 2005; Gordon et al., 2008; Jia and Jeon, 2016) are no longer required to achieve high cell density cultures. The newly isolated hosts offer an attractive regulation of the T7 expression system applicable not only in structural biology but also in synthetic biology, either for the expression of complex metabolic pathways or for production of metabolites such as fatty acids (Xu et al., 2013).

Even if C41(DE3) and C43(DE3) strains have boosted the structural biology of bacterial MP, they have failed to improve production of eukaryotic MP (Hattab et al., 2015). Only few examples of mammalian proteins expressed at the *E. coli* membrane are known. For example, Skretas and co-workers (Skretas et al., 2012) screened libraries of genomic fragments to improve the production of G protein-coupled receptors (GPCRs) in *E. coli*. They have identified three different gene fragments (*nlpDΔ(349-380), nagD* and *ptsN-yhbJ-npr*) that enhance indirectly the production of membrane-embedded and properly folded GPCRs. Instead, Sarkar and co-workers (Sarkar et al., 2008) enhanced the amount of the neurotensin receptor (NTR1) by expressing plasmids encoding random library of the NTR1. Using fluorescent ligands and cell sorting, they have identified mutants of NTR1 exhibiting a higher level of production, not only in *E. coli* but also in yeast and mammalian cells.

It has often been advocated that the speed of coupled transcription and translation in bacteria is too fast and therefore incompatible with the proper folding

and production at high levels of eukaryotic MPs. For example, Jackson and co-workers (Jackson et al., 2012) used chaperones from cold adapted bacteria together with a decrease of the culture growth temperature after induction with IPTG up to 15°C to circumvent inclusion bodies formation of the h-SQR.

In order to obtain a good yields of foreign MP, it is often necessary to monitor genetic and growth parameters to overcome stresses induced by proteins expression, not only in bacteria but also in yeast or eukaryotic HEK cells. It has been shown that, in some cases, toxicity problems are linked to the specific biochemical and biophysical protein properties (Gubellini et al., 2011). In most cases, however, the high transcription rate of the T7 expression system overloads the translocation and membrane insertion machineries very early after induction, causing an accumulation of the unfolded target protein that aggregates rapidly as inclusion bodies (Geertsma et al., 2008). The activation of the heat-shock response, as indicated by the increment of proteolytic proteins in combination with higher levels of chaperons, and the inactivation of energy metabolism appears as a desperate attempt of cells to circumvent this problem (Freigassner et al., 2009).

Wagner and co-workers (Wagner et al., 2007) have illustrated this view by analysing the *E. coli* proteome upon overexpression of three different MPs: YidC, YedZ and LepI. They have shown that MPs over-expression critically affect BL21(DE3)pLysS cell growth and that overload of the translocon leads to the accumulation of inclusion bodies seizing chaperones (DnaJ/K, GroEL/S) and proteases (HslU/V, ClpXP, ClpB).

Interestingly, the same authors noticed that YidC was not produced as inclusion bodies in C41(DE3) or C43(DE3) mutant hosts, and that chaperone and protease levels were decreased in this hosts (Wagner et al., 2008). This example revealed that fine-tuning of the target gene expression not only suppresses the toxicity associated with MPs production and consequently heat-shock and protease responses, but also allows foreign proteins, including eukaryotic proteins, to fold and accumulate in bacterial membranes.

Several groups have developed approaches to circumvent the toxicity associated with eukaryotic MP overproduction (Schlegel et al., 2015). Using beta-lactamase bradykinin receptor fusion protein, Skretas and Georgiou found that co-expression of *ybaB*, which encodes a putative-DNA binding protein of unknown function, significantly

enhances the production of bradykinin receptor and several other MP by up to 10-fold (Skretas and Georgiou, 2010).

More recently, they have used the same receptor, either fused to GFP or the kanamycin resistance gene, to identify two new "toxicity suppressors", namely, the membrane-bound DnaK co-chaperone DjlA and the inhibitor of the mRNA-degrading activity of the *E. coli* RNase E, RraA (Gialama et al., 2017). By co-expressing *djlA* or *rraA* genes, they managed to suppress the toxicity of several MPs.

Here we have demonstrated that the h-SQR is readily targeted to the bacterial membrane in C45(DE3) without the need of external chaperones or extremely low temperature growth conditions. The h-SQR enhanced production highlights the importance of a low rate MP production in the cell to improve folding and targeting of the protein.

The data obtained in the first part of the Ph.D. about C44(DE3) and C45(DE3) strongly suggest that they will further impact the field, and together with direct evolution of the target gene (Sarkar et al., 2008) and codon usage adaptation (Pechmann and Frydman, 2012; Norholm et al., 2013; Pechmann et al., 2014; Boël et al., 2016), will likely extend the use of *E. coli* for eukaryotic MP production. It could be also interesting to combine our regulatory mutations of the T7 system with genetic factors (ybaB, DnaK/J) to see whether they can act in synergy to improve the folding and expression levels of eukaryotic MPs.

Chapter 2: Formation of intracellular membranes (ICMs)

1 Importance of MP folding-shape in the membrane bilayer

The large amounts of ICMs formation upon over-expression of the F-ATPase b subunit in both C43(DE3) and C44(DE3) showed that the sequence and structural properties of the protein, combined with lipid bilayer composition, are key factors in this phenomenon. It is worth noting that F-ATPase b subunit expression plasmid is toxic in both C41(DE3) and C45(DE3) hosts, which compromises ICMs formation.

The F-ATPase b subunit has an extending cytoplasmic domain that is considerably larger than its transmembrane regions. This is a common feature with MP causing ICMs formation, for example the complete ATP synthase (Von Meyenburg et al., 1984), the fumarate-reductase complex (Weiner et al., 1984b), the chemotaxis Tsr receptor, the MtlA permease (Weeghel et al., 1990; Lefman et al., 2004) and the FtsY-receptor-ribosome complex (Herskovits et al., 2002). The lateral pressure applied from the external domains of these large protein complexes might force the membrane to bend, which in turn would induce invaginations of the inner membrane (Eriksson et al., 2009).

In particular, the membrane network organization observed under overexpression of the Tsr receptor (Weis et al., 2003b) seems to be similar to the F-ATPase b subunit one composed by tubules and vesicles. Lefman and co-workers (Lefman et al., 2004) suggest a mechanism of membrane proliferation (Figure 47). They have proposed that at low levels, the cytoplasmic membrane would have not small local invaginations. As the levels of Tsr increase, the inner membrane surface would also increase in order to accommodate the protein, which in turn induces membrane invagination. Proximal membrane bumps would allow interactions between the long cytoplasmic domains, which would lead to the creation of zipper-like structures. In addition, the periplasmic domain of the protein could also interact with the other side of the membrane.

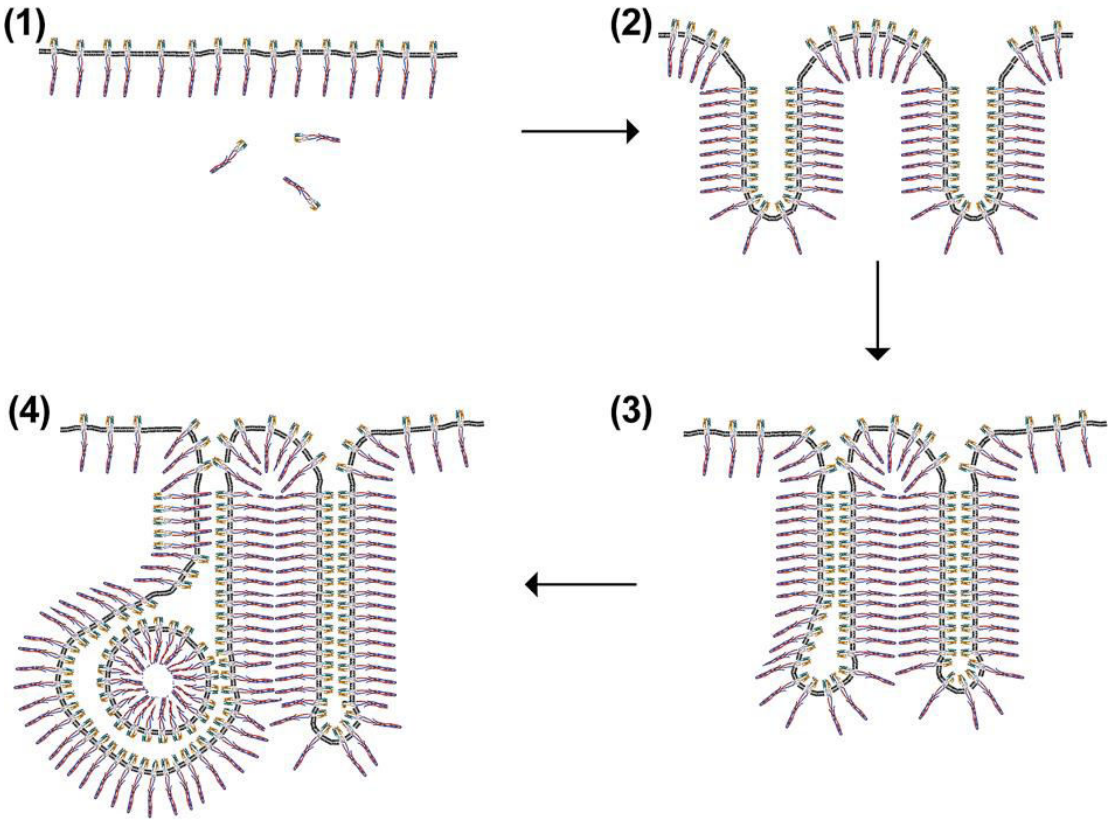


Figure 47. Mechanism of membrane proliferation upon accumulation of Tsr protein. **(1)** Cells producing low levels of Tsr would present no or small local invaginations. **(2)** Increasing levels of Tsr would be accompanied by an increase in the membrane surface. **(3)** The long cytoplasmic domains of Tsr in adjacent membranes would interact creating zipper-like structures. **(4)** Similar structures would be created by the interactions of Tsr periplasmic domains. Sporadically, the end regions of the membrane bumps are large enough to board invaginations from different region of the membrane, which will protrude from the backside through the boundaries of surrounding membranes. These rounded tubules would have the appearance of a vesicle in a cross-sectional view. (Lefman et al., 2004).

The mechanisms of membrane formation by the F-ATPase b subunit would be similar to those induced by Tsr. However, in contrast to Tsr, the F-ATPase b subunit does not have a periplasmic region able to interact with the other side of the membrane and, therefore, membrane invaginations would only be stabilised by the interactions of the cytoplasmic domains (Arechaga, 2013).

We also observed that the vesicles contain a large amount of cytoplasmic protein and few periplasmic proteins (Figure 45). This is not what one can expect from the Tsr invagination model. It cannot be excluded that vesicles were broken during cells disruption, however, gentle lysis of the cells with lysozyme and potter showed a very similar pattern of protein in the ICMs (Jorge Royes, personal communication).

Although high levels of protein expression are required for the formation of the ICMs network, this observation suggests that the capability to bend the cytoplasmic membrane is an intrinsic property of the protein. Supporting this view, here we also

showed that C-terminal fusion of F-ATPase b subunit with the BFP fluorescent protein is deleterious to ICMs formation (Figure 31).

In contrast to the transmembrane proteins, several monotopic membrane proteins also induce ICMs formation; showing that the membranes network formation is a process subject to the change of numerous factors. Some examples are the *E. coli* glycerol-3-phosphate acyltransferase PlsB (Wilkison et al., 1986b), the peptidoglycan precursor glycosyltransferase MurG (Ha et al., 2000; Laan et al., 2003) and the lipid glycosyltransferases alMGS and alDGS (Eriksson et al., 2009). The glycosyltransferase ability to remodel the membrane has been studied. The mechanism by which alMGS stimulates ICMs formation is connected to the protein insertion in the lipid bilayer interface (Ge et al., 2014).

2 Role of CL in ICMs formation

Arechaga and co-workers (Arechaga et al., 2000) observed an increment of CL from 2-4 to 14% in F-ATPase b subunit induced ICM at the expense of its biosynthetic precursor phosphatidylglycerol (PG).

The function of CL in mitochondria and other eukaryotic organelles such as the endoplasmic reticulum has been extensively studied (Mileykovskaya and Dowhan, 2009; Schlame and Ren, 2009). In mitochondria CL are mainly present in the inner membrane cristae where they are important to maintain the proton electrochemical gradient for the production of ATP (Mannella et al., 2013). It has been shown that the interaction between CL and ATP synthase dimers is also important to control mitochondrial cristae morphology (Paumard et al., 2002). Acehan and co-workers (Acehan et al., 2011) employed cryo-electron tomography to compare the structural organization of the complex V in *Drosophila* muscle mitochondria from wild-type and CL-synthase mutant flies. Their findings suggest that CL are critical for the degree of oligomerization and the degree of order in ATP synthase assembly, which are likely to affect cristae morphology and the efficiency in energy production. In addition, mutants lacking subunits *e* and *g* of the ATP synthase would not dimerize, which in turn results in the disruption of lamellar cristae (Arnold et al., 1998, 1999). The absence of these subunits in the bacterial ATP synthase could explain why there are not cristae-like invaginations in bacteria under

physiological conditions. It would be interesting then to produce the *e* and *g* subunits of the mitochondrial F-ATPase to test whether low amounts of those subunits are able to trigger ICM.

Even if there is still a lack of information about the CL contribution in dynamics and morphology of bacterial membranes, numerous examples of interaction between CL and proteins in the bacterial network have been reported (Schlame et al., 1991; Palsdottir and Hunte, 2004). In cells overproducing the fumarate reductase (Weiner et al., 1984b), it was observed that cells produce ICMs to maintain a constant lipid/protein ratio.

Several indications suggest a direct interaction between the overexpressed glycosyltransferase MurG and CL in proliferating membranes (Laan et al., 2003). They observed a CL enrichment following MurG protein overexpression and a higher enzymatic activity of MurG in vesicles containing CL than in vesicles containing PG.

The magnesium transporter A, MgtA, needs CL to be active *in vitro*, while over-expressed MgtA in C43(DE3) cells co-localizes with CL *in vivo* (Subramani et al., 2016).

Recently, Danne and co-workers (Danne et al., 2017) demonstrated that the membrane-remodelling capacity of the phospholipid N-methyltransferase PmtA form the plant pathogen *Agrobacterium tumefaciens*, depends on the presence of CL. The absence of CL causes the abrogation of vesicle-like structure in CL-deletion mutants, suggesting their importance in both tubulation and vesiculation.

In this work, and in collaboration with Gerardo Carranza and Ignacio Arechaga (IBBTEC, Santander, Spain), we have shown that CL depletion in C43(DE3) cells overproducing F-ATPase b subunit have a dramatic effect on membrane morphology (Carranza et al., 2017).

In C43(DE3)*ΔclsABC* strain, lamellar structures are observed upon over-production of F-ATPase b subunit but remain unstructured. Similarly, in *S.cerevisiae* onion-like structures were observed in mutants with reduced levels of CL (Mileykovskaya and Dowhan, 2009). Thus, suggesting that an interaction lipid-protein is necessary to membranes structuration.

Stabilization of bacterial ICMs could be mediated by electrostatic interactions between F-ATPase b subunit and CL. At neutral pH, CL is negatively charged and, as any of the phosphate moieties of its head group, is able to pick up a hydrogen and make a

H-bond with the free OH of the glycerol (Kates et al., 1993; Haines, 2009). Electrostatic interactions between basic residues of integral membrane proteins and CL are common (Epand et al., 2007). For instance, solid-state NMR experiments have shown electrostatic interactions of CL with a basic residue in F-ATPase c subunit (Laage et al., 2015). Given that F-ATPase b subunit contains 16 lysines in its soluble part, it is likely that such interactions also play an important role in ICMs stabilization.

To achieve total CL depletion, all three genes (*clsA, clsB* and *clsC)* had to be removed, as previously described by Tan and co-workers (Tan et al., 2012). In single *clsA* deletion, which has long been known to be the only gene responsible for CL synthesis, no CL was detected but only in cells grown at low salt and harvested in logarithmic phase (Hiraoka et al., 1993). However, if cells were harvested in stationary phase, CL traces were still detectable (Nishijima et al., 1988). The C43(DE3) *cls* mutants confirm what has been described in previous research. C43(DE3) *cls* triple mutant over-producing F-ATPase b subunit shows a total lack of CL. To compensate this lack bacteria shows an increase in PG and PE levels compared to wild-type C43(DE3) cells over-producing the F-ATPase b subunit. These results indicate that the effects of *cls* deletion on lipid metabolism are enhanced by the F-ATPase b subunit overproduction (Carranza et al., 2017).

The characteristic conical CL structure conferred it polymorphic features (Dahlberg, 2007). This property allows different structuration of CL molecules as micelle, lamella or hexagonal membrane structures, depending on its acyl composition (Sankaram et al., 1989) or the presence of divalent cations (Cullis et al., 1978; Vail and Stollery, 1979; de Kruijff and Cullis, 1980; Seddon et al., 1983; Rand et al., 1990). The polymorphic CL behaviour caused by divalent cations has been suggested to be essential in different fusion and fission membrane events (Ortiz et al., 1999).

Furthermore, McGarrity and co-workers (McGarrity and Armstrong, 1981) observed that CL fatty acid composition fluctuates based on different growth conditions. In particular, the number of acyl chain unsaturation increases proportionally with decreasing of temperature (Garrett et al., 2012).

Although, under physiological conditions CL has been found in the bilayer phase, when the latter is perturbed CL tends to form hexagonal phases (Lewis and McElhaney,

2009; Schlame and Ren, 2009). This hexagonal pattern can be observed in C43(DE3) cells upon overexpression of F-ATPase b subunit (Figure 48).

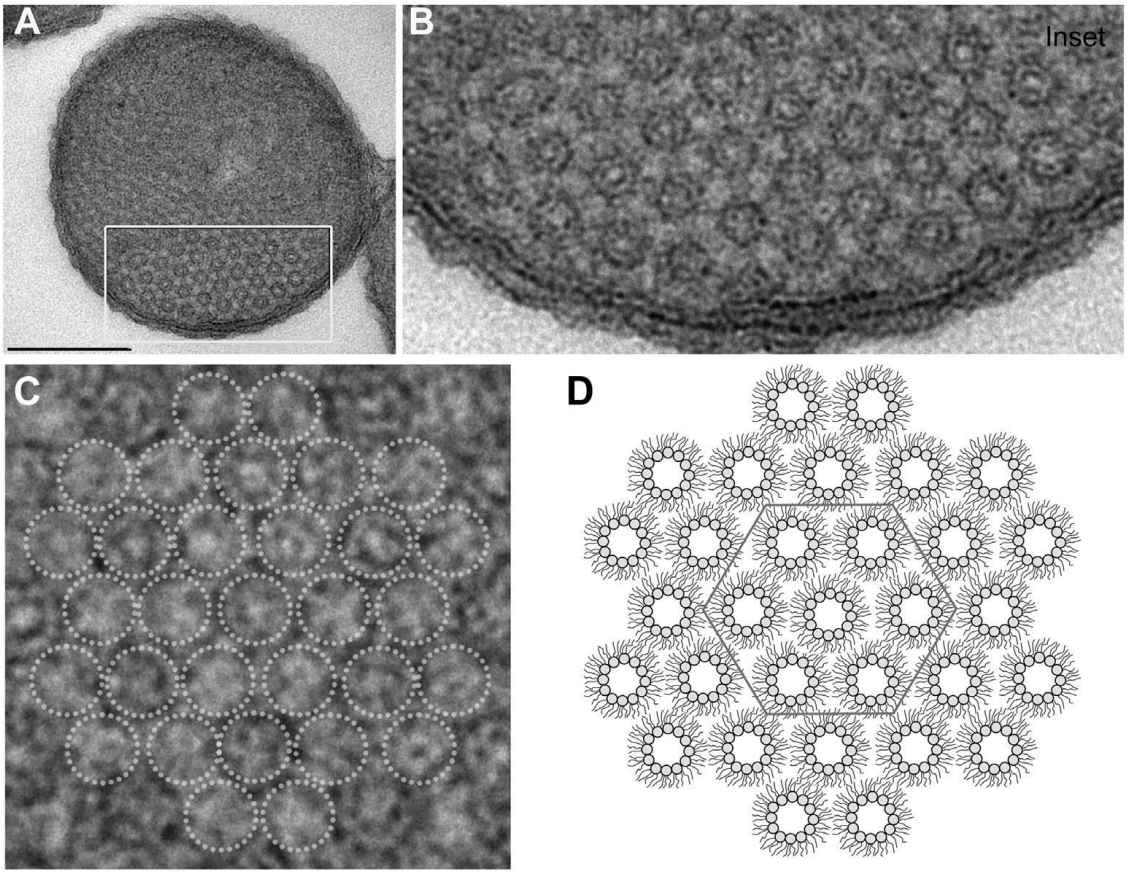


Figure 48. Non-lamellar morphology of ICMs. (A) Orthogonal view of thin sections of *E. coli* C43(DE3) cells over-producing F-ATPase b subunit. Scale bar, 0.25 μM. (B) Magnified view of the boxed area marked in A. (C) Magnification of B, dotted circles indicate ICMs hexagonal phase. (D) Schematic representation of inverted hexagonal phases as described in (Carranza et al., 2017).

It has been demonstrated that CL are one of the factors that cause bilayer lamellar phase transformation (Ortiz et al., 1999; Pan et al., 2014) and promote formation of inverted hexagonal II phases (Rand and Sengupta, 1972; de Kruijff and Cullis, 1980). Also, lipids in inverted hexagonal phases self-assemble in long tubes arranged in a hexagonal lattice (Seddon, 1990). All this suggests that the peculiar physical properties of CL play an essential role in ICMs formation observed under F-ATPase b subunit over-expression.

The same behaviour was observed upon addition of dynamin related protein 1 (Drp1) to CL containing lipid in *in-vitro* systems (Stepanyants et al., 2015). Authors suggested that Drp1 association with CL increases the propensity from lamellar to HII phase transition, which could be essential in mitochondrial division.

Due to its geometry, CL has been found to localize in negative curved regions in *E. coli* membranes (Huang and Ramamurthi, 2010; Renner and Weibel, 2011). In this region, lateral pressure of lipids in a monolayer decreases upon insertion of a CL molecule, which, in turn, contributes to the creation of membrane folds (Nichols-Smith et al., 2004; Schlame and Ren, 2009; Unsay et al., 2013). A similar lipid packing modulation by CL was observed in inner mitochondrial cristae (Khalifat et al., 2011; Zeczycki et al., 2014).

3 ICMs molecular regulation

From a metabolic point of view, the F-ATPase b subunit over-expression revealed an intricate scenario. Only few minutes after protein expression we observed a high exploitation of the carbohydrate metabolism, membrane transport and amino acids biosynthesis pathways caused by IPTG induction. In contrast, the genes involved in transcription and translation pathways were down regulated. According to Durfee and co-workers (Durfee et al., 2008), this metabolic response suggests the presence of an early stringent response in the cell. However, we found no significant stimulation of ppGpp synthesis upon induction of F-ATPase b subunit (Figure 43). Consequently, we proposed two possible explanations for this phenomenon: (*i*) the strain C43(DE3) is affected by the partial disruption of the outer membrane porin OmpC (Kwon et al., 2015). Genomic sequencing data strongly suggest that full-length OmpC expression is seriously compromised. However, OmpC reads have been detected in RNA-sequencing experiment but it is unlikely that the full-length protein is produced. Whether OmpC deletion induces a metabolic stress remains to be investigated; (*ii*) the F-ATPase b subunit over-expression induces in bacteria a major metabolic change that is not connected to carbon source starvation and which has not been described before upon the over-production of other MPs (Gubellini et al., 2011; Marciniak et al., 2012).

Additionally, one cannot rule out that the phenomenon explanation arise from the mix of the two theories. However, further control experiments are required, with overexpression of soluble or MPs in C43(DE3), to distinguish which metabolic response is specific to the formation of ICM.

As mentioned in results section, after 5 minutes of induction we observed a great over-expression of operons involved in nutrients uptake, like maltose, glucitol, mannose, glycerol etc.. These nutrients fall into four groups, summarised in Table 17 based on the molecular mechanism of their transport across the membranes. Briefly, fructose, glucitol, mannitol, mannose and N-acetylglucosamine (NAG) pass through the outer membrane *via* the porin OmpF and then cross the inner membrane *via* the phosphotransferase system (PTS); galactose, glycerol, lactose and melibiose pass the outer membrane through OmpF, but their transport across the inner membrane is performed by other nutrient-specific proteins. While, trehalose is a LamB/PTS nutrient and maltose is a LamB/non-PTS nutrient (Travisano and Lenski, 1996).

		Outer membrane	
		OmpF	**LamB**
Inner membrane	**PTS**	Glucose Fructose Glucitol Mannitol Mannose NAG	Trehalose
	non - PTS	Galactose Glycerol Lactose Melibiose	Maltose

Table 17. Mechanism of nutrients import classification. Table adapted from (Travisano and Lenski, 1996).

We hypothesize that following the OmpC production impairment, in *E.coli* C43(DE3) there might be an increased amount of the porin OmpF. The massive nutrients uptake, generally imported inside the cell through the porins OmpF and LamB, could reflect the improved OmpF presence and LamB over-expression.

Among the operons most represented, the *malKlamBmalM* operon stands out. This operon forms a divergent operon complex with the *malEFG* operon, the activation of which depends on *MalT* and on the global activator cyclic AMP (cAMP) receptor

protein (Johansson et al., 1998). In particular, the outer membrane porin LamB works as a channel for maltodextrins and the bacteriophage lambda receptor (Randall-Hazelbauer and Schwartz, 1973; Freundlieb et al., 1988). However, it has been demonstrated that the high-level expression of LamB is toxic (Carlson and Silhavy, 1993). This toxicity is expressed by maltose sensitivity, altered outer membrane permeability and induction of the pspA stress protein (Cosma et al., 1995). Since elevated expression of LamB causes cell death upon deletion of the osmoregulator OmpR (Reimann and Wolfe, 2011), we speculated that the cell should implement a quick-stress response mechanism. Among the stress response systems known to sense damage in *E. coli*, the Cpx two-component system (TCS) responds to periplasmic or inner membrane (IM) protein misfolding (Raivio, 2014). The TCS consists of the protein CpxA, the IM sensory histidine kinase and CpxR, which is the DNA-binding response regulator (**Figure 3**).

In vitro analyses showed that the reconstituted phosphorylation and dephosphorylation abilities of CpxA are controlled by changes in lipid bilayer properties of proteoliposomes. In addition, an increase in anionic surface charges and likely bilayer thickness of the vesicles resulted in increased levels of CpxR and the inhibitor CpxP (Keller et al., 2015).

According to Price and Raivio (Price and Raivio, 2009) 50 genes are known to be regulated by Cpx. In our RNA-sequencing experiment, the majority of these genes are similarly regulated 5 minutes and 3 hours after induction, suggesting an activation of the Cpx response, particularly in terms of envelope components (**Table 18**). While, 22 hours after induction there is no evidence of this response suggesting that once formed, ICMs are well tolerated.

Gene	5 minutes	3 hours	22 hours	Price and Raivio (2008)
degP	0.9	0.8	- 0.5	Envelope protein maintenance
rdoA-dsbA	0.2	0.8	- 0.3	
ppiA	0.3	0.3	0	
ppiD	- 0.3	- 0.1	- 0.1	
psd	- 0.3	0.3	- 0.3	
secA	- 0.1	- 0.5	0	
spy	1.3	1.8	- 0.1	
ompC	disrupted from insertion in C43(DE3)			Envelope components
ompF	- 0.1	0.5	0	
nanc	0.7	- 5.6	0	

acrD	- 0.5	1.5	0	
mdtABDC		NC		
efeU	- 0.5	- 2.3	0.9	
yccA	0.3	1.1	- 0.7	
ycfS	0.6	1.8	- 0.5	
	0			
yqjA	0.1	0.8	0	
yebE	0.8	2.4	- 0.5	
cpxP	0.8	1.7	- 0.5	Signal transduction
cpxRA	0.5	0.6	- 0.3	
rpoE-rseABC	0.3	0.6	0	
				Bacterial appendages and chemotaxis
motAB-cheAW	0	0.3	0.5	
tsr		NC		
aer	0.1	0.3	- 0.3	
csgBAC	1.0	- 0.1	- 0.1	
csgDEFG	- 0.1	0.6	- 0.1	
pap	1.0	2.0	NC	
aroK	- 1.0	- 0.3	- 0.1	Unrelated to envelope components or stress or unknown function
ftnB	0.8	0.9	0	
htpX	0.5	0.3	- 0.3	
mviM		NC		
ung	- 0.3	0	0.1	
ybaJ		NC		
ydeH	- 0.3	1.7	- 0.5	

Table 18. Genes and operons proposed to be under Cpx regulation by Price and Raivio. In green genes regulated positively or negatively by the Cpx system according to literature and in red genes, which do not fall in the Cpx system regulation. The log2 FC was used to screen differently regulated genes.

The expression of the porin OmpC, disrupted from insertion in C43(DE3), is generally positively regulated from the Cpx response. Unfortunately, in literature there are no reports linking OmpC deletion and the Cpx response.

Among the most regulated genes of the Cpx response are *spy* and *cpxP*. Spheroplast protein Y (Spy) protects the cell from environmental changes that might lead to protein aggregation. Its expression is transcriptionally regulated by the Cpx stress response system coupled with the Bae system (Srivastava et al., 2014). Similarly to Cpx system, Bae is a TCS in which an IM protein, BaeS act as histidine kinase to sense membrane stress. Upon activation, BaeS gets autophosphorylated and transfers this phosphate group to its regulatory partner, BaeR (Nishino et al., 2005).

The periplasmic CpxP protein is the negative-inhibitor of CpxA (Price and Raivio, 2009). From the 3' UTR of its mRNA is expressed a sRNA known to be highly abundant during Cpx stress, CpxQ (Grabowicz and Silhavy, 2017). Unfortunately, the level of this

sRNA is missing in our analyses due to lack of information about its sequence and position in BL21(DE3) at the time of the analysis. Indeed, a detailed map of BL21(DE3) has been published only in May of this year (2017) (Kim et al., 2017).

It is worth noting that the sRNA MicA might has an influence on LamB expression. Indeed, it has been demonstrated that this sRNA down regulates the *lamB* gene expression in *Salmonella* (Bossi and Figueroa-Bossi, 2007). Similarly, we noticed that 5 minutes after induction, the MicA level was low (log_2 FC = -1.0) and the LamB expression was extremely high, whereas, 3 hours after induction we could observe an inversion of sRNA and LamB expression levels.

The main reason to tight regulate folding and insertion of proteins by activating stress response pathway and producing sRNA is to maintain the proton-motive force (PMF) across the IM, which is a source of cellular energy in Gram-negative bacteria (Grabowicz and Silhavy, 2017).

In *E. coli* the level of PMF is maintained by the phage-shock-protein (Psp) activated by IM damages. An inducing signal in the IM activates the PspB-PspC complex resulting in recruitment of PspA causing the release of PspF that activate the psp operon (Flores-Kim and Darwin, 2016). We observed that all the Psp operon proteins are up regulated upon over-expression of F-ATPase b subunit, according to Carlson and Silhavy (Carlson and Silhavy, 1993) who reported the strong stimulation of the Psp stress response upon production of the mutant LamB(A23D).

In our analysis the *psp* operon starts to be express just after induction, remains over-expressed 3 hours after induction and proteomics analysis indicates the presence of proteins PspA, PspB and PspC in sucrose gradient fractions containing the F-ATPase b subunit. It would be interesting to follow the PMF *in vivo* upon overproduction of F-ATPase b subunit to assess whether drop in membrane potential correlates with F-ATPase b subunit accumulation and the activation of the Psp response pathway. An attractive hypothesis would be that PspABC proteins physically interact with ICM and thus participate to their regulation. This hypothesis is supported from the observations that PspA forms high-order oligomers *in vivo* (Engl et al., 2009) and creates connections with the CL-containing polar regions of the cell along with PspBC (Jovanovic et al., 2014). The authors proposed a functional link between CL, PspBC-dependent signalling and polar IM localization of PspA, in order to maintain its integrity.

In addition, PspA seems to have significant similarity with Vipp1, a protein involved in thylakoid membrane biogenesis and protection in cyanobacteria and higher plants (Otters et al., 2013; Zhang et al., 2014). They show common N-terminal amphipathic helices that play an essential role in bilayer association (McDonald et al., 2015). It seems that PspA and VIPP1 aggregate together when a stored curvature elastic (SCE) stress occurs to stabilize the membrane and for both were observed a higher affinity for vesicles-containing anionic lipids (Figure 49).

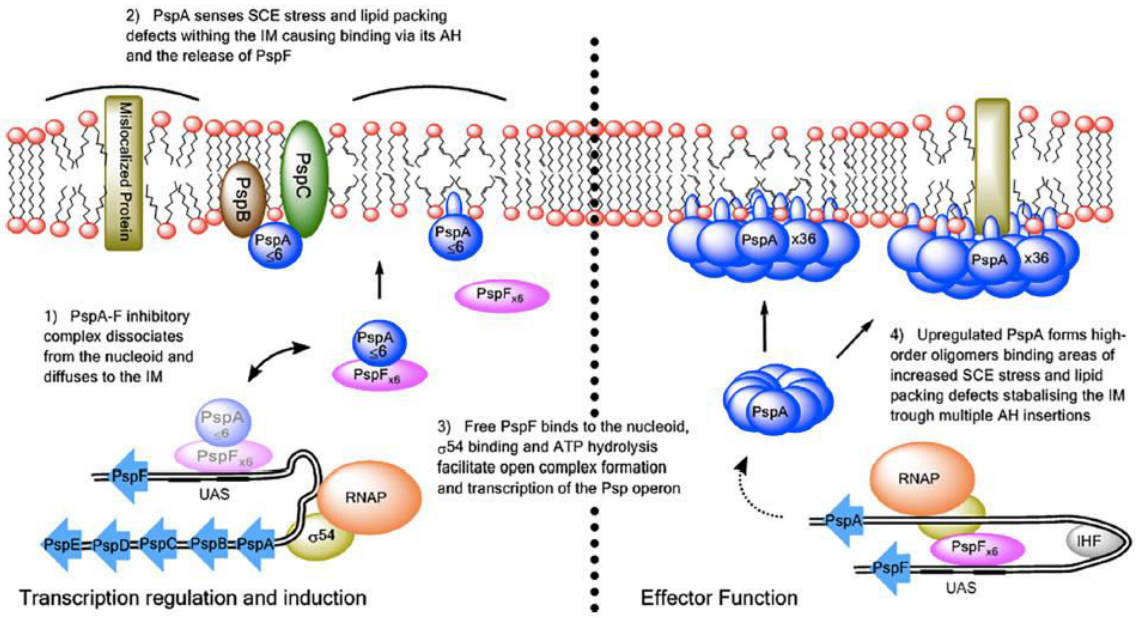


Figure 49. Mechanism of the Psp response of *E. coli* proposed by McDonald and co-workers (2015). The PspA-PspF inhibitory complex diffuses to the IM upon stress damage induction. PspA can reach the IM either exploiting the PspBC sensors or by directly binding to the IM at locations of SCE stress and membrane packing defects. The binding of PspA to the IM allows PspF to initiate transcription of the psp operon and activate the Psp response. The up regulated PspA forms high-order oligomers and multiple helix insertions into the IM from PspA reduces some of the SCE stress.

So far, one possible scenario may be that induction of F-ATPase b subunit induces membrane curvature interacting with anionic lipids, such as CL. Either the high-curvature of membranes or IM breaks to produce small vesicles, can compromise the membrane PMF activating the Cpx and Psp responses.

