UNIVERSITY OF TRENTO

International PhD Program in Biomolecular Sciences

XXVIII Cycle

Molecular communication between artificial cells
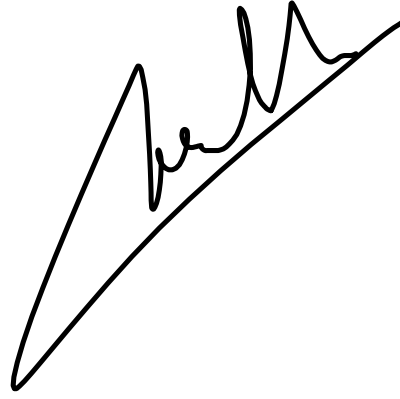

Tutor

Prof. Sheref S. Mansy

Candidate

Dario Cecchi

Armenise-Harvard Laboratory of synthetic and reconstructive biology

CIBIO (Centre for Integrative Biology)

Academic Year 2015-2016

Declaration

I, Dario Cecchi, confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Abstract

Genetic engineering has been widely used to reprogram cells for a variety of purposes, suggesting a wide range of possible applications in industrial and academic research. Although the techniques available are very well established, the mechanisms of cellular life are not completely understood. Therefore, despite offering a versatile tool, engineered cells are prone to possible unexpected behaviors.

Investigations of more controllable systems are partially focused on the creation of cellular mimics assembled from discrete components with defined properties. The controlled assembly of molecules allows the creation of entities able to form compartments in water solutions and to carry out enzymatic reactions or gene expression.[1–6] These artificial cells are able to establish communication pathways with natural cells[1,4] and may be further developed to fight pathogens or cancer cells, for example. Despite these promising results, technological applications based on cellular mimics necessitate further technical improvements.

A considerable defect of artificial cells is the lack of some mechanisms for self-sustainment that are instead present in engineered living cells. Besides few strategies aimed at energy restoration,[7,8] artificial cells are not yet able to efficiently use the available resources in their environment. Considering these technical limitations, this thesis proposes to improve communication pathways between artificial and natural cells by exploiting multiple kinds of cellular mimics. Artificial cells can vary in composition and if engineered to coordinate activity, could be capable of overcoming individual weaknesses.

To investigate the possibility of creating communities of artificial cells that collaborate with each other, the work described here was focused on establishing molecular communication pathways between two kinds of artificial cells. The designed communication was based on the exchange of chemical messages between two cellular mimics resulting either in genetic regulation or enzymatic reactions. On one side, lipid vesicles carrying gene expression through *in vitro* transcription and translation reactions and on the other side a novel structure composed of modified proteins, named proteinosome, to carry out enzymatic reactions. Each part of the communication pathway was separately investigated. Some efforts were put into the characterization of genetic switches so as to be able to better tune gene expression. All the other components were then singularly tested before combining together.

One way that artificial cells, either alone or in a community, can function as a useful technology is if the artificial cells are able to sense and respond to environmental changes. The sensing functionality can be conferred by natural or synthetic transcriptional regulators. It is possible to modify biological macromolecules to interact with chemical

messages released by natural cells. The second part of the thesis summarizes two distinct works aimed at developing two kinds of biosensors with potential applications within artificial cells.

Several technical problems arose while testing the communication pathway, and it was necessary to change the initial strategy to include engineered cells. Nonetheless, the work presented here offers a method for the establishment of molecular communication pathways within communities of artificial cells that could serve as the basis for future implementation in more efficient communication systems.

Table of contents

Abbreviations list

3OC6 HSL HSL = N-3-(oxohexanoyl) homoserine lactone

αHL = alpha-hemolysin

ACP = acyl carrier protein

AHLs = acyl homoserine lactones

APS = ammonium persulfate

aTc = anhydrotetracycline

DEPC = Diethyl pyrocarbonate

DFHBI = (Z)-4-(3,5-difluoro-4-hydroxybenzylidene)-1,2-dimethyl-1H-imidazol-5(4H)-one

DMF = N,N-dimethylformamide

DTT = dithiothreitol

FDEL = Freeze-Dried Empty Liposomes

GOx = glucose oxidase

HRP = horseradish peroxidase

IPA = indole-3-propionic acid

IPTG = isopropyl ß-D-1-thiogalactopyranoside

LLO = listeryolisin O

MBP = maltose binding protein

MG = malachite green

MGI = malachite green isothiocyanate

mRNA = messenger RNA

Ni-NTA = nickel-nitrilotriacetic acid

PCR = polymerase chain reaction

PEG = polyethylene glycol

PFO = perfringolysin O

PNIPAAm = poly(*N*-isopropylacrylamide)

POPC = 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine

QS = quorum sensing

RBS = ribosome binding site

SAM = S-adenosylmethionine

TEMED = N,N,N′,N′-Tetramethylethylenediamine

tRNA = transfer RNA

TX/TL = transcription/translation

UTR = untranslated region

w/o = water-in-oil

w. t. = wild type

Chapter 1

Building life-like entities

Our current understanding of the chemistry and physics of life is far from complete, thus complicating efforts in formulating an explicit set of rules that describe life. Nonetheless, what is known is sufficient to engineer new phenotypes of existing, living cells. That is, biology is understood well enough to allow for some type of intervention with limited predictability, because collateral effects may arise from any direct modification, but much room for improvement remains.

Cells engineered to sense and respond to the environment have some useful application. Any organism can actively sense external stimuli at a cellular level, e.g. chemical messages released from other cells or changes in pH or temperature. Cells can consequently respond to these variations by releasing other signals that can be intercepted by other cells. The regulation of these signaling pathways can be enzymatically driven and regulated at a genetic level. By knowing the molecular details of these processes, it is possible to rationally design and exploit these communication pathways. Some attempts in this direction produced engineered bacteria that are able to fight cancer cells or pathogens.[9–14]

Despite these promising applications, our incomplete understanding of biology results in an inability to predict the activity of engineered pathways. The newly engineered device could lead to unexpected and possibly detrimental effects. One reason is that the engineered pathways typically exploit biological parts that are normally used for other purposes inside of the cell. As a consequence, there may be crosstalk or a competition of resources, leading to reduced efficacy.

One approach to avoiding any possible cross-interference with the host genetic pathways is to modify natural cells with orthogonal elements. Novel nucleic acids polymers might be exploited for genetic inheritance[15] and DNA plasmids containing new base pairs can be successfully replicated in bacteria.[16] Unnatural amino acids can be included in polypeptides by means of orthogonal ribosomes[17,18] and modifications of tRNAs based either on a suppressed stop codon to carry a novel amino acid[19] or on the recognition of a quadruplet codon.[20]

Problems can also arise from interactions with the environment. A cell engineered to fight a disease may be stopped by the action of the immune system, for example. One recent attempt to circumvent such problems exploited the implantation in mice of engineered cells in a semipermeable compartment that protects the engineered cells from the action of immune cells.[21]

The promising examples described above are not yet sufficient to offer an efficient system free of any possible unexpected effects. The orthogonal elements available for gene expression are not completely independent from the host mechanisms and the mice

implants only allow the use of cells for a period of time limited to their life cycle. These aspects strongly limit the control of designed interactions with other cells.

To reduce the number of unmanageable outcomes, alternative devices may be constructed by the assembly of simpler components with defined properties. Cellular mimics can be assembled from individually purified molecules to perform interactions with natural cells similarly to what was described for engineered cells. The use of a simplified system, whose properties are well characterized, may result in better control. In other words, reducing the complexity of the system should also reduce the number of collateral effects. The elements composing the artificial cells are chosen to have no interference with natural cells. Furthermore, the artificial cells are designed to carry specific functions and are not able for self-replication nor have any homeostatic mechanism to react to environmental changes. For these reasons the artificial cells ensure a higher degree of control and should not produce any unwanted interactions with natural cells.

An artificial cell is conceived to contain all the elements required for molecular communication with natural cells confined to a compartment that allows also for some exchange with the external environment. All of the components may be assembled by taking inspiration from what can be found in nature but do not require to be an exact reproduction. Communication pathways can be regulated at the genetic level through the use of cell-free transcription and translation systems and compartments can be based either on lipid bilayers as natural cells or on other elements able to create a selectively permeable membrane.

1.1 In vitro *transcription and translation systems*

Genetic regulation can be established in artificial compartments through reaction mixtures for *in vitro* transcription and translation (TX/TL). Cell-free protein synthesis was first described in the 1950s as a tool for the characterization of the mechanisms involved in gene expression[22,23] before modern technologies for the introduction of recombinant DNA in *E. coli* were available.[24–26] To date, several optimized protocols were developed based either on eukaryotic or on prokaryotic cells,[27,28] and systems based on *E. coli* are among the most widely used in synthetic biology.[29]

There are mainly two different reaction mixtures based on *E. coli* machinery, one based on a combination of each single component needed for gene expression and the other on a crude cell extract. The first was initially developed by Ueda and colleagues[30] as a mixture composed of enzymes, transcription and translation factors recombinantly overproduced and purified. The system was named "Protein synthesis Using Recombinant Elements" (PURE) and further optimized independently by other research

groups and biotech companies. George Church developed stable *E. coli* strains to facilitate the purification process by grouping more than a single component in the same strain.[31] An *E. coli* extract capable of protein synthesis was shown to be functional much prior to the development of the PURE system[22] and the original protocol underwent several optimization processes.[24] The protocol for the preparation of an *E. coli* extract is usually referred to as S30, for the centrifugation conditions (30,000 g for 30 min) initially used to separate the clarified cytosolic content from the membrane debris after cell lysis.[22]

The two systems offer different advantages and disadvantages. A system composed of individually purified elements is free from endogenous nucleases, metabolic enzymes and energy-consuming factors.[29] Nonetheless, the presence of extra factors confers added value to the S30 system. All the enzymes involved in glycolysis can help the restoration of energy resources and allow for longer-lasting gene expression.[8] Enzymes involved in fatty acid synthesis resulted useful in a pathway described in the following chapters (see § 3.2.1). Even nucleases and proteases turned out to increase the final protein yield, presumably because of a higher turnover of resources.[32]

Being a fully defined system composed of reconstituted minimal components, it has been argued that the PURE system can be more easily modeled.[33,34] Despite the higher information on the reaction components of the PURE system compared to the S30, it is possible to define some basic rules for the design of genetic elements both for the PURE system[35,36] and for the S30 extract.[37,38] Furthermore, even if a better control degree can be obtained for the PURE system because of the higher probability to get a complete model of the reaction, the S30 extract leaves room for specific modifications by the use of different *E. coli* strains that can result in different factor's mixtures.[24] The systems used in this work were the PURE system from New England BioLabs and a home made S30 extract prepared according to the protocols of Noireaux and colleagues.[39]