Another possibility could be that the localization of the F-ATPase b subunit in the IM is perceived from cell as a mislocalized protein or that a small fraction of the protein is misfolded. Thus, may cause SCE stress affecting the PMF of membranes and activate envelope stress responses as the Cpx response and the Psp response. The up-regulated pspA may forms high-order oligomers by binding to CL or anionic phospholipids around the mislocalized F-ATPase b subunit, causing invagination of the

membranes in a similar way to the one proposed for the caveolae-induces vesicles formation (Figure 50). Caveolae are endocytic molecule located in mammalian cell surfaces where they play roles in endocytosis, oncogenesis and the uptake of pathogenic bacteria. It has been shown that heterologous caveolae formed inside *E. coli* have a size and morphology similar to those of mammalian cells (Walser et al., 2012).

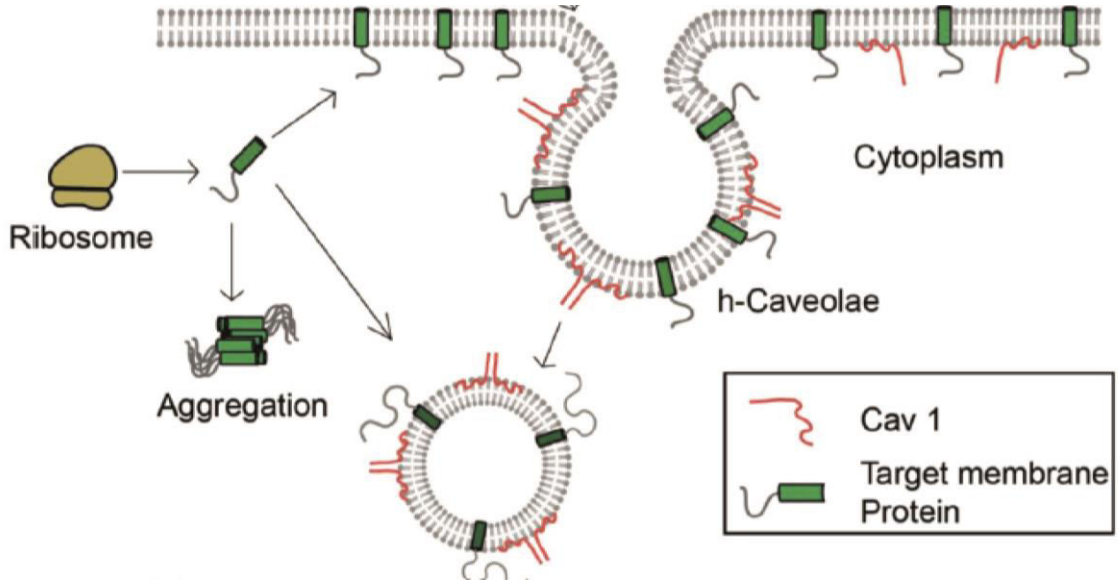


Figure 50. Caveolae-induced vesicles formation mechanism. Caveolae molecules can bind to IM triggering vesicles formation.

Conclusions & Perspectives

1 T7 expression system regulation as a tool to improve MP production in *E. coli*

Membrane proteins account for more than 50% of the total cell membrane (Engelman, 2005), it is therefore not surprising that they represent the majority of drug targets (Overington et al., 2005; Terstappen and Reggiani, 2001). However, there are an extremely low number of unique MP structures available in the PDB. Out of the ~90,000 available protein structures to date about 600 represent membrane proteins with only a sixty of those being human membrane protein structures.

The T7 based expression system has had a great success in structural biology of proteins. The C41(DE3) and C43(DE3) T7 bacterial expression hosts, isolated from the parental strain BL21(DE3) provided the means to solve membrane-protein structures (Miroux and Walker, 1996). It is worth noting that, in contrast to its parental strain, in C43(DE3) the *lon* activity is restored. Since its function is to clear the cell from misfolded protein, it would be interesting to study C43(DE3) *lon* deficient strain, to detect or not changes in cell behaviour and recombinant production yield. The success of C43(DE3), including those using arabinose promoter-based expression plasmid may rely on the ability of the cell to remove misfolded proteins that often contaminate the purified recombinant protein.

Despite their success, these mutant hosts are poorly regulated by IPTG and the first objective of my Ph.D. was to further improve the regulation of the T7 expression system by selecting new bacterial hosts. We have indeed isolated by GFP genetic screening two new mutant hosts, namely C44(DE3) and C45(DE3). Their most remarkable feature is that the target MPs was well folded and localized in *E. coli* membranes. In addition, DDM widely used for 3D crystallization can readily solubilize them. C45(DE3) strain challenged with a human membrane protein, as the h-SQR, showed a great expression, which breaks the barrier of producing functional human MPs in *E. coli*.

Besides, a wider MP production screening can give us insights about the influence of protein properties on expression in C44(DE3) and C45(DE3). It has been demonstrated that proteins with a large number of predicted TMs or a large MW were generally less successful (Bernaudat et al., 2011).

Even if the new hosts, C44(DE3) and C45(DE3) are better regulated than the previous ones, the sfGFP production remains down-regulated compared to C41(DE3). The optimal condition would be to isolate or build a strain perfectly regulated: completely silent before addition of IPTG and capable of reaching C41(DE3) levels of recombinant proteins after induction.

The stop codon mutation identified in both strains was the only peculiar mutation of C44(DE3). In the strain C45(DE3), however, other mutations have arisen, which would be better to study individually to gain perhaps important information, like it was done for C41(DE3) (Schlegel et al., 2015) and C43(DE3) (Kwon et al., 2015).

Furthermore, Schlegel and co-workers (Schlegel et al., 2015) demonstrate that the genotype of C41(DE3) reflects the stresses and conditions used during its isolation and the long term culture needed to cure the expression plasmid from the host. Does it happen the same for the others strains and is it correlated with the kind of protein used?

To further support the high value of these strains a future direction of research would be to challenge the new hosts with direct evolution of the target gene (Sarkar et al., 2008) and codon usage adaptation (Boël et al., 2016). Indeed, next to host optimization it seems that being able to precisely and stably control protein translation rates could be the key to optimize MP production. It would be interesting then, to modify translation initiation regions or using N-terminal tags that can results in higher yields of protein.

2 Dynamic and molecular features of membrane invaginations upon over-expression of the F-ATPase b subunit

The discovery that the over-expression of the F-ATPase b subunit can induce ICMs proliferation in C43(DE3) (Arechaga et al., 2000) was among the forerunner of this issue. Especially because it allows bettering understand MP gene over-expression and membrane biogenesis in *E. coli*.

In addition to express the protein at low level in the cytoplasmic membrane or as inclusion bodies, it is also possible that the over-expression can trigger membranes proliferation to store higher amount of protein. It would be exciting to discover a trigger-signal capable of direct the mechanism behind.

In both C43(DE3) and C44(DE3) the formation of membrane network is accompanied by a CL level increases. Thanks to the collaboration of the IBBTEC institute, in Spain, we gained new insights about the role of CL and protein topology in membrane biogenesis. ICMs formation even in CL-deficient mutant shows that the lipid presence is necessary to the structuration of membranes but it is not essential to produce higher amount of proteins. It could be interesting to understand if and in which way the other phospholipid interact with the F-ATPase b subunit upon CL depletion. Additionally, identifying the minimal fragment of the F-ATPase b subunit capable to trigger ICMs formation could be interesting in order to use it as a tag to address any foreign MP to the ICM. Indeed, it has already been demonstrated that the fragment from the residue 24 until 156 of the F-ATPase b subunit owns dimerization features at 20°C and is essential for ICM formation (Revington et al., 1999). Shorter version of the soluble part of the F-ATPase b subunit should be tested for ICM formation in C43(DE3).

ICMs in *E. coli* could be a suitable model system to study not only the dynamics of membranes formation in bacteria but also the evolutionary mechanism of protein-lipid interaction and membrane biogenesis in bacteria, cyanobacteria, mitochondria and chloroplast.

From a molecular point of view, most of the work still remains to be done. At a descriptive level we have gained insight about the mRNA expression at different time points after IPTG induction. However, exploring if the additional mutations that have arisen in C43(DE3) are linked to ICMs formation could be a fair question, especially because C44(DE3) and C43(DE3) have common genetic features. An interesting experiment to be done, could be the transcriptome analysis between 5 minutes and 3 hours after induction by employing as control C43(DE3) cells over-producing another MP that cannot trigger ICMs formation.

In C43(DE3) strain, the pattern of expression 5 minutes after induction seems to be linked to the high production of LamB and the deletion of the sRNA micF, which

cannot negatively regulate OmpF. It should be investigated if in a C43(DE3) deleted from *ompF* gene the same happens.

Furthermore, F-ATPase b subunit over-expression in C43(DE3) *pspABC* deletion strains could give other information about the importance of the stress response. A more in-depth analysis of the C43(DE3) proteome is ongoing, as well, in order to cross transcriptomic and quantitative proteomic data.

So far, a well-studied case of ICMs formation in bacteria is the Alphaproteobacterium *Magnetospirillum gryphiswaldense* which contains well-ordered nanocrystals for magnetic navigation (Raschdorf et al., 2016). The authors identified the gene complement controlling membranes formation by using a cryo-electron tomography. In a similar fashion, this technique could be used to gain insights about where ICMs formation starts and what are the very first wheels to be activated. It might be really interesting, in addition, from an evolutionary point of view, if some genes identified in *M. gryphiswaldense* find their counterparts in *E. coli* C43(DE3).

Bibliography

1. Acehan, D., Malhotra, A., Xu, Y., Ren, M., Stokes, D.L., and Schlame, M. (2011). Cardiolipin Affects the Supramolecular Organization of ATP Synthase in Mitochondria. Biophys. J. *100*, 2184–2192.

2. Ades, S.E., Connolly, L.E., Alba, B.M., and Gross, C.A. (1999). The Escherichia coli çE-dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti-ç factor. Genes Dev. *13*, 2449–2461.

3. Alfasi, S., Sevastsyanovich, Y., Zaffaroni, L., Griffiths, L., Hall, R., and Cole, J. (2011). Use of GFP fusions for the isolation of Escherichia coli strains for improved production of different target recombinant proteins. J. Biotechnol. *156*, 11–21.

4. Almén, M., Nordstr?m, K.J., Fredriksson, R., and Schi?th, H.B. (2009). Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. BMC Biol. *7*, 50.

5. Angius, F., Ilioaia, O., Uzan, M., and Miroux, B. (2016). Membrane Protein Production in Escherichia coli: Protocols and Rules. Methods Mol. Biol. Clifton NJ *1432*, 37–52.

6. Arditti, R.R., Scaife, J.G., and Beckwith, J.R. (1968). The nature of mutants in the lac promoter region. J. Mol. Biol. *38*, 421–426.

7. Arechaga, I. (2013). Membrane Invaginations in Bacteria and Mitochondria: Common Features and Evolutionary Scenarios. J. Mol. Microbiol. Biotechnol. *23*, 13–23.

8. Arechaga, I., Miroux, B., Karrasch, S., Huijbregts, R., de Kruijff, B., Runswick, M.J., and Walker, J.E. (2000). Characterisation of new intracellular membranes in Escherichia coli accompanying large scale over-production of the b subunit of F1Fo ATP synthase. FEBS Lett. *482*, 215–219.

9. Armour, G.A., and Brewer, G.J. (1990). Membrane morphogenesis from cloned fragments of bacteriophage PM2 DNA that contain the sp6.6 gene. FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. *4*, 1488–1493.

10. Arnold, I., Pfeiffer, K., Neupert, W., Stuart, R.A., and Schägger, H. (1998). Yeast mitochondrial F1F0-ATP synthase exists as a dimer: identification of three dimer-specific subunits. EMBO J. *17*, 7170–7178.

11. Arnold, I., Pfeiffer, K., Neupert, W., Stuart, R.A., and Schägger, H. (1999). ATP Synthase of Yeast Mitochondria ISOLATION OF SUBUNIT j AND DISRUPTION OF THE ATP18GENE. J. Biol. Chem. *274*, 36–40.

12. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. *2*.

13. Bäcklund, E., Ignatushchenko, M., and Larsson, G. (2011). Suppressing glucose uptake and acetic acid production increases membrane protein overexpression in Escherichia coli. Microb. Cell Factories *10*, 35.

14. Bader, J., and Teuber, M. (1973). Action of polymyxin B on bacterial membranes. 1. Binding to the O-antigenic lipopolysaccharide of Salmonella typhimurium. Z. Naturforschung Teil C Biochem. Biophys. Biol. Virol. *28*, 422–430.

15. Bakowski, M.A., Braun, V., Lam, G.Y., Yeung, T., Heo, W.D., Meyer, T., Finlay, B.B., Grinstein, S., and Brumell, J.H. (2010). The phosphoinositide phosphatase SopB manipulates membrane surface charge and trafficking of the Salmonella-containing vacuole. Cell Host Microbe *7*, 453–462.

16. Baneyx, F., and Georgiou, G. (1990). In vivo degradation of secreted fusion proteins by the Escherichia coli outer membrane protease OmpT. J. Bacteriol. *172*, 491–494.

17. Barchinger, S.E., and Ades, S.E. (2013). Regulated proteolysis: control of the Escherichia coli σ(E)-dependent cell envelope stress response. Subcell. Biochem. *66*, 129–160.

18. Bastiaanse, E.M., Höld, K.M., and Van der Laarse, A. (1997). The effect of membrane cholesterol content on ion transport processes in plasma membranes. Cardiovasc. Res. *33*, 272–283.

19. Baumgarten, T., Schlegel, S., Wagner, S., Löw, M., Eriksson, J., Bonde, I., Herrgård, M.J., Heipieper, H.J., Nørholm, M.H.H., Slotboom, D.J., et al. (2017). Isolation and characterization of the E. coli membrane protein production strain Mutant56(DE3). Sci. Rep. *7*, 45089.

20. Bernaudat, F., Frelet-Barrand, A., Pochon, N., Dementin, S., Hivin, P., Boutigny, S., Rioux, J.-B., Salvi, D., Seigneurin-Berny, D., Richaud, P., et al. (2011). Heterologous Expression of Membrane Proteins: Choosing the Appropriate Host. PLoS ONE *6*.

21. Bochtler, M., Hartmann, C., Song, H.K., Bourenkov, G.P., Bartunik, H.D., and Huber, R. (2000). The structures of HslU and the ATP-dependent protease HslU–HslV. Nature *403*, 800–805.

22. Boedeker, C., Schüler, M., Reintjes, G., Jeske, O., Teeseling, M.C.F. van, Jogler, M., Rast, P., Borchert, D., Devos, D.P., Kucklick, M., et al. (2017). Determining the bacterial cell biology of Planctomycetes. Nat. Commun. *8*, ncomms14853.

23. Boël, G., Letso, R., Neely, H., Price, W.N., Wong, K.-H., Su, M., Luff, J.D., Valecha, M., Everett, J.K., Acton, T.B., et al. (2016). Codon influence on protein expression in E. coli correlates with mRNA levels. Nature *529*, 358–363.

24. Bogdanov, M., Mileykovskaya, E., and Dowhan, W. (2008). Lipids in the assembly of membrane proteins and organization of protein supercomplexes: implications for lipid-linked disorders. Subcell. Biochem. *49*, 197–239.

25. Bonner, G., Patra, D., Lafer, E.M., and Sousa, R. (1992). Mutations in T7 RNA polymerase that support the proposal for a common polymerase active site structure. EMBO J. *11*, 3767.

26. Bossi, L., and Figueroa-Bossi, N. (2007). A small RNA downregulates LamB maltoporin in Salmonella. Mol. Microbiol. *65*, 799–810.

27. Bowen, W.R. (2006). Biomimetic separations – learning from the early development of biological membranes. Desalination *199*, 225–227.

28. van den Brink-van der Laan, E., Killian, J.A., and de Kruijff, B. (2004). Nonbilayer lipids affect peripheral and integral membrane proteins via changes in the lateral pressure profile. Biochim. Biophys. Acta *1666*, 275–288.

29. Brosse, A., Korobeinikova, A., Gottesman, S., and Guillier, M. (2016). Unexpected properties of sRNA promoters allow feedback control via regulation of a two-component system. Nucleic Acids Res. *44*, 9650–9666.

30. Brown, M.F. (2012). Curvature forces in membrane lipid-protein interactions. Biochemistry (Mosc.) *51*, 9782–9795.

31. Bujard, H., Gentz, R., Lanzer, M., Stueber, D., Mueller, M., Ibrahimi, I., Haeuptle, M.T., and Dobberstein, B. (1987). A T5 promoter-based transcription-translation system for the analysis of proteins in vitro and in vivo. Methods Enzymol. *155*, 416–433.

32. Calloni, G., Chen, T., Schermann, S.M., Chang, H.-C., Genevaux, P., Agostini, F., Tartaglia, G.G., Hayer-Hartl, M., and Hartl, F.U. (2012). DnaK functions as a central hub in the E. coli chaperone network. Cell Rep. *1*, 251–264.

33. Carlson, J.H., and Silhavy, T.J. (1993). Signal sequence processing is required for the assembly of LamB trimers in the outer membrane of Escherichia coli. J. Bacteriol. *175*, 3327–3334.

34. Carranza, G., Angius, F., Ilioaia, O., Solgadi, A., Miroux, B., and Arechaga, I. (2017). Cardiolipin plays an essential role in the formation of intracellular membranes in Escherichia coli. Biochim. Biophys. Acta BBA - Biomembr. *1859*, 1124–1132.

35. Casiraghi, M., Damian, M., Lescop, E., Point, E., Moncoq, K., Morellet, N., Levy, D., Marie, J., Guittet, E., Banères, J.-L., et al. (2016). Functional Modulation of a G Protein-Coupled Receptor Conformational Landscape in a Lipid Bilayer. J. Am. Chem. Soc. *138*, 11170–11175.

36. Castanié-Cornet, M.-P., Bruel, N., and Genevaux, P. (2014). Chaperone networking facilitates protein targeting to the bacterial cytoplasmic membrane. Biochim. Biophys. Acta *1843*, 1442–1456.

37. Chandrangsu, P., and Helmann, J.D. (2014). Sigma Factors in Gene Expression. In ELS, John Wiley & Sons Ltd, ed. (Chichester, UK: John Wiley & Sons, Ltd), p.

38. Chen, G., Hayhurst, A., Thomas, J.G., Harvey, B.R., Iverson, B.L., and Georgiou, G. (2001). Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS). Nat. Biotechnol. *19*, 537–542.

39. Chen, Y., Song, J., Sui, S., and Wang, D.-N. (2003). DnaK and DnaJ facilitated the folding process and reduced inclusion body formation of magnesium transporter CorA overexpressed in Escherichia coli. Protein Expr. Purif. *32*, 221–231.

40. Chung, H. j., Bang, W., and Drake, M. a. (2006). Stress Response of Escherichia coli. Compr. Rev. Food Sci. Food Saf. *5*, 52–64.

41. Collinson, I. (2017). SecA-a New Twist in the Tale. J. Bacteriol. *199*, e00736-16.

42. Cooper, V.S., Schneider, D., Blot, M., and Lenski, R.E. (2001). Mechanisms causing rapid and parallel losses of ribose catabolism in evolving populations of Escherichia coli B. J. Bacteriol. *183*, 2834–2841.

43. Cormack, B.P., Valdivia, R.H., and Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP).

44. Cornejo, E., Subramanian, P., Li, Z., Jensen, G.J., and Komeili, A. (2016). Dynamic Remodeling of the Magnetosome Membrane Is Triggered by the Initiation of Biomineralization. MBio *7*, e01898-15.

45. Cornelius, F. (1991). Functional reconstitution of the sodium pump. Kinetics of exchange reactions performed by reconstituted Na/K-ATPase. Biochim. Biophys. Acta BBA-Rev. Biomembr. *1071*, 19–66.

46. Cosma, C.L., Danese, P.N., Carlson, J.H., Silhavy, T.J., and Snyder, W.B. (1995). Mutational activation of the Cpx signal transduction pathway of Escherichia coli suppresses the toxicity conferred by certain envelope-associated stresses. *Molecular Microbiology*.

47. Cronan, J.E. (2003). Bacterial Membrane Lipids: Where Do We Stand? Annu. Rev. Microbiol. *57*, 203–224.

48. Cullis, P.R., Verkleij, A.J., and Ververgaert, P.H.J.T. (1978). Polymorphic phase behaviour of cardiolipin as detected by 31P NMR and freeze-fracture techniques. Effects of calcium, dibucaine and chlorpromazine. Biochim. Biophys. Acta BBA - Biomembr. *513*, 11–20.

49. Cymer, F., and von Heijne, G. (2013). Cotranslational folding of membrane proteins probed by arrest-peptide-mediated force measurements. Proc. Natl. Acad. Sci. U. S. A. *110*, 14640–14645.

50. Cymer, F., von Heijne, G., and White, S.H. (2015). Mechanisms of integral membrane protein insertion and folding. J. Mol. Biol. *427*, 999–1022.

51. Dahlberg, M. (2007). Polymorphic Phase Behavior of Cardiolipin Derivatives Studied by Coarse-Grained Molecular Dynamics. J. Phys. Chem. B *111*, 7194–7200.

52. Dalbey, R.E., Kuhn, A., Zhu, L., and Kiefer, D. (2014). The membrane insertase YidC. Biochim. Biophys. Acta *1843*, 1489–1496.

53. Dalebroux, Z.D., and Swanson, M.S. (2012). ppGpp: magic beyond RNA polymerase. Nat. Rev. Microbiol. *10*, 203–212.

54. Danne, L., Aktas, M., Unger, A., Linke, W.A., Erdmann, R., and Narberhaus, F. (2017). Membrane Remodeling by a Bacterial Phospholipid-Methylating Enzyme. MBio *8*, e02082-16.

55. Darwin, A.J. (2005). The phage-shock-protein response. Mol. Microbiol. *57*, 621–628.

56. De Wulf, P., McGuire, A.M., Liu, X., and Lin, E.C.C. (2002). Genome-wide profiling of promoter recognition by the two-component response regulator CpxR-P in Escherichia coli. J. Biol. Chem. *277*, 26652–26661.

57. Delarue, M., Poch, O., Tordo, N., Moras, D., and Argos, P. (1990). An attempt to unify the structure of polymerases. Protein Eng. Des. Sel. *3*, 461–467.

58. Delhaye, A., Collet, J.-F., and Laloux, G. (2016). Fine-Tuning of the Cpx Envelope Stress Response Is Required for Cell Wall Homeostasis in Escherichia coli. MBio *7*, e00047-16.

59. Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A., and Bukau, B. (1999). Trigger factor and DnaK cooperate in folding of newly synthesized proteins. Nature *400*, 693–696.

60. Devos, D.P., and Reynaud, E.G. (2010). Evolution. Intermediate steps. Science *330*, 1187–1188.

61. Diaz-Villanueva, J., D?az-Molina, R., and Garc?a-Gonz?lez, V. (2015). Protein Folding and Mechanisms of Proteostasis. Int. J. Mol. Sci. *16*, 17193–17230.

62. Dowhan, W., and Bogdanov, M. (2012). Molecular genetic and biochemical approaches for defining lipid-dependent membrane protein folding. Biochim. Biophys. Acta *1818*, 1097–1107.

63. Dragosits, M., and Mattanovich, D. (2013). Adaptive laboratory evolution–principles and applications for biotechnology. Microb. Cell Factories *12*, 64.

64. Drew, D., Slotboom, D.-J., Friso, G., Reda, T., Genevaux, P., Rapp, M., Meindl-Beinker, N.M., Lambert, W., Lerch, M., Daley, D.O., et al. (2005). A scalable, GFP-based pipeline for membrane protein overexpression screening and purification. Protein Sci. *14*, 2011–2017.

65. Drew, D., Lerch, M., Kunji, E., Slotboom, D.-J., and de Gier, J.-W. (2006). Optimization of membrane protein overexpression and purification using GFP fusions. Nat. Methods *3*, 303–313.

66. Drew, D., Newstead, S., Sonoda, Y., Kim, H., von Heijne, G., and Iwata, S. (2008). GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in Saccharomyces cerevisiae. Nat. Protoc. *3*, 784–798.

67. Drew, D.E., von Heijne, G., Nordlund, P., and de Gier, J.-W.L. (2001). Green fluorescent protein as an indicator to monitor membrane protein overexpression in Escherichia coli. FEBS Lett. *507*, 220–224.

68. Dumon-Seignovert, L., Cariot, G., and Vuillard, L. (2004). The toxicity of recombinant proteins in Escherichia coli: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). Protein Expr. Purif. *37*, 203–206.

69. Durfee, T., Hansen, A.-M., Zhi, H., Blattner, F.R., and Jin, D.J. (2008). Transcription Profiling of the Stringent Response in Escherichia coli. J. Bacteriol. *190*, 1084–1096.

70. Engelman, D.M. (2005). Membranes are more mosaic than fluid. Nature *438*, 578–580.

71. Engl, C., Jovanovic, G., Lloyd, L.J., Murray, H., Spitaler, M., Ying, L., Errington, J., and Buck, M. (2009). In vivo localizations of membrane stress controllers PspA and PspG in Escherichia coli. Mol. Microbiol. *73*, 382–396.

72. Epand, R.F., Tokarska-Schlattner, M., Schlattner, U., Wallimann, T., and Epand, R.M. (2007). Cardiolipin Clusters and Membrane Domain Formation Induced by Mitochondrial Proteins. J. Mol. Biol. *365*, 968–980.

73. Epshtein, V., and Nudler, E. (2003). Cooperation Between RNA Polymerase Molecules in Transcription Elongation. Science *300*, 801–805.

74. Eriksson, H.M., Wessman, P., Ge, C., Edwards, K., and Wieslander, Å. (2009). Massive Formation of Intracellular Membrane Vesicles in Escherichia coli by a Monotopic Membrane-bound Lipid Glycosyltransferase. J. Biol. Chem. *284*, 33904.

75. Esposito, D., and Chatterjee, D.K. (2006). Enhancement of soluble protein expression through the use of fusion tags. Curr. Opin. Biotechnol. *17*, 353–358.

76. Facey, S.J., Neugebauer, S.A., Krauss, S., and Kuhn, A. (2007). The mechanosensitive channel protein MscL is targeted by the SRP to the novel YidC membrane insertion pathway of Escherichia coli. J. Mol. Biol. *365*, 995–1004.

77. Fairman, J.W., Noinaj, N., and Buchanan, S.K. (2011). The structural biology of β-barrel membrane proteins: a summary of recent reports. Curr. Opin. Struct. Biol. *21*, 523–531.

78. Fekkes, P., and Driessen, A.J.M. (1999). Protein Targeting to the Bacterial Cytoplasmic Membrane. Microbiol. Mol. Biol. Rev. *63*, 161–173.

79. Finkel, S.E., and Johnson, R.C. (1992). The Fis protein: it's not just for DNA inversion anymore. Mol. Microbiol. *6*, 3257–3265.

80. Flores-Kim, J., and Darwin, A.J. (2016). The Phage Shock Protein Response. Annu. Rev. Microbiol. *70*, 83–101.

81. Folch, J., Lees, M., Sloane-Stanley, G.H., and others (1957). A simple method for the isolation and purification of total lipids from animal tissues. J Biol Chem *226*, 497–509.

82. Forterre, P., and Gribaldo, S. (2010). Bacteria with a eukaryotic touch: A glimpse of ancient evolution? Proc. Natl. Acad. Sci. U. S. A. *107*, 12739–12740.

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

84. Freundlieb, S., Ehmann, U., and Boos, W. (1988). Facilitated diffusion of p-nitrophenyl-alpha-D-maltohexaoside through the outer membrane of Escherichia coli. Characterization of LamB as a specific and saturable channel for maltooligosaccharides. *J.Biol.Chem.*, 314–320.

85. Fröbel, J., Rose, P., and Müller, M. (2012). Twin-arginine-dependent translocation of folded proteins. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *367*, 1029–1046.

86. Fuerst, J.A., and Sagulenko, E. (2011). Beyond the bacterium: planctomycetes challenge our concepts of microbial structure and function. Nat. Rev. Microbiol. *9*, 403–413.

87. Gallet, P.F., Maftah, A., Petit, J.-M., Denis-Gay, M., and Julien, R. (1995). Direct Cardiolipin Assay in Yeast Using the Red Fluorescence Emission of 10-N-Nonyl Acridine Orange. Eur. J. Biochem. *228*, 113–119.

88. Ganz, T., and Lehrer, R.I. (1998). Antimicrobial peptides of vertebrates. Curr. Opin. Immunol. *10*, 41–44.

89. Garcia Fernandez, M.I., Ceccarelli, D., and Muscatello, U. (2004). Use of the fluorescent dye 10-N-nonyl acridine orange in quantitative and location assays of cardiolipin: a study on different experimental models. Anal. Biochem. *328*, 174–180.

90. Garrett, T.A., O'Neill, A.C., and Hopson, M.L. (2012). Quantification of cardiolipin molecular species in Escherichia coli lipid extracts using liquid chromatography/electrospray ionization mass spectrometry. Rapid Commun. Mass Spectrom. *26*, 2267–2274.

91. Ge, C., Gómez-Llobregat, J., Skwark, M.J., Ruysschaert, J.-M., Wieslander, Å., and Lindén, M. (2014). Membrane remodeling capacity of a vesicle-inducing glycosyltransferase. FEBS J. *281*, 3667–3684.

92. Geertsma, E.R., Groeneveld, M., Slotboom, D.-J., and Poolman, B. (2008). Quality control of overexpressed membrane proteins. Proc. Natl. Acad. Sci. *105*, 5722–5727.

93. Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. *5*, R80.

94. Georgescauld, F., Popova, K., Gupta, A.J., Bracher, A., Engen, J.R., Hayer-Hartl, M., and Hartl, F.U. (2014). GroEL/ES chaperonin modulates the mechanism and accelerates the rate of TIM-barrel domain folding. Cell *157*, 922–934.

95. Giacalone, M.J., Gentile, A.M., Lovitt, B.T., Berkley, N.L., Gunderson, C.W., and Surber, M.W. (2006). Toxic protein expression in Escherichia coli using a rhamnose-based tightly regulated and tunable promoter system. BioTechniques *40*, 355–364.

96. Gialama, D., Kostelidou, K., Michou, M., Delivoria, D.C., Kolisis, F.N., and Skretas, G. (2017). Development of *Escherichia coli* Strains That Withstand Membrane Protein-Induced Toxicity and Achieve High-Level Recombinant Membrane Protein Production. ACS Synth. Biol. *6*, 284–300.

97. Goldbeck, C.P., Jensen, H.M., TerAvest, M.A., Beedle, N., Appling, Y., Hepler, M., Cambray, G., Mutalik, V., Angenent, L.T., and Ajo-Franklin, C.M. (2013). Tuning Promoter Strengths for Improved Synthesis and Function of Electron Conduits in Escherichia coli. ACS Synth. Biol. *2*, 150–159.

98. González-Gil, G., Bringmann, P., and Kahmann, R. (1996). FIS is a regulator of metabolism in Escherichia coli. Mol. Microbiol. *22*.

99. Gordon, E., Horsefield, R., Swarts, H.G.P., de Pont, J.J.H.H.M., Neutze, R., and Snijder, A. (2008). Effective high-throughput overproduction of membrane proteins in Escherichia coli. Protein Expr. Purif. *62*, 1–8.

100. Grabowicz, M., and Silhavy, T.J. (2017). Envelope Stress Responses: An Interconnected Safety Net. Trends Biochem. Sci. *42*, 232–242.

101. Graham, T.R., and Kozlov, M.M. (2010). Interplay of proteins and lipids in generating membrane curvature. Curr. Opin. Cell Biol. *22*, 430–436.

102. Greene, S.E., and Komeili, A. (2012). Biogenesis and subcellular organization of the magnetosome organelles of magnetotactic bacteria. Curr. Opin. Cell Biol. *24*, 490–495.

103. Greenspan, P., Mayer, E.P., and Fowler, S.D. (1985). Nile red: a selective fluorescent stain for intracellular lipid droplets. J. Cell Biol. *100*, 965–973.

104. Grogan, D.W., and Cronan, J.E. (1997). Cyclopropane ring formation in membrane lipids of bacteria. Microbiol. Mol. Biol. Rev. *61*, 429–441.

105. Grünberg, K., Wawer, C., Tebo, B.M., and Schüler, D. (2001). A large gene cluster encoding several magnetosome proteins is conserved in different species of magnetotactic bacteria. Appl. Environ. Microbiol. *67*, 4573–4582.

106. Gubellini, F., Verdon, G., Karpowich, N.K., Luff, J.D., Boel, G., Gauthier, N., Handelman, S.K., Ades, S.E., and Hunt, J.F. (2011). Physiological Response to Membrane Protein Overexpression in E. coli. Mol. Cell. Proteomics *10*, M111.007930-M111.007930.

107. Guillier, M., and Gottesman, S. (2006). Remodelling of the Escherichia coli outer membrane by two small regulatory RNAs. Mol. Microbiol. *59*, 231–247.

108. Gul, N., Linares, D.M., Ho, F.Y., and Poolman, B. (2014). Evolved Escherichia coli Strains for Amplified, Functional Expression of Membrane Proteins. J. Mol. Biol. *426*, 136–149.

109. Gur, E., Biran, D., and Ron, E.Z. (2011). Regulated proteolysis in Gram-negative bacteria--how and when? Nat. Rev. Microbiol. *9*, 839–848.

110. Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J. Bacteriol. *177*, 4121–4130.

111. Ha, S., Walker, D., Shi, Y., and Walker, S. (2000). The 1.9 Å crystal structure of Escherichia coli MurG, a membrane-associated glycosyltransferase involved in peptidoglycan biosynthesis. Protein Sci. *9*, 1045–1052.

112. Haines, T.H. (2009). A new look at Cardiolipin. Biochim. Biophys. Acta BBA - Biomembr. *1788*, 1997–2002.

113. Hainrichson, M., Nudelman, I., and Baasov, T. (2008). Designer aminoglycosides: the race to develop improved antibiotics and compounds for the treatment of human genetic diseases. Org. Biomol. Chem. *6*, 227–239.

114. Hammon, J., Palanivelu, D.V., Chen, J., Patel, C., Minor, D.L., and * (2009). A green fluorescent protein screen for identification of well-expressed membrane proteins from a cohort of extremophilic organisms. Protein Sci. Publ. Protein Soc. *18*, 121.

115. Hancock, R.E. (1987). Role of porins in outer membrane permeability. J. Bacteriol. *169*, 929–933.

116. Hatahet, F., Boyd, D., and Beckwith, J. (2014). Disulfide bond formation in prokaryotes: History, diversity and design. Biochim. Biophys. Acta BBA - Proteins Proteomics *1844*, 1402–1414.

117. Hattab, G., Warschawski, D.E., Moncoq, K., and Miroux, B. (2015). Escherichia coli as host for membrane protein structure determination: a global analysis. Sci. Rep. *5*.

118. Heijne, G. von (1986). The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. EMBO J. *5*, 3021–3027.

119. Hengge-Aronis, R. (2002). Signal Transduction and Regulatory Mechanisms Involved in Control of the σS (RpoS) Subunit of RNA Polymerase. Microbiol. Mol. Biol. Rev. *66*, 373–395.

120. Herman, C., Prakash, S., Lu, C.Z., Matouschek, A., and Gross, C.A. (2003). Lack of a robust unfoldase activity confers a unique level of substrate specificity to the universal AAA protease FtsH. Mol. Cell *11*, 659–669.

121. Herskovits, A.A., Shimoni, E., Minsky, A., and Bibi, E. (2002). Accumulation of endoplasmic membranes and novel membrane-bound ribosome–signal recognition particle receptor complexes in *Escherichia coli*. J. Cell Biol. *159*, 403–410.

122. Hirano, Y., Hossain, M.M., Takeda, K., Tokuda, H., and Miki, K. (2007). Structural studies of the Cpx pathway activator NlpE on the outer membrane of Escherichia coli. Struct. Lond. Engl. 1993 *15*, 963–976.

123. Hiraoka, S., Matsuzaki, H., and Shibuya, I. (1993). Active increase in cardiolipin synthesis in the stationary growth phase and its physiological significance in Escherichia coli. FEBS Lett. *336*, 221–224.

124. Hjelm, A., Karyolaimos, A., Zhang, Z., Rujas, E., Vikström, D., Slotboom, D.J., and de Gier, J.-W. (2017). Tailoring Escherichia coli for the l-Rhamnose PBAD Promoter-Based Production of Membrane and Secretory Proteins. ACS Synth. Biol. *6*, 985–994.