1.2 Compartments

Natural cells have evolved within compartments mainly composed of lipids, proteins and carbohydrates. A simpler compartment can be obtained in water by the mere use of amphipathic lipids, consisting of hydrophilic ("head") and hydrophobic ("tail") parts (A). The hydrophobic force generally results in the hydrophobic chains aggregating together away from water molecules and the hydrophilic moieties mediating contacts with water.[40] This kind of interactions can give rise to the formation of two different supramolecular assemblies: micelles or vesicles (Figure 1 - 1). Micelles are spherical structures composed of a polar surface and a hydrophobic internal core where hydrophobic tails interact with each other. The assembly of vesicles instead occurs when

lipids form a double-layered membrane. In this structure, lipid tails point towards the center of the double layer and polar heads interact with water molecule present both on the external and on the internal side (lumen) of vesicles. This membrane is able to separate a water compartment from the external environment. Lipid vesicles can be prepared by several methods and with several lipid compositions.[41] Different kinds of lipids influence the formation of the bilayer and what determines the formation of micelles or vesicles is related to factors such as the surface of the polar head group and the length of the hydrophobic tail.[42] Among the methods described, there are two which are mainly exploited with *in vitro* TX/TL systems.
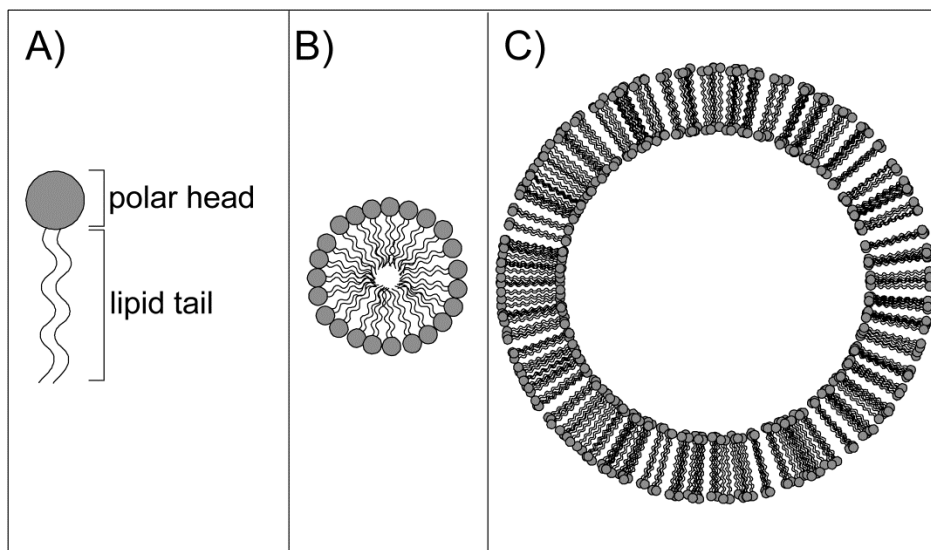
Figure 1 - 1. Schematic representation of an amphipathic lipid and the possible supramolecular assembly in water solutions. A) An amphipathic lipid is mainly composed of a polar group (named "head") and a lipid moiety (named "tail"). When dispersed in water, lipid tails tend to interact with each other to exclude water molecules, resulting in different kinds of supramolecular assembly. B) A micelle is formed when lipid tails point towards the center of a spherical structure. C) A liposome is a spherical vesicle formed when lipids assemble in a bilayer sheet where the tails point towards the center of the sheet. A typical biological membrane mainly contains phospholipids, composed of a phosphate group (head) joined to two fatty acid chains (tails).

The Freeze-Dried Empty Liposomes (FDEL)[43] method (Figure 1 - 2) involves the formation of a thin lipid film after the evaporation of the organic solvent where lipids are dispersed, typically chloroform. The film is resuspended in water to allow for the formation of vesicles containing multiple lipid bilayers and homogenized by mechanical stirring or extrusion. Vesicles are then lyophilized and later resuspended with the TX/TL reaction.[1,44,45]
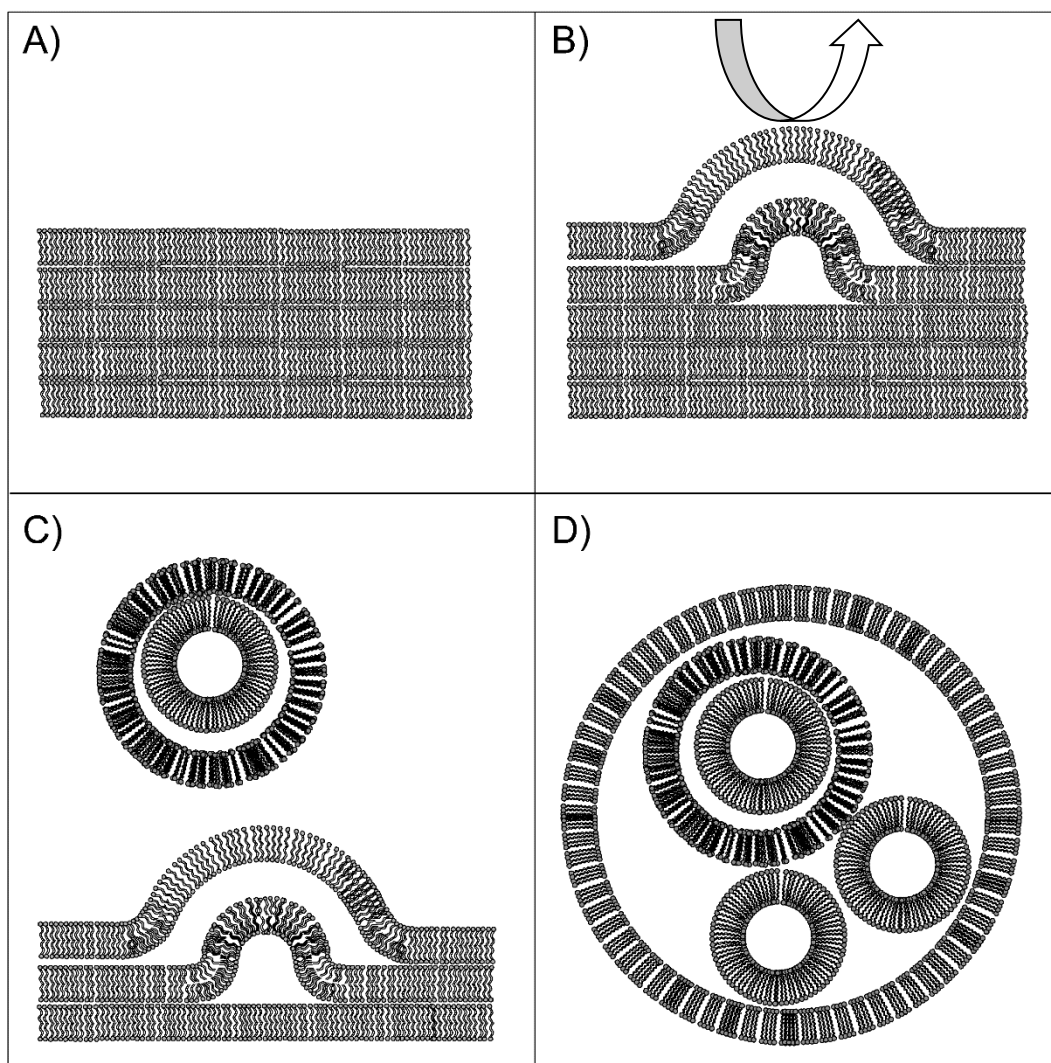
Figure 1 - 2. Freeze-Dried Empty Liposomes (FDEL). A) A thin film of lipids consisting of multiple bilayers is created by solvent evaporation. Typically, lipids are dispersed in organic solvents like chloroform and then evaporated to allow the formation of a homogeneous film. B) The resuspension of the film with water induces the swelling of lipid bilayers, from which C) multilamellar vesicles detach, to D) generate a heterogeneous dispersion of lipid vesicles.

The other method was first described by Pautot *et al.* and involves the formation of one lipid leaflet at a time (Figure 1 - 3).[46] A solution of mineral oil containing phospholipids is deposited on top of a water solution. Phospholipids partition to the interface of the two phases with polar head groups facing the water phase and the hydrophobic tails submerged in the oil phase. Separately, a water-in-oil (w/o) emulsion is created by mechanical agitation of a mixture of water and mineral oil that is stabilized by lipids. The w/o emulsion is later placed on top of the biphasic solution. The water droplets are then forced from the oil to the water phase by mild centrifugation. The passage through the interface allows the formation of a second lipid leaflet surrounding the monolayer that stabilize the w/o emulsion droplet, thus generating a lipid bilayer. The process is compatible with the encapsulation of transcription-translation machinery.[7,47–50]

The contents of each vesicle may vary regardless of the vesicle generation protocol used. Therefore, gene expression through TX/TL will likely vary among the lipid vesicles. For the method developed by Pautot *et al.*, a heterogeneous distribution of components was confirmed and that the DNA template concentration inside each vesicle was indicated as one of the most critical factors influencing gene expression.[51] Therefore, variability between compartments can be partially overcome by increasing the DNA concentration to ensure that the minimal amount of genetic elements required for function are present in each droplet.
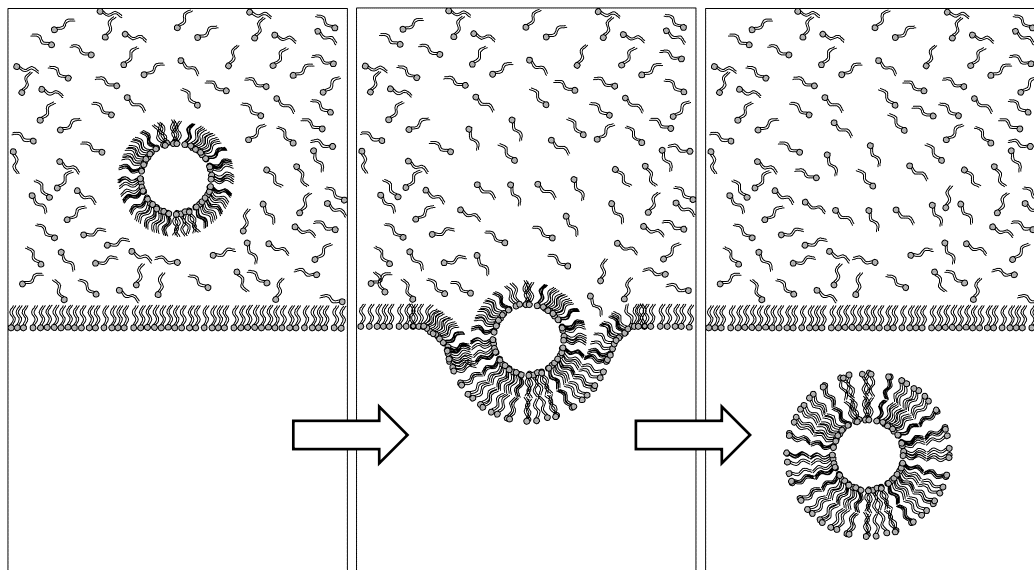
Figure 1 - 3. A method for the creation of liposomes based on emulsion droplets. Water-in-oil emulsion droplets are stabilized by a phospholipid monolayer and forced to cross a phospholipid interphase between oil and water. Another phospholipid leaflet is formed around the droplet during this passage. Lipid vesicles carrying enzymatic reactions and *in vitro* TX/TL systems can be formed with this method by the encapsulation of the desired reaction mixture in lipid emulsion droplets.