125. Huang, K.C., and Ramamurthi, K.S. (2010). Macromolecules that prefer their membranes curvy. Mol. Microbiol. *76*, 822–832.

126. Huber, D., Boyd, D., Xia, Y., Olma, M.H., Gerstein, M., and Beckwith, J. (2005). Use of thioredoxin as a reporter to identify a subset of Escherichia coli signal sequences that promote signal recognition particle-dependent translocation. J. Bacteriol. *187*, 2983–2991.

127. Huvet, M., Toni, T., Sheng, X., Thorne, T., Jovanovic, G., Engl, C., Buck, M., Pinney, J.W., and Stumpf, M.P.H. (2011). The evolution of the phage shock protein response system: interplay between protein function, genomic organization, and system function. Mol. Biol. Evol. *28*, 1141–1155.

128. Iost, I., Guillerez, J., and Dreyfus, M. (1992). Bacteriophage T7 RNA polymerase travels far ahead of ribosomes in vivo. J. Bacteriol. *174*, 619–622.

129. Ippen, K., Miller, J.H., Scaife, J., and Beckwith, J. (1968). New Controlling Element in the Lac Operon of E. coli. Nature *217*, 825–827.

130. Israelachvili, J.N., Mitchell, D.J., and Ninham, B.W. (1977). Theory of self-assembly of lipid bilayers and vesicles. Biochim. Biophys. Acta BBA - Biomembr. *470*, 185–201.

131. Jackson, M.R., Melideo, S.L., and Jorns, M.S. (2012). Human Sulfide:Quinone Oxidoreductase Catalyzes the First Step in Hydrogen Sulfide Metabolism and Produces a Sulfane Sulfur Metabolite. Biochemistry (Mosc.) *51*, 6804–6815.

132. Jahn, T.R., and Radford, S.E. (2005). The Yin and Yang of protein folding: The Yin and Yang of protein folding. FEBS J. *272*, 5962–5970.

133. Jarsch, I.K., Daste, F., and Gallop, J.L. (2016). Membrane curvature in cell biology: An integration of molecular mechanisms. J. Cell Biol. *214*, 375–387.

134. Jia, B., and Jeon, C.O. (2016). High-throughput recombinant protein expression in *Escherichia coli* : current status and future perspectives. Open Biol. *6*, 160196.

135. Johansson, J., Dagberg, B., Richet, E., and Uhlin, B.E. (1998). H-NS and StpA proteins stimulate expression of the maltose regulon in Escherichia coli. J. Bacteriol. *180*, 6117–6125.

136. Joly, N., Engl, C., Jovanovic, G., Huvet, M., Toni, T., Sheng, X., Stumpf, M.P.H., and Buck, M. (2010). Managing membrane stress: the phage shock protein (Psp) response, from molecular mechanisms to physiology. FEMS Microbiol. Rev. *34*, 797–827.

137. Jovanovic, G., Mehta, P., Ying, L., and Buck, M. (2014). Anionic lipids and the cytoskeletal proteins MreB and RodZ define the spatio-temporal distribution and function of membrane stress controller PspA in Escherichia coli. Microbiology *160*, 2374–2386.

138. Kaewsuya, P., Danielson, N.D., and Ekhterae, D. (2007). Fluorescent determination of cardiolipin using 10-N-nonyl acridine orange. Anal. Bioanal. Chem. *387*, 2775–2782.

139. Kalivoda, K.A., Steenbergen, S.M., and Vimr, E.R. (2013). Control of the Escherichia coli Sialoregulon by Transcriptional Repressor NanR. J. Bacteriol. *195*, 4689–4701.

140. Kaneda, T. (1991). Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance. Microbiol. Rev. *55*, 288–302.

141. Kates, M., Syz, J.-Y., Gosser, D., and Haines, T.H. (1993). pH-dissociation characteristics of cardiolipin and its 2'-deoxy analogue. Lipids *28*, 877–882.

142. Katz, C., and Ron, E.Z. (2008). Dual Role of FtsH in Regulating Lipopolysaccharide Biosynthesis in Escherichia coli. J. Bacteriol. *190*, 7117–7122.

143. Kawate, T., and Gouaux, E. (2006). Fluorescence-Detection Size-Exclusion Chromatography for Precrystallization Screening of Integral Membrane Proteins. Structure *14*, 673–681.

144. Ke, N., Landgraf, D., Paulsson, J., and Berkmen, M. (2016). Visualization of Periplasmic and Cytoplasmic Proteins with a Self-Labeling Protein Tag. J. Bacteriol. *198*, 1035–1043.

145. Keller, R., Ariöz, C., Hansmeier, N., Stenberg-Bruzell, F., Burstedt, M., Vikström, D., Kelly, A., Wieslander, Å., Daley, D.O., and Hunke, S. (2015). The Escherichia coli Envelope Stress Sensor CpxA Responds to Changes in Lipid Bilayer Properties. Biochemistry (Mosc.) *54*, 3670–3676.

146. Keren, K. (2011). Cell motility: the integrating role of the plasma membrane. Eur. Biophys. J. *40*, 1013–1027.

147. Khalifat, N., Fournier, J.-B., Angelova, M.I., and Puff, N. (2011). Lipid packing variations induced by pH in cardiolipin-containing bilayers: The driving force for the cristae-like shape instability. Biochim. Biophys. Acta BBA - Biomembr. *1808*, 2724–2733.

148. Kihara, A., Akiyama, Y., and Ito, K. (1996). A protease complex in the Escherichia coli plasma membrane: HflKC (HflA) forms a complex with FtsH (HflB), regulating its proteolytic activity against SecY. EMBO J. *15*, 6122–6131.

149. Kim, H.K., Choi, S.I., and Seong, B.L. (2010). 5S rRNA-assisted DnaK refolding. Biochem. Biophys. Res. Commun. *391*, 1177–1181.

150. Kim, S., Jeong, H., Kim, E.-Y., Kim, J.F., Lee, S.Y., and Yoon, S.H. (2017). Genomic and transcriptomic landscape of Escherichia coli BL21(DE3). Nucleic Acids Res. *45*, 5285.

151. Kim, Y.E., Hipp, M.S., Bracher, A., Hayer-Hartl, M., and Hartl, F.U. (2013). Molecular chaperone functions in protein folding and proteostasis. Annu. Rev. Biochem. *82*, 323–355.

152. Komeili, A. (2012). Molecular mechanisms of compartmentalization and biomineralization in magnetotactic bacteria. FEMS Microbiol. Rev. *36*, 232–255.

153. Komeili, A., Li, Z., Newman, D.K., and Jensen, G.J. (2006). Magnetosomes are cell membrane invaginations organized by the actin-like protein MamK. Science *311*, 242–245.

154. König, E., Schlesner, H., and Hirsch, P. (1984). Cell wall studies on budding bacteria of the Planctomyces/Pasteuria group and on a Prosthecomicrobium sp. Arch. Microbiol. *138*, 200–205.

155. Kremers, G.-J., Goedhart, J., van den Heuvel, D.J., Gerritsen, H.C., and Gadella, T.W.J. (2007). Improved Green and Blue Fluorescent Proteins for Expression in Bacteria and Mammalian Cells † , ‡. Biochemistry (Mosc.) *46*, 3775–3783.

156. de Kruijff, B., and Cullis, P.R. (1980). Cytochrome c specifically induces non-bilayer structures in cardiolipin-containing model membranes. Biochim. Biophys. Acta BBA - Biomembr. *602*, 477–490.

157. Kudva, R., Denks, K., Kuhn, P., Vogt, A., Müller, M., and Koch, H.-G. (2013). Protein translocation across the inner membrane of Gram-negative bacteria: the Sec and Tat dependent protein transport pathways. Res. Microbiol. *164*, 505–534.

158. Kwon, S.-K., Kim, S.K., Lee, D.-H., and Kim, J.F. (2015). Comparative genomics and experimental evolution of Escherichia coli BL21(DE3) strains reveal the landscape of toxicity escape from membrane protein overproduction. Sci. Rep. *5*.

159. Laage, S., Tao, Y., and McDermott, A.E. (2015). Cardiolipin interaction with subunit c of ATP synthase: solid-state NMR characterization. Biochim. Biophys. Acta *1848*, 260–265.

160. Laan, E. van den B. der, Boots, J.-W.P., Spelbrink, R.E.J., Kool, G.M., Breukink, E., Killian, J.A., and Kruijff, B. de (2003). Membrane Interaction of the Glycosyltransferase MurG: a Special Role for Cardiolipin. J. Bacteriol. *185*, 3773–3779.

161. Lange, R., and Hengge-Aronis, R. (1994). The cellular concentration of the sigma S subunit of RNA polymerase in Escherichia coli is controlled at the levels of transcription, translation, and protein stability. Genes Dev. *8*, 1600–1612.

162. Laubacher, M.E., and Ades, S.E. (2008). The Rcs phosphorelay is a cell envelope stress response activated by peptidoglycan stress and contributes to intrinsic antibiotic resistance. J. Bacteriol. *190*, 2065–2074.

163. Lee, A.G. (2004). How lipids affect the activities of integral membrane proteins. Biochim. Biophys. Acta *1666*, 62–87.

164. Lee, A.G. (2011). Biological membranes: the importance of molecular detail. Trends Biochem. Sci. *36*, 493–500.

165. Lee, H.C., and Bernstein, H.D. (2001). The targeting pathway of Escherichia coli presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. Proc. Natl. Acad. Sci. U. S. A. *98*, 3471–3476.

166. Lefman, J., Zhang, P., Hirai, T., Weis, R.M., Juliani, J., Bliss, D., Kessel, M., Bos, E., Peters, P.J., and Subramaniam, S. (2004). Three-Dimensional Electron Microscopic Imaging of Membrane Invaginations in Escherichia coli Overproducing the Chemotaxis Receptor Tsr. J. Bacteriol. *186*, 5052–5061.

167. Lehrer, R.I., Barton, A., Daher, K.A., Harwig, S.S., Ganz, T., and Selsted, M.E. (1989). Interaction of human defensins with Escherichia coli. Mechanism of bactericidal activity. J. Clin. Invest. *84*, 553–561.

168. Lenfant, F., Labia, R., and Masson, J.M. (1990). Probing the active site of β-lactamase R-TEM1 by informational suppression. Biochimie *72*, 495–503.

169. Lewis, M. (2005). The lac repressor. C. R. Biol. *328*, 521–548.

170. Lewis, R.N.A.H., and McElhaney, R.N. (2009). The physicochemical properties of cardiolipin bilayers and cardiolipin-containing lipid membranes. Biochim. Biophys. Acta BBA - Biomembr. *1788*, 2069–2079.

171. Li, G., and Young, K.D. (2012). Isolation and identification of new inner membrane-associated proteins that localize to cell poles in *Escherichia coli*: Identification of polar proteins. Mol. Microbiol. *84*, 276–295.

172. Lima, S., Guo, M.S., Chaba, R., Gross, C.A., and Sauer, R.T. (2013). Dual molecular signals mediate the bacterial response to outer-membrane stress. Science *340*, 837–841.

173. Lin, K.-F., Sun, C.-S., Huang, Y.-C., Chan, S.I., Koubek, J., Wu, T.-H., and Huang, J.J.-T. (2012). Cotranslational Protein Folding within the Ribosome Tunnel Influences Trigger-Factor Recruitment. Biophys. J. *102*, 2818–2827.

174. Linde, K., Gröbner, G., and Rilfors, L. (2004). Lipid dependence and activity control of phosphatidylserine synthase from Escherichia coli. FEBS Lett. *575*, 77–80.

175. Lindsay, M.R., Webb, R.I., and Fuerst, J.A. (1997). Pirellulosomes: A New Type of Membrane-Bounded Cell Compartment in Planctomycete Bacteria of the Genus Pirellula. Microbiology *143*, 739–748.

176. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

177. Luirink, J., Yu, Z., Wagner, S., and de Gier, J.-W. (2012). Biogenesis of inner membrane proteins in Escherichia coli. Biochim. Biophys. Acta BBA - Bioenerg. *1817*, 965–976.

178. Lundstrom, K. (2007). Structural genomics and drug discovery. J. Cell. Mol. Med. *11*, 224–238.

179. Lyons, J.A., Shahsavar, A., Paulsen, P.A., Pedersen, B.P., and Nissen, P. (2016). Expression strategies for structural studies of eukaryotic membrane proteins. Curr. Opin. Struct. Biol. *38*, 137–144.

180. Mahillon, J., Léonard, C., and Chandler, M. (1999). IS elements as constituents of bacterial genomes. Res. Microbiol. *150*, 675–687.

181. Majdalani, N., and Gottesman, S. (2005). The Rcs phosphorelay: a complex signal transduction system. Annu. Rev. Microbiol. *59*, 379–405.

182. Makino, T., Skretas, G., Kang, T.-H., and Georgiou, G. (2011). Comprehensive engineering of Escherichia coli for enhanced expression of IgG antibodies. Metab. Eng. *13*, 241–251.

183. Mannella, C.A., Lederer, W.J., and Jafri, M.S. (2013). The connection between inner membrane topology and mitochondrial function. J. Mol. Cell. Cardiol. *62*, 51–57.

184. Marciniak, B.C., Trip, H., van-der Veek, P.J., and Kuipers, O.P. (2012). Comparative transcriptional analysis of Bacillus subtilis cells overproducing either secreted proteins, lipoproteins or membrane proteins. Microb. Cell Factories *11*, 66.

185. Marius, P., Alvis, S.J., East, J.M., and Lee, A.G. (2005). The Interfacial Lipid Binding Site on the Potassium Channel KcsA Is Specific for Anionic Phospholipids. Biophys. J. *89*, 4081–4089.

186. Markiewicz, P., Kleina, L.G., Cruz, C., Ehret, S., and Miller, J.H. (1994). Genetic Studies of the lac Repressor. XIV. Analysis of 4000 Altered Escherichia coli lac Repressors Reveals Essential and Non-essential Residues, as well as “Spacers” which do not Require a Specific Sequence. J. Mol. Biol. *240*, 421–433.

187. Massey-Gendel, E., Zhao, A., Boulting, G., Kim, H.-Y., Balamotis, M.A., Seligman, L.M., Nakamoto, R.K., and Bowie, J.U. (2009). Genetic selection system for improving recombinant membrane protein expression in E. coli. Protein Sci. Publ. Protein Soc. *18*, 372–383.

188. McDonald, C., Jovanovic, G., Ces, O., and Buck, M. (2015). Membrane Stored Curvature Elastic Stress Modulates Recruitment of Maintenance Proteins PspA and Vipp1. MBio *6*.

189. McGarrity, J.T., and Armstrong, J.B. (1981). The effect of temperature and other growth conditions on the fatty acid composition of Escherichia coli. Can. J. Microbiol. *27*, 835–840.

190. McMahon, H.T., and Gallop, J.L. (2005). Membrane curvature and mechanisms of dynamic cell membrane remodelling. Nature *438*, 590–596.

191. Mechold, U., Murphy, H., Brown, L., and Cashel, M. (2002). Intramolecular Regulation of the Opposing (p)ppGpp Catalytic Activities of RelSeq, the Rel/Spo Enzyme from Streptococcus equisimilis. J. Bacteriol. *184*, 2878.

192. van Meer, G., Voelker, D.R., and Feigenson, G.W. (2008). Membrane lipids: where they are and how they behave. Nat. Rev. Mol. Cell Biol. *9*, 112–124.

193. Merz, F., Hoffmann, A., Rutkowska, A., Zachmann-Brand, B., Bukau, B., and Deuerling, E. (2006). The C-terminal Domain of Escherichia coli Trigger Factor Represents the Central Module of Its Chaperone Activity. J. Biol. Chem. *281*, 31963–31971.

194. Mileykovskaya, E., and Dowhan, W. (2000). Visualization of Phospholipid Domains in Escherichia coli by Using the Cardiolipin-Specific Fluorescent Dye 10-N-Nonyl Acridine Orange. J. Bacteriol. *182*, 1172.

195. Mileykovskaya, E., and Dowhan, W. (2009). Cardiolipin membrane domains in prokaryotes and eukaryotes. Biochim. Biophys. Acta BBA - Biomembr. *1788*, 2084–2091.

196. Mim, C., and Unger, V.M. (2012). Membrane curvature and its generation by BAR proteins. Trends Biochem. Sci. *37*, 526–533.

197. Miroux, B., and Walker, J.E. (1996). Over-production of Proteins inEscherichia coli: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. J. Mol. Biol. *260*, 289–298.

198. Model, P., Jovanovic, G., and Dworkin, J. (1997). The Escherichia coli phage-shock-protein (psp) operon. Mol. Microbiol. *24*, 255–261.

199. Mogk, A., Deuerling, E., Vorderwülbecke, S., Vierling, E., and Bukau, B. (2003). Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. Mol. Microbiol. *50*, 585–595.

200. Morita, M., Kanemori, M., Yanagi, H., and Yura, T. (1999). Heat-Induced Synthesis of ς32 in Escherichia coli: Structural and Functional Dissection of rpoH mRNA Secondary Structure. J. Bacteriol. *181*, 401–410.

201. Moulin, M., Solgadi, A., Veksler, V., Garnier, A., Ventura-Clapier, R., and Chaminade, P. (2015). Sex-specific cardiac cardiolipin remodelling after doxorubicin treatment. Biol. Sex Differ. *6*.

202. Mukherjee, S., Raghuraman, H., and Chattopadhyay, A. (2007). Membrane localization and dynamics of Nile Red: Effect of cholesterol. Biochim. Biophys. Acta BBA - Biomembr. *1768*, 59–66.

203. Muthukrishnan, A.-B., Kandhavelu, M., Lloyd-Price, J., Kudasov, F., Chowdhury, S., Yli-Harja, O., and Ribeiro, A.S. (2012). Dynamics of transcription driven by the tetA

promoter, one event at a time, in live Escherichia coli cells. Nucleic Acids Res. *40*, 8472–8483.

204. Nagamori, S., Smirnova, I.N., and Kaback, H.R. (2004). Role of YidC in folding of polytopic membrane proteins. J. Cell Biol. *165*, 53–62.

205. Nannenga, B.L., and Baneyx, F. (2011). Reprogramming chaperone pathways to improve membrane protein expression in Escherichia coli. Protein Sci. Publ. Protein Soc. *20*, 1411–1420.

206. Natale, P., Brüser, T., and Driessen, A.J.M. (2008). Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. Biochim. Biophys. Acta *1778*, 1735–1756.

207. Nevo, R., Charuvi, D., Shimoni, E., Schwarz, R., Kaplan, A., Ohad, I., and Reich, Z. (2007). Thylakoid membrane perforations and connectivity enable intracellular traffic in cyanobacteria. EMBO J. *26*, 1467–1473.

208. Nichols-Smith, S., Teh, S.-Y., and Kuhl, T.L. (2004). Thermodynamic and mechanical properties of model mitochondrial membranes. Biochim. Biophys. Acta BBA - Biomembr. *1663*, 82–88.

209. Nicolson, G.L. (2014). The Fluid—Mosaic Model of Membrane Structure: Still relevant to understanding the structure, function and dynamics of biological membranes after more than 40 years. Biochim. Biophys. Acta BBA - Biomembr. *1838*, 1451–1466.

210. Nieboer, M., Vis, A.J., and Witholt, B. (1996). Overproduction of a foreign membrane protein in Escherichia coli stimulates and depends on phospholipid synthesis. Eur. J. Biochem. *241*, 691–696.

211. Nielsen, C., Goulian, M., and Andersen, O.S. (1998). Energetics of inclusion-induced bilayer deformations. Biophys. J. *74*, 1966–1983.

212. Nishijima, S., Asami, Y., Uetake, N., Yamagoe, S., Ohta, A., and Shibuya, I. (1988). Disruption of the Escherichia coli cls gene responsible for cardiolipin synthesis. J. Bacteriol. *170*, 775–780.

213. Nishino, K., Honda, T., and Yamaguchi, A. (2005). Genome-Wide Analyses of Escherichia coli Gene Expression Responsive to the BaeSR Two-Component Regulatory System. J. Bacteriol. *187*, 1763–1772.

214. Nonaka, G., Blankschien, M., Herman, C., Gross, C.A., and Rhodius, V.A. (2006). Regulon and promoter analysis of the E. coli heat-shock factor, σ32, reveals a multifaceted cellular response to heat stress. Genes Dev. *20*, 1776–1789.

215. Norholm, M.H.H., Toddo, S., Virkki, M.T.I., Light, S., von Heijne, G., and Daley, D.O. (2013). Improved production of membrane proteins in Escherichia coli by selective codon substitutions. FEBS Lett. *587*, 2352–2358.

216. Ochs, M., Veitinger, S., Kim, I., Weiz, D., Angerer, A., and Braun, V. (1995). Regulation of citrate-dependent iron transport of Escherichia coli: FecR is required for transcription activation by FeeI. Mol. Microbiol. *15*, 119–132.

217. Odoi, K.A., Huang, Y., Rezenom, Y.H., and Liu, W.R. (2013). Nonsense and Sense Suppression Abilities of Original and Derivative Methanosarcina mazei Pyrrolysyl-tRNA Synthetase-tRNAPyl Pairs in the Escherichia coli BL21(DE3) Cell Strain. PLoS ONE *8*, e57035.

218. Ohta, N., Kato, Y., Watanabe, H., Mori, H., and Matsuura, T. (2016). In vitro membrane protein synthesis inside Sec translocon-reconstituted cell-sized liposomes. Sci. Rep. *6*, srep36466.

219. Oliver, P.M., Crooks, J.A., Leidl, M., Yoon, E.J., Saghatelian, A., and Weibel, D.B. (2014). Localization of Anionic Phospholipids in Escherichia coli Cells. J. Bacteriol. *196*, 3386–3398.

220. Ortiz, A., Killian, J.A., Verkleij, A.J., and Wilschut, J. (1999). Membrane Fusion and the Lamellar-to-Inverted-Hexagonal Phase Transition in Cardiolipin Vesicle Systems Induced by Divalent Cations. Biophys. J. *77*, 2003–2014.

221. Osumi-Davis, P.A., Sreerama, N., Volkin, D.B., Middaugh, C.R., Woody, R.W., and Woody, A.-Y.M. (1994). Bacteriophage T7 RNA Polymerase and Its Active-site Mutants. J. Mol. Biol. *237*, 5–19.

222. Otters, S., Braun, P., Hubner, J., Wanner, G., Vothknecht, U.C., and Chigri, F. (2013). The first α-helical domain of the vesicle-inducing protein in plastids 1 promotes oligomerization and lipid binding. Planta *237*, 529–540.

223. Otto, K., and Silhavy, T.J. (2002). Surface sensing and adhesion of Escherichia coli controlled by the Cpx-signaling pathway. Proc. Natl. Acad. Sci. U. S. A. *99*, 2287–2292.

224. Overington, J.P., Al-Lazikani, B., and Hopkins, A.L. (2005). How many drug targets are there? Nat. Rev. *5*.

225. Palsdottir, H., and Hunte, C. (2004). Lipids in membrane protein structures. Biochim. Biophys. Acta BBA - Biomembr. *1666*, 2–18.

226. Pan, R., Jones, A.D., and Hu, J. (2014). Cardiolipin-Mediated Mitochondrial Dynamics and Stress Response in *Arabidopsis*. Plant Cell Online *26*, 391–409.

227. Papanastasiou, M., Orfanoudaki, G., Koukaki, M., Kountourakis, N., Sardis, M.F., Aivaliotis, M., Karamanou, S., and Economou, A. (2013). The Escherichia coli Peripheral Inner Membrane Proteome. Mol. Cell. Proteomics MCP *12*, 599.

228. Park, S.-C., Jia, B., Yang, J.-K., Van, D.L., Shao, Y.G., Han, S.W., Jeon, Y.-J., Chung, C.H., and Cheong, G.-W. (2006). Oligomeric structure of the ATP-dependent protease La (Lon) of Escherichia coli. Mol. Cells *21*, 129–134.

229. Parsons, J.B., and Rock, C.O. (2013). Bacterial lipids: metabolism and membrane homeostasis. Prog. Lipid Res. *52*, 249–276.

230. Paumard, P., Vaillier, J., Coulary, B., Schaeffer, J., Soubannier, V., Mueller, D.M., Brèthes, D., di Rago, J.-P., and Velours, J. (2002). The ATP synthase is involved in generating mitochondrial cristae morphology. EMBO J. *21*, 221–230.

231. Pechmann, S., and Frydman, J. (2012). Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. Nat. Struct. Mol. Biol. *20*, 237–243.

232. Pechmann, S., Chartron, J.W., and Frydman, J. (2014). Local slowdown of translation by nonoptimal codons promotes nascent-chain recognition by SRP in vivo. Nat. Struct. Mol. Biol. *21*, 1100–1105.

233. PETIT, J.-M., MAFTAH, A., RATINAUD, M.-H., and JULIEN, R. (1992). 10N-Nonyl acridine orange interacts with cardiolipin and allows the quantification of this phospholipid in isolated mitochondria. Eur. J. Biochem. *209*, 267–273.

234. du Plessis, D.J.F., Nouwen, N., and Driessen, A.J.M. (2011). The Sec translocase. Biochim. Biophys. Acta BBA - Biomembr. *1808*, 851–865.

235. Pohnert, G. (2002). Phospholipase A2 activity triggers the wound-activated chemical defense in the diatom Thalassiosira rotula. Plant Physiol. *129*, 103–111.

236. Powl, A.M., East, J.M., and Lee, A.G. (2007). Different effects of lipid chain length on the two sides of a membrane and the lipid annulus of MscL. Biophys. J. *93*, 113–122.

237. Powl, A.M., East, J.M., and Lee, A.G. (2008a). Anionic Phospholipids Affect the Rate and Extent of Flux through the Mechanosensitive Channel of Large Conductance MscL. Biochemistry (Mosc.) *47*, 4317–4328.

238. Powl, A.M., East, J.M., and Lee, A.G. (2008b). Importance of direct interactions with lipids for the function of the mechanosensitive channel MscL. Biochemistry (Mosc.) *47*, 12175–12184.

239. Price, N.L., and Raivio, T.L. (2009). Characterization of the Cpx Regulon in Escherichia coli Strain MC4100. J. Bacteriol. *191*, 1798–1815.

240. Raetz, C.R. (1978). Enzymology, genetics, and regulation of membrane phospholipid synthesis in Escherichia coli. Microbiol. Rev. *42*, 614–659.

241. Raffa, R.G., and Raivio, T.L. (2002). A third envelope stress signal transduction pathway in Escherichia coli. Mol. Microbiol. *45*, 1599–1611.

242. Raivio, T.L. (2014). Everything old is new again: An update on current research on the Cpx envelope stress response. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1843*, 1529–1541.

243. Raivio, T.L., and Silhavy, T.J. (1997). Transduction of envelope stress in Escherichia coli by the Cpx two-component system. J. Bacteriol. *179*, 7724–7733.

244. Rand, R.P., and Sengupta, S. (1972). Cardiolipin forms hexagonal structures with divalent cations. Biochim. Biophys. Acta BBA - Biomembr. *255*, 484–492.

245. Rand, R.P., Fuller, N.L., Gruner, S.M., and Parsegian, V.A. (1990). Membrane curvature, lipid segregation, and structural transitions for phospholipids under dual-solvent stress. Biochemistry (Mosc.) *29*, 76–87.

246. Randall-Hazelbauer, L., and Schwartz, M. (1973). Isolation of the Bacteriophage Lambda Receptor from Escherichia coli. J. Bacteriol. *116*, 1436–1446.

247. Rapp, M., Drew, D., Daley, D.O., Nilsson, J., Carvalho, T., Melén, K., De Gier, J.-W., and Von Heijne, G. (2004). Experimentally based topology models for E. coli inner membrane proteins. Protein Sci. *13*, 937–945.

248. Raschdorf, O., Forstner, Y., Kolinko, I., Uebe, R., Plitzko, J.M., and Sch?ler, D. (2016). Genetic and Ultrastructural Analysis Reveals the Key Players and Initial Steps of Bacterial Magnetosome Membrane Biogenesis. PLOS Genet. *12*, e1006101.

249. Redelberger, D., Seduk, F., Genest, O., Méjean, V., Leimkühler, S., and Iobbi-Nivol, C. (2011). YcdY Protein of Escherichia coli, an Atypical Member of the TorD Chaperone Family. J. Bacteriol. *193*, 6512–6516.

250. Reimann, S.A., and Wolfe, A.J. (2011). Constitutive Expression of the Maltoporin LamB in the Absence of OmpR Damages the Cell Envelope. J. Bacteriol. *193*, 842–853.

251. Renner, L.D., and Weibel, D.B. (2011). Cardiolipin microdomains localize to negatively curved regions of Escherichia coli membranes. *108*, 6264–6269.

252. Revington, M., McLachlin, D.T., Shaw, G.S., and Dunn, S.D. (1999). The Dimerization Domain of the b Subunit of theEscherichia coli F1F0-ATPase. J. Biol. Chem. *274*, 31094–31101.

253. Rhodius, V.A., Suh, W.C., Nonaka, G., West, J., and Gross, C.A. (2006). Conserved and variable functions of the sigmaE stress response in related genomes. PLoS Biol. *4*, e2.

254. Ringquist, S., Shinedling, S., Barrick, D., Green, L., Binkley, J., Stormo, G.D., and Gold, L. (1992). Translation initiation in Escherichia coli: sequences within the ribosome-binding site. Mol. Microbiol. *6*, 1219–1229.

255. Roepe, P.D., and Kaback, H.R. (1990). Isolation and functional reconstitution of soluble melibiose permease from Escherichia coli. Biochemistry (Mosc.) *29*, 2572–2577.

256. Ruiz, N., and Silhavy, T.J. (2005). Sensing external stress: watchdogs of the Escherichia coli cell envelope. Curr. Opin. Microbiol. *8*, 122–126.

257. Sachelaru, I., Petriman, N.A., Kudva, R., Kuhn, P., Welte, T., Knapp, B., Drepper, F., Warscheid, B., and Koch, H.-G. (2013). YidC occupies the lateral gate of the SecYEG

translocon and is sequentially displaced by a nascent membrane protein. J. Biol. Chem. *288*, 16295–16307.

258. saiSree, L., Reddy, M., and Gowrishankar, J. (2001). IS186 insertion at a hot spot in the lon promoter as a basis for lon protease deficiency of Escherichia coli B: identification of a consensus target sequence for IS186 transposition. J. Bacteriol. *183*, 6943–6946.

259. Samuelson, J.C., Chen, M., Jiang, F., Möller, I., Wiedmann, M., Kuhn, A., Phillips, G.J., and Dalbey, R.E. (2000). YidC mediates membrane protein insertion in bacteria. Nature *406*, 637–641.

260. Sanchez-Garcia, L., Martín, L., Mangues, R., Ferrer-Miralles, N., Vázquez, E., and Villaverde, A. (2016). Recombinant pharmaceuticals from microbial cells: a 2015 update. Microb. Cell Factories *15*.

261. Sankaram, M.B., Powell, G.L., and Marsh, D. (1989). Effect of acyl chain composition on salt-induced lamellar to inverted hexagonal phase transitions in cardiolipin. Biochim. Biophys. Acta BBA - Biomembr. *980*, 389–392.

262. Santos, A.X.S., and Riezman, H. (2012). Yeast as a model system for studying lipid homeostasis and function. FEBS Lett. *586*, 2858–2867.

263. Sarkar, C.A., Dodevski, I., Kenig, M., Dudli, S., Mohr, A., Hermans, E., and Plückthun, A. (2008). Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. Proc. Natl. Acad. Sci. *105*, 14808–14813.

264. Sarramegna, V., Talmont, F., de Roch, M.S., Milon, A., and Demange, P. (2002). Green fluorescent protein as a reporter of human μ-opioid receptor overexpression and localization in the methylotrophic yeast Pichia pastoris. J. Biotechnol. *99*, 23–39.

265. Sauer, R.T., and Baker, T.A. (2011). AAA+ proteases: ATP-fueled machines of protein destruction. Annu. Rev. Biochem. *80*, 587–612.

266. Schein, C.H. (2004). A cool way to make proteins. Nat. Biotechnol. *22*, 826–827.

267. Schlame, M., and Ren, M. (2009). The role of cardiolipin in the structural organization of mitochondrial membranes. Biochim. Biophys. Acta BBA - Biomembr. *1788*, 2080–2083.

268. Schlame, M., Beyer, K., Hayer-Hartl, M., and Klingenberg, M. (1991). Molecular species of cardiolipin in relation to other mithocondrial phospholipids. Is there an acyl specificty of the interaction between cardiolipin and the ADP/ATP carrier? *FEBS*, 459–466.

269. Schlegel, S., Löfblom, J., Lee, C., Hjelm, A., Klepsch, M., Strous, M., Drew, D., Slotboom, D.J., and de Gier, J.-W. (2012). Optimizing Membrane Protein Overexpression in the Escherichia coli strain Lemo21(DE3). J. Mol. Biol. *423*, 648–659.

270. Schlegel, S., Genevaux, P., and de Gier, J.-W. (2015). De-convoluting the Genetic Adaptations of E. coli C41(DE3) in Real Time Reveals How Alleviating Protein Production Stress Improves Yields. Cell Rep. *10*, 1758–1766.

271. Schlegel, S., Genevaux, P., and Gier, J.-W. de (2017). Isolating Escherichia coli strains for recombinant protein production. Cell. Mol. Life Sci. *74*, 891–908.

272. Schleif, R. (2010). AraC protein, regulation of the l-arabinose operon in Escherichia coli, and the light switch mechanism of AraC action. FEMS Microbiol. Rev. *34*, 779–796.

273. Schmitt, R. (1968). Analysis of melibiose mutants deficient in alpha-galactosidase and thiomethylgalactoside permease II in Escherichia coli K-12. J. Bacteriol. *96*, 462–471.

274. Schneider, D., Duperchy, E., Depeyrot, J., Coursange, E., Lenski, R.E., and Blot, M. (2002). Genomic comparisons among Escherichia coli strains B, K-12, and O157:H7 using IS elements as molecular markers. BMC Microbiol. *2*, 18.

275. Schulze, R.J., Komar, J., Botte, M., Allen, W.J., Whitehouse, S., Gold, V.A.M., Nijeholt, J.A.L. a, Huard, K., Berger, I., Schaffitzel, C., et al. (2014). Membrane protein insertion and proton-motive-force-dependent secretion through the bacterial holo-translocon SecYEG–SecDF–YajC–YidC. Proc. Natl. Acad. Sci. *111*, 4844–4849.

276. Scotti, P.A., Urbanus, M.L., Brunner, J., de Gier, J.W., von Heijne, G., van der Does, C., Driessen, A.J., Oudega, B., and Luirink, J. (2000). YidC, the Escherichia coli homologue of mitochondrial Oxa1p, is a component of the Sec translocase. EMBO J. *19*, 542–549.

277. Seddon, J.M. (1990). Structure of the inverted hexagonal (HII) phase, and non-lamellar phase transitions of lipids. Biochim. Biophys. Acta BBA - Rev. Biomembr. *1031*, 1–69.

278. Seddon, J.M., Kaye, R.D., and Marsh, D. (1983). Induction of the lamellar-inverted hexagonal phase transition in cardiolipin by protons and monovalent cations. Biochim. Biophys. Acta BBA - Biomembr. *734*, 347–352.

279. Sener, M.K., Olsen, J.D., Hunter, C.N., and Schulten, K. (2007). Atomic-level structural and functional model of a bacterial photosynthetic membrane vesicle. Proc. Natl. Acad. Sci. U. S. A. *104*, 15723–15728.

280. Sezonov, G., Joseleau-Petit, D., and D'Ari, R. (2007). Escherichia coli Physiology in Luria-Bertani Broth. J. Bacteriol. *189*, 8746–8749.

281. Sharma, S.K., De los Rios, P., Christen, P., Lustig, A., and Goloubinoff, P. (2010). The kinetic parameters and energy cost of the Hsp70 chaperone as a polypeptide unfoldase. Nat. Chem. Biol. *6*, 914–920.