In addition to lipid membranes, it is possible to create other kinds of compartments in aqueous solutions where enzymatic reactions or in vitro gene expression can occur. Compartmentalization can be driven by diverse forces that can be divided into two main groups. One type of compartments consists of semipermeable membranes obtained by means of diverse interactions such as ion bonds or Pickering emulsions. The second type of compartments is obtained by the separation of two liquid phases so that droplets of a phase are dispersed in the other. These droplets typically contain a content in water much lower than the external phase and offer a membrane-free compartment where several kinds of small and big molecules can be absorbed. The number of possible systems available constitutes an interesting scenario of potential artificial cells with different physical and chemical properties.

Pickering emulsions can be generated through a process in which emulsion droplets are stabilized by the localization of particles to the interface of two immiscible phases.[52] Particles showing an equilibrium between hydrophobic and hydrophilic properties can behave as surfactants and isolate water droplets in an organic phase. These includes colloidal nanoparticles made of a mixture of hydrophobic and hydrophilic molecules to create colloidosomes[3,53] and amphipathic protein-polymer conjugates can be made to create structures called proteinosomes.[5,6] The basic procedure for the creation of these compartments involves an emulsification of aqueous solutions in an organic phase together with colloidal particles or amphipathic protein-polymer conjugates. The surfactant behavior of these particles creates a porous membrane surrounding the water droplet that is later stabilized by intermolecular covalent bonding. The cross-linked membrane can be safely transferred to water solutions, where it could potentially be in contact with natural cells.

Porous membranes can also be formed by the aggregation of molecules through ionic interactions occurring at the interface of two aqueous solutions. One example is given by alginate microcapsules, resulting from the precipitation of anionic polysaccharide alginate in presence of specific counterions. Complexes with sodium alginate are soluble, while other cations such as calcium or the polysaccharide chitosan induce precipitation. The addition of sodium alginate solution to a solution containing one of these cations triggers salt precipitation at the interface, resulting in the separation of two aqueous phase through an ionic membrane. The dimension of the compartments can be controlled by controlling the dimensions of sodium alginate droplets added to the cation solution. Enzymes or TX/TL machineries can be mixed together with sodium alginate to obtain cellular mimics capable of catalysis or gene expression.[4,54] Similarly to phospholipid bilayers, despite the different composition and chemical properties, this type of membrane allow for the passage of small molecules while retaining macromolecules,[4] and some methods are available to adjust the permeability to a certain extent.[55]

A good example of membrane-free compartments is given by the process of coacervation, which involves the separation of two liquid phases as a result of the attraction of oppositely charged macro-ions. The interaction is strong enough to reduce the binding with water molecules but do not result in an arrangement as stable and ordered as salt crystals. Therefore, the attractive forces give rise to phase separation rather than salt precipitation. Coacervate droplets were initially proposed as a model of protocells crucial for the origins of life,[56] and some applications as cellular mimics were recently reported.[57,58] The absence of a membrane and the charge of these compartments allows for the uptake of small and big molecules from external solution.[58,59] Furthermore,

the molecularly crowded environment of coacervate droplets has been reported to enhance enzymatic activity[60] and gene expression.[57]

It is also possible to have cellular mimics composed of multiple compartments, that allows for the specific localization of protein expression or enzymatic reactions. Water solutions differing in density can be mixed together in a multiphasic system, and water-in-oil emulsion droplets or lipid vesicles can be made to include these multiple compartments.[61,62] In aqueous two-phase systems, for example, it has been shown how a protein synthesized in these systems can preferentially localize to one of the two phases.[63] Multiple compartmentalizations can also be achieved in liposomes obtained by Pautot's method where different emulsion droplets are merged together. Every compartment can carry separate gene expression reactions.[48] Separate compartments can be put in contact to communicate with each other and allow the free diffusion of small molecules, resulting in a cascade of enzymatic reactions where the major players involved are confined to different spaces.[64] The possibility to physically separate each single component of an enzymatic reaction could allow for the optimization of each single step with a reduced waste in resources. Therefore, multiple compartments offer a good tool for the improvement of artificial cells.

1.3 Molecular communication

Although gene expression and enzymatic reactions have been widely shown to occur in confined compartments, there are only a few examples of artificial cells actively interacting with their environment. A chitosan-alginate capsule can synthesize and release by passive diffusion quorum sensing molecules and deliver a signal to bacteria.[4] Small molecules can induce a response in liposome-based artificial cells through genetic regulatory elements, such as riboswitches[1,2] or transcriptional repressors[65,66].