282. Shimada, T., Fujita, N., Yamamoto, K., and Ishihama, A. (2011). Novel Roles of cAMP Receptor Protein (CRP) in Regulation of Transport and Metabolism of Carbon Sources. PLoS ONE *6*, e20081.

283. Shineberg, B., and Zipser, D. (1973). The lon Gene and Degradation of β-Galactosidase Nonsense Fragments. J. Bacteriol. *116*, 1469–1471.

284. Siguier, P., Gourbeyre, E., and Chandler, M. (2014). Bacterial insertion sequences: their genomic impact and diversity. FEMS Microbiol. Rev. *38*, 865–891.

285. Sikdar, R., Simmons, A.R., and Doerrler, W.T. (2013). Multiple Envelope Stress Response Pathways Are Activated in an Escherichia coli Strain with Mutations in Two Members of the DedA Membrane Protein Family. J. Bacteriol. *195*, 12–24.

286. Silhavy, T.J., Kahne, D., and Walker, S. (2010). The Bacterial Cell Envelope. Cold Spring Harb. Perspect. Biol. *2*, a000414.

287. Skerra, A. (1994). Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in Escherichia coli. Gene *151*, 131–135.

288. Skjærven, L., Cuellar, J., Martinez, A., and Valpuesta, J.M. (2015). Dynamics, flexibility, and allostery in molecular chaperonins. FEBS Lett. *589*, 2522–2532.

289. Skretas, G., and Georgiou, G. (2009). Genetic analysis of G protein-coupled receptor expression in Escherichia coli: inhibitory role of DnaJ on the membrane integration of the human central cannabinoid receptor. Biotechnol. Bioeng. *102*, 357–367.

290. Skretas, G., and Georgiou, G. (2010). Simple Genetic Selection Protocol for Isolation of Overexpressed Genes That Enhance Accumulation of Membrane-Integrated Human G Protein-Coupled Receptors in Escherichia coli. Appl. Environ. Microbiol. *76*, 5852–5859.

291. Skretas, G., Makino, T., Varadarajan, N., Pogson, M., and Georgiou, G. (2012). Multi-copy genes that enhance the yield of mammalian G protein-coupled receptors in Escherichia coli. Metab. Eng. *14*, 591–602.

292. Snyder, W.B., and Silhavy, T.J. (1995). Beta-galactosidase is inactivated by intermolecular disulfide bonds and is toxic when secreted to the periplasm of Escherichia coli. J. Bacteriol. *177*, 953–963.

293. Sousa, M.C., Trame, C.B., Tsuruta, H., Wilbanks, S.M., Reddy, V.S., and McKay, D.B. (2000). Crystal and solution structures of an HslUV protease-chaperone complex. Cell *103*, 633–643.

294. Srivastava, S.K., Lambadi, P.R., Ghosh, T., Pathania, R., and Navani, N.K. (2014). Genetic regulation of spy gene expression in Escherichia coli in the presence of protein unfolding agent ethanol. Gene *548*, 142–148.

295. Steitz, T.A. (2008). A structural understanding of the dynamic ribosome machine. Nat. Rev. Mol. Cell Biol. *9*, 242–253.

296. Stepanyants, N., Macdonald, P.J., Francy, C.A., Mears, J.A., Qi, X., and Ramachandran, R. (2015). Cardiolipin's propensity for phase transition and its reorganization by

dynamin-related protein 1 form a basis for mitochondrial membrane fission. Mol. Biol. Cell *26*, 3104–3116.

297. Studier, F.W. (1991). Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. J. Mol. Biol. *219*, 37–44.

298. Studier, F.W. (2005). Protein production by auto-induction in high-density shaking cultures. Protein Expr. Purif. *41*, 207–234.

299. Studier, F.W., and Moffatt, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. *189*, 113–130.

300. Subramani, S., Perdreau-Dahl, H., and Morth, J.P. (2016). The magnesium transporter A is activated by cardiolipin and is highly sensitive to free magnesium in vitro. ELife *5*, e11407.

301. Sugiyama, N., Minami, N., Ishii, Y., and Amano, F. (2013). Inhibition of Lon protease by bacterial lipopolisaccharide (LPS) though inhibition of ATPase. Adv. Biosci. Biotechnol. *04*, 590–598.

302. Suno, R., Niwa, H., Tsuchiya, D., Zhang, X., Yoshida, M., and Morikawa, K. (2006). Structure of the whole cytosolic region of ATP-dependent protease FtsH. Mol. Cell *22*, 575–585.

303. Svitil, A.L., Cashel, M., and Zyskind, J.W. (1993). Guanosine tetraphosphate inhibits protein synthesis in vivo. A possible protective mechanism for starvation stress in Escherichia coli. J. Biol. Chem. *268*, 2307–2311.

304. Swart, E.C., Serra, V., Petroni, G., and Nowacki, M. (2016). Genetic Codes with No Dedicated Stop Codon: Context-Dependent Translation Termination. Cell *166*, 691–702.

305. Tan, B.K., Bogdanov, M., Zhao, J., Dowhan, W., Raetz, C.R.H., and Guan, Z. (2012). Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates. Proc. Natl. Acad. Sci. *109*, 16504–16509.

306. Tanaka, M., Arakaki, A., and Matsunaga, T. (2010). Identification and functional characterization of liposome tubulation protein from magnetotactic bacteria. Mol. Microbiol. *76*, 480–488.

307. Tanford, C. (1978). The Hydrophobic Effect and the Organization of Living Matter. Science *200*, 1012–1018.

308. Tate, C.G. (2012). A crystal clear solution for determining G-protein-coupled receptor structures. Trends Biochem. Sci. *37*, 343–352.

309. Terstappen, G.C., and Reggiani, A. (2001). In silico research in drug discovery. Trends Pharmacol. Sci. *22*, 23–26.

310. Teter, S.A., Houry, W.A., Ang, D., Tradler, T., Rockabrand, D., Fischer, G., Blum, P., Georgopoulos, C., and Hartl, F.U. (1999). Polypeptide Flux through Bacterial Hsp70. Cell *97*, 755–765.

311. Travers, A.A., and Burgess, R.R. (1969). Cyclic Re-use of the RNA Polymerase Sigma Factor. Nature *222*, 537–540.

312. Travisano, M., and Lenski, R.E. (1996). Long-term experimental evolution in Escherichia coli. IV. Targets of selection and the specificity of adaptation. Genetics *143*, 15–26.

313. Traxler, M.F., Summers, S.M., Nguyen, H.-T., Zacharia, V.M., Hightower, G.A., Smith, J.T., and Conway, T. (2008). The global, ppGpp-mediated stringent response to amino acid starvation in Escherichia coli. Mol. Microbiol. *68*, 1128–1148.

314. Traxler, M.F., Zacharia, V.M., Marquardt, S., Summers, S.M., Nguyen, H.-T., Stark, S.E., and Conway, T. (2011). Discretely calibrated regulatory loops controlled by ppGpp partition gene induction across the 'feast to famine' gradient in Escherichia coli: Architecture of the stringent response. Mol. Microbiol. *79*, 830–845.

315. Unsay, J.D., Cosentino, K., Subburaj, Y., and García-Sáez, A.J. (2013). Cardiolipin Effects on Membrane Structure and Dynamics. Langmuir *29*, 15878–15887.

316. Vail, W.J., and Stollery, J.G. (1979). Phase changes of cardiolipin vesicles mediated by divalent cations. Biochim. Biophys. Acta BBA - Biomembr. *551*, 74–84.

317. Vellanoweth, R.L., and Rabinowitz, J.C. (1992). The influence of ribosome-binding-site elements on translational efficiency in Bacillus subtilis and Escherichia coli in vivo. Mol. Microbiol. *6*, 1105–1114.

318. Vieux, E.F., Wohlever, M.L., Chen, J.Z., Sauer, R.T., and Baker, T.A. (2013). Distinct quaternary structures of the AAA+ Lon protease control substrate degradation. Proc. Natl. Acad. Sci. U. S. A. *110*, E2002-2008.

319. Viklund, H., Granseth, E., and Elofsson, A. (2006). Structural classification and prediction of reentrant regions in alpha-helical transmembrane proteins: application to complete genomes. J. Mol. Biol. *361*, 591–603.

320. Villalobo, A. (1990). Reconstitution of ion-motive transport ATPases in artificial lipid membranes. Biochim. Biophys. Acta BBA - Bioenerg. *1017*, 1–48.

321. Von Meyenburg, K., Jørgensen, B.B., and Van Deurs, B. (1984). Physiological and morphological effects of overproduction of membrane-bound ATP synthase in Escherichia coli K-12. EMBO J. *3*, 1791.

322. Wagner, M., and Horn, M. (2006). The Planctomycetes, Verrucomicrobia, Chlamydiae and sister phyla comprise a superphylum with biotechnological and medical relevance. Curr. Opin. Biotechnol. *17*, 241–249.

323. Wagner, S., Baars, L., Ytterberg, A.J., Klussmeier, A., Wagner, C.S., Nord, O., Nygren, P.-Å., Wijk, K.J. van, and Gier, J.-W. de (2007). Consequences of Membrane Protein Overexpression in Escherichia coli. Mol. Cell. Proteomics *6*, 1527–1550.

324. Wagner, S., Klepsch, M.M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Högbom, M., van Wijk, K.J., Slotboom, D.J., Persson, J.O., et al. (2008). Tuning Escherichia coli for membrane protein overexpression. Proc. Natl. Acad. Sci. U. S. A. *105*, 14371–14376.

325. Wahlberg, J.M., and Spiess, M. (1997). Multiple determinants direct the orientation of signal-anchor proteins: the topogenic role of the hydrophobic signal domain. J. Cell Biol. *137*, 555–562.

326. Waldo, G.S., Standish, B.M., Berendzen, J., and Terwilliger, T.C. (1999). Rapid protein-folding assay using green fluorescent protein. Nat. Biotechnol. *17*, 691–695.

327. Walker, J.E., and Miroux, B. (1999). Selection of Escherichia coli host that are optimized for the overexpression of proteins. *Manual of Industrial Microbiology and Biotechnology*.

328. Walser, P.J., Ariotti, N., Howes, M., Ferguson, C., Webb, R., Schwudke, D., Leneva, N., Cho, K.-J., Cooper, L., Rae, J., et al. (2012). Constitutive Formation of Caveolae in a Bacterium. Cell *150*, 752–763.

329. Wang, H.H., Isaacs, F.J., Carr, P.A., Sun, Z.Z., Xu, G., Forest, C.R., and Church, G.M. (2009). Programming cells by multiplex genome engineering and accelerated evolution. Nature *460*, 894–898.

330. Wanner, B.L., Kodaira, R., and Neidhardt, F.C. (1977). Physiological regulation of a decontrolled lac operon. J. Bacteriol. *130*, 212–222.

331. Way, M., Pope, B., Gooch, J., Hawkins, M., and Weeds, A.G. (1990). Identification of a region in segment 1 of gelsolin critical for actin binding. EMBO J. *9*, 4103–4109.

332. Weber, S., Granzow, H., Weiland, F., and Marquardt, O. (1996). Intracellular membrane proliferation in E. coli induced by foot-and-mouth disease virus 3A gene products. Virus Genes *12*, 5–14.

333. Weeghel, R.P. van, Keck, W., and Robillard, G.T. (1990). Regulated high-level expression of the mannitol permease of the phosphoenolpyruvate-dependent sugar phosphotransferase system in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. *87*, 2613.

334. Weiner, J.H., Lemire, B.D., Elmes, M.L., Bradley, R.D., and Scraba, D.G. (1984a). Overproduction of fumarate reductase in Escherichia coli induces a novel intracellular lipid-protein organelle. J. Bacteriol. *158*, 590–596.

335. Weiner, J.H., Lemire, B.D., Elmes, M.L., Bradley, R.D., and Scraba, D.G. (1984b). Overproduction of fumarate reductase in Escherichia coli induces a novel intracellular lipid-protein organelle. J. Bacteriol. *158*, 590.

336. Weis, R.M., Hirai, T., Chalah, A., Kessel, M., Peters, P.J., and Subramaniam, S. (2003a). Electron Microscopic Analysis of Membrane Assemblies Formed by the Bacterial Chemotaxis Receptor Tsr. J. Bacteriol. *185*, 3636–3643.

337. Weis, R.M., Hirai, T., Chalah, A., Kessel, M., Peters, P.J., and Subramaniam, S. (2003b). Electron Microscopic Analysis of Membrane Assemblies Formed by the Bacterial Chemotaxis Receptor Tsr. J. Bacteriol. *185*, 3636–3643.

338. Welch, A.K., Claggett, S.B., and Cain, B.D. (2008). The b (arg36) contributes to efficient coupling in F(1)F (O) ATP synthase in Escherichia coli. J. Bioenerg. Biomembr. *40*, 1–8.

339. Westphal, S., Heins, L., Soll, J., and Vothknecht, U.C. (2001). Vipp1 deletion mutant of Synechocystis: a connection between bacterial phage shock and thylakoid biogenesis? Proc. Natl. Acad. Sci. U. S. A. *98*, 4243–4248.

340. Wickstrum, J.R., Santangelo, T.J., and Egan, S.M. (2005). Cyclic AMP receptor protein and RhaR synergistically activate transcription from the L-rhamnose-responsive rhaSR promoter in Escherichia coli. J. Bacteriol. *187*, 6708–6718.

341. Wikström, M., Kelly, A.A., Georgiev, A., Eriksson, H.M., Klement, M.R., Bogdanov, M., Dowhan, W., and Wieslander, A. (2009). Lipid-engineered Escherichia coli membranes reveal critical lipid headgroup size for protein function. J. Biol. Chem. *284*, 954–965.

342. Wilkison, W.O., Walsh, J.P., Corless, J.M., and Bell, R.M. (1986a). Crystalline arrays of the Escherichia coli sn-glycerol-3-phosphate acyltransferase, an integral membrane protein. J. Biol. Chem. *261*, 9951–9958.

343. Wilkison, W.O., Walsh, J.P., Corless, J.M., and Bell, R.M. (1986b). Crystalline arrays of the Escherichia coli sn-glycerol-3-phosphate acyltransferase, an integral membrane protein. J. Biol. Chem. *261*, 9951–9958.

344. Williamson, S.I., Wannemuehler, M.J., Jirillo, E., Pritchard, D.G., Michalek, S.M., and McGhee, J.R. (1984). LPS regulation of the immune response: separate mechanisms for murine B cell activation by lipid A (direct) and polysaccharide (macrophage-dependent) derived from Bacteroides LPS. J. Immunol. Baltim. Md 1950 *133*, 2294–2300.

345. Xu, P., Gu, Q., Wang, W., Wong, L., Bower, A.G.W., Collins, C.H., and Koffas, M.A.G. (2013). Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. Nat. Commun. *4*, ncomms2425.

346. Xu, Z., Horwich, A.L., and Sigler, P.B. (1997). The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. Nature *388*, 741–750.

347. Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G., and Court, D.L. (2000). An efficient recombination system for chromosome engineering in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. *97*, 5978–5983.

348. Zambolin, S., Clantin, B., Chami, M., Hoos, S., Haouz, A., Villeret, V., and Delepelaire, P. (2016). Structural basis for haem piracy from host haemopexin by Haemophilus influenzae. Nat. Commun. *7*.

349. Zeczycki, T.N., Whelan, J., Hayden, W.T., Brown, D.A., and Shaikh, S.R. (2014). Increasing levels of cardiolipin differentially influence packing of phospholipids found in the mitochondrial inner membrane. Biochem. Biophys. Res. Commun. *450*, 366–371.

350. Zhang, Y.-M., and Rock, C.O. (2008). Membrane lipid homeostasis in bacteria. Nat. Rev. Microbiol. *6*, 222–233.

351. Zhang, S., Shen, G., Li, Z., Golbeck, J.H., and Bryant, D.A. (2014). Vipp1 Is Essential for the Biogenesis of Photosystem I but Not Thylakoid Membranes in Synechococcus sp. PCC 7002. J. Biol. Chem. *289*, 15904–15914.

352. Zhang, Z., Kuipers, G., Niemiec, Ł., Baumgarten, T., Slotboom, D.J., de Gier, J.-W., and Hjelm, A. (2015). High-level production of membrane proteins in E. coli BL21(DE3) by omitting the inducer IPTG. Microb. Cell Factories *14*.

353. Zhao, J.-B., Wei, D.-Z., and Tong, W.-Y. (2007). Identification of Escherichia coli host cell for high plasmid stability and improved production of antihuman ovarian carcinoma × antihuman CD3 single-chain bispecific antibody. Appl. Microbiol. Biotechnol. *76*, 795–800.

354. Zhao, K., Liu, M., and Burgess, R.R. (2005). The Global Transcriptional Response of Escherichia coli to Induced σ32 Protein Involves σ32 Regulon Activation Followed by Inactivation and Degradation of σ32 in Vivo. J. Biol. Chem. *280*, 17758–17768.

355. Zimmerberg, J., and Kozlov, M.M. (2006). How proteins produce cellular membrane curvature. Nat. Rev. Mol. Cell Biol. *7*, 9–19.

Résumé substantiel en français:

La biologie structurale des protéines membranaires se heurte à une difficulté majeur: le manqué d'un système d'expression universel. Malgré l'absence de compartimentalisation, *Escherichia coli* est encore aujourd'hui le moyen le plus simple de produire des protéines membranaires. La moitié des structures atomiques de protéines membranaires présentes dans la PDB ont été obtenues à partir d' *Escherichia coli*. Le système d'expression le plus utilisé pour la transcription des protéines membranaires, est le système basé sur l'ARN polymérase T7 (ARNpol T7) (Hattab et al., 2015). L'inconvénient de ce système est néanmoins que la vitesse de transcription de l'ARNpol T7 est dix fois plus rapide que celle de l'enzyme bactérienne. Les souches bactériennes mutantes C43(DE3) et C41(DE3) ont été isolées à partir de la souche parentale BL21l(DE3). Il apparaît clairement que la toxicité provoquée par la surproduction des protéines membranaires est liée à la quantité trop élevée d'ARNpol T7 dans la cellule (Wagner et al., 2008; Kwon et al., 2015). Les protéines membranaires ont besoin d'une vitesse de transcription/traduction plus basse pour se replier correctement dans la membrane de la bactérie. Dans la premiere partie de la thèse ont a isolé et caractérisé les nouvelles souches C44(DE3) et C45(DE3) dans lesquelles le niveau d'ARNpol T7 est efficacement régulé par un mécanisme non transcriptionnel très favorable à l'expression des protéines membranaires du à l'existance d' un codone de stop d'ambre sur C44(DE3) et un codone d'ocre sur C45(DE3).

Dans le système d'expression fondé sur l'utilisation de l' ANR polymerase du phage T7, la surproduction de certaines protéines membranaires s'accompagne d'une prolifération de membranes internes dont la morphologie ressemble aux crêtes des membranes internes mitochondriales. Nous avons donce etudié les bases moléculaires de la plasticité métabolique d'Escherichia coli dans le contexte de la synthèse lipidique et de l'assemblage des membranes internes. Le projet se décomposé en trois parties:

1. Identifier le programme spatiaux temporel de la formation de ces membranes par microscopie optique associé à la cytométrie de flux ;
2. Identifier, par des approaches protéomiques et transcriptomiques, les éléments fonctionnels régulateurs dans ce modèle de proliferation membranaire;

3. Mettre au point un nouveau crible génétique, fondé sur le contenu en cardiolipides, pour isoler de nouveaux mutants bactériens.

Pour mieux comprendre les voies métaboliques impliquées dans la biogenèse, la prolifération et l'organisation des membranes, nous avons utilisé une approche de séquençage d'ARN à haut débit à différents temps après induction de la surexpression de la sous-unité AtpF dans la souche C43(DE3). Ensuite, et en collaboration avec Gerardo Carranza and Ignacio Arechaga (Université de Cantabria, Espagne), nous avons construit et étudié des mutants de C43(DE3) déficients pour les trois gènes codants pour des enzymes de la biosynthèse des cardiolipides afin d'évaluer leur participation dans la biogénèse des membranes intracellulaires.

Annexes

Chapter 3

Membrane Protein Production in *Escherichia coli*: Protocols and Rules

Federica Angius, Oana Ilioaia, Marc Uzan, and Bruno Miroux

Abstract

Functional and structural studies on membrane proteins are limited by the difficulty to produce them in large amount and in a functional state. In this review, we provide protocols to achieve high-level expression of membrane proteins in *Escherichia coli*. The T7 RNA polymerase-based expression system is presented in detail and protocols to assess and improve its efficiency are discussed. Protocols to isolate either membrane or inclusion bodies and to perform an initial qualitative test to assess the solubility of the recombinant protein are also included.

Key words Production of recombinant proteins, *E. coli*, T7 RNA polymerase

1 Introduction

Membrane protein (MP) production is still a challenge for biochemists and biophysicists. Over the last decade, eukaryotic expression systems have emerged and have proven to be very useful for structural studies of eukaryotic MP such as G-protein-coupled receptors [1]. However bacterial expression systems remain widely used. We have recently conducted a global survey of the protein data bank (PDB) and found that half of unique MP structures deposited in the PDB have been produced in *E. coli* [2]. Provided that the recombinant MP is well folded within the membrane of the host, bacteria can produce, at very low cost, sufficient amount of the target MP for X-ray crystallization or NMR studies. *E. coli* is also the most versatile host for specific isotopic labeling of proteins required for NMR studies. In this review, we focus on the T7 RNA polymerase (T7 RNAP) bacterial expression system which is, so far, the most efficient in producing large amount of membrane proteins for structural studies [2]. Figure 1 provides an overview of how the expression system works in the bacterial host BL21λ(DE3). The gene encoding the T7 RNAP is inserted in the lambda DE3 under the control of the *lacUV5* promoter. Upon addition of

Isabelle Mus-Veteau (ed.), *Heterologous Expression of Membrane Proteins: Methods and Protocols,* Methods in Molecular Biology, vol. 1432, DOI 10.1007/978-1-4939-3637-3_3,

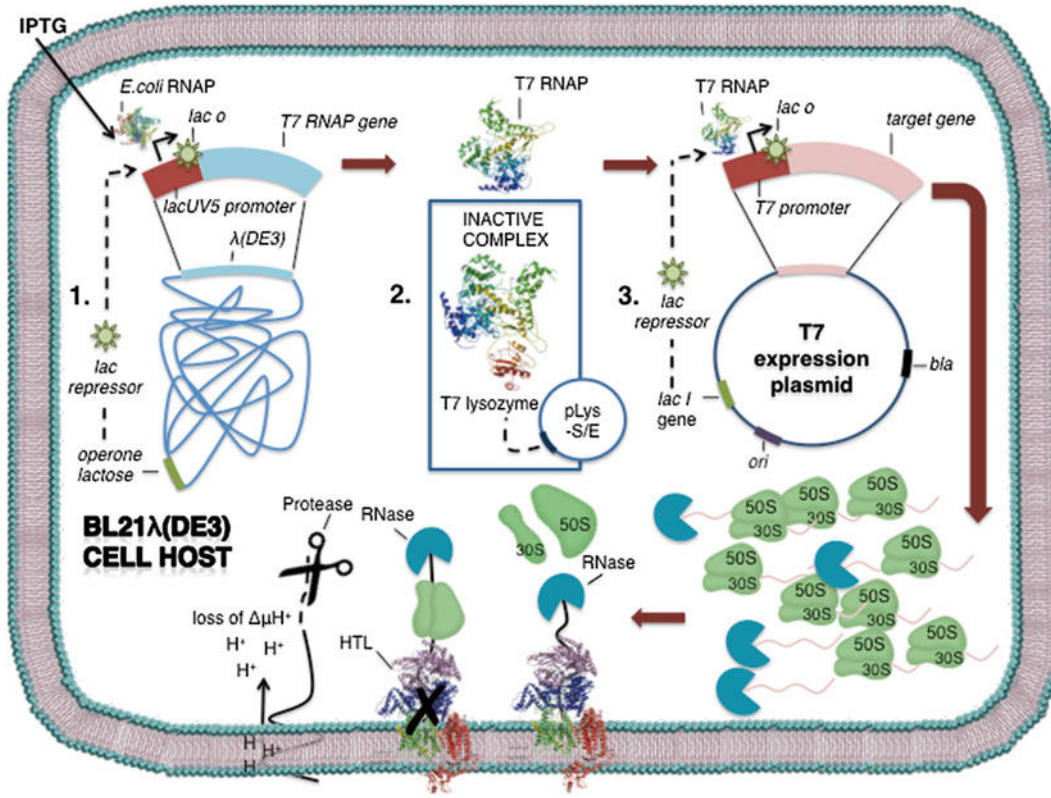


Fig. 1 Global view of the T7-based expression system in BL21λ(DE3). The T7 RNA polymerase gene is lysogenic in the genome and its expression is under the control of the IPTG-inducible *lacUV5* promoter. Upon addition of IPTG, the T7RNAP will specifically transcribe the target gene inserted in the T7 expression plasmid and the target MP might be produced at very high levels. However, overexpression of the target mRNA is, most of the time, toxic to the cell because it overloads the translation machinery and uncouple transcription from translation. The newly synthetized membrane protein might also overload the folding and secretion machineries causing mistargeting of the overproduced MP, protein aggregation, and ultimately proton leak and loss of energy homeostasis. To circumvent these difficulties, the expression system can be regulated by several ways: *1*. Repressing the *lacUV5* promoter using a *lac* repressor; *2*. expressing the T7 lysozyme from a pLys-S/E plasmid, which will inhibit its activity; *3*. inserting the *lac* repressor in the T7 multi-copy expression plasmid

IPTG, isopropyl β-D-1-thiogalactopyranoside, a non-metabolized derivative of lactose, the T7 RNAP is produced and will specifically transcribe the target gene inserted in a T7 expression vector downstream of the T7 promoter. The mRNA of the target MP is highly expressed because the T7 RNAP transcriptional elongation rate is ten times faster than the *E. coli* enzyme. In addition, the T7 expression vector is present in multiple copies. In many cases, the target mRNA overloads the translation machinery triggering ribosome destruction and growth arrest [3]. Naked un-translated mRNA are rapidly degraded by RNases and, in some cases, the RNA degradation is faster than the transcriptional activity of the T7 RNAP leading to lower yield of the target than expected [4]. For this

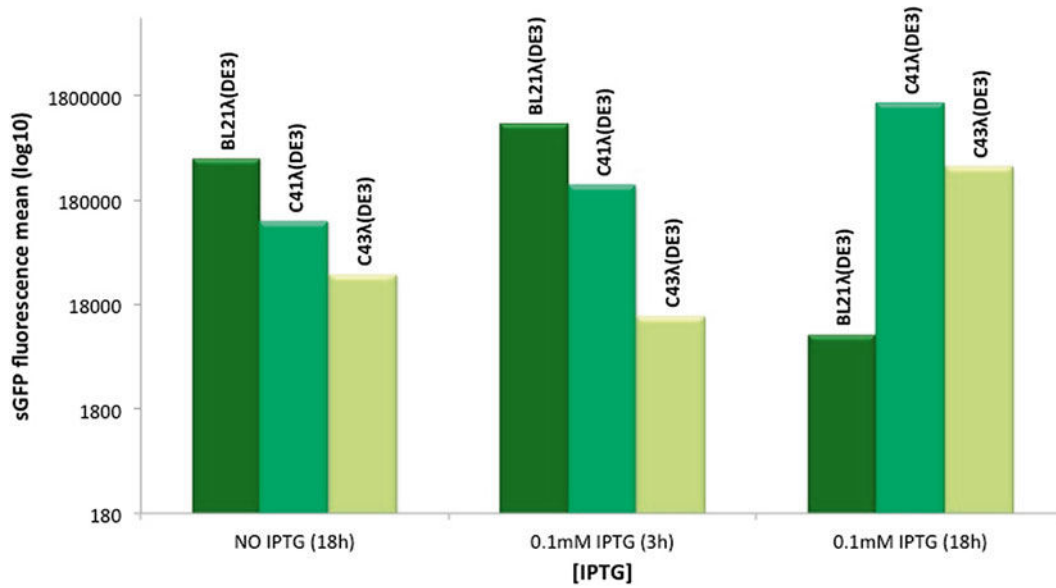


Fig. 2 Analysis of GFP fluorescence by flow cytometry. The pHis17-sGFP T7 expression plasmid has been transformed in the BL21λ(DE3) (*green*), C41λ(DE3) (*light green*), and C43λ(DE3) (*yellow*) bacterial hosts. The fluorescence has been recorded with the flow cytometer Accuri C6 3 h and 18 h after induction with 0.1 mM of IPTG. To assess the basal level of expression of sGFP in each host, cells were grown for 18 h with no addition of IPTG

reason, we provide here a rapid protocol to assess the levels of your target mRNA when the yield of the corresponding MP is low. Insertion of the target MP at the *E. coli* membrane can also overload the translocation and secretion machineries. The recombinant MP will then be not only misfolded and produced at low level but proton permeability of the bacterial membrane might be compromised leading to cell death. Over the last 20 years, the T7 expression system has been optimized to improve its regulation and extend its ability to produce large amount of MP. For instance the use of lysozyme has been shown to strongly inhibit the activity of the T7 RNAP, thus providing a means to decrease the basal activity of the LacUV5 promoter and to tune the activity of the T7-RNAP upon induction [5, 6]. Other groups have isolated mutant hosts from the parental strain BL21λ(DE3) [7, 8]. Some of them namely C41λ(DE3) and C43λ(DE3) have proven to be extremely useful for structural biologists; these mutant hosts contributed to 28% of non-*E. coli* unique MP structures and 19% of *E. coli* unique MP structures deposited into the PDB [2]. To illustrate that the mutant hosts are better regulated than the parental strain, we used the green fluorescent protein (superfold version, sGFP) as gene reporter. After transformation with the pRSET-sGFP expression plasmid, cells have been induced at $OD_{600nm} = 0.6$ with 0.1 mM IPTG. Figure 2 shows the mean green fluorescent intensity analyzed by flow cytometry. In all three bacterial hosts, the basal level of sGFP fluorescence after an overnight culture in 2*TY-rich

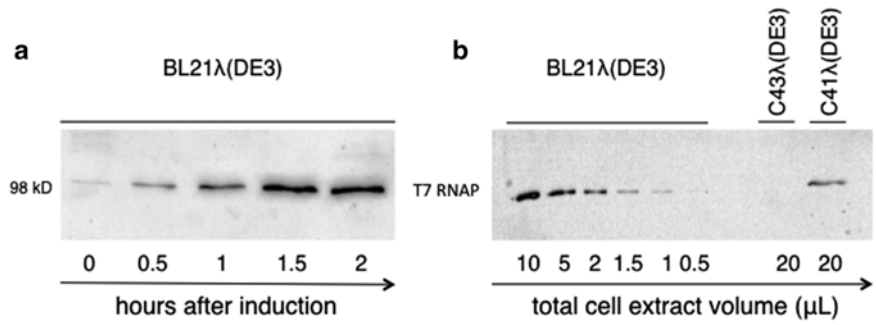


Fig. 3 Immunodetection of the T7 RNAP enzyme in T7 expression hosts. Total cell extracts were loaded on SDS-PAGE and proteins were transferred on nitrocellulose membrane. The T7 RNAP protein was revealed using the anti-T7RNAP from Novagene and a second antibody coupled to peroxidase. Peroxidase activity was detected by chemiluminescence. (**a**) Time course of the T7 RNAP protein expression in BL21λ(DE3) upon addition of 0.7 mM IPTG. (**b**) Expression levels of T7 RNAP 2 h after 0.7 mM IPTG induction in BL21λ(DE3), C43λ(DE3), C41λ(DE3). For BL21λ(DE3) host, decreasing amounts of total cell extract have been loaded to compare the intensity of the signal with the mutant hosts

medium is very high showing that the expression system is leaky. However, basal fluorescence intensities are 4 and 13 times decreased in C41λ(DE3) and C43λ(DE3), respectively. Upon addition of IPTG, fluorescence intensity increased twice in BL21λ(DE3) host 3 h after induction and decreased strongly after overnight induction. This is due to loss of the expression plasmid, cell death, and lysis. In contrast, sGFP production reached a maximal value after overnight induction in both C41λ(DE3) and C43λ(DE3) hosts (13- and 11-fold induction, respectively). At the molecular level, Wagner et al. have shown that, in C41λ(DE3) and C43λ(DE3) hosts, the strong *lacUV5* promoter recombined with the wild-type genomic copy of the *lac* promoter. Consequently, the amount of T7 RNAP enzyme produced upon addition of IPTG is ten times reduced in C41λ(DE3) and undetectable in C43λ(DE3) using the commercially available anti-T7 RNAP antibody from Novagene (Fig. 3). In this chapter we provide protocols to design your construct and choose the appropriate host/vector combination, isolate new bacterial hosts, set up growth conditions, assess your expression system by flow cytometry, fractionate bacterial cells, and perform a first biochemical analysis.

2 Materials

2.1 Materials for RNA Isolation and Sucrose Gradient

1. Tips, plastic tubes, glass, gloves, water must be RNase free.
2. A dry bath to warm up the samples to 65 °C.
3. A laboratory fume hood.
4. A spectrophotometer.

5. QIAGEN RNase-Free DNase I Set.
6. Lysis solution for 700 µL of culture: 35.5 µL 20% SDS + 7 µL of 200 mM Na-EDTA + 500 µL water-saturated phenol.
7. Water-saturated phenol.
8. Phenol/chloroform solution v/v 1:1. The chloroform should contain isoamyl alcohol in a proportion v/v 24:1.
9. 3 M Na-acetate pH 5.
10. Ethanol 100% RNase free.
11. Ethanol 70% RNase free.
12. Gradient maker.
13. 10 mM Tris–HCl, pH 8.
14. Sucrose solutions: 50% and 5% (w/v) prepared in 10 mM Tris–HCl, pH 8.
15. Beckman Coulter Ultra-Clear™ centrifuge tubes.

2.2 Media, Buffers, and Chemicals

1. 1*LB Medium: 10 g Bacto Tryptone, 5 g Bacto Yeast Extract, and 5 g NaCl. Add ultrapure water to 900 mL. Adjust the pH to 7.2 with NaOH. Add water at a final volume of 1 L and autoclave for 20 min at 121 °C.
2. 2*TY Medium: 16 g Bacto Tryptone, 10 g Bacto Yeast Extract, and 5 g NaCl. Add ultrapure water to 800 mL. Adjust the pH to 7.2 with NaOH, adjust the final volume to 900 mL, and autoclave for 20 min at 121 °C.
3. 2*TY with glucose: 2 g Glucose in 100 mL of final volume of water, filter sterilize. Add the glucose solution in the autoclaved medium. Adding glucose could be useful if you wish to repress further the expression vector before induction.
4. Isopropyl-beta-D-galactoside (IPTG): Prepare 100 mM, 500 mM, 700 mM, and 1 M stock solutions in ultrapure water, sterilize with 0.22 µm filter, aliquot, and store at –20 °C.
5. Antibiotics: Prepare 1000 times stock solutions of antibiotics. Ampicillin (100 mg/mL) can be prepared in ultrapure water and stored at –20 °C. Tetracycline (12,5 mg/mL) and kanamycin (30 mg/mL) are freely soluble in water but in time the solutions can turn turbid due to precipitation. It is thus recommended to prepare it in 95% ethanol. Dilute 1000 times the stock solution in medium prior to use.
6. Phosphate-buffered saline (PBS): 10 mM Phosphate, 150 mM NaCl, pH 7.4 (tablets are commercially available).
7. TEP buffer: 10 mM Tris–HCl, pH 8, 1 mM EDTA, and 0.001% PMSF.
8. Triton X-100.
9. Dodecyl-maltoside (DDM).