Despite the limited number of mechanisms described, some attempts to control communication between artificial and natural cells have been accomplished. In a work carried out in our group,[1] it was shown how it is possible to translate an inert chemical message for *E. coli* into a meaningful signal able to induce a response. The communication pathway involved the expression of a pore-forming protein inside of lipid vesicles, under the control of the theophylline riboswitch. In the presence of theophylline, the riboswitch bound to the molecule undergoes a conformational change unveiling a hidden ribosome binding site (RBS) thus allowing for the expression of α-hemolysin. The pores formed by this protein released IPTG outside of liposomes and activated the expression of a reporter gene in *E. coli*. Similar to artificial cells based on lipid vesicles, a communication pathway was demonstrated between bacteria and cellular mimics

obtained by a Pickering emulsion of a phospholipid monolayer surrounding water-in-oil emulsion droplets. These phospholipid emulsion droplets were tested for their ability to set communication with bacteria exploiting genetic regulatory elements derived from quorum sensing machinery.[67] Although the system is not so versatile because it requires bacteria to be included in emulsion droplets, it demonstrates the feasibility of establishing communication between artificial cells and bacteria in both directions, either by sensing or by sending a chemical message.

1.4 The need for artificial cells communities

The potential interactions between natural and artificial cells are limited by several constraints. Artificial cells displaying more complex behavior require more energy demanding circuitry beyond what is easily possible with current technology. Natural cells have evolved pathways for energy resource management and regeneration, while artificial systems are still far from complete self-sufficiency. Nonetheless, some progress has been made.

Attempts at restoring energy resources for TX/TL reactions in cellular mimics resulted in increased protein yield and prolonged activity of the system. The selectivity of the membranes is chosen to retain all the elements required for protein synthesis inside of the artificial cell so that small molecules cannot be directly supplemented from the external solution to replenish the reaction. To increase the availability of resources some elements can be added to the TX/TL reaction, such as maltose that was reported for glucose regeneration and ATP production through the Krebs cycle.[8] Amino acids and nucleotides can also be regenerated by the degradation of nascent RNA and protein. It was shown how an increased turnover of these resources by the addition of specific RNases and proteases increases the final yield of the protein of interest.[32] It is also possible to allow the passage of small molecules, like amino acids and nucleotides, that are not permeable to certain type of lipid membranes such as 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) through α-hemolysin pores resulting in prolonged protein expression.[7]

If artificial cells were able to self-assemble, all the elements required to perform a specific function could be easily generated thus providing a much more efficient system. A complete duplication process in a cellular mimic has not been described yet but some examples of cellular division processes were conducted in liposomes. The process of growth and division was first shown to occur through physical processes consequent to the addition of micelles[68] and change in osmotic pressure[62] or redox status.[69] It is clear that extreme conditions like these can be hardly combined with a gene expression

system, but some attempts on the reconstitution of the natural machinery responsible for *E. coli* cell divisions have been made as well. A combination of the purified proteins involved in the mechanism are able to form rings at the center of liposomes,[70] eventually leading to division.[71] It is also possible to reconstitute basic elements of the eukaryotic cytoskeleton that are involved in cellular division such as actin filaments.[72,73] Although there are no reports of division, it is possible to modulate the shape of the membrane through the combination with myosin.[74] Further, it has been shown that by combining PCR and ionic interactions between DNA and lipid membrane, it is possible also to force the segregation of vesicles with an equal distribution of DNA molecules. Artificial cells may then duplicate together with genetic information.[75]

Despite the absence of a completely self-sustainable cellular mimic, the scenario of artificial cell is broadening over the years offering an increasing number of artificial cells suitable for different kind of applications. Every system has a diverse set of advantages and disadvantages and their combination may complement reciprocal defects. We can then envision a mixed population of artificial cells assigned with different tasks to be fulfilled as the result of a collaboration. The success of the desired function will be entrusted to the community rather than to the individual cellular mimic.

The creation of artificial cell communities is strictly related to the construction of molecular communication pathways. In collaboration with Prof. Stephen Mann at the University of Bristol, we are currently in the process of creating interactions in a mixed population of cells consisting of liposomes carrying *in vitro* gene expression systems, proteinosomes with enzymatic activity and *E. coli*. In the meantime, his group managed to show how it is possible to have interactions between different kinds of artificial cells in a predator-prey system where a protease-loaded coacervate seeks and destroys a proteinosome.[58] The project carried in this collaboration is instead focused on cooperative interactions between artificial cells, possibly expanding the range of communication pathways between artificial and natural cells by means of the exchange of chemical messages.

The PhD project described in this dissertation is organized in two main sections further divided into chapters. The first part reports the work aimed at establishing a communication pathway between two different populations of artificial cells: liposomes able for gene synthesis and proteinosomes with catalytic activity. The designed network foresaw the delivery of a chemical message from liposomes under the control of a genetic switch. Proteinosomes would have triggered gene expression by enzymatic synthesis of the inducer molecule and performed enzymatic activity using the chemical message coming from liposomes as a substrate. A consistent part of the project was dedicated to

the research of a genetic circuit able for the tightest control on the chemical message delivery from liposomes. Several transcriptional regulators were tested in *in vitro* TX/TL reactions and the best was integrated within artificial cells. Consequently, all of the steps of the molecular communication pathway were separately analyzed. Both the delivery of a chemical message from liposomes and the enzymatic production of the molecule responsible for gene induction were tested separately. The number of problems arose while setting the artificial cells network led to reconsider the whole design so to include some engineered bacteria.