10. Phosphododecylcholine (Fc12).
11. Sodium dodecyl sulfate (SDS).

2.3 Web Resources

2.3.1 Sequence Analysis and Molecular Biology Tools

1. **SMART** (protein domain database): http://smart.embl-heidelberg.de/
2. **Jpred** (secondary structure prediction): http://www.compbio.dundee.ac.uk/jpred/index.html
3. **ExPasy**: http://www.expasy.ch/
4. **Amplify** (PCR simulation, oligonucleotide design) http://engels.genetics.wisc.edu/amplify/

Vector design:

5. **Dna20:** https://www.dna20.com/resources/bioinformatics-tools
6. **Serial-Cloner**: http://serialbasics.free.fr/Serial_Cloner-Download.html
7. **APE**: http://biologylabs.utah.edu/jorgensen/wayned/ape/

Molecular and structural biology websites

8. Steewe White: http://blanco.biomol.uci.edu/mpstruc/
9. Dror Warschawski: http://www.drorlist.com/nmr/MPNMR.html

Academic expression plasmid resources

10. Protein Science Initiative. http://psimr.asu.edu/about.html

3 Methods

3.1 Designing Constructs for Expression

1. Before starting molecular cloning experiments, check if your target MP is already available in an expression vector (*see* Protein Science Initiative Web site) and search the literature to see if your MP target or related proteins have been produced in recombinant systems (*see* **Note 1**).
2. Be aware that *E. coli* cannot produce at high levels proteins larger than 90 kDa. Ribosomes drop off very long mRNA leading to incomplete synthesis products. If possible break up your protein into smaller fragments. Use SMART (protein domain identification) or Jpred (secondary structure prediction) to define the boundaries carefully.
3. Addition of purification Tag: For N-terminal constructs start protein synthesis with three amino acids before the Tag. In pRSET vector (Invitrogen), the N-terminal sequence is **MRGS**-(His)6 which gives a very good yield of recombinant protein. Consider adding more than 6 histidines (up to 12 but then preferentially in C-terminal position) to achieve a stronger

binding on Nickel column. There is no generic rule regarding cleavage sequences after the Tag but TEV cleavage sequence is widely used for MP as the TEV protease is still active in the presence of the most commonly used detergents [9].

4. Your target MP might not spontaneously go to the inner membrane of the bacteria. Consider making a C-terminal fusion with periplasmic maltose-binding protein (MBP), which contains a periplasmic signal sequence, to target your MP to the *E. coli* membrane [10, 11].
5. If possible engineer dua-ribosome-binding site (RBS) expression vectors like pET-Duet (Novagen) so that you can clone a fluorescent protein (FP) gene downstream of your target MP gene. This allows you to follow the cell population of bacteria by flow cytometry, to assess the stability and toxicity of your expression vector quickly, and to establish the optimal induction conditions (*see* Subheading 3.5.1). FP fusion with your target MP is also an option developed successfully by several laboratories [12].

3.2 Selecting the Optimal Expression Vector/ Bacterial Host

Selecting the right combination of vector/bacterial host is an essential step to achieve the optimal production of your MP (*see* **Note 2**). The following rules apply to the T7 RNAP-based expression system:

1. In combination with C41λ(DE3) and C43λ(DE3) bacterial hosts use high-copy-number plasmids like those containing the pMB1 origin of replication (200–600 copies/cell). A non-exhaustive list is pMW7 and derivatives (pHis and pRun) [13, 14], pGEM (Promega), pRSET and pDEST (Invitrogen), pIVEX (5prime), and pPR-IBA (IBA). Avoid *lacI* and *lacO* sequences in the plasmid.
2. In combination with BL21λ(DE3), use preferentially medium-copy-number vectors and those containing *lacI* and *lacO* sequences like pET (3, 9, 14, 17, 20, 23 from Novagen), to reduce the amount of T7 RNAP before induction. Consider using the companion plasmid pLyS to inhibit the T7 RNAP after induction.
3. The BL21AI host, which contains the T7RNA polymerase gene under the control of the arabinose promoter, and the Lemo21 host [6], which contains a companion plasmid expressing the lysozyme under the rhamnose promoter, may also be useful to titrate the amount or activity of T7 RNA polymerase (*see* **Note 3**).

3.3 Viability Test on Agar Plate

1. Preparing the agar plates: Melt 500 mL (or less) 2*YT agar medium in water bath at 100 °C. When the solution is clear switch the temperature to 55 °C. Add 500 μL 0.7 M sterile IPTG and mix vigorously. Add the required antibiotic.

Pour the plates and wait for 1 h till the agar is solid. Incubate the plate upside down O/N at 37 °C.

2. Cell transformation: Take out of the −80 °C freezer a stock of competent cells and thaw them on ice. Add 10 ng of plasmid to 50 μL of cells and leave it on ice for 20 min. Place the microcentrifuge tube on a water bath set at 42 °C for 90 s. Replace the tube on ice for 5 min. Add 500–700 μL of SOC or LB media (without antibiotic) and allow the culture to grow for 45 min at 37 °C. Plate 100 μL of transformation mix on 2*TY agar plate containing only the antibiotic and 100 μL on agar plate with both antibiotic and IPTG. Incubate O/N at 37 °C.
3. Analysis of the cell population: Count the number of colonies on both plates. If you see no colony on IPTG-containing plates then expression of your target MP compromises cell growth. Transform the cells with the empty plasmid to verify if the plasmid is also toxic for the cell. If you have the same number of colonies in both conditions, then expression of your target MP is not toxic. Usually the size of the colonies is reduced in the presence of IPTG.

3.4 Selecting a Host Strain Adapted to the Expression of Your Target MP

The protocol below allows you to genetically isolate low-frequency mutants when your bacterial host/vector expression system compromises cell growth upon induction (*see* **Note 4**). The protocol uses the green fluorescent protein from Aequora Victoria [15] but can be performed without FP marker [7].

3.4.1 Selection Procedure

1. It is essential to work in a sterile environment. Pre-warm five 250 mL flasks containing 50 mL of 2*TY media. Add antibiotic prior to use. Autoclave 40 microcentrifuge tubes and fill them with 900 μL of sterile water or 2*TY media.
2. Transform the parental strain with your target MP expression vector, ideally in co-expression with a fluorescent protein. Inoculate a freshly transformed colony in 50 mL 2*TY medium. Grow cells until OD_{600nm} has reached 0.4–0.6. Add IPTG to 0.7 mM final concentration. One, two, and three hours after induction, take 1 mL of culture. After a low-speed centrifugation (300 × *g* for 2 min), gently resuspend the pellet in 1 mL of sterile water.
3. Perform serial dilutions from 1/10 to $1/10^4$ and immediately plate 100 μL of all dilutions on the IPTG/antibiotic-containing plates. Incubate the plates O/N at 37 °C.
4. Check the presence of green fluorescent colonies (Fig. 4) under UV light (above 300 nm to avoid the mutagenic effect of UV). Select ten small fluorescent colonies (*see* **Note 5**) for each selection experiment and make over-day cultures in tubes with 1 mL 2*TY containing ampicillin. After 2–3 h (when the

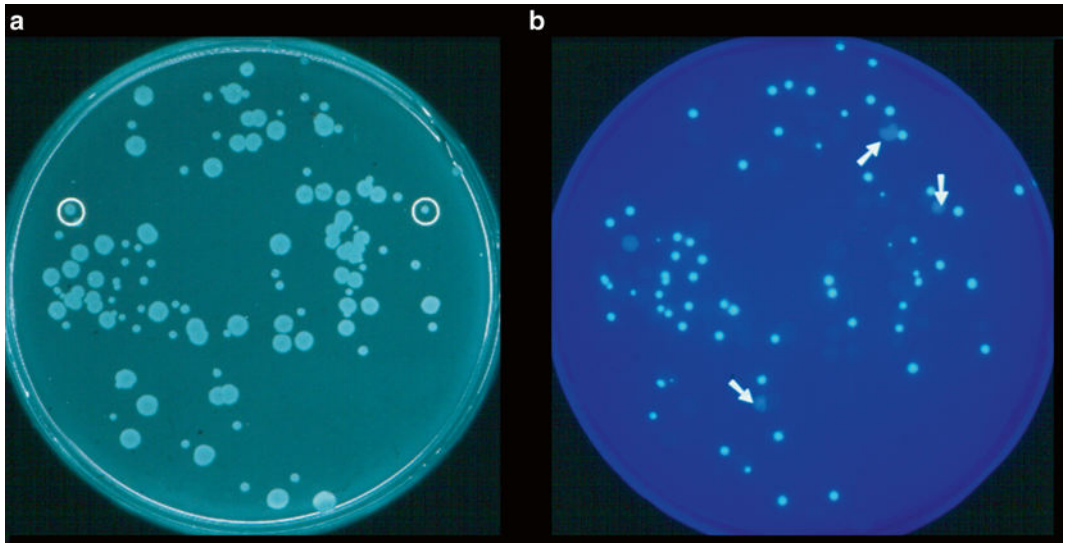


Fig. 4 Selection of bacterial hosts using GFP as gene reporter. After transformation of the pMW7-GFP expression plasmid, cells are grown in 2*TY medium and induced at $OD_{600\ nm} = 0.6$. Two hours after induction, serial dilutions of the culture are plated on IPTG-containing plates. The next day, plates were illuminated under a normal light (**a**) or UV light (**b**)

culture becomes turbid) save the mutants on 2*TY ampicillin agar plates with and without IPTG. Incubate the plates O/N at 37 °C.

5. Check that all mutants now grow on IPTG-containing plate. The size of the colonies in the presence of IPTG can be considered as a characteristic feature for a couple of vector/host.

3.4.2 Localization of the Mutation

You need to check if the mutation is in the bacterial genome or in the plasmid.

1. Plasmid rescue: Purify the plasmid DNA from each clone and transform the initial host, i.e., BL21λ(DE3). Plate 100 μL of the transformation mix on 2*TY plates with antibiotic and with or without IPTG. Incubate O/N at 37 °C. Check the presence of colonies on IPTG-containing agar plates. No colony means that the mutation is not in the expression vector and consequently most likely in the bacterial host genome.
2. Curing the bacterial host from the plasmid: Grow the bacterial mutant host in a 250 mL flask containing 50 mL of 2*TY without antibiotic. Make daily serial 10 times dilutions of the culture and plate 100 μL of the $1/10^8$ and $1/10^7$ dilutions on 2*TY agar plates containing IPTG but no antibiotic. After O/N incubation at 37 °C, check the fluorescence under UV light. Usually, after 3–5 days of culture in the absence of antibiotic, large colonies that have lost their green fluorescence and therefore the expression plasmid will appear.

3. Isolate a large nonfluorescent colony, make a glycerol stock, and prepare calcium-competent cells. Transform the cured host with the original expression vector (not the cured one) and verify that the "colony size phenotype" on an IPTG-containing plate is restored.

3.5 Optimization of Growth Conditions

1. General rules: Choose expression hosts where the T7 RNAP expression of activity is tightly controlled. Systematically test induction temperature below 25 °C. Check the optimal IPTG concentration for your target MP (*see* Subheading 3.5.1).
2. Specific rules for BL21λ(DE3) host: Do not induce the target MP with IPTG or follow Alfasi's protocol [8] by adding extremely low concentration of IPTG (10 μM). Use pLysS as companion plasmid to downregulate the activity of the T7RNAP.
3. Specific condition for C41λ(DE3) and C43λ(DE3) bacterial mutant hosts: Test cell viability on IPTG plates. If the expression of your target MP is not toxic in C41λ(DE3) then keep this mutant host and induce with 0.1 and 0.7 mM IPTG. Test 3-h and O/N induction. Use C43λ(DE3) when expression of your target gene is toxic for C41λ(DE3). Add IPTG at 0.7 mM O/N.

3.5.1 Exploring Induction Conditions and Mutant Hosts by Flow Cytometry

If you co-express a fluorescent protein, superfold GFP for instance, then you can detect GFP fluorescence on FL1 graph (*see* **Note 6**) and also check the size (FSC) and the granularity (SSC) of your cells by using a density plot. As illustrated in Fig. 2, this can be extremely useful to compare rapid expression hosts as well as expression conditions.

1. Native conditions: Collect cells by centrifugation (300 × *g* for 2 min) and remove the supernatant. Resuspend cells in 0.5–1 mL PBS. Repeat the washing step three times. Dilute 1/1000 the cells in PBS before loading the sample on the cytometer.
2. Analysis on fixed cells: Collect cells by centrifugation and aspirate the supernatant. Resuspend cells in 0.5–1 mL PBS. Add formaldehyde to 4%. Fix for 10 min at 37 °C. Wash the cells three times with PBS. Dilute 1/1000 the cells in PBS before loading the sample on the cytometer.

3.5.2 Testing the mRNA Stability: Phenol/Chloroform RNA Extraction from E. coli

If your target membrane protein is not produced in several vector/host combinations then you should check the mRNA stability of the target gene either by quantitative real-time PCR or by "Northern blot" analysis. Obtaining high-quality RNA is the first and often most critical step. Phenol/chloroform extraction is an easy way to remove proteins from nucleic acid samples: nucleic acids remain in the aqueous phase while proteins separate into the

organic phase or lie at the phase interface. This protocol can be used also to extract RNA from bacteria grown in a rich medium. Phenol is a dangerous poison that burns the skin and the lungs upon inhalation. You must wear gloves and manipulate carefully under fume hood. A solution of PEG400 is recommended for first aid.

1. First phenol extraction: Add 700 μL of cell culture into the lysis solution maintained at 65 °C. Keep at 65 °C and vortex vigorously intermittently about ten times for 10 s. Cool the tubes on ice and centrifuge for 2 min, 15,000 × *g*, at 4 °C. Transfer the aqueous phase in a new Eppendorf tube, being careful not to contaminate with the interface phase.
2. Second phenol extraction: Add an equal volume of water-saturated phenol. Place the tubes at 65 °C and vortex vigorously intermittently about ten times for 10 s. Cool the tubes on ice and centrifuge for 2 min, 15,000 × *g*, at 4 °C. Transfer the aqueous phase in a new Eppendorf tube, being careful not to contaminate with the interface phase.
3. First phenol/chloroform extraction: Add an equal volume of phenol/chloroform and vortex vigorously intermittently about ten times for 10 s; this step can be done at room temperature. Centrifuge for 2 min, 15,000 × *g*, at 4 °C. Transfer the aqueous phase in a new Eppendorf tube, being careful not to contaminate with the interface phase.
4. First RNA precipitation: Add 1/10 the volume of 3 M Na-acetate, pH 5 (or 5 M NaCl) and 2.5 volume ethanol. Mix and place at −20 °C for 1–2 h (*see* **Note 7**). Centrifuge at 15,000 × *g* for 30–60 min at 4 °C. Remove carefully the supernatant and wash the pellet with 1 mL of 70% ethanol. Centrifuge at 15,000 × *g* for 30 min at 4 °C. Dry pellet in air and resuspend in 400 μL of sterile water.
5. Treatment with RNase-free DNase I (*see* **Note 8**; protocol provided from QIAGEN).
6. Second phenol/chloroform extraction as described in **step 3**.
7. Second RNA precipitation as described in **step 4**.
8. Measure the RNA concentration with a spectrophotometer: 1 A_{260} corresponds to 40 μg/mL RNA. Check the purity of the RNA by estimating the ratio 260/280 nm which must be 2.0.

3.6 Collecting Membranes or Inclusion Bodies from E. coli

This section provides protocols for isolation of inclusion bodies and bacterial membranes when internal membrane proliferation occurs within the cell (*see* **Note 10**).

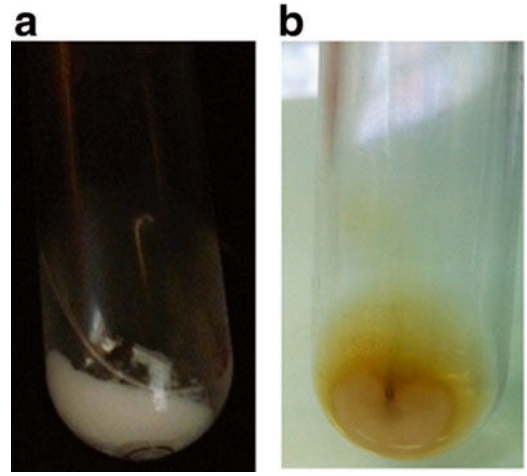


Fig. 5 Pictures of inclusion bodies and intracellular membrane pellets. (**a**) Inclusion bodies of OmpF protein 10,000 × *g* pellet, (**b**) intracellular membrane 100,000 × *g* pellet containing the b subunit of the ATP-synthase

3.6.1 Check the Presence of Inclusion Bodies

1. Breaking the cells: Harvest the culture by centrifugation at 7000 × *g*, 10 min, at 4 °C. Resuspend the pellet in 25 mL of TEP buffer. Disrupt the bacteria by passing the suspension twice in a French Press or cell disruptor.
2. Differential centrifugation: Pellet the cell debris at 600 × *g* for 10 min. Keep the supernatant. Collect the putative inclusion bodies by centrifuging the supernatant at 10,000 × *g* for 15 min at 4 °C. You should see a white brawny pellet. Collect the bacterial membranes by centrifugation of the 10,000 × *g* supernatant at 100,000 × *g* for 1 h.
3. Wash of the inclusion bodies: Wash the first pellet obtained at 10,000 × *g* with 25 mL TEP buffer supplemented with 2% Triton X-100. Centrifuge at 10,000 × *g* for 30 min. Inclusion body pellet is usually white (*see* Fig. 5a). Repeat the wash. Resuspend the pellet in 25 mL TEP buffer without detergent and centrifuge at 10,000 × *g*. Repeat the wash in order to remove all traces of detergent.
4. Resuspend the pellet in 2 mL TEP buffer and proceed to protein assay.

3.6.2 Collecting E. coli Membranes in the Absence of Inclusion Bodies

1. Breaking the cells: Harvest the culture by centrifugation at 7000 × *g*, 10 min, at 4 °C. Resuspend the pellet in 25 mL of TEP buffer. Disrupt the bacteria by passing the suspension twice in a French Press or cell disruptor.
2. Differential centrifugation: Collect the P1 pellet of internal membranes by low-speed centrifugation: 2500 × *g* for 10 min (*see* **Note 11**). Centrifuge the supernatant (S1) at 100,000 × *g* for 1 h to recover the inner and outer membranes.

3. Wash P1 with 25 mL of TEP buffer and centrifuge at 2500 × *g* for 10 min at 4 °C to remove unbroken cells (P2).
4. The supernatant (S2) contains the washed internal membranes.
5. Centrifuge for 1 h at 100,000 × *g* in order to pellet the internal membranes. You should see a brown pellet (*see* Fig. 5b).
6. Resuspend the pellet in 2 mL of buffer and assay the protein concentration.

3.6.3 Sucrose Gradient Protocol

The purpose of sucrose gradient is to concentrate and separate membrane vesicles according to their specific density. For high-purity requirements, continuous gradients are used. If you do not have access to gradient maker, then use step gradients.

1. Setting up the gradient maker: Attach capillary tubes to the end of the tubing emerging from the gradient maker. Gradient maker tubing must be clean; otherwise the sucrose gradient will not flow correctly. Close the mixer between the two compartments of the maker. Add the higher percentage sucrose solution to the outlet side of the maker. Start the stirring and add the lower percentage sucrose solution to the other compartment. Place the capillary tube on the top of the ultracentrifuge tube. Switch on the peristaltic pump and open the mixer between the two sucrose solutions. Check the flow rate of the pump to ensure that the gradient is poured drop by drop. When the gradient is completed, stop the stirrer and carefully remove the capillary tube.
2. Sample loading and centrifugation: Gently load 1 mL of 2 mg/mL protein sample on the top of the gradient paying attention not to mix the sample with the gradient. Make sure to fill up the ultracentrifuge tube (12 mL) and balance tubes with 10 mM Tris–HCl, pH 8. Centrifuge for at least 18 h at 100,000 × *g* and 4 °C. Collect 1 mL fraction in Eppendorf tubes from top to bottom.
3. Run an SDS-PAGE gel with all the fractions to detect your target MP in the different types of membranes.

3.6.4 Testing the Solubility of Your Target MP

Usually, folded MP in native membranes can be solubilized with detergent. However after production in heterologous membranes, it is frequently occurring that the target MP is difficult to solubilize. Although it is associated to the membrane fraction it might be misfolded and therefore behaves like inclusion bodies. A simple test is to compare the solubility of your target MP in three different detergents: dodecyl-maltoside (DDM), phosphododecylcholine (Fc12), and SDS.

1. Solubilization of the target MP: Prepare three Eppendorf tubes with 100 µg of your target MP in TEP buffer supplemented with 150 mM NaCl. In each tube add separately one

detergent to 1% final concentration for DDM and Fc12 and 2% for SDS (at least ten times above the critical micelle concentration). Adjust the final volume to 100 μL with buffer to perform the solubilization at 1 mg/mL. Incubate for 1 h at 4 °C and ultracentrifuge at 100,000 × *g* for 30 min.

2. Run an SDS-PAGE to check if the target MP is in the supernatant (solubilized) or in the pellet.
3. If the target MP is solubilized only by SDS, then you have inclusion bodies. If it is solubilized by all three detergents then it is likely to be well folded. If DDM cannot solubilize your target MP then try other detergents but it is likely that your target MP is misfolded.

4 Notes

1. Expression protocols are usually poorly described and you will have to go to several previous publications to find out the exact expression vector or host that was used. Try to find out the exact final yield of purified MP target per liter of culture. Below 1 mg/L you may spend 90% of your time growing cells to perform a single biophysical analysis.
2. There is a plethora of vectors and expression systems commercially available. A systematic analysis of expression protocols in bacteria [2] showed that for 80% of membrane protein structures, the two main expression systems use the T7 and arabinose promoter-based expression plasmids. The distribution of secondary structures among the different expression system is asymmetrical. For instance beta-barrel membrane protein structures were preferentially obtained using the arabinose promoter-based expression system or the T7 system with BL21λ(DE3) as expression host. In contrast, alpha helical integral membrane proteins (IMP) were almost all produced in the T7 system. The C41λ(DE3) and C43λ(DE3) bacterial hosts succeeded in producing 50% of heterologous IMP.
3. Lysozyme is a natural inhibitor of T7 RNA polymerase. If you chose to downregulate your expression system by using companion plasmids that express lysozyme (pLyS/E), take into consideration that it requires the addition of a second antibiotic which could affect considerably the cell growth.
4. When the production of the target MP is toxic, the bacteria are unable to form colonies on plates containing the inducer. The selection of new bacterial hosts (mutation in the expression vector is rare) is based upon their ability to form colonies on plate in the presence of inducer, here IPTG. By analyzing their presence, number, and size you can determine the degree of toxicity of the expression of the target protein.

5. In order to isolate and select the new mutant hosts, it is critical to have no more than 200 colonies on the plate. The frequency of occurrence of mutant hosts varies from 10^{-4} to 10^{-6}, hence the importance of diluting the culture. Figure 4 shows a fairly good correlation between the size of the colonies and the intensity of fluorescence. Most of the normal-size colonies do not exhibit fluorescence; they have lost the ability to express the gene. The smaller colonies on the other hand are almost all highly fluorescent.
6. Flow cytometers measure a variety of cellular characteristics such as relative cell size, internal complexity/granularity, cell surface properties/refractive indices, levels of autofluorescence, presence or absence of an exogenous fluorescent probe, and relative fluorescence intensities. Here we used the BD Accuri™ C6 Cytometer. The C6 cytometer is equipped with a blue and a red laser (488 and 640 nm, respectively) and four fluorescent detectors. Standard optical filters are FL1 533/30 nm (e.g., FITC/GFP), FL2 585/40 nm (e.g., PE/PI), FL3 > 670 nm (e.g., PerCP, PerCP-Cy™5.5, PE-Cy7), and FL4 675/25 nm (e.g., APC). Follow the wash procedure of the cytometer before and after the analysis to avoid cells aggregating within the cytometer.
7. If you are in hurry, you can stop the protocol at this step and store the samples at −20 °C. You can also shorten the protocol by omitting the first phenol/chloroform extraction but this is possible only if you have not grown your cells in a rich medium.
8. The treatment with DNase I-RNase free is necessary because samples from high-density *E. coli* culture may be contaminated with DNA.
9. After the 4 °C centrifugation, leave the samples at room temperature for few seconds. This clarifies the solution and helps to transfer the aqueous phase in a new clean tube.
10. On some occasions, upon overexpression of a membrane protein in *E. coli* membrane proliferation has been observed [16–20]. For instance, the overproduction of the *E. coli* b-subunit of the $F1F_o$ ATP synthase resulted in the development of a large network of intracytoplasmic membranes (ICM, Fig. 5b). The bacterial host responds to the overproduction of a membrane protein by synthesizing lipids and by converting phosphatidyl glycerol into cardiolipids at the stationary phase [18, 19].
11. It is highly unusual to collect membranes at 2500 × *g*. It might be due to the high number of cells (1 L culture at $OD_{600nm} = 8$ concentrated in 25 mL), or the high density of the membranes and their association with DNA and cell debris. However, after washing of the 2500 × *g* pellet, membranes will not anymore pellet at 2500 × *g* but, as expected, at 100,000 × *g*.

Acknowledgments

This work was supported by the Centre National de la Recherche Scientifique, INSERM, and by the "Initiative d'Excellence" program from the French State (Grant "DYNAMO," ANR-11-LABEX-0011-01). FA is supported by a DYNAMO PhD fellowship.

References

1. Tate CG (2012) A crystal clear solution for determining G-protein-coupled receptor structures. Trends Biochem Sci 37:343–352. doi:10.1016/j.tibs.2012.06.003
2. Hattab G, Warschawski DE, Moncoq K, Miroux B (2015) Escherichia coli as host for membrane protein structure determination: a global analysis. Sci Rep 5:12097
3. Dong H, Nilsson L, Kurland CG (1995) Gratuitous overexpression of genes in Escherichia coli leads to growth inhibition and ribosome destruction. J Bacteriol 177:1497–1504
4. Arechaga I, Miroux B, Runswick MJ, Walker JE (2003) Over-expression of Escherichia coli F1F(o)-ATPase subunit a is inhibited by instability of the uncB gene transcript. FEBS Lett 547:97–100
5. Moffatt BA, Studier FW (1987) T7 lysozyme inhibits transcription by T7 RNA polymerase. Cell 49:221–227
6. Wagner S, Klepsch MM, Schlegel S et al (2008) Tuning Escherichia coli for membrane protein overexpression. Proc Natl Acad Sci 105:14371–14376. doi:10.1073/pnas.0804090105
7. Miroux B, Walker JE (1996) Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260:289–298. doi:10.1006/jmbi.1996.0399
8. Alfasi S, Sevastsyanovich Y, Zaffaroni L et al (2011) Use of GFP fusions for the isolation of Escherichia coli strains for improved production of different target recombinant proteins. J Biotechnol 156:11–21. doi:10.1016/j.jbiotec.2011.08.016
9. Gräslund S, Nordlund P, Weigelt J et al (2008) Protein production and purification. Nat Methods 5:135–146. doi:10.1038/nmeth.f.202
10. Shibata Y, White JF, Serrano-Vega MJ et al (2009) Thermostabilization of the neurotensin receptor NTS1. J Mol Biol 390:262–277. doi:10.1016/j.jmb.2009.04.068
11. Egloff P, Hillenbrand M, Klenk C et al (2014) Structure of signaling-competent neurotensin receptor 1 obtained by directed evolution in Escherichia coli. Proc Natl Acad Sci U S A 111:E655–E662. doi:10.1073/pnas.1317903111
12. Drew D, Lerch M, Kunji E et al (2006) Optimization of membrane protein overexpression and purification using GFP fusions. Nat Methods 3:303–313. doi:10.1038/nmeth0406-303
13. Way M, Pope B, Gooch J et al (1990) Identification of a region in segment 1 of gelsolin critical for actin binding. EMBO J 9:4103–4109
14. Orriss GL, Runswick MJ, Collinson IR et al (1996) The delta- and epsilon-subunits of bovine F1-ATPase interact to form a heterodimeric subcomplex. Biochem J 314(Pt 2):695–700
15. Walker J, Miroux B (1999) Selection of Escherichia coli hosts that are optimized for the overexpression of proteins. Man. Ind. Microbiol. Biotechnol. 2nd Ed. MIMB2.
16. von Meyenburg K, Jørgensen BB, van Deurs B (1984) Physiological and morphological effects of overproduction of membrane-bound ATP synthase in Escherichia coli K-12. EMBO J 3:1791–1797
17. Wilkison WO, Walsh JP, Corless JM, Bell RM (1986) Crystalline arrays of the Escherichia coli sn-glycerol-3-phosphate acyltransferase, an integral membrane protein. J Biol Chem 261:9951–9958
18. Weiner JH, Lemire BD, Elmes ML et al (1984) Overproduction of fumarate reductase in Escherichia coli induces a novel intracellular lipid-protein organelle. J Bacteriol 158:590–596
19. Arechaga I, Miroux B, Karrasch S et al (2000) Characterisation of new intracellular membranes in Escherichia coli accompanying large scale over-production of the b subunit of F(1)F(o) ATP synthase. FEBS Lett 482:215–219
20. Eriksson HM, Wessman P, Ge C et al (2009) Massive formation of intracellular membrane vesicles in Escherichia coli by a monotopic membrane-bound lipid glycosyltransferase. J Biol Chem 284:33904–33914. doi:10.1074/jbc.M109.021618

Biochimica et Biophysica Acta 1859 (2017) 1124–1132

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbamem

Cardiolipin plays an essential role in the formation of intracellular membranes in *Escherichia coli*

Gerardo Carranza [a], Federica Angius [b], Oana Ilioaia [b], Audrey Solgadi [c], Bruno Miroux [b,*], Ignacio Arechaga [a,*]

[a] Departamento de Biología Molecular and Instituto de Biomedicina y Biotecnología de Cantabria (IBBTEC), Universidad de Cantabria – CSIC - SODERCAN, Santander, Spain
[b] Laboratoire de Biologie Physico-Chimique des Protéines Membranaires, Institut de Biologie Physico-Chimique, CNRS, Univ Paris Diderot, Sorbonne Paris Cité, PSL Research University, Paris, France
[c] Université Paris-Saclay, Institut Paris Saclay d'Innovation Thérapeutique, INSERM, CNRS, – Plateforme SAMM - CHATENAY-MALABRY, France

ARTICLE INFO

Article history:
Received 15 December 2016
Received in revised form 22 February 2017
Accepted 7 March 2017
Available online 09 March 2017

Keywords:
Bacteria
Cardiolipin
Electron microscopy
F1Fo-ATPase
Flow cytometry
Fluorescence
Membrane biogenesis
Phospholipid

ABSTRACT

Mitochondria, chloroplasts and photosynthetic bacteria are characterized by the presence of complex and intricate membrane systems. In contrast, non-photosynthetic bacteria lack membrane structures within their cytoplasm. However, large scale over-production of some membrane proteins, such as the fumarate reductase, the mannitol permease MtlA, the glycerol acyl transferase PlsB, the chemotaxis receptor Tsr or the ATP synthase subunit b, can induce the proliferation of intra cellular membranes (ICMs) in the cytoplasm of *Escherichia coli*. These ICMs are particularly rich in cardiolipin (CL). Here, we have studied the effect of CL in the generation of these membranous structures. We have deleted the three genes (*clsA*, *clsB* and *clsC*) responsible of CL biosynthesis in *E. coli* and analysed the effect of these mutations by fluorescent and electron microscopy and by lipid mass spectrometry. We have found that CL is essential in the formation of non-lamellar structures in the cytoplasm of *E. coli* cells. These results could help to understand the structuration of membranes in *E. coli* and other membrane organelles, such as mitochondria and ER.

1. Introduction

Gram negative bacteria are characterized by the presence of two membranes separated by a periplasmic space. However, observations of intracellular cytoplasmic membranes (ICMs) in *E. coli* have been reported, usually associated to the over-expression of membrane proteins [1–8]. The most dramatic effect has been observed in *E. coli* cells over-expressing the F-ATPase membrane subunit b [1]. The morphology of these membranes resembles that found in mitochondrial inner *cristae* membrane. We have previously shown that the soluble part of subunit b, which contains a positively charged coiled-coiled domain is essential for ICM formation [1]. Moreover, as in the case of mitochondria, these *E. coli* ICMs are characterized by an increased content in cardiolipin (CL) [1,2,9].

Cardiolipin is an anionic phospholipid mostly found in energy-transducing membranes such as the inner membranes of bacteria and mitochondria [10–12]. Because of its anionic nature, CL has been suggested to act as a proton trap that enhances the activity of respiratory complexes [13]. CL is a large phospholipid composed of two diacylphosphatidyl moieties bound to a glycerol molecule [14]. The small size of the hydroxyl head group relative to the large hydrophobic side chains confers to CL peculiar physical properties [15,16] that are essential in the formation of membrane domains in bacteria and mitochondria [17]. Fusion and fission events in these membranes are partially driven by the propensity of CL to disturb the bilayer lamellar phase [18–21] and by their ability to interact strongly with membrane proteins [8,22]. CL has been found to promote Lα to HII phase transitions and, in some cases, it has been associated to the formation of lipid cubic structures [23–26].

CL metabolism in bacteria is different to that in mitochondria [27]. Whereas in mitochondria CL levels remain constant, in bacteria these levels oscillate depending on the physiological state [28]. Bacterial CL levels increase in stationary phase [29] or in response to osmotic stress and energy depravation [30]. CL has been found to be dispensable in *E. coli* mutants, providing that its function is restored by a second protein or that the bacteria are gown in laboratory conditions in the presence of divalent cations [31,32].

CL synthesis in *E. coli* is a reversible reaction catalysed by the product of three *cls* genes [33,34]. ClsA, encoded by the *clsA* gene [35,36] is a phospholipase D enzyme which catalyses the transfer of the phosphatidyl moiety between two phosphatidyl-glycerol molecules (PG) to form CL and glycerol [27,37,38]. Deletion of *clsA* in *E. coli* prevents CL

* Corresponding author at: Institut de Biologie Physico-Chimique, CNRS, Univ Paris Diderot, Sorbonne Paris Cité, PSL Research University, Paris, France, and Departamento de Biología Molecular and Instituto de Biomedicina y Biotecnología de Cantabria (IBBTEC), Universidad de Cantabria- CSIC- SODERCAN, Santander, Spain.
E-mail addresses: miroux@ibpc.fr (B. Miroux), arechagai@unican.es (I. Arechaga).

http://dx.doi.org/10.1016/j.bbamem.2017.03.006
0005-2736/

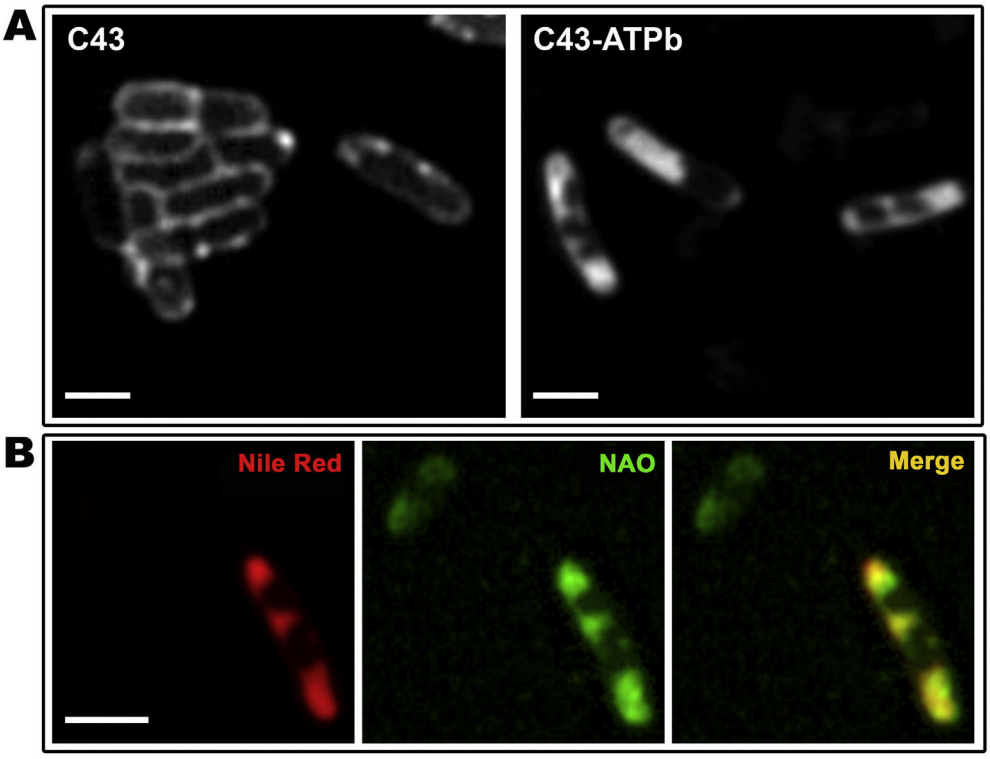


Fig. 1. Membrane staining of *E. coli C43(DE3)* overexpressing F-ATPase subunit b. *A*, *E. coli* cells were labelled with Nile Red (2 μM) and examined by fluorescence microscopy. Wild type C43(DE3) cells harbour either an empty plasmid (*left panel*) or produce F-ATPase subunit b (*right panel*). *B*, Co-localization of Nile Red and NAO staining on C43(DE3) cells producing F-ATPase subunit b. Scale bar, 1 μm.

synthesis in the logarithmic phase but not the stationary phase [39]. CL synthesis in this phase is catalysed by the product of *clsA* as well as by two additional enzymes encoded by *clsB* (also known as *ybhO*) [40] and *clsC* (also known as *ymdC*) genes [34]. Only upon deletion of the three *cls* genes can CL be totally eliminated [34]. CL synthesis catalysed by ClsB and ClsC differs to the reaction catalysed by ClsA, as they use PE (phosphatidyl-ethanolamine) and PG for the formation of CL and ethanolamine [34].