The second part of the thesis summarizes the trials carried in engineering two novel biosensors that could possibly be integrated within artificial cells. A DNA-based sensor was designed to carry two aptamer sequences, one for the binding of an analyte and another one for the binding of a fluorescent molecule. The fluorescence quantum yield of the fluorophore would be consistently increased when bound to the aptamer. The folding of this aptamer would depend on the presence of the analyte: the binding of the analyte to the respective aptamer would induce a conformational change on the DNA sensor allowing the binding of the fluorophore. In order to develop a DNA aptamer able for binding a fluorophore, three *in vitro* evolution strategies were performed but none of them succeeded, therefore it was not possible to develop the designed biosensor.

The second biosensor was based on the rational design of a transcriptional regulator able to control gene expression according to the presence of an analyte. The protein was inspired to a natural transcriptional repressor whose activity is regulated by the presence of a small molecule. The engineering process was based on the change of the binding affinity to a molecule similar to the natural ligand. The protein would have been consequently integrated into a genetic circuit and allow the detection of the analyte, i. e. the novel ligand, through the regulation of the expression of a fluorescent reporter protein. Attempts at changing the affinity of the regulator to a novel ligand were performed after *in silico* analysis of the interactions between the protein and the natural ligand. Mutants were designed and tested *in vitro* for gene regulation under the control of both the natural and the novel ligand. Finally, some information on the interactions between protein and ligand were acquired but no efficient regulator has been yet developed.

Aims of the thesis

The work carried in this PhD project is divided into 4 experimental sections described in the following chapters.

Chapter 2, "Genetic engineering for use in artificial cell", describes a part of the project aimed at finding the best regulatory system for use in artificial cells composed of liposomes carrying *in vitro* gene expression systems. Four genetic circuits were assembled to carry a transcriptional regulator and a reporter gene. The efficacy of each regulator was tested by means of *in vitro* TX/TL reactions;

Chapter 3, "Towards artificial cells consortia", is centered on a step-by-step analysis of a molecular communication pathway designed between two kinds of artificial cells: liposomes carrying *in vitro* transcription and translation reaction and proteinosomes carrying enzymatic reactions. Liposomes were tested for the ability to perform gene expression and for the delivery of a chemical message to proteinosomes through fluorescent reporters. One of the enzymes carried by proteinosomes was tested for the ability to produce a chemical message capable of triggering gene expression in liposomes;

Chapter 4, "Selection for a malachite green DNA aptamer for use in a sensor molecule", reports some attempts at developing a biosensor based on DNA. Engineering this biosensor foresaw the *in vitro* evolution of a DNA aptamer able to bind and increase the fluorescence yield of the fluorophore malachite green. Three strategies aimed at evolving the aptamer were designed and tested;

Chapter 5, "Engineering TrpR to sense the neurotransmitter serotonin", summarizes some trials focused on the development of another biosensor based on the transcriptional repressor TrpR. Some mutants were rationally designed to switch the affinity of the protein from the original ligand tryptophan to serotonin, that has a very similar chemical structure. Similarly to what was performed for the transcriptional regulators in chapter 2, these mutants were cloned in a genetic circuit and tested through *in vitro* TX/TL reactions.

Part A

Communication pathways for artificial cells

Chapter 2

Genetic engineering for use in artificial cell

One basic requirement for an artificial cell with potential application is tunability. A desired function must be activated only when needed, so these systems must include a switch mechanism capable of changing from an "on" to an "off" state. An artificial cell based on lipid vesicles encapsulating *E. coli* transcription and translation machinery can be engineered by following the same methods applied to natural cells, by intervening at a genetic level.

Gene expression can be regulated through several mechanisms, many of which occur through interactions between small molecules and macromolecules, such as proteins and nucleic acids. The effect of these molecules can be both positive or negative on gene expression. Previous work carried out in our group demonstrated the possibility to control gene expression in artificial cells by means of a riboswitch.[1,2] These sequences at the 5'-untranslated region (UTR) of mRNAs can mask or unveil the ribosome binding site according to the presence or absence of a ligand specific for the aptameric domain of the riboswitch. Although effective, this regulation turned out to be not strict enough and showed much background expression in the off state, in the absence of the inducer molecule.

One of the major disadvantages of the system was thought to be due to the fact that the regulation occurs at a translational level. The gene of interest is actively transcribed under the control of a highly processive polymerase (T7 RNA Polymerase)[76] and even a small percentage of mRNA may result in a high background expression. Therefore, regulation at the level of transcription seemed a better alternative to obtain a stricter off state. One very simple example of this kind of genetic regulation is given by transcriptional repressors. The major feature shared by these protein factors is to bind specific DNA sequences, called operators, that are close or enclosed in promoter regions and prevent the binding of RNA polymerases by steric hindrance. The action of these proteins is regulated by the presence of a cofactor, a small molecule able to induce an allosteric change that can forbid or allow DNA binding (Figure 2 - 1). In order to find a good genetic regulation for use in artificial cells, three repressors from *E. coli* were taken into consideration: 1) LacI, involved in the transport and metabolism of lactose through the regulation of the Lac operon;[77] 2) TetR, regulating the expression of a transmembrane pump aimed at expelling the antibiotic tetracycline;[78] 3) TrpR, regulating the synthesis of the amino acid tryptophan.[79]