In this report we have investigated the role of CL in the formation of ICMs in *E. coli* C43(DE3) cells over-producing F-ATPase subunit b. We have generated single (*ΔclsA*), double (*ΔclsAC*) and triple mutants (*ΔclsABC*) of CL synthase genes and we have analysed the effect on membrane morphology and lipid metabolism by fluorescent, electron microscopy and lipid mass spectrometry, respectively. The results suggest that CL-promoted Lα/HII phase transitions are favoured by over-expression of F-ATPase subunit b.

2. Materials and methods

2.1. Plasmids and molecular biology techniques

E. coli atpF gene, which encodes F-ATPase subunit b, was cloned into a pMW7vector as previously described [1]. The genes encoding TrwB and TrwC were cloned into pET3a plasmid (Novagene) using conjugative R388 plasmid as template, resulting in plasmids pMEC20 and pJR03, respectively. Genetic material was amplified by PCR using Phusion DNA Polymerase (Thermo Scientific) following the manufacturer's protocols. Primers were synthesized by Integrated DNA Technology. Colony PCR was used to screen for appropriate colonies, and then the PCR products were sequenced by Sanger sequencing (Stab vida). Plasmid transformations were accomplished using electroporation. Genomic gene deletions and plasmid constructs were verified by PCR and DNA sequencing. All strains are listed in Table S1.

2.2. Protein over-expression

Bacteria were grown in LB liquid medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) at 37 °C to an optical density of 0.6 at 600 nm. Then isopropyl-2-D-thio-galactopyranoside (IPTG) was added to a final concentration of 0.7 mM. The cells were grown for 3 h at 37 °C with appropriate antibiotics (100 μg/mL of ampicillin and/or 25 μg/mL of kanamycin). Viability experiment was performed as follow: serial 10 times dilutions were performed to reach 1 to 10^8 dilution of the culture. A portion (100 μl) of the last three dilutions was spread on ampicillin (50 μg/ml) agar plates. The number of viable cells was determined by counting the colonies on the most suitable plate with 100 to 300 colonies per plate.

2.3. Construction of chromosomal mutants lacking cls genes

C43(DE3) mutants containing deletions of the *cls* genes were constructed using the Keio collection of mutants [41]. Keio mutants harbour a kanamycin cassette flanked by two FLP recombinase sites, allowing for the excision of the cassette. P1vir transduction and calcium availability to regulate infectivity were used to transfer these cassettes from the

Table 1
Phospholipid composition in control cells (total fraction).

Strain	PE	PG	CL	LPE
C43(DE3)	75	21	4	*n.d.*
C43(DE3) *ΔclsA*	73	25	2	*n.d.*
C43(DE3) *ΔclsAC*	69	24	2	5
C43(DE3) *ΔclsABC*	68	27	*n.d.*	4

Control cells harbouring the empty pMW7 plasmid were harvested and membranes were purified as described in M&M. PE, phosphatidyl-ethanolamine; PG, phosphatidyl-glycerol; CL, cardiolipin; LPE, lyso-phosphatidyl ethanolamine. Amounts of phospholipids are given as molar percentage. *n.d.*, no detectable.

Table 2
Phospholipid composition in ICMs from cells expressing *F-ATPase subunit b*.

Strain	PE	PG	CL	LPE
C43(DE3)	77 ± 4	14 ± 6	9 ± 0.5	*n.d.*
C43(DE3) *ΔclsA*	50	50	*n.d.*	*n.d.*
C43(DE3) *ΔclsAC*	56	44	*n.d.*	*n.d.*
C43(DE3) *ΔclsABC*	55 ± 3	43 ± 4	*n.d.*	2

Cells containing pMW7-b plasmid, which encodes for *atpF* gene were harvested after overnight induction at 25 °C. Membranes were fractioned and processed for lipid analysis as described in M&M. Amounts of phospholipids are given as molar percentage. *n.d.*, no detectable.

mutants to C43(DE3) strain. The excision of the kanamycin cassette was carried out using the plasmid pCP20, which is a plasmid with an ampicillin resistance gene that shows temperature-sensitive replication and thermal induction of FLP synthesis [42]. FLP recombinase mediates the recombination of the two flanking FLP sequences ejecting the kanamycin sequence leaving only a FLP scar sequence. Transformants for pCP20 were selected for ampicillin resistance at 30 °C. Colonies were purified at 42 °C without antibiotics to cure the pCP20 plasmid. C43(DE3) mutants were replicated on plates without antibiotics and plates with ampicillin or kanamycin. Colonies that did not grow on plates with antibiotics were verified as deletions using PCR and DNA sequencing. The derived strains were subjected to another round of P1vir transduction and elimination of the kanamycin marker to introduce additional null mutations.

2.4. General fluorescence microscopy

Cells were immobilized on microscope slides covered with a thin film of 1.2% agarose in water. Membranes were visualized with Nile Red (Sigma-Aldrich) (1.5 μM) and 10-nonyl bromide acridine orange (NAO, Sigma-Aldrich) (2 μM). The fluorescent dyes were excited using a mercury lamp. Red fluorescence (with excitation at 550 nm and emission at 640 nm) from Nile Red was detected by using a standard Rhodamine filter unit (Ex. 546/12–Em. 608/65). Green fluorescence (with excitation at 495 nm and emission at 525 nm) from NAO was detected by using a standard GFP filter unit (Ex. 470/40–Em. 525/50). Standard fluorescence microscopy was carried out using Zeiss Axio Imager M1 upright fluorescence microscope (Zeiss Plan-Neofluar _100/1.30 NAOil Ph3objective) equipped with a 12 bits B&W camera (AxioCam MRm). The image analysis was carried out using ImageJ v.1.38 (National Institutes of Health).

2.5. Time-lapse microscopy

Cells were seeded by adding diluted samples onto individual 1.5% (wt/vol) low-melt agarose pads in M9 minimal medium using glucose as the carbon source as described in [43]. Membranes were visualized with NAO or with *N*-(3-triethylammoniumpropyl)-4-(p-diethylaminophenyl-hexatrienyl) pyridinium dibromide (FM 4-64, Thermo-Fisher) (1.5 μM) vital membrane stain. Fluorescence from FM4-64 (excitation at 515 nm and emission at 720 nm) was detected by using a set of filters TRICT-Cy5 (Ex. 560/25 nm–Em. at 700/25 nm). Minimizing exposure of bacteria to fluorescence was essential for assuring their survival during observation, so short exposure times were used (10–50 ms for NAO and FM4-64, respectively). The analysis was carried out using a Nikon Eclipse Ti (Nikon CFI APO TIRF_100/1.49 NA Oil objective) equipped with hardware autofocus module (PFS), an environmental chamber set to 37 °C and with a Hamamatsu OrcaR2 CCD camera. Images were taken every minute for 1 h and the image analysis was carried out using ImageJ v.1.38 (National Institutes of Health).

2.6. Electron microscopy

Cells were harvested, pelleted and rinsed in 0.12 M phosphate buffer, pH 7.4, fixed with 3% glutaraldehyde, and then with 2% osmium tetroxide. Fixed cells were dehydrated in acetone, and embedded in araldite

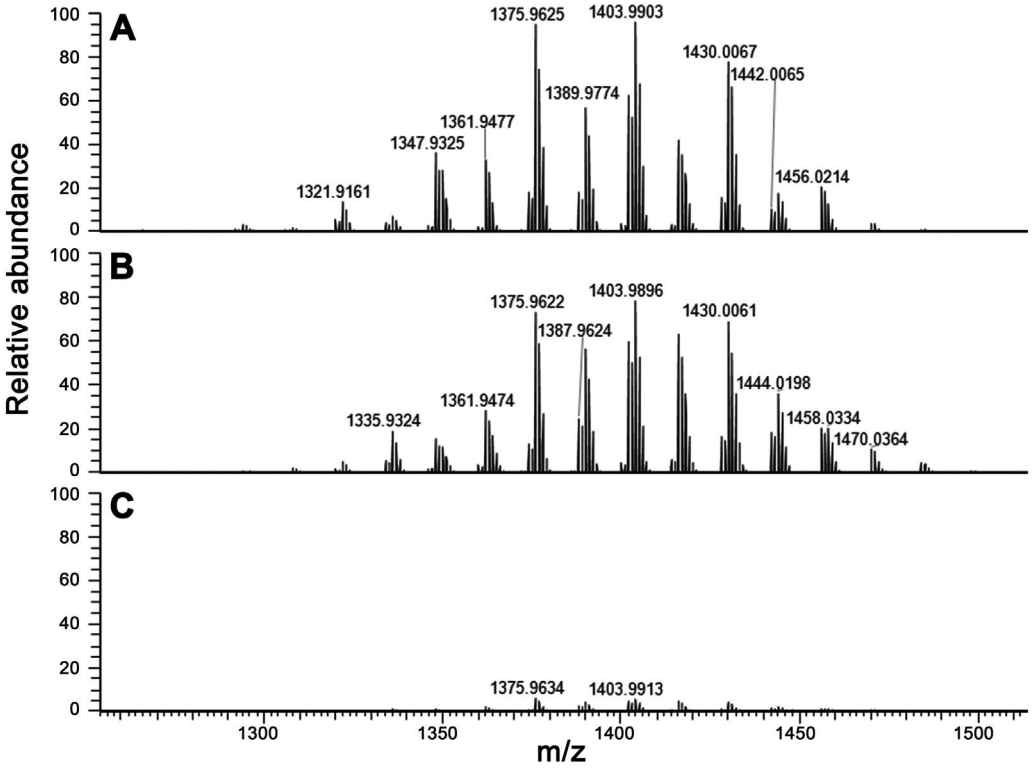


Fig. 2. Identification of Cardiolipin molecular species by mass spectrometry. After lipid extraction, phospholipid analysis was performed by LC-CAD-ESI/MS. Cardiolipin molecular species identified in: *A*, C43(DE3) cells over-producing F-ATPase subunit b; *B*. C43(DE3) control cells harbouring the empty pHis17 plasmid; *C*, C43 (DE3)Δcls(ABC) triple mutant cells over-expressing F-ATPase subunit b. The identification of the fatty acids in CL compatible with these molecular masses are shown in Table 3.

Table 3
Cardiolipin molecular species identified in *E. coli* C43(DE3) upon overproduction of F-ATPase subunit b.

Experimental mass $[M-H]^-$	Exact mass $[M-H]^-$	CL molecular species	Molecular formula $[M-H]^-$
1321.9161	1321.9180	62:1	$C_{71}H_{135}O_{17}P_2^-$
1333.9179	1333.9179	63:2	$C_{72}H_{135}O_{17}P_2^-$
1335.9327	1335.9336	63:1	$C_{7l}H_{137}O_{17}P_2^-$
1347.9325	1347.9336	64:2	$C_{73}H_{137}O_{17}P_2^-$
1349.9342	1349.9493	64:1	$C_{73}H_{139}O_{17}P_2^-$
1361.9477	1361.9493	65:2	$C_{74}H_{139}O_{17}P_2^-$
1373.9683	1373.9505	66:3	$C_{75}H_{139}O_{17}P_2^-$
1375.9625	1375.9649	66:2	$C_{75}H_{141}O_{17}P_2^-$
1387.9630	1387.9649	67:3	$C_{76}H_{141}O_{17}P_2^-$
1389.9774	1389.9806	67:2	$C_{76}H_{143}O_{17}P_2^-$
1401.9775	1401.9806	68:3	$C_{77}H_{143}O_{17}P_2^-$
1403.9903	1403.9962	68:2	$C_{77}H_{145}O_{17}P_2^-$
1415.9920	1415.9962	69:3	$C_{78}H_{145}O_{17}P_2^-$
1427.9920	1427.9962	70:4	$C_{79}H_{145}O_{17}P_2^-$
1430.0067	1430.0119	70:3	$C_{79}H_{147}O_{17}P_2^-$
1442.0065	1442.0131	71:4	$C_{80}H_{147}O_{17}P_2^-$
1444.0205	1444.0275	71:3	$C_{80}H_{149}O_{17}P_2^-$
1456.0214	1456.0275	72:4	$C_{81}H_{149}O_{17}P_2^-$

(Durcupan, Fluka, Switzerland). Cross sections were obtained by cutting embedded cells using a Leica ultracut UCT ultramicrotome. Ultrathin sections were stained with uranyl acetate in the dark and examined with a JEOL 2010 electron microscope.

2.7. Lipid analysis

Bacteria were grown in 2*TY medium (16 g/L tryptone, 10 g/L yeast extract, 5 g/L NaCl, pH 7.2) at 37 °C to an optical density of 0.6 at 600 nm. IPTG was added to a final concentration of 0.7 mM and cells were thereafter grown overnight at 25 °C. ICM from C43(DE3) bacterial host and C43(DE3)*Δcls* derivatives were isolated as described in [1]. In control cells harbouring pMW7 plasmid, ICM are not present and therefore total membranes (both outer and inner) were isolated by100,000*g* ultracentrifugation for 1 h at 4 °C. Lipid analysis was performed by mass spectrometry at the SAMM mass spectrometry facility following the protocol described in [44].

2.8. Flow cytometry

Cell were stained with NAO (2 μM) 16 h after induction with IPTG in 2*TY medium and incubated for 1h at 37 °C. Excess NAO was removed by washing the cells with PBS (10 mM Phosphate, 150 mM NaCl, pH 7.4). Cells were diluted in PBS to 10 $mOD_{600\ nm}$ prior cytometry analysis. Cells were counted on a C6 Accury cytometer with a threshold of 10000 on forward scatter (FSC) and of 3000 on side scatter (SSC) to remove electronic noise and small contaminants. NAO green fluorescence was analysed on a FL1 detector (533/30 nm).

3. Results

3.1. Intracellular membranes in cells overproducing F-ATPase subunit b analysed by fluorescent microscopy

Intracellular membrane (ICM) formation in *E. coli* is usually monitored by EM analysis of cross sections of fixed cells [1,2,4,5,7]. However, this procedure is not applicable for studying dynamic properties of ICMs. Therefore we aimed at examining membrane formation by fluorescent optical microscopy. First, we generated chimeric fusion proteins of F-ATPase subunit b with blue fluorescent proteins. However, over-expression of these fusion proteins resulted in the absence of ICMs, as observed by EM analysis of cross-sections of cells expressing the fusion protein F-ATPase subunit b-BFP (Fig. S1), which confirms that the structure of the membrane network is dependent on the soluble part of F-ATPase subunit b [1]. To overcome this limitation, three fluorescent lipid probes (NAO, Red-Nile and FM4-64) widely used for membranes and organelles staining were employed. Image analysis by fluorescence microscopy of *E. coli* C43(DE3) cells over-producing F-ATPase subunit b stained with Nile Red (2 μM) revealed the presence of multiple internal and cytoplasmic foci, which were absent in control C43(DE3) cells

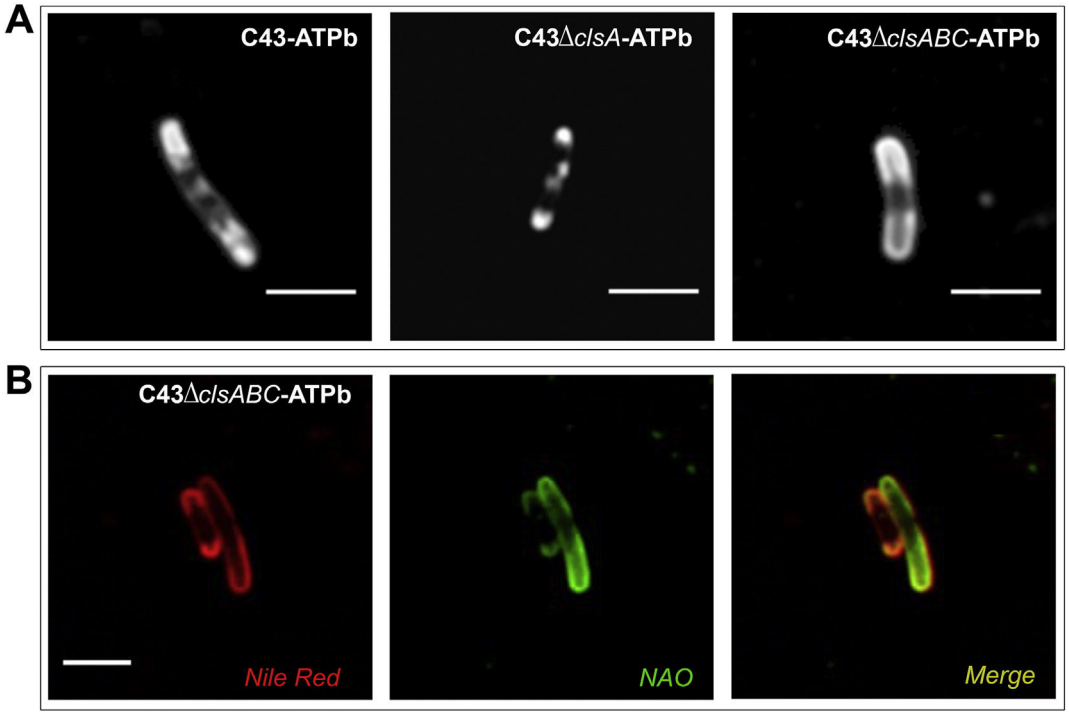


Fig. 3. Membrane staining of cardiolipin synthase *cls* mutants. *A*, Nile Red staining of cells over-producing F-ATPase subunit b. *Left panel*: C43(DE3) *Middle panel*: C43(DE3)*ΔclsA* and *right panel*: C43(DE3)*ΔclsABC*. *B*, Co-localization of Nile Red and NAO staining in C43(DE3)*ΔclsABC* cells over-producing F-ATPase subunit b.

harbouring an empty plasmid (Fig. 1A). The multiple internal foci pattern was also absent (Fig. S2) in additional controls consisting of cells over-producing the conjugative coupling membrane protein TrwB, or in cells over-expressing a soluble protein such as the conjugative relaxase TrwC, two proteins that do not induce ICM formation. The stain pattern observed by fluorescence microscopy in C43(DE3) cells over-expressing F-ATPase b subunit matches that observed by electron microscopy (Fig. S3), thus suggesting that the observed *foci* corresponded, in fact, to the ICMs. To confirm this result, C43(DE3) cells overproducing F-ATPase subunit b were also stained with NAO (10-*N*-nonyl acridine orange), a probe that stains preferentially anionic phospholipids [45] and that has been used to measure CL levels in mitochondria [46–49] and bacteria [17,50]. As observed in Fig. 1B, NAO stain (2 μM) co-localized with Nile Red, demonstrating the presence of anionic phospholipids (CL and PG) in the ICMs.

3.2. Phospholipid composition of cells lacking cardiolipin synthase genes

In order to assess the role of cardiolipin in ICMs formation we deleted sequentially the three *cls* genes responsible of CL biosynthesis in *E. coli* [34] and we grew them in C43(DE3) cells. Lipid content in total cell membranes from each mutant strain was analysed by mass-spectrometry in the absence and in the presence of recombinant F-ATPase subunit b. First, we analysed cells that were not over-producing F-ATPase subunit b. As shown in Table 1, deletion of *clsA* gene hardly affects phospholipid composition of cells harvested at stationary phase. However, phospholipid composition of *ΔclsAC* double mutant is affected, resulting in a reduction of PE and CL levels (69% and 2% respectively) and concomitant increase of LPE (5%, lyso-phosphatidylethanolamine). Triple deletion of *cls* genes (*ΔclsABC*) results in no detectable CL. In a second step, F-ATPase subunit b was overproduced in all three *cls* C43(DE3) mutants (Fig. S4). After overnight induction at 25°C, membrane fractions were isolated and their phospholipid composition was determined. In all three *cls* deletion mutants CL level was below the detection limit, which was concomitant with a large increase in PG (+307%) and reduction (−140%) of PE levels (Table 2). In contrast, ICM isolated from C43(DE3) cells over-producing F-ATPase subunit b showed the highest content of CL (4.7%).

The fatty acid composition of CL in control cells and in cells over-producing F-ATPase subunit b was analysed by LC-MS (Fig. 2 and Table 3).

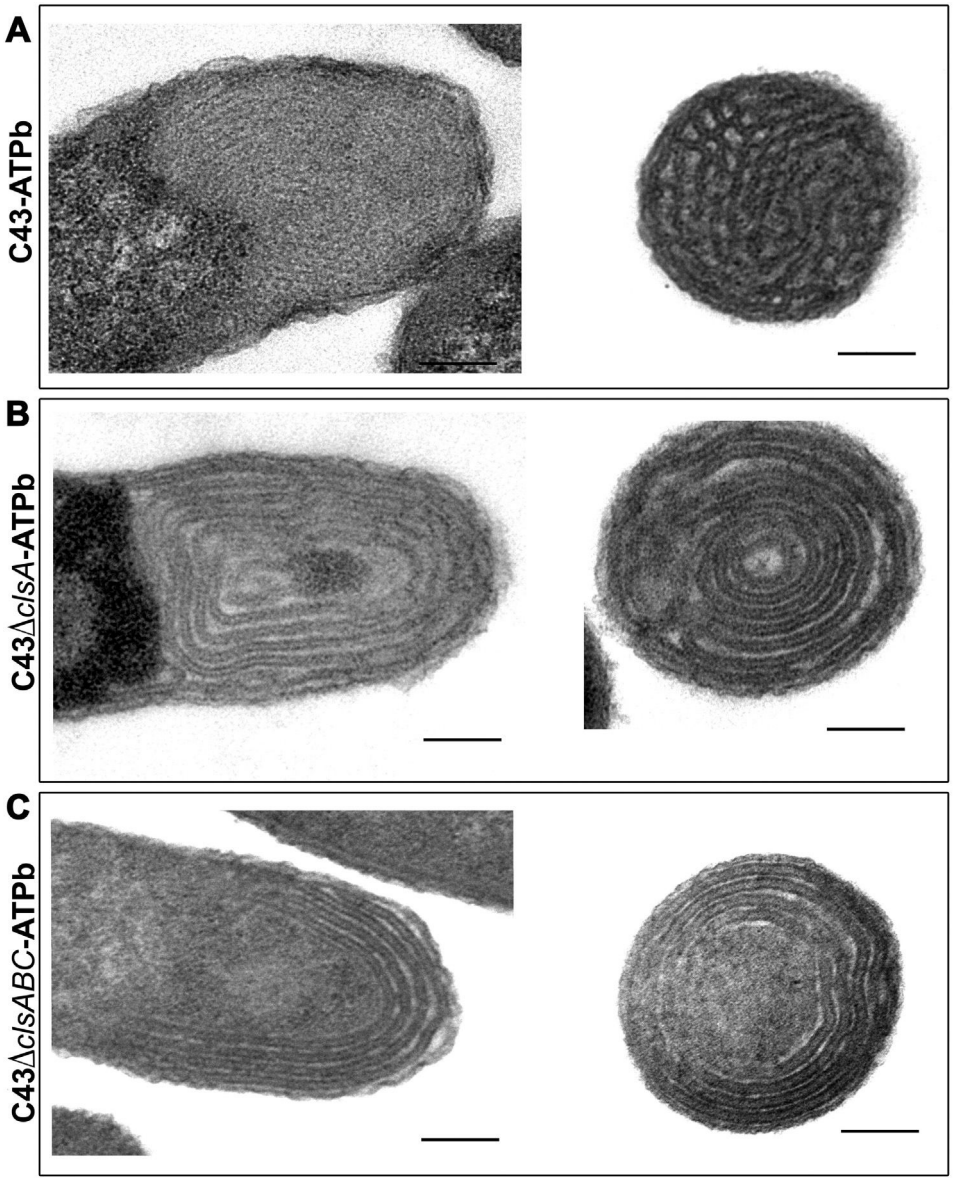


Fig. 4. Electron microscopy of C43(DE3) cls mutants. Electron micrographs of thin sections of C43(DE3) (*A*), C43(DE3)*ΔclsA* (*B*) and C43(DE3)*ΔclsABC* (*C*) cells over-producing F-ATPase subunit b. Control cells with an empty vector do not develop any ICM [1] Scale bar, 0.2 μm.

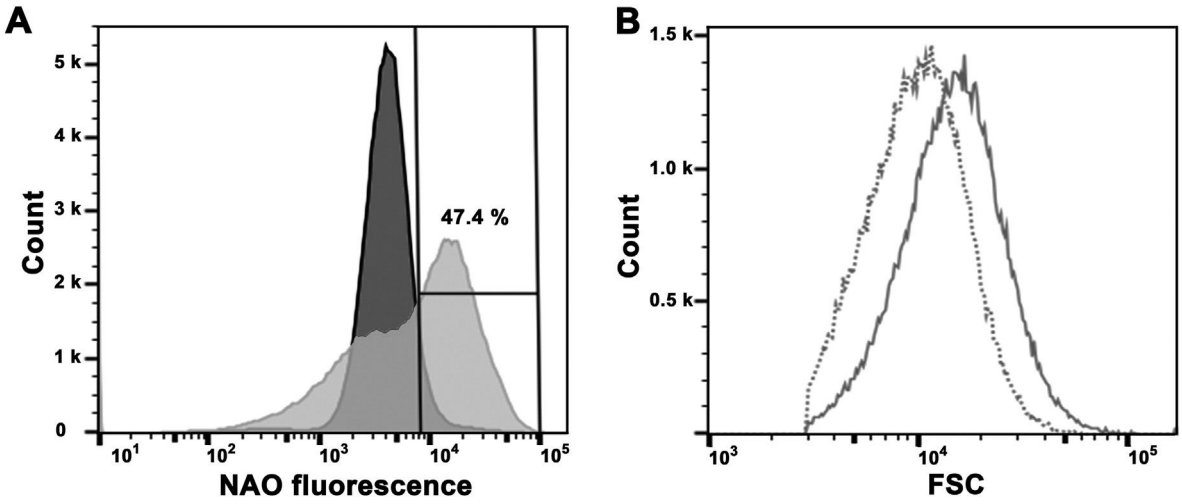


Fig. 5. Flow cytometry analysis of C43(DE3) cells upon ICMs formation. Cells were grown at 37 °C and left OVN at 25 °C after addition of IPTG. After NAO staining (2 mM, OD600 nm = 0.1), cells were analysed on a C6 cytometer as described in experimental procedure. *A*, FL1 green fluorescence intensity of NAO stained cells harbouring a control empty plasmid (*dark gray*) or cells overproducing F-ATPase subunit b (*light gray*). *B*, forward scatter analysis of both NAO negative (*dashed line*) and positive populations (*continuous line*) of C43(DE3) cells overproducing F-ATPase subunit b.

We found that the over-expression of F-ATPase subunit b did not significantly affect the fatty acid composition of CL, being palmitic (16:0) and vaccinic (18:1) acids, the most abundant fatty acids, which was similar to previously reported findings [51].

3.3. Cardiolipin depletion in C43(DE3) cells disrupts ICMs formation

C43(DE3) *cls* mutants were stained with Nile Red and analysed by optical microscopy (Fig. 3A). Comparison of the staining pattern between C43(DE3), single *ΔclsA* mutant and *ΔclsABC* triple mutant cells over-producing F-ATPase subunit b revealed the absence of internal *foci* in the triple mutant. Interestingly, in the case of single deletion mutant (*ΔclsA*) some inner foci could be observed, albeit in much lower amounts than wild type. Co-localization experiment with both NAO and Nile Red showed that PG accumulates at the poles of C43(DE3)*ΔclsABC* triple mutant cells (Fig. 3B).

In order to obtain a higher resolution view of the membranes, cross-sections of cells were analysed by electron microscopy. Close inspection of C43(DE3) cells over-producing F-ATPase subunit b (Fig. 4A) showed the characteristic internal membrane morphology previously described [1]. In contrast, and as observed in the fluorescence experiments, the morphology of the ICMs in the cls mutants was affected. Analysis of C43(DE3)*ΔclsA* cells over-producing F-ATPase subunit b revealed the presence of multi lamellar structures (Fig. 4B). In the case of the triple

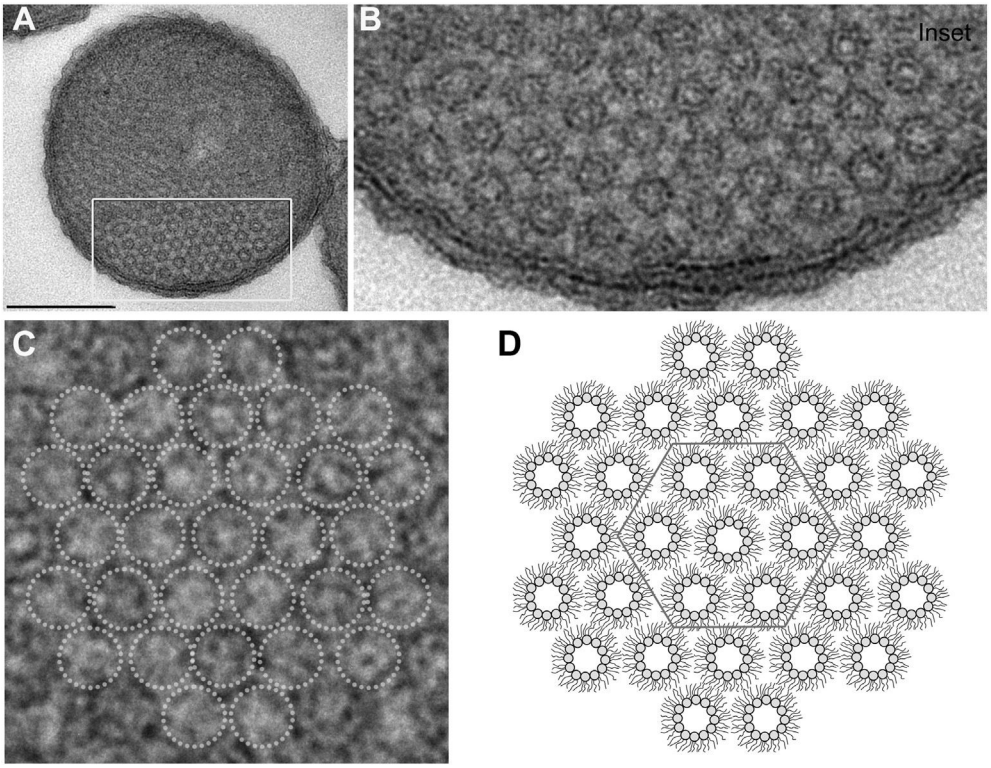


Fig. 6. Nonlamellar morphology of ICMs. *A*, Orthogonal view of thin sections of *E. coli* C43(DE3) cells over-producing F-ATPase subunit b. Scale bar, 0.25 μM. *B*, Magnified view of the boxed area marked in A. *C*, Magnification of B, dotted circles indicate ICMs hexagonal phase. *D*, Schematic representation of inverted hexagonal phases as described in [73–76]

mutant (*ΔclsABC*) similar lamellar phenotype could still be observed, but to lower extent than in the single mutant (Fig. 4C). Analysis of the double cardiolipin synthase mutant C43(DE3)*ΔclsAC* revealed a morphology similar to that found in the triple mutant (data not shown). Control C43(DE3) cells do not present any ICM [1].

3.4. Dynamics of ICM formation upon over-production of F-ATPase subunit b

The dynamics of ICMs formation was analysed by time-lapse fluorescent microscopy and by flow cytometry. Video image analysis of C43(DE3) cells over-producing the F-ATPase subunit b stained with NAO suggested that ICM proliferation starts at the cell poles and then progresses along the cell volume (Fig. S5A and Movie S1). Analysis of time-lapse images of cells stained with FM4-64 (Fig. S5B bottom panel and Movie S2) was in agreement with the dynamic of internal foci formation observed by NAO. In this experiment cells did not duplicate over the one hour time course experiment. This growth arrest was associated with ICM formation and not with the staining, as control cells are able to grow in the presence of FM4-64 (Fig. S5B upper panel and Movie S3), suggesting that ICM formation was deleterious for cell growth and duplication. Close inspection of time lapse images revealed that only a fraction (15–20%) of the cells exhibited ICMs as reflected by the strong fluorescence signal. To confirm this observation, cells over-producing the F-ATPase subunit b or harbouring the control pHis17 empty plasmid were analysed by flow cytometry after NAO staining. After overnight induction 99% of the control cells remained NAO negative (Fig. 5A, mean fluorescence of 4733). In contrast, the NAO mean fluorescence of cells overproducing F-ATPase subunit b increased 2.5 times. Cell population analysis showed that 52.6% of the cells remained NAO negative while 47.4% of cells exhibited a mean fluorescence intensity of 21324 (Fig. 5A), which conforms with the fact that only a fraction of the bacterial population is highly fluorescent, as observed by video-microscopy. In addition, the size (FSC parameter) of both NAO negative and positive cells was compared. Fig. 4B shows that NAO positive cells are 40% larger (11903 and 17206 FSC mean values for NAO negative and positive cells, respectively) which is consistent with the size of the NAO positive cell observed by microscopy. To address the toxicity phenotype observed by video microscopy upon overproduction of F-ATPase subunit b, viability experiments were performed as described in experimental procedures. Before induction both control cells and cells harbouring the F-ATPase subunit-b plasmid exhibited the same viability (suppl. Fig. S6). In contrast, 3 h after the overproduction of F-ATPase subunit b, viability of the cells was on average 15 times decreased suggesting that the fitness of cells overproducing ICM is affected.

3.5. ICMs morphology suggests the formation of inverted hexagonal (HII) phases

Close inspection of ICMs associated with over-production of F-ATPase subunit b (Fig. 6) revealed an inverted hexagonal morphology similar to that found in liposomes made from pure CL [52], The formation of H_{II} phases in liposomes is favoured by addition of divalent cations [23]. The hexagonal phases found in C43(DE3) cells over-producing F-ATPase subunit b were absent in control cells and in cells in which *cls* genes were deleted. These results suggested that the formation of hexagonal phases was dependent on the charge and topology of the over-expressed protein and in the presence of CL.

4. Discussion

Cardiolipin (CL) plays an essential role in membrane biogenesis in bacteria and mitochondria. The function of CL in mitochondria and other eukaryotic organelles such as the endoplasmic reticulum has been extensively studied [17,53] showing an essential role of CL in the cristae formation. However, there is still a lack of information on CL contribution to the dynamics and morphology of bacterial membranes. Here, we have shown that, CL depletion in C43(DE3) cells over-producing F-ATPase subunit b had a dramatic effect on membrane morphology.