All of the repressors can exist in a form bound to their ligand or in a free form. The first form is known as holorepressor while the second is the aporepressor. The effects induced by the switch from apo- to holorepressor is different for the four repressors taken under examination. LacI and TetR are subjected to a negative regulation, while TrpR

22

undergoes a positive regulation. In the absence of the cofactors, LacI and TetR repress gene expression while the binding to the ligand forbids the binding to the operators (Figure 2 - 1 A). TrpR has instead an opposite mechanism, being active only as a holorepressor (Figure 2 - 1 B). The natural ligand for LacI is allolactose, a catabolite created by the enzyme β-galactosidase in presence of lactose.[80] The catabolite induces the expression of all the genes required for its metabolism. TetR is instead involved in the resistance to the antibiotic tetracycline. When the antibiotic is present, *E. coli* cells react with the production of a pump to expel the drug.[78] TrpR regulates a whole operon involved in the synthesis of tryptophan shutting down the expression of the enzymes required in presence of high levels of the amino acid. Interestingly, this operon is subjected to a secondary regulation that involves the coding sequence *trpL* whose peculiarity is to induce different folding of the mRNA sequence according to the levels of tryptophan. The mRNA includes two tryptophan codons responsible for this change: in presence of high levels of tryptophan the mRNA forms a transcriptional terminator downstream of the peptide and none of the following genes is transcribed, while low levels of tryptophan causes a stall in ribosome that induces an alternative folding in the downstream RNA preventing the formation of a terminator and allowing the transcription of the downstream genes.[81–83]
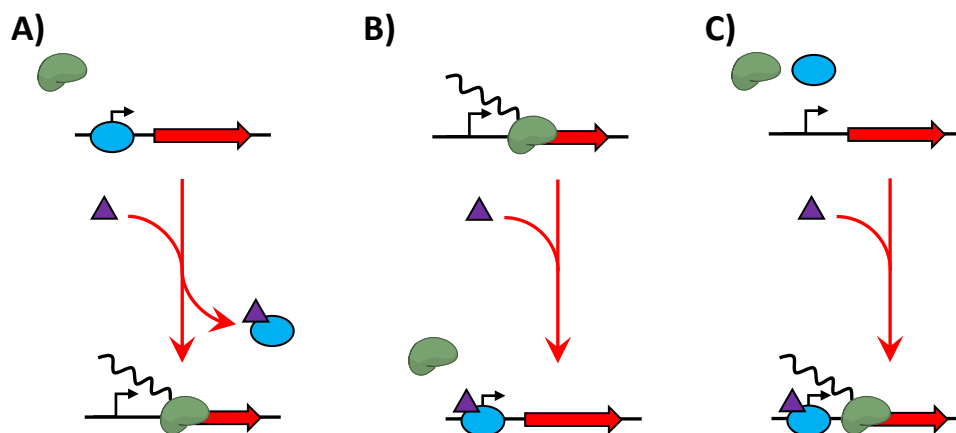
Figure 2 - 1. Transcriptional regulatory pathways tested. A) Transcriptional repressor with negative regulation (TetR, EsaR, LacI). The repressor binds to the operator sequences close to the promoter thus blocking RNA polymerase binding through steric hindrance. The binding to the ligand induces a conformational change that clears the promoter sequence to start gene transcription. B) Transcriptional repressors with positive regulation are active for repression only in presence of the ligand (TrpR). C) Transcriptional activators promote the binding of RNA polymerase rather than blocking it. The activity depends on the presence of small cofactors that induce the necessary allosteric change to lead RNA polymerase binding (LuxR).

Not all of the natural cofactors can be exploited for *in vitro* analysis, but alternative ligands are commercially available. Allolactose can be substituted by the cheaper and more efficient isopropyl β-D-1-thiogalactopyranoside (IPTG) which is not hydrolyzable by

β-galactosidase and has been regularly used for the induction of recombinant protein expression.[84] Tetracycline can be substituted by the non-toxic analog anhydrotetracycline (aTc), which was shown to have a 30-fold increased binding affinity.[85] For TrpR, several analogs are described in the literature either with a positive or a negative action on the repressor.[86] Among these, indole-3-propionic acid (IPA) was chosen because of the ability to induce gene expression *in vivo*.[87]

The transcriptional regulators were individually tested in TX/TL reactions for their ability to control the expression of a reporter gene. The repressor and the reporter genes were cloned separately under the control of a T7 promoter, and the reporter gene contained the specific operator sequence for the repressor right downstream of the promoter. Both translation and transcription levels were simultaneously monitored by means of a fluorescent protein and an aptamer joint at the 3'-UTR of the mRNA able to bind and increase the fluorescence quantum yield of a fluorophore. In order to individuate the regulator with the best difference between on and off state, the experiments were conducted in presence or absence of the repressor and in presence or absence of the specific inducer molecule.

2.1 Materials and methods

2.1.1 Reagents and general supplies

PURExpress® *In vitro* Protein Synthesis Kit, DpnI, RNase inhibitor was purchased from New England Biolabs; Wizard® SV Gel and PCR Clean-Up System Wizard® Plus SV Minipreps DNA Purification System were purchased from Promega; indole 3-propionic acid (IPA), isopropyl β-D-1-thiogalactopyranoside (IPTG), anhydrotetracycline (aTc