Previous studies have found that it is necessary to delete the three *E. coli cls* genes to achieve total CL depletion [34]. Single deletion of *clsA* gene results in CL depletion only in cells grown at low salt and harvested in logarithmic phase [39]. However, in cells harvested in stationary phase CL traces were still detectable [32]. Our results with C43(DE3) cell extracts harbouring single (*ΔclsA*), double (*ΔclsAC*) or triple (*ΔclsABC*) cardiolipin synthase mutants confirm those previous findings (Table I). For instance, deletion of *clsB* and *clsC* genes, resulted in an increase in lyso-phosphatidyl-ethanolamine levels, which is in accordance with the reaction catalyzed by both enzymes. Moreover, *cls* mutants over-producing F-ATPase subunit b lack any of CLs. In contrast, we found a dramatic increase in PG levels and a decrease in PE. On the other hand, wild type C43(DE3) cells over-producing F-ATPase subunit b had normal PE and PG contents but a higher content of CL. These results indicate that the effects of *cls* deletion on lipid metabolism are enhanced by the overproduction of F-ATPase subunit b.

CL is a polymorphic phospholipid [52] that can be assembled in micellar, lamellar or hexagonal membrane structures depending on its acyl composition [54] or in the presence of divalent cations [23–26,55]. CL polymorphic behavior has been suggested to be essential in different fusion and fission membrane events [18]. Depending on the growth conditions *E. coli* CL fatty acid composition fluctuates [56], increasing the number of acyl chain unsaturations as temperature decreases [51]. Under physiological conditions CL is mostly found in the bilayer phase. However, in membrane contact areas and in other regions where the bilayer is perturbed CL tends to form hexagonal phases [15,53]. Due to its geometry CL has been found to localize in negative curved regions in *E. coli* membranes [57,58]. Lateral pressure of lipids in a monolayer decreases upon insertion of a CL molecule which, in turn, contributes to the creation of membrane folds [53,59,60].

Lipid packing modulation by CL [61,62] has been reported to play an important role in reshaping inner mitochondrial morphology. In particular, CL has been found to disturb bilayer lamellar phase [18,19] and promote the formation of inverted hexagonal II phases [23,63]. Lipids in inverted hexagonal phases self-assemble in long tubes arranged in a hexagonal lattice [64]. Interestingly, this hexagonal pattern can be observed in cells over-producing F-ATPase subunit b (Fig. 5). It is likely that the peculiar CL physical properties also play an essential role in the creation of ICMs reported here.

On the other hand, our results point out that the interactions of F-ATPase subunit b with CL mediate membrane remodelling and lattice formation. In the C43(DE3) *ΔclsABC* strain, lamellar structures are observed upon over-production of F-ATPase subunit b but remain unstructured. This has also been observed upon addition of dynamin-related protein 1(Drp1) to CL containing lipid in *in vitro* systems [65]. Authors suggested that Drp1 association with CL increases the propensity for lamellar to HII phase transitions, which could be essential in mitochondrial division.

In addition to CL physical properties, the topology of the protein is also important in the formation of the hexagonal phases. Overexpression of membrane proteins with different topology, such as the conjugative coupling protein TrwB results in the lack of ICMs (Suppl. Fig. S2). In a previous study we showed that the soluble coiled-coiled domain of F-ATPase subunit b was essential for ICMs formation [1]. Here we also showed that C-terminal fusion of F-ATPase subunit b with the blue fluorescent protein (BFP) is deleterious for ICMs formation (Suppl. Fig. S1). This dependence on protein topology for the formation of ICMs in *E. coli* has been thoroughly discussed in a previous report [9]. It is worth noting that mitochondrial F-ATPase subunit b has also been shown to be important for the stability of inner *cristae* morphology [66], and that mutations on the protein results in the destabilization of the *cristae* [67].

Stabilization of bacterial ICMs could be mediated by electrostatic interactions between F-ATPase subunit b and cardiolipin. CL is able to interact with a large number of membrane proteins [68,69]. At neutral pH, CL is negatively charged as any of the phosphate moieties of its head group is able to pick up an electron and make a H bond with the free OH of the glycerol [70,71] (Fig. S7A). Electrostatic interactions between basic residues of integral membrane proteins and CL are common [22]. For instance, solid state NMR experiments have shown electrostatic interactions of CL with a basic residue in F-ATPase subunit c [72]. Given that F-ATPase subunit b contains 16 lysines in its soluble part (Fig. S7B), it is likely that such interactions also play an important role in ICM stabilization.

5. Conclusions

The biophysical properties of cardiolipin (CL) and its role in shaping membrane morphology have been characterized in a natural biological context. CL forms inverted hexagonal phases (HII) only in C43(DE3) cells over-producing membrane proteins. This morphology is dependent on the topology of the over-expressed protein and on CL synthesis as modification of the topology of the recombinant protein and mutation of the enzymes responsible of CL biosynthesis in *E. coli* affects the morphology of ICM. Intracellular membranes formation in *E. coli* could be a suitable model system to study not only the dynamics of membrane formation in bacteria but also other CL containing biological organelles such as the inner mitochondrial membrane or ER.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbamem.2017.03.006.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Transparency document

The Transparency document associated with article can be found, in online version.

Acknowledgments

This work was supported by the Spanish Ministerio de Economía y Competitividad (MINECO) grant BFU2013-49486-EXP (to I A), by the Centre National de la Recherche Scientifique, INSERM, and by the "Initiative d'Excellence" program from the French State (Grant "DYNAMO", ANR-11-LABEX-0011-01) (to BM). FA is supported by a DYNAMO PhD fellowship. The authors acknowledge Jorge Mata, Dr. F. Madrazo, Dr. M.T. Berciano and Dr. M. Lafarga at the IDIVAL institute in Santander, Spain, for help with the cross-sections of cells. We also like to acknowledge Région Ile de France for co-funding the SAMM MS Facility at IPSIT.

References

[1] I. Arechaga, B. Miroux, S. Karrasch, R. Huijbregts, B. de Kruijff, M.J. Runswick, J.E. Walker, Characterisation of new intracellular membranes in *Escherichia coli* accompanying large scale over-production of the b subunit of F(1)F(o) ATP synthase, FEBS Lett. 482 (2000) 215–219.
[2] J.H. Weiner, B.D. Lemire, M.L. Elmes, R.D. Bradley, D.G. Scraba, Overproduction of fumarate reductase in *Escherichia coli* induces a novel intracellular lipid-protein organelle, J. Bacteriol. 158 (1984) 590–596.
[3] R.P. van Weeghel, W. Keck, G.T. Robillard, Regulated high-level expression of the mannitol permease of the phosphoenolpyruvate-dependent sugar phosphotransferase system in *Escherichia coli*, Proc. Natl. Acad. Sci. U. S. A. 87 (1990) 2613–2617.
[4] W.O. Wilkison, J.P. Walsh, J.M. Corless, R.M. Bell, Crystalline arrays of the *Escherichia coli* sn-glycerol-3-phosphate acyltransferase, an integral membrane protein, J. Biol. Chem. 261 (1986) 9951–9958.
[5] R.M. Weis, T. Hirai, A. Chalah, M. Kessel, P.J. Peters, S. Subramaniam, Electron microscopic analysis of membrane assemblies formed by the bacterial chemotaxis receptor Tsr, J. Bacteriol. 185 (2003) 3636–3643.
[6] J. Lefman, P. Zhang, T. Hirai, R.M. Weis, J. Juliani, D. Bliss, M. Kessel, E. Bos, P.J. Peters, S. Subramaniam, Three-dimensional electron microscopic imaging of membrane invaginations in *Escherichia coli* overproducing the chemotaxis receptor Tsr, J. Bacteriol. 186 (2004) 5052–5061.
[7] K. von Meyenburg, B.B. Jorgensen, B. van Deurs, Physiological and morphological effects of overproduction of membrane-bound ATP synthase in *Escherichia coli* K-12, Embo J 3 (1984) 1791–1797.
[8] H.M. Eriksson, P. Wessman, C. Ge, K. Edwards, A. Wieslander, Massive formation of intracellular membrane vesicles in *Escherichia coli* by a monotopic membrane-bound lipid glycosyltransferase, J. Biol. Chem. 284 (2009) 33904–33914.
[9] I. Arechaga, Membrane invaginations in bacteria and mitochondria: common features and evolutionary scenarios, J. Mol. Microbiol. Biotechnol. 23 (2013) 13–23.
[10] E. Mileykovskaya, M. Zhang, W. Dowhan, Cardiolipin in energy transducing membranes, Biochemistry (Mosc) 70 (2005) 154–158.
[11] W. Dowhan, Molecular basis for membrane phospholipid diversity: why are there so many lipids? Annu. Rev. Biochem. 66 (1997) 199–232.
[12] F.L. Hoch, Cardiolipins and biomembrane function, Biochim. Biophys. Acta 1113 (1992) 71–133.
[13] T.H. Haines, N.A. Dencher, Cardiolipin: a proton trap for oxidative phosphorylation, FEBS Lett. 528 (2002) 35–39.
[14] J. Lecocq, C.E. Ballou, On the structure of cardiolipin, Biochemistry 3 (1964) 976–980.
[15] R.N. Lewis, R.N. McElhaney, The physicochemical properties of cardiolipin bilayers and cardiolipin-containing lipid membranes, Biochim. Biophys. Acta 1788 (2009) 2069–2079.
[16] M. Schlame, M. Ren, Y. Xu, M.L. Greenberg, I. Haller, Molecular symmetry in mitochondrial cardiolipins, Chem Phys Lipids 138 (2005) 38–49.
[17] E. Mileykovskaya, W. Dowhan, Cardiolipin membrane domains in prokaryotes and eukaryotes, Biochim. Biophys. Acta 1788 (2009) 2084–2091.
[18] A. Ortiz, J.A. Killian, A.J. Verkleij, J. Wilschut, Membrane fusion and the lamellar-to-inverted-hexagonal phase transition in cardiolipin vesicle systems induced by divalent cations, Biophys. J. 77 (1999) 2003–2014.
[19] R. Pan, A.D. Jones, J. Hu, Cardiolipin-mediated mitochondrial dynamics and stress response in Arabidopsis, Plant Cell 26 (2014) 391–409.
[20] T. Doan, J. Coleman, K.A. Marquis, A.J. Meeske, B.M. Burton, E. Karatekin, D.Z. Rudner, FisB mediates membrane fission during sporulation in *Bacillus subtilis*, Genes Dev. 27 (2013) 322–334.
[21] L.V. Chernomordik, M.M. Kozlov, Mechanics of membrane fusion, Nat. Struct. Mol. Biol. 15 (2008) 675–683.
[22] R.F. Epand, M. Tokarska-Schlattner, U. Schlattner, T. Wallimann, R.M. Epand, Cardiolipin clusters and membrane domain formation induced by mitochondrial proteins, J. Mol. Biol. 365 (2007) 968–980.
[23] R.P. Rand, S. Sengupta, Cardiolipin forms hexagonal structures with divalent cations, Biochim. Biophys. Acta 255 (1972) 484–492.
[24] B. De Kruijff, A.J. Verkleij, J. Leunissen-Bijvelt, C.J. Van Echteld, J. Hille, H. Rijnbout, Further aspects of the Ca2 +-dependent polymorphism of bovine heart cardiolipin, Biochim. Biophys. Acta 693 (1982) 1–12.
[25] P.R. Cullis, A.J. Verkleij, P.H. Ververgaert, Polymorphic phase behaviour of cardiolipin as detected by 31P NMR and freeze-fracture techniques. Effects of calcium, dibucaine and chlorpromazine, Biochim. Biophys. Acta 513 (1978) 11–20.
[26] J.M. Seddon, R.D. Kaye, D. Marsh, Induction of the lamellar-inverted hexagonal phase-transition in cardiolipin by protons and mono-valent cations, Biochim. Biophys. Acta 734 (1983) 347–352.
[27] M. Schlame, Cardiolipin synthesis for the assembly of bacterial and mitochondrial membranes, J. Lipid Res. 49 (2008) 1607–1620.
[28] I. Shibuya, Metabolic regulations and biological functions of phospholipids in *Escherichia coli*, Prog. Lipid Res. 31 (1992) 245–299.
[29] S. Hiraoka, H. Matsuzaki, I. Shibuya, Active increase in cardiolipin synthesis in the stationary growth phase and its physiological significance in *Escherichia coli*, FEBS Lett. 336 (1993) 221–224.
[30] T. Romantsov, Z. Guan, J.M. Wood, Cardiolipin and the osmotic stress responses of bacteria, Biochim. Biophys. Acta 1788 (2009) 2092–2100.
[31] S. Kikuchi, I. Shibuya, K. Matsumoto, Viability of an *Escherichia coli* pgsA null mutant lacking detectable phosphatidylglycerol and cardiolipin, J. Bacteriol. 182 (2000) 371–376.
[32] S. Nishijima, Y. Asami, N. Uetake, S. Yamagoe, A. Ohta, I. Shibuya, Disruption of the *Escherichia coli* cls gene responsible for cardiolipin synthesis, J. Bacteriol. 170 (1988) 775–780.
[33] B.E. Tropp, Cardiolipin synthase from *Escherichia coli*, Biochim. Biophys. Acta 1348 (1997) 192–200.
[34] B.K. Tan, M. Bogdanov, J. Zhao, W. Dowhan, C.R. Raetz, Z. Guan, Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 16504–16509.
[35] G. Pluschke, Y. Hirota, P. Overath, Function of phospholipids in *Escherichia coli*. Characterization of a mutant deficient in cardiolipin synthesis, J. Biol. Chem. 253 (1978) 5048–5055.
[36] A. Ohta, T. Obara, Y. Asami, I. Shibuya, Molecular cloning of the cls gene responsible for cardiolipin synthesis in *Escherichia coli* and phenotypic consequences of its amplification, J Bacteriol 163 (1985) 506–514.
[37] C.B. Hirschberg, E.P. Kennedy, Mechanism of the enzymatic synthesis of cardiolipin in *Escherichia coli*, Proc. Natl. Acad. Sci. U. S. A. 69 (1972) 648–651.
[38] E. Tunaitis, J.E. Cronan Jr., Characterization of the cardiolipin synthetase activity of *Escherichia coli* envelopes, Arch. Biochem. Biophys. 155 (1973) 420–427.
[39] S. Hiraoka, K. Nukui, N. Uetake, A. Ohta, I. Shibuya, Amplification and substantial purification of cardiolipin synthase of *Escherichia coli*, J. Biochem. 110 (1991) 443–449.

[40] D. Guo, B.E. Tropp, A second *Escherichia coli* protein with CL synthase activity, Biochim. Biophys. Acta 1483 (2000) 263–274.
[41] T. Baba, T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K.A. Datsenko, M. Tomita, B.L. Wanner, H. Mori, Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection, Mol. Syst. Biol. 2 (2006) (2006) 0008.
[42] D. Yu, H.M. Ellis, E.C. Lee, N.A. Jenkins, N.G. Copeland, D.L. Court, An efficient recombination system for chromosome engineering in *Escherichia coli*, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 5978–5983.
[43] J.W. Young, J.C. Locke, A. Altinok, N. Rosenfeld, T. Bacarian, P.S. Swain, E. Mjolsness, M.B. Elowitz, Measuring single-cell gene expression dynamics in bacteria using fluorescence time-lapse microscopy, Nat. Protoc. 7 (2012) 80–88.
[44] M. Moulin, A. Solgadi, V. Veksler, A. Garnier, R. Ventura-Clapier, P. Chaminade, Sex-specific cardiac cardiolipin remodelling after doxorubicin treatment, Biol. Sex Differ. 6 (2015) 20.
[45] P.M. Oliver, J.A. Crooks, M. Leidl, E.J. Yoon, A. Saghatelian, D.B. Weibel, Localization of anionic phospholipids in *Escherichia coli* cells, J. Bacteriol. 196 (2014) 3386–3398.
[46] J.M. Petit, A. Maftah, M.H. Ratinaud, R. Julien, 10N-nonyl acridine orange interacts with cardiolipin and allows the quantification of this phospholipid in isolated mitochondria, Eur. J. Biochem. 209 (1992) 267–273.
[47] P.F. Gallet, A. Maftah, J.M. Petit, M. Denis-Gay, R. Julien, Direct cardiolipin assay in yeast using the red fluorescence emission of 10-N-nonyl acridine orange, Eur. J. Biochem. 228 (1995) 113–119.
[48] M.I. Garcia Fernandez, D. Ceccarelli, U. Muscatello, Use of the fluorescent dye 10-N-nonyl acridine orange in quantitative and location assays of cardiolipin: a study on different experimental models, Anal. Biochem. 328 (2004) 174–180.
[49] P. Kaewsuya, N.D. Danielson, D. Ekhterae, Fluorescent determination of cardiolipin using 10-N-nonyl acridine orange, Anal. Bioanal. Chem. 387 (2007) 2775–2782.
[50] E. Mileykovskaya, W. Dowhan, Visualization of phospholipid domains in *Escherichia coli* by using the cardiolipin-specific fluorescent dye 10-N-nonyl acridine orange, J. Bacteriol. 182 (2000) 1172–1175.
[51] T.A. Garrett, A.C. O'Neill, M.L. Hopson, Quantification of cardiolipin molecular species in *Escherichia coli* lipid extracts using liquid chromatography/electrospray ionization mass spectrometry, Rapid Commun. Mass Spectrom. 26 (2012) 2267–2274.
[52] M. Dahlberg, Polymorphic phase behavior of cardiolipin derivatives studied by coarse-grained molecular dynamics, J. Phys. Chem. B 111 (2007) 7194–7200.
[53] M. Schlame, M. Ren, The role of cardiolipin in the structural organization of mitochondrial membranes, Biochim. Biophys. Acta 1788 (2009) 2080–2083.
[54] M.B. Sankaram, G.L. Powell, D. Marsh, Effect of acyl chain composition on salt-induced lamellar to inverted hexagonal phase transitions in cardiolipin, Biochim. Biophys. Acta 980 (1989) 389–392.
[55] W.J. Vail, J.G. Stollery, Phase changes of cardiolipin vesicles mediated by divalent cations, Biochim. Biophys. Acta 551 (1979) 74–84.
[56] J.T. McGarrity, J.B. Armstrong, The effect of temperature and other growth conditions on the fatty acid composition of *Escherichia coli*, Can. J. Microbiol. 27 (1981) 835–840.
[57] L.D. Renner, D.B. Weibel, Cardiolipin microdomains localize to negatively curved regions of *Escherichia coli* membranes, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 6264–6269.
[58] K.C. Huang, K.S. Ramamurthi, Macromolecules that prefer their membranes curvy, Mol. Microbiol. 76 (2010) 822–832.
[59] S. Nichols-Smith, S.Y. Teh, T.L. Kuhl, Thermodynamic and mechanical properties of model mitochondrial membranes, Biochim. Biophys. Acta 1663 (2004) 82–88.
[60] J.D. Unsay, K. Cosentino, Y. Subburaj, A.J. Garcia-Saez, Cardiolipin effects on membrane structure and dynamics, Langmuir 29 (2013) 15878–15887.
[61] N. Khalifat, J.B. Fournier, M.I. Angelova, N. Puff, Lipid packing variations induced by pH in cardiolipin-containing bilayers: the driving force for the cristae-like shape instability, Biochim. Biophys. Acta 1808 (2011) 2724–2733.
[62] T.N. Zeczycki, J. Whelan, W.T. Hayden, D.A. Brown, S.R. Shaikh, Increasing levels of cardiolipin differentially influence packing of phospholipids found in the mitochondrial inner membrane, Biochem. Biophys. Res. Commun. 450 (2014) 366–371.
[63] B. de Kruijff, P.R. Cullis, Cytochrome c specifically induces non-bilayer structures in cardiolipin-containing model membranes, Biochim. Biophys. Acta 602 (1980) 477–490.
[64] J.M. Seddon, Structure of the inverted hexagonal (HII) phase, and non-lamellar phase transitions of lipids, Biochim. Biophys. Acta 1031 (1990) 1–69.
[65] N. Stepanyants, P.J. Macdonald, C.A. Francy, J.A. Mears, X. Qi, R. Ramachandran, Cardiolipin's propensity for phase transition and its reorganization by dynamin-related protein 1 form a basis for mitochondrial membrane fission, Mol. Biol. Cell 26 (2015) 3104–3116.
[66] P. Paumard, J. Vaillier, B. Coulary, J. Schaeffer, V. Soubannier, D.M. Mueller, D. Brethes, J.P. di Rago, J. Velours, The ATP synthase is involved in generating mitochondrial cristae morphology, Embo J. 21 (2002) 221–230.
[67] V. Soubannier, J. Vaillier, P. Paumard, B. Coulary, J. Schaeffer, J. Velours, In the absence of the first membrane-spanning segment of subunit 4(b), the yeast ATP synthase is functional but does not dimerize or oligomerize, J. Biol. Chem. 277 (2002) 10739–10745.
[68] M. Schlame, K. Beyer, M. Hayer-Hartl, M. Klingenberg, Molecular species of cardiolipin in relation to other mitochondrial phospholipids. Is there an acyl specificity of the interaction between cardiolipin and the ADP/ATP carrier? Eur. J. Biochem. 199 (1991) 459–466.
[69] H. Palsdottir, C. Hunte, Lipids in membrane protein structures, Biochim. Biophys. Acta 1666 (2004) 2–18.
[70] M. Kates, J.Y. Syz, D. Gosser, T.H. Haines, pH-dissociation characteristics of cardiolipin and its 2′-deoxy analogue, Lipids 28 (1993) 877–882.
[71] T.H. Haines, A new look at cardiolipin, Biochim. Biophys. Acta 1788 (2009) 1997–2002.
[72] S. Laage, Y. Tao, A.E. McDermott, Cardiolipin interaction with subunit c of ATP synthase: solid-state NMR characterization, Biochim. Biophys. Acta 1848 (2015) 260–265.
[73] R.P. Rand, N.L. Fuller, S.M. Gruner, V.A. Parsegian, Membrane curvature, lipid segregation, and structural transitions for phospholipids under dual-solvent stress, Biochemistry 29 (1990) 76–87.
[74] D.P. Siegel, Energetics of intermediates in membrane fusion: comparison of stalk and inverted micellar intermediate mechanisms, Biophys. J. 65 (1993) 2124–2140.
[75] V.S. Malinin, B.R. Lentz, On the analysis of elastic deformations in hexagonal phases, Biophys. J. 86 (2004) 3324–3328.
[76] J. Jouhet, Importance of the hexagonal lipid phase in biological membrane organization, Front. Plant Sci. 4 (2013) 494.

A novel regulation mechanism of the T7 RNA polymerase based expression system improves overproduction and folding of membrane proteins

Federica Angius[1], Oana Ilioaia[1], Amira Amrani[1], Annabelle Suisse[1,2], Lindsay Rosset[1], Amélie Legrand[1], Abbas Abou-Hamdan[1,3], Marc Uzan[1], Francesca Zito[1] and Bruno Miroux[1]

[1] Laboratoire de Biologie Physico-Chimique des Protéines Membranaires, CNRS, University Paris Diderot, Sorbonne Paris Cité, Institut de Biologie Physico-Chimique PSL Research University, Paris, France [2] Present address: Helen L. and Martin S. Kimmel Center at the Skirball Institute for Biomolecular Medicine and Department of Cell Biology, NYU School of Medicine, New York, USA [3] Present address: Institut de Biologie Intégrative de la Cellule, CNRS, Gif sur Yvette, France

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Membrane protein (MP) overproduction is one of the major bottlenecks in structural genomics and biotechnology. Despite the emergence of eukaryotic expression systems, bacteria remain a cost effective and powerful tool for protein production. The T7 RNA polymerase based expression system is a successful and efficient expression system, which achieves high-level production of proteins. However it is not adapted to the overproduction of foreign MP, which require a fine-tuning of their expression to minimize the toxicity associated with MP production. Here we report a novel regulation mechanism for the T7 expression system. We have isolated two bacterial hosts, namely C44(DE3) and C45(DE3), harboring a stop codon in the T7 RNA polymerase gene. Translation of the full length T7 RNA polymerase enzyme is under the control of the basal nonsense suppressive activity of the BL21(DE3) host. Evaluation of hosts with superfolder green fluorescent protein (sfGFP) revealed an unprecedented tighter control of transgene expression. Analysis of a collection of twenty MP fused to GFP by flow cytometry and Fluorescent detection coupled to size exclusion chromatography showed an improved production yield and quality of several bacterial membrane proteins. A human monotopic MP, that was produced as inclusion bodies in all existing hosts, is now produced in an active form at the *E. coli* membrane of C45(DE3) host, indicating that the newly identified hosts have the potential to break the barrier of producing functional eukaryotic MP in bacteria.

Membrane protein production | E. coli | T7 RNA polymerase

-

Introduction

Membrane proteins (MP) are essential for several biological processes such as cell signaling, metabolite transport, energy conversion, and ion homeostasis. Approximately 30% of genes encode MP, and a growing number of human diseases are found to be due to their dysfunctions; in fact, MP represent about 50% of drug discovery targets(1, 2). Despite their importance, structural and biochemical investigations of MP are limited by the difficulty to produce and purify them in a functional state. Despite an increasing use of eukaryotic expression systems, about half of the unique MP structures (both eukaryotic and prokaryotic) present in the Protein Data Bank (PDB) have been obtained after production in *E. coli,* showing that it remains a powerful vehicle in structural biology of MP. We have previously showed the T7 and arabinose promoters, have contributed to 80% of all unique MP structures (63% and 17%, respectively)(3). The T7 RNA polymerase (T7RNAP)-based expression system was designed to maximize the production levels of the target protein(4). Bacteriophage lambda was modified by inserting the T7 RNA polymerase gene under the control of the *lac L8 UV5* promoter, a variant of the lactose promoter that is insensitive to catabolic repression. The resulting DE3 phage was then inserted into a BL21 host, a protease deficient *E. coli* B strain. Upon addition of IPTG, the T7RNAP enzyme was produced at high levels and subsequently transcribed the target gene which was cloned downstream of a T7 promoter in the expression plasmid. This expression system is powerful for at least three main reasons: one, the T7RNAP enzyme is ten times faster than the *E. coli* RNA polymerase; two, the enzyme is highly specific for the T7 promoter; and three, the high copy number of the expression plasmid does not titrate the *E. coli* RNA polymerase. Nevertheless, the T7 based expression system also has at least two major drawbacks. First, the *lacUV5* promoter is leaky and difficult to control with its non-metabolized inducer IPTG in a proportional manner. Moreover the T7 promoter's basal transcriptional activity is repressed only with overproduction of the T7RNAP inhibitor, T7 lysozyme(5), in a companion plasmid. But despite this ability to inhibit the T7RNAP, some expression plasmids, especially those with high copy numbers(6), are difficult to maintain in rich media even in the absence of inducer(7). Second, Dong and coworkers have shown that overproduction of the T7RNAP enzyme triggers ribosome destruction and cell death (8). This intrinsic toxicity of the T7 system was used by Miroux and Walker in a genetic screen to isolate two mutant hosts, namely C41(DE3) and C43(DE3) that were adapted for the production of soluble and membrane proteins(9). In a BL21(DE3) host that is harboring a high copy number plasmid, the mRNA level of the target gene can reach ribosomal RNA levels, a level that is beyond the cell's translational abilities. Under induction conditions, the level of target mRNA is 10-fold lower in C41(DE3) than in the parental strain, leading to a slower accumulation of the target protein. Accumulation of the target protein is further delayed in C43(DE3). The net results were: an improved fitness of the cells, an increased production of the target MP on a long-term induction experiment, and in some cases, metabolic adaptation to stimulate internal membrane proliferation(10). At the molecular level, Wagner and colleagues showed that in C41(DE3) and C43(DE3), the sequence of the *lacUV5* promoter upstream of the T7RNAP gene in the genome of lambda DE3, reversed to wild-type. However, the L8 mutation in the CRP-binding site was maintained in both strains. This mutation considerably weakens the lac promoter as it cannot be activated by the CRP-cyclic AMP complex(11). An additional mutation in C43(DE3) was identified in the *lacI* gene which resulted in an amino acid change from valine to phenylalanine at residue 192(12). The *lacI* gene codes for the *lac* operon repressor, an allosteric protein which binds allolactose or IPTG. This mutation decreased the binding of inducer to the repressor and conferred a super repressor I^s phenotype(13), which further inhibits the transcription initiation of the T7RNAP gene. Given that both mutant hosts have had a major impact on the field of MP structural biology (19% of non *E. coli* membrane protein structures were produced in these T7 expression hosts - 50% of

which were complex α-helical MP structures(14)), we aimed at isolating more efficiently regulated hosts with a lower basal level of target gene expression, without compromising the promoter strength. To achieve this goal, we used the green fluorescent protein (GFP) as reporter gene and we isolated two mutant hosts, namely C44(DE3) and C45(DE3). By identifying sequence changes in the C44(DE3) and C45(DE3) genomes, we described a radically new T7RNAP regulation mechanism which is based on the nonsense suppressive activity of the BL21(DE3) host. Both C44(DE3) and C45(DE3) mutant hosts extend the promoter strength coverage of the T7RNAP expression system in a unique way. Complementary to the previous expression strains, C41(DE3) and C43(DE3), they open a new avenue for increasing the space coverage of large scale MP production.

Results and Discussion

Isolation and characterization of BL21(DE3) derivatives. The bacterial mutants C44(DE3) and C45(DE3) were isolated from two independent experiments as described in the materials and methods section. The relative fluorescent intensity (RFI) of sfGFP(15) was measured by flow cytometry to assess the regulation of the T7 system in all BL21(DE3) derivatives. Three hours after IPTG addition, sfGFP RFI increased 4 and 7-fold in C44(DE3) and C45(DE3), respectively (Figure 1A). A subsequent 22 and 18-fold increase of RFI in C44(DE3) and C45(DE3) respectively were observed between 3h and 22h after induction. This suggests that the recombinant protein (which in the case of superfolder sfGFP is well reflected by the measured RFI(15)) continues to accumulate in the stationary phase. Both mutant hosts were compared to C41(DE3) and C43(DE3). Figure 1B shows that the rate of sfGFP production is the highest in C41(DE3) and decreased in C43(DE3), as previously described (9). In C44(DE3) and C45(DE3) hosts, the rate of sfGFP production is strongly decreased. In contrast to BL21(DE3), which is difficult to titrate with IPTG (Figure S1), both C44(DE3) and C45(DE3) could be titrated over an unprecedented range of IPTG concentration starting from 10μM and up to 0.7 mM (Figure 1C). Analysis of cell populations showed that in BL21(DE3), the highest RFI was obtained without adding IPTG (Figure 2A) as previously reported(16). Addition of IPTG to the culture was toxic, as reflected by the appearance of a large number of cells (80% of the bacterial

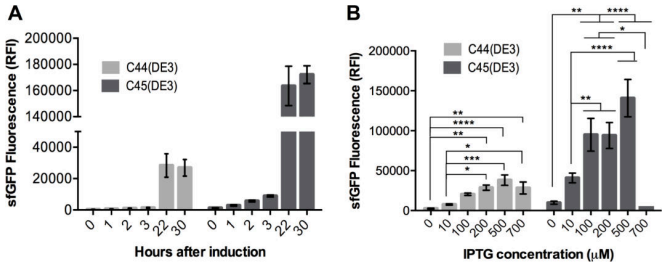


Fig. 1. Regulation of the T7 expression system in C44(DE3) and C45(DE3) bacterial hosts. Cells were analyzed by flow cytometry for sfGFP RFI. **(A)** Time courseexperiment after addition of 0.7mM IPTG **(B)** IPTG titration experiment. Cells were induced with increasing concentration of IPTG and analyzed after overnight induction. Each plot shows a mean value of three independent experiments. Standard error of the mean is indicated. Statistical significancewas analyzed using One-way ANOVA test. F values are: F(5,30)=9.907 for C44(DE3) and F(5,30)=19.06 for C45(DE3).

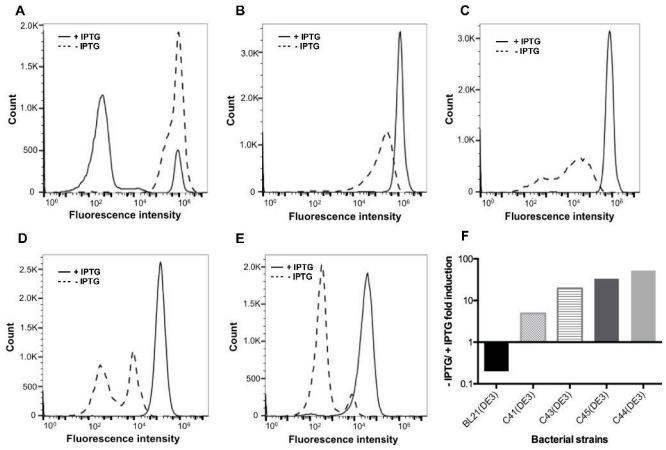


Fig. 2. Assessment of the T7 expression hosts using superfolder GFP. Flow cytometry analysis of an overnight culture of BL21(DE3) **(A)** C41(DE3) **(B)**C43(DE3) **(C)** C44(DE3) **(D)** and C45(DE3) **(E)** transformed with pHis17-sfGFPvector. Dashed line indicates RFI of non-induced cells while continuous lineshows the RFI of cells induced at OD600=0.4 with 0.7 mM IPTG. **(F)** RFI meanvalue ratio of induced over non-induced cells.

Significance

Understanding the function of membrane proteins (MP) at molecular level requires to produce them in large amounts. *Escherichia coli* remains a popular vehicle for the overexpression of proteins. It is however not adapted for the production of membrane proteins which often precipitate in the bacterial cytoplasm. Here we present a new mode of regulation of the T7 expression system in which the T7 RNA polymerase is controlled at both transcriptional and translational levels due to the insertion of a stop codon in the T7 RNA polymerase gene. In the newly isolated hosts, the foreign MP folds and accumulates in an unprecedented way, thus breaking the barrier of producing eukaryotic MP in large amounts for biochemical and structural studies.

Reserved for Publication Footnotes

population) that have lost sfGFP fluorescence (Table S1). In contrast, the basal level of sfGFP fluorescence was decreased 3.5-fold in C41(DE3) and addition of IPTG resulted in a homogeneous population of cells (93%) exhibiting the highest level of sfGFP RFI (Figure 2B and Table S1). In C43(DE3), two populations were observed in the absence of IPTG with a global sfGFP mean RFI of 36,615 +/- 346 (Table S1). Addition of IPTG resulted in a single fluorescent population showing that all cells had kept the ability to produce sfGFP (Figure 2C). The most remarkable feature of both C44(DE3) and C45(DE3) mutant hosts was their extremely low levels of sfGFP RFI after overnight culture in the absence of IPTG (Figure 2D-E). For C44(DE3) mutant, 89% of the cell population exhibited a sfGFP RFI of 359, which is only about twice the natural fluorescence of bacterial cells (Table S1). Consequently, the induction fold change of sfGFP RFI estimated by flow cytometry increased up to 48 times in C44(DE3) and 33 times in C45(DE3) after 22 hours of induction with IPTG (Figure 2F). To summarize, the two bacterial expression hosts isolated in this study contain three important features that were not present in any of the existing T7 expression bacterial hosts: one, tight repression of the target gene in rich medium at 37°C; two, a tunable level of target protein expression with increasing IPTG concentration; three, a slow and continuous target protein accumulation in both the exponential and stationary phase at

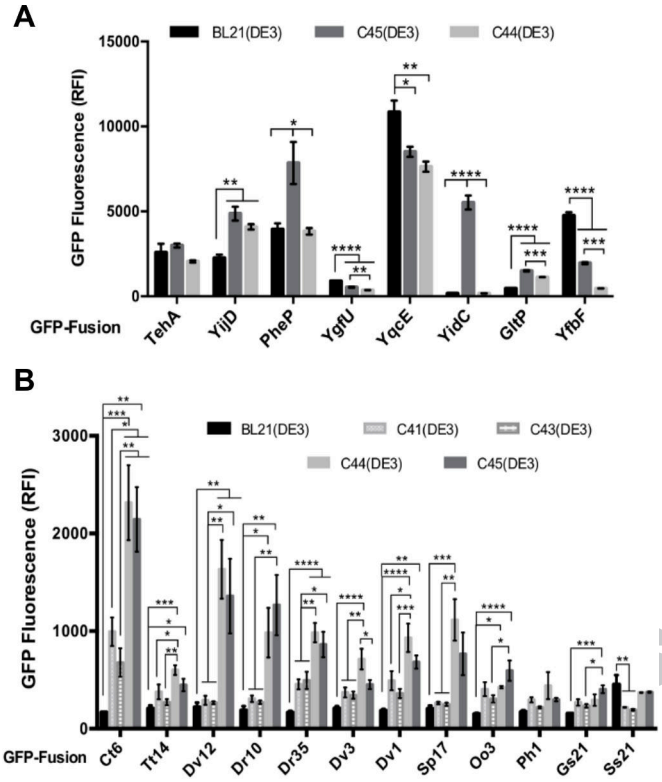


Fig. 3. Flow cytometry analysis of MP-GFP production levels. BL21(DE3) cells were grown overnight in LB medium at 30°C. All other BL21(DE3) derivatives were grown in in 2*YT medium, induced at OD600 =0.4 with 0.7 mM IPTG and further grown overnight at 37°C. GFP RFI was recorded asdescribed in materials and methods. **(A)** Triplicate cultures producing *E.coli*MP-GFP fusions(16). **(B)** Six-time replica cultures producing non *E. coli* MPGFP(19). All F values listed in supplementary Table S8.

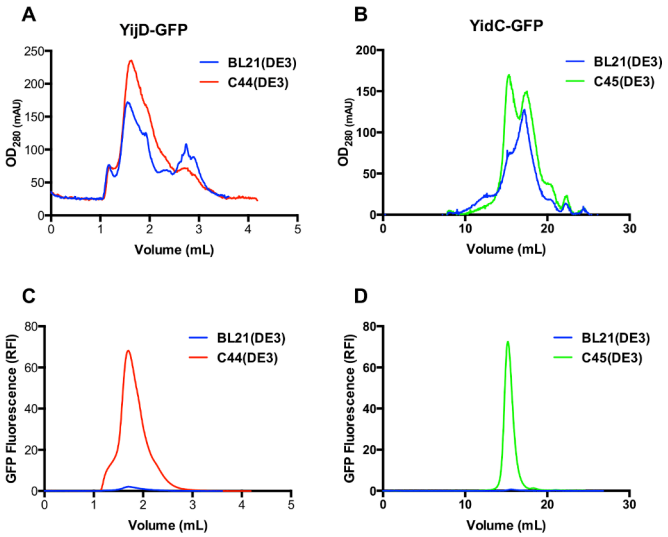


Fig. 4. Quantification of MP-GFP production levels by Fluorescence detection size-exclusion chromatography. BL21(DE3), C44(DE3) and C45(DE3)host strains were transformed either with YijD-GFP **(A, C)** or YidC-GFP **(B, D)** expression vectors. Cell pellets from 200 ml culture were solubilizedwith DDM (1% final concentration) and 100 µl of solubilized material wasinjected on a Superdex-200 5/150GL column for GFP-YijD or on Superose-6 10/300GL column for GFP-YidC. **(C, F).** Absorbance at 280nm **(A-B)** was recorded while GFP fluorescence **(C, D)** was analyzed using a Jasco detector.

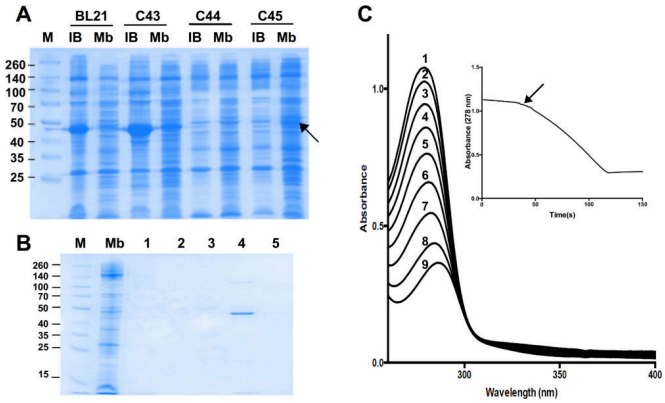


Fig. 5. Functional expression of recombinant human SQR in C45(DE3). **(A)** SDS-PAGE analysis of hSQR inclusion bodies formation in BL12(DE3) andBL21(DE3) derivatives. 30 µg of inclusion bodies (IB) or membrane fraction(Mb) proteins were analyzed by SDS-PAGE stained with Coomassie Blue solution. Lane M: molecular weight marker. **(B)** SDS-PAGE analysis of hSQR purification. Membrane fraction (Mb), flow through (1) column wash with 20 mM (2) and 40 mM (3) Imidazol, column elution with 200 mM (4) and 500 mM (5) Imidazol. **(C)** Spectral course of SQR catalytic assays. The reaction mixture contained 100 µM CoQ1 and 400 µM of Na2 S . The activity of 0.2 µg SQR was recorded during 600 s at 287 nm. Curves 1 to 9 were recorded every 10 s after hSQR addition. The inset shows the time course of absorbance changes at 287 nm, black arrow show hSQR addition.

37°C. This is achieved without titration of lysozyme to inhibit the T7RNAP as in the Lemo21(DE3)(17) strain, and without downregulating the T7RNAP gene expression as in C41(DE3) or C43(DE3). The newly isolated hosts offer an unprecedented control of the target gene production levels and together with C41(DE3) and C43(DE3) hosts they allow to produce the target gene either at *E. coli* physiological levels (Table S1, GFP RFI in induced C44(DE3) or up to 30 times (in C41(DE3)), above the maximum production levels of sfGFP in C44(DE3). This makes the T7 expression system attractive not only in structural biology but also in synthetic biology, either for the expression of complex metabolic pathways or for production of metabolites such as fatty acids(18)

Validation of C44(DE3) and C45(DE3) for MP production.

Next, C44(DE3) and C45(DE3) mutant hosts were challenged with two collections of expression plasmids fused to GFP and encoding MP from *E. coli* and other bacterial species (MP-GFP, see Table S2). First, the toxicity associated with the overproduction of the MP-GFP fusion was assessed on solid medium. For each construct, cells were diluted 10^7-fold and were plated on 2xYT agar medium with or without IPTG. The surface area (SA) of colonies was measured from photographic pictures using Image J software. The relative colony surface area (RCSA) upon induction was calculated (SA(+IPTG)/SA(-IPTG)*100) and is presented in Table S3. In the BL21(DE3) host, all expression vectors, including the soluble protein sfGFP, prevented colony formation on 2xYT plates supplemented with 0.7 mM IPTG (RCSA=0, Table S3). In contrast, almost none of the MP-GFP fusions prevented colony formation on IPTG containing plates in either C44(DE3) or C45(DE3), with the exception of Ph1-EGFP and GS21-EGFP (Table S3). The RCSA values were generally lower in C45(DE3) for most of the MP-GFP fusions as well as for the pHis-sfGFP control plasmid. Next, GFP fluorescence was measured by flow cytometry in liquid cultures. As shown in Figure 3A, among eight *E. coli* MP-GFP expression plasmids tested in this study, four target proteins, namely YijD, PheP, YidC and GltP were produced at significantly higher levels in either C44(DE3) or C45(DE3) bacterial hosts, while three constructs (YgfU, YfbF, and YqcE) remained better produced in the BL21(DE3) host. To expand our results, twelve non *E.*

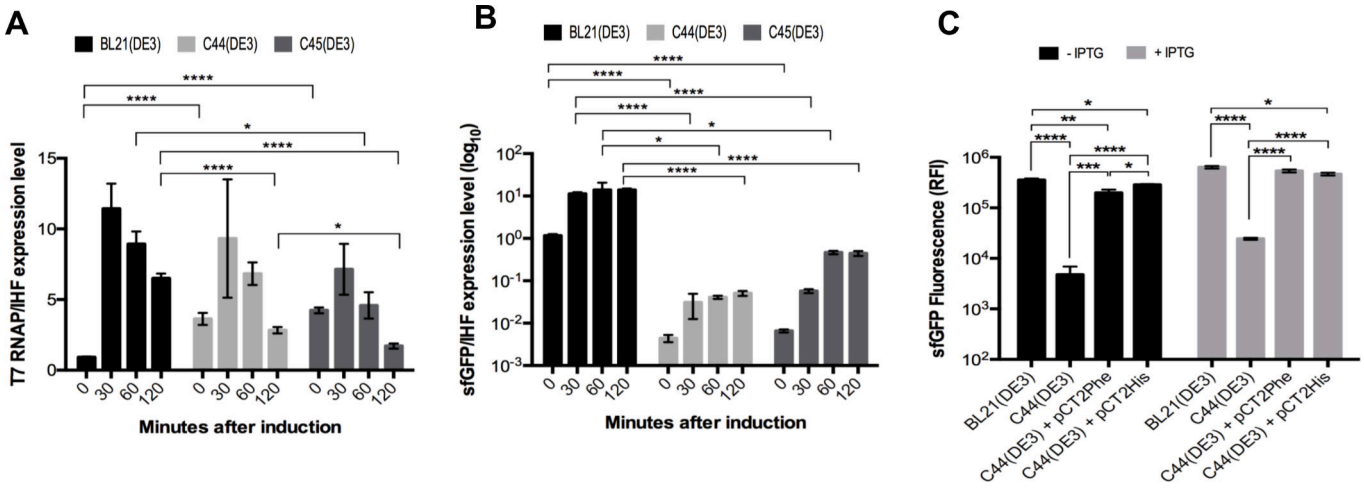


Fig. 6. Translational regulation of the T7RNAP gene in C44(DE3) and C45(DE3) hosts. RT-qPCR was performed as described in materials and methods. IHF was chosen as reference gene. **(A)** T7RNAP/IHF expression levels **(B)** sfGFP/IHF expression levels. **(C)** Cells were transformed either with sfGFP as control or with sfGFP and pCT2Phe or pCT2His plasmid. The graph represents sfGFP RFI analyzed by flow cytometry after overnight cultures with (gray columns) or without addition of IPTG (black columns). Three independent experiments were performed with technical triplicate. Standard error of the mean is indicated on each plot. Statistical significance was analyzed using One-way ANOVA test.

coli MP fused to GFP(19), were produced in BL21(DE3) and BL21(DE3) derivatives. Almost all constructs were produced at significantly higher levels in either C44(DE3) or C45(DE3) bacterial hosts (Figure 3B). Whole cell GFP fluorescence analysis can be misleading due to GFP cleavage from the MP and does not reflect the exact fraction of MP that can be extracted from bacterial(20) or yeast(21) membranes. In order to assess more precisely the quality and quantity of MP-GFP fusions produced, we used fluorescence-detection size-exclusion chromatography (F-sec) method(22). Two *E.coli* proteins, namely YijD and YidC were produced in 200mL volume cultures; membrane preparations were solubilized in presence of DDM and subjected to F-sec analysis as described in materials and methods. Figure 4 shows that GFPe-YijD RFI increased 34-fold when extracted from C44(DE3) membranes (Figure 4C), while GFPd-YidC RFI increased 105-fold in C45(DE3) solubilized membranes (Figure 4D). To solubilize MP from both bacterial hosts with the same detergent-protein ratio, we repeated the experiment with a fixed amount of proteins (18 mg with 1% DDM in 9 mL final volume). The F-sec analysis showed that the fusion protein level increased 20-fold in C45(DE3) bacterial membranes (data not shown). Since the total amount of MP obtained from a 1 L culture of C45(DE3) cells quadrupled (126 mg in C45(DE3) versus 32 mg in BL21(DE3)), the yield per liter of culture of YidC-GFP increased 70-fold, a fold change close to the one measured per volume in Figure 4E. In order to test if C44(DE3) and C45(DE3) mutant hosts could expand the sequence coverage of MP to eukaryotic proteins, we tested the human sulfide quinone oxydoreductase (SQR). SQR is a monotopic MP that has been previously produced in *E. coli* at moderate levels (250 µg/L of culture of cells grown OVN at 15°C) only in presence of the companion plasmid pCPN10/60, which produces cold-adapted chaperon proteins from *Oleispira antarctica*(23). The same synthetic gene used in(23) encoding for the SQR was introduced in pHis17 and transformed either in BL21(DE3) or in the four other mutant hosts. The recombinant protein was produced as inclusion bodies in BL21(DE3), C41(DE3) and C43(DE3) but was readily targeted to the membrane of C45(DE3) (Figure 5A). To confirm that the protein was correctly folded in C45(DE3) bacterial membranes, SQR was purified from a 1 L culture (Figure 5B, 600 µg of purified protein) and assayed for activity using Na_2S as a substrate. The resulting activity curves (Figure 5C) were similar to those previously published(23), and confirmed that the protein is correctly folded, binds its cofactor FAD, and is able to oxidize Na_2S (Km=11 mM). Membrane protein production often induces a stress to the host cell. In some cases it has been shown that the target membrane protein is by itself toxic to the cell(24). In most cases, however, the T7 expression system overloads the translocation and membrane insertion machineries very early after induction, causing an excess of the unfolded target protein that aggregates rapidly as inclusion bodies(25). Heat-shock protein, unfolded stress response, and proteases often appear as a desperate attempt to circumvent this problem (for review see(26)). Wagner and colleagues have nicely illustrated this view by conducting a global proteomic study with YidC as a model protein partially produced as inclusion bodies(27). They show that overproduction of the YidC membrane protein critically affects BL21(DE3)pLysS cell growth and increases chaperone (DNAJ/K, GROEL/S) and protease (HslU/V, ClpXP, ClpB) levels. Interestingly, the same authors found that inclusion bodies formation was greatly decreased in C41(DE3) or C43(DE3) mutant hosts, and that toxicity in the form of chaperone and protease induction was strongly reduced(17). In C45(DE3) host, toxicity linked to YidC production was low an comparable to the one observed with the soluble protein sfGFP (Table S3) and solubilized YidC levels increased 100 times in C45(DE3) membranes as compared to the BL21(DE3) host (Figure 4). Similarly, we have shown that the human SQR is readily targeted to the bacterial membrane in C45(DE3) without the need of chaperones or growth conditions adaptations. Both examples revealed that fine-tuning of the regulation of target gene expression not only suppresses the toxicity associated with MP production and consequently heat shock and protease responses, but also allows foreign proteins, including eukaryotic proteins, to fold and accumulate in bacterial membranes.

Analysis of the T7 gene regulation in C44(DE3) and C45(DE3). In order to identify the mechanisms that govern the induction of the T7RNAP by IPTG in the isolated mutant hosts, the genomes of BL21(DE3) and its derivatives were entirely sequenced as described in materials and methods section. All mutations were confirmed by PCR analysis (Table S4, Figure S2). Our data (Table S5) confirmed the genomes analyses of both C41(DE3) and C43(DE3) strains previously published(12, 28). In C44(DE3) we found two mutations (Table S6). The first mutation was positioned in the *rbsD* gene, which had also been identified by others(12, 28) in C41(DE3) and C43(DE3) genomes. In BL21(DE3), the *rbsD* gene encoding D-ribose pyranase is

interrupted by the insJK5 insertion. According to Schlegel and co-workers(28), the presence of an IS element in *rbsD* prevents efficient growth of BL21(DE3) on ribose as its sole carbon and energy source. Deletion of the insJK5 insertion restores the wild-type *rbsD* gene in C41(DE3), C43(DE3), C44(DE3), and C45(DE3) derivatives, enabling efficient growth on ribose. The second mutation consists of an *amber* stop codon at the position Glu_{656} of the T7RNAP. Interestingly, we found that C45(DE3) also carries an *ochre* stop codon at position Gln_{359} of the T7RNAP, which highlights a common genetic trait between both C44(DE3) and C45(DE3) mutants (Table S6). However, the pattern of genetic changes in C45(DE3) bacterial mutant is more complex because of three additional deletions in *gltL, B-lom* and *ycdX loci.* The T7RNAP enzyme contains three C-terminal domains (A, B and C) that are critical for the catalytic activity of nucleotide polymerases(29). In C45(DE3) host, the truncated T7RNAP lacks all three domains and is therefore most likely inactive. In C44(DE3) host, the T7RNAP enzyme is truncated just after the second domain and the 228 amino-acid deletion removes the C-terminal domain. Although C-terminal deletions of this enzyme have not been systematically studied, the C motif is conserved in all nucleotide polymerases and the point mutation $Asp_{812}Asn$ in the C-terminal domain completely abolishes its activity(30). To demonstrate our hypothesis T7RNAP and sfGFP mRNA levels were measured by RT-qPCR experiments. Figure 6A shows only moderate reductions in T7RNAP mRNA levels (which cannot account for T7RNAP protein levels falling below its antibody detection levels (see and Figure S3)) in C44(DE3) and C45(DE3) as compared to BL21(DE3) host. In contrast, the target sfGFP mRNA levels, controlled by the T7RNAP activity, were dramatically decreased in both mutant hosts (Figure 6B) as compared to BL21(DE3) host for two reasons. First, the basal levels of sfGFP mRNA, before addition of IPTG, showed a 200–fold decrease in C45(DE3) and a 300-fold decrease in C44(DE3), when compared to BL21(DE3). Second, the maximal levels of sfGFP mRNA after IPTG addition showed a 270-fold and 30-fold decrease in C44(DE3) and C45(DE3), respectively. Next, we co-expressed *amber* suppressors in C44(DE3) and tested whether sfGFP expression phenotypes could be reversed to that of the parental strain, BL21(DE3). We used the pCT2Phe and pCT2His amber suppressor vectors(31), which mediate the insertion of a phenylalanine or a histidine at the amber stop codon, respectively. C44(DE3) cells transformed with both pHis-sfGFP and pCT2Phe or pCT2His plasmids were unable to form colonies on IPTG containing plates (data not shown). Flow cytometry analysis showed that in C44(DE3) cells expressing *amber* suppressor, sfGFP RFI reached the level of the BL21(DE3) parental host in both overnight induced and non-induced cultures (Figure 6C). Thus, we were able to fully demonstrate the mechanism of regulation of the T7 system in C44(DE3). The production of the target protein depends on the translation of the full-length T7RNAP enzyme, which is possible only when ribosomes read through the stop codon mutations introduced in the T7RNAP gene. In fact, a very weak suppressive stop codon activity has been incidentally observed in BL21(DE3)(32), likely accounting for the synthesis of minute amounts of the T7RNAP enzyme all over the bacterial cell growth.

Several groups have developed genetic screens and synthetic biology approaches to explain protein production in more detail and ultimately to circumvent the associated toxicity (2, 33, 34). They identified stress response genes (*ybaB, dnaK, dnaJ, djlA, rraA*) as toxicity suppressors upon MP overproduction. In contrast, the study of Baumgarten and colleagues, who used YidC-GFP fusion to isolate a mutant host in which a mutation in the T7RNAP strongly decreases its binding to the T7 promoter(35), and our results show that fine tuning of target MP transcription is also a powerful and complementary approach to suppress toxicity and to achieve high cell density and high levels of MP production. Twenty years ago, we anticipated that C41(DE3) and C43(DE3) mutant hosts would impact the field of MP structural biology and our recent overview of the PDB showed that they significantly contributed to the number of MP structures(14). The data presented here strongly suggest that C44(DE3) and C45(DE3) bacterial hosts will further impact the field, and together with direct evolution of the target gene(36) and codon usage adaptation(37–40), will most likely extend the use of *E. coli* for eukaryotic MP production.

Materials and Methods

Bacterial strains and plasmids. Escherichia coli strains C41(DE3) and C43(DE3) have been previously described(9, 41). All plasmids are listed in Table S3. The pHis17 and pMW7-GFP-Xa plasmids were derived from pMW7(42). The pHis17 plasmid was modified by addition of three 6*HIS-Tag immediately in frame after the BamHI, HindIII and EcoRI cloning restriction sites. The pMW7-GFP-Xa vector, contains GFP with a factor Xa cleavage sequence in frame of the C-terminal end of GFP. The gene encoding the sfGFP protein was amplified from the vector pDHL1029(43) using the forward primer 5'CAGGGATCCAGTAAAGGTGAAGAACTG3' and the reverse primer 5'CAGAAGCTTTTATTTGTAGAGTTCATC 3' and was subcloned into pHis17. A synthetic optimized version of the gene encoding human sulfide-quinone oxydoreductase (SQR)(23) was obtained from Eurofins genomics, and subcloned between the BamHI and EcoRI sites of the pHis17 plasmid. Amber suppressor plasmids pCT2Phe and pCT2His (cam, Su(UAG)Phe or His), constructed in pACYC184, are described by Lenfant *et al.*(31).

Mutants selection. The *E.coli* strains C44(DE3) and C45(DE3) were isolated in two independent selection experiments according to using pMW7-GFP-Xa as gene reporter(3, 7). GFP-expressing colonies were revealed under far UV light (300 nm UV lamp to avoid additional mutations) and subcloned onto 2xYT ampicillin plates. The next day, only colonies not exhibiting visible green fluorescence were further analyzed. The expression plasmid was removed from the isolated bacterial mutants according to (9).

Genome analysis. Chromosomal DNA was prepared with the Wizard® Genomic DNA Purification Kit (Promega, U.S.A.). Whole genome analysis of C44(DE3) and C45(DE3) was performed at the genomic platform of the Pasteur Institute (Paris) using an Illumina Miseq sequencing. The reads were mapped onto the reference genome sequence of BL21(DE3) (GenBank ID AM946981).

Production of the recombinant proteins. The C41(DE3), C43(DE3), C44(DE3) and C45(DE3) bacterial hosts were grown overnight in 2xYT broth supplemented with ampicillin (100µg/mL) when harboring the pHis17 or pMW7 plasmids or with kanamycin (50µg/mL) for all the other plasmids. Overnight pre-cultures were diluted 1:100 and cultures were grown at 37°C in an Infors HT Minitron at 130 rpm. For all bacterial hosts except for BL21(DE3), protein production was induced by adding 0.7 mM IPTG to the media when the culture had reached an OD_{600} of 0.4. The BL21(DE3) host harboring either pET28a+ or pLIC expression plasmids was cultivated 16H without addition of IPTG according to Zhang and coworkers(16) to optimize the protein production in Luria-Bertani (LB) broth at 30°C. For the human SQR, temperature of growth was lowered to 30°C for the BL21(DE3), C44(DE3), and C45(DE3) strains, and to 25°C for C43(DE3). Cells were harvested by centrifugation (2500 x g for 10 min.) and cell pellet was suspended in 30 mL of 50 mM Tris-HCl buffer, pH 7.4 supplemented with 1x protease inhibitor cocktail (Thermo scientific). The cells were lysed using the Cell Distruptor at a pressure of 2 kBar (Constant Systems LTD). Cell debris were removed by centrifugation at 2,500 *x g* for 10 min. Inclusion bodies were recovered at 10,000 *x g* for 20 min.. Membranes were recovered after centrifugation at 100,000 *x g* of the low speed supernatant for 1h.

Flow cytometry. After overnight induction, the cells were washed with PBS (10 mM Phosphate, 150 mM NaCl, pH 7.4) and diluted to an OD_{600nm} of 0.01. They were analyzed and counted on a C6 Accury flow cytometer with a threshold of 10,000 on forward scatter (FSC) and of 3,000 on side scatter (SSC) to remove electronic noise and small contaminants. Their fluorescence was measured on a FL1 detector (emission detection 533/±30 nm) with forward and side scatter values above 20,000 and 12,000 respectively.

Purification and activity assay of the human SQR. All purification steps were conducted at 4°C. The membrane fractions were diluted at 2 mg/mL proteins in 50 mM Tris-HCl buffer (pH 7.4), 50 mM NaCl, 10 % glycerol, 1x protease inhibitor cocktail and solubilized in presence of 2% DDM for 1h, at 4 °C. Insolubilized material was removed by centrifugation at 100,000 *x g* at 4 °C for 1 hour. The high speed supernatant was collected and diluted in 100 mL final volume of buffer A (100 mM Tris-HCl, pH 8, 150 mM NaCl, 10 % glycerol) with 5 mM Imidazol. The sample was loaded onto a His Pur™ Ni-NTA Chromatography column (Thermo Scientific) equilibrated with buffer A with 5 mM Imidazol and 0.05 % DDM. After 10 volumes column washes with buffer A containing 20mM and 40mM Imidazol and 0.05 % DDM, SQR was eluted with buffer A containing 0.05 % DDM and 200 mM Imidazol. Activity assays were conducted under anaerobic conditions at room temperature using a cuvette with a screw-cap equipped with a Teflon-silicon septum.

The sulfide-quinone oxidoreductase activity was measured by monitoring the reduction of Coenzyme Q_1 at 278 nm according to (23). The CoQ_1 reduction rates were determined after subtraction of the initial slope obtained before addition of hSQR. The CoQ_1 absorption at 278 nm was corrected at 700 nm.

Fluorescence-detection size-exclusion chromatography. Bacterial membranes of 200 mL volume culture were resuspended in 5 mL solubilization buffer (20mM Tris-Hcl, pH 8, 150 mM NaCl and 2% DDM). After one hour solubilization at 4°C, insoluble material was discarded by ultracentrifugation for 30 min at 100,000 *x g*. To avoid different protein/detergent ratios between different strains, 18 mg (2mg/mL) of membranes were solubilized with 1% DDM in 20mM Tris-Hcl, pH 8.0, 150 mM NaCl buffer for 1h at 4°C. Solubilized MP (100 µL) were loaded on a Superose-6 10/300GL column (GE-Healthcare) for YidC-GFP or onto a Superdex-200 5/150GL column (GE-Healthcare) for YijD-GFP. Size-exclusion chromatography was performed in presence of TBS supplemented with 0.03% DDM. Intelligent fluorescence detector (FP-2020, Jasco) was used to measure GFP relative fluorescent intensity with an excitation wavelength of 488 nm and emission wavelength of 512 nm.

Real time quantitative PCR gene expression analysis. Total RNAs were extracted from cells using a phenol-chloroform protocol(3). The cDNA synthesis was performed using the QuantiTect reverse transcription Kit (Qiagen) according to manufacturer's protocol. RT-qPCR was carried out in triplicate using the SsoAdvanced™ Universal SYBR® Green Supermix (Biorad) and the CFX96 Touch™ Real-Time PCR Detection System (Biorad). IHF mRNA served as internal control for standardization. Primers used are listed in Table S7.

Statistical analyses. All the experiments, unless specified, were made in triplicate. Error bars were calculated taking into account the standard error of the mean (SEM). The null hypothesis was tested by using the one-way ANOVA. P values < 0.05 were considered statistically significant. * p value<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

Abbreviation: IPTG, IsoPropyl β-D-1-ThioGalactopyranoside; sfGFP, superfolder Green Fluorescent Protein; MP, Membrane Protein; GPCR, G Protein-Coupled Receptors; PDB, Protein Data Bank; T7RNAP, T7 RNA Polymerase

PBS, Phosphate Buffered Saline; TBS, Tris Buffered Saline; FSC, Forward Scatter Channel; SSC, Side Scatter Channel; SA, Surface Area; RT-qPCR, Real Time; quantitative Polymerase Chain Reaction; RFI, Relative Fluorescent Intensity; Ni-NTA, Nickel-NitriloTriacetic Acid; DDM, *n*-Dodecyl β-D-Maltoside ; DMSO DiMethyl SulfOxide

ACKNOWLEDGMENTS. This work was supported by the Centre National de la Recherche Scientifique, INSERM, and by the "Initiative d'Excellence" program from the French State (Grant "DYNAMO", ANR-11-LABEX-0011-01). FA is supported by a DYNAMO PhD fellowship. Authors wish to thank Christiane Bouchier (genomic sequencing platform, Pasteur Institute, Paris) for DNA sequencing, Corrine Cordier for stimulating discussions and Nisha Gopal for critical reading of the manuscript. MP-GFP fusion plasmids were kindly provided by Jan Whilem de Gier and Anna Helm (Stockholm university, Sweden) and Dan Minor (UCSF, USA). The pMW7-GFP and pHis17 plasmids were given by Mike Runswick and JE Walker (MRC, Cambridge, UK). The pCT2Phe and pCT2His plasmids were kindly provided by Mathias Springer (IBPC, Paris).

1. Lundstrom K (2007) Structural genomics and drug discovery. *J Cell Mol Med* 11(2):224–238.
2. Skretas G, Georgiou G (2009) Genetic analysis of G protein-coupled receptor expression in Escherichia coli: inhibitory role of DnaJ on the membrane integration of the human central cannabinoid receptor. *Biotechnol Bioeng* 102(2):357–367.
3. Angius F, Ilioaia O, Uzan M, Miroux B (2016) Membrane Protein Production in Escherichia coli: Protocols and Rules. *Methods Mol Biol Clifton NJ* 1432:37–52.
4. Studier FW, Moffatt BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* 189(1):113–130.
5. Moffatt BA, Studier FW (1987) T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* 49(2):221–227.
6. Jones KL, Kim SW, Keasling JD (2000) Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. *Metab Eng* 2(4):328–338.
7. Walker JE, Miroux B (1999) Selection of Escherichia coli Hosts That Are Optimized for the Overexpression of Proteins. *In Manual of Industrial Microbiology and Biotechnology*, MIMB2. (A.L. Demain and J.E. Davies), pp 575–584. ASM press.
8. Dong H, Nilsson L, Kurland CG (1995) Gratuitous overexpression of genes in Escherichia coli leads to growth inhibition and ribosome destruction. *J Bacteriol* 177(6):1497–1504.
9. Miroux B, Walker JE (1996) Over-production of Proteins inEscherichia coli: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. *J Mol Biol* 260(3):289–298.
10. Arechaga I, et al. (2000) Characterisation of new intracellular membranes in Escherichia coli accompanying large scale over-production of the b subunit of F(1)F(o) ATP synthase. *FEBS Lett* 482(3):215–219.
11. Ippen K, Miller JH, Scaife J, Beckwith J (1968) New controlling element in the Lac operon of E. coli. *Nature* 217(5131):825–827.
12. Kwon S-K, Kim SK, Lee D-H, Kim JF (2015) Comparative genomics and experimental evolution of Escherichia coli BL21(DE3) strains reveal the landscape of toxicity escape from membrane protein overproduction. *Sci Rep* 5. doi:10.1038/srep16076.
13. Markiewicz P, Kleina LG, Cruz C, Ehret S, Miller JH (1994) Genetic Studies of the lac Repressor. XIV. Analysis of 4000 Altered Escherichia coli lac Repressors Reveals Essential and Non-essential Residues, as well as "Spacers" which do not Require a Specific Sequence. *J Mol Biol* 240(5):421–433.
14. Hattab G, Warschawski DE, Moncoq K, Miroux B (2015) Escherichia coli as host for membrane protein structure determination: a global analysis. *Sci Rep* 5. doi:10.1038/srep12097.
15. Pédelacq J-D, Cabantous S, Tran T, Terwilliger TC, Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* 24(1):79–88.
16. Zhang Z, et al. (2015) High-level production of membrane proteins in E. coli BL21(DE3) by omitting the inducer IPTG. *Microb Cell Factories* 14. doi:10.1186/s12934-015-0328-z.
17. Wagner S, et al. (2008) Tuning Escherichia coli for membrane protein overexpression. *Proc Natl Acad Sci U S A* 105(38):14371–14376.
18. Xu P, et al. (2013) Modular optimization of multi-gene pathways for fatty acids production in E. coli. *Nat Commun* 4:1409.
19. Hammon J, Palanivelu DV, Chen J, Patel C, Minor DL (2009) A green fluorescent protein screen for identification of well-expressed membrane proteins from a cohort of extremophilic organisms. *Protein Sci Publ Protein Soc* 18(1):121–133.
20. García-Fruitós E, et al. (2005) Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. *Microb Cell Factories* 4:27.
21. Sarramegna V, Talmont F, Seree de Roch M, Milon A, Demange P (2002) Green fluorescent protein as a reporter of human mu-opioid receptor overexpression and localization in the methylotrophic yeast Pichia pastoris. *J Biotechnol* 99(1):23–39.
22. Kawate T, Gouaux E (2006) Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. *Struct Lond Engl 1993* 14(4):673–681.
23. Jackson MR, Melideo SL, Jorns MS (2012) Human sulfide:quinone oxidoreductase catalyzes the first step in hydrogen sulfide metabolism and produces a sulfane sulfur metabolite. *Biochemistry (Mosc)* 51(34):6804–6815.
24. Gubellini F, et al. (2011) Physiological response to membrane protein overexpression in E. coli. *Mol Cell Proteomics MCP* 10(10):M111.007930.
25. Geertsma ER, Groeneveld M, Slotboom D-J, Poolman B (2008) Quality control of overexpressed membrane proteins. *Proc Natl Acad Sci U S A* 105(15):5722–5727.
26. Freigassner M, Pichler H, Glieder A (2009) Tuning microbial hosts for membrane protein production. *Microb Cell Factories* 8:69.
27. Wagner S, et al. (2007) Consequences of Membrane Protein Overexpression in Escherichia coli. *Mol Cell Proteomics* 6(9):1527–1550.
28. Schlegel S, Genevaux P, de Gier J-W (2015) De-convoluting the Genetic Adaptations of E. coli C41(DE3) in Real Time Reveals How Alleviating Protein Production Stress Improves Yields. *Cell Rep*. doi:10.1016/j.celrep.2015.02.029.
29. Bonner G, Patra D, Lafer EM, Sousa R (1992) Mutations in T7 RNA polymerase that support the proposal for a common polymerase active site structure. *EMBO J* 11(10):3767–3775.
30. Osumi-Davis PA, et al. (1994) Bacteriophage T7 RNA polymerase and its active-site mutants. Kinetic, spectroscopic and calorimetric characterization. *J Mol Biol* 237(1):5–19.
31. Lenfant F, Labia R, Masson JM (1990) Probing the active site of beta-lactamase R-TEM1 by informational suppression. *Biochimie* 72(6–7):495–503.
32. Odoi KA, Huang Y, Rezenom YH, Liu WR (2013) Nonsense and sense suppression abilities of original and derivative Methanosarcina mazei pyrrolysyl-tRNA synthetase-tRNA(Pyl) pairs in the Escherichia coli BL21(DE3) cell strain. *PloS One* 8(3):e57035.
33. Schlegel S, Genevaux P, de Gier J-W (2016) Isolating Escherichia coli strains for recombinant protein production. *Cell Mol Life Sci CMLS*. doi:10.1007/s00018-016-2371-2.
34. Gialama D, et al. (2016) Development of Escherichia coli strains that withstand membrane protein-induced toxicity and achieve high-level recombinant membrane protein production. *ACS Synth Biol*. doi:10.1021/acssynbio.6b00174.
35. Baumgarten T, et al. (2017) Isolation and characterization of the E. coli membrane protein production strain Mutant56(DE3). *Sci Rep* 7:45089.
36. Sarkar CA, et al. (2008) Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc Natl Acad Sci U S A* 105(39):14808–14813.
37. Boël G, et al. (2016) Codon influence on protein expression in E. coli correlates with mRNA levels. *Nature* 529(7586):358–363.
38. Nørholm MHH, et al. (2013) Improved production of membrane proteins in Escherichia coli by selective codon substitutions. *FEBS Lett* 587(15):2352–2358.
39. Pechmann S, Frydman J (2013) Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. *Nat Struct Mol Biol* 20(2):237–243.
40. Pechmann S, Chartron JW, Frydman J (2014) Local slowdown of translation by nonoptimal codons promotes nascent-chain recognition by SRP in vivo. *Nat Struct Mol Biol* 21(12):1100–1105.
41. Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* 185:60–89.
42. Way M, Pope B, Gooch J, Hawkins M, Weeds AG (1990) Identification of a region in segment 1 of gelsolin critical for actin binding. *EMBO J* 9(12):4103–4109.
43. Ke N, Landgraf D, Paulsson J, Berkmen M (2016) Visualization of Periplasmic and Cytoplasmic Proteins with a Self-Labeling Protein Tag. *J Bacteriol* 198(7):1035–1043.

Please review all the figures in this paginated PDF and check if the figure size is appropriate to allow reading of the text in the figure.

If readability needs to be improved then resize the figure again in 'Figure sizing' interface of Article Sizing Tool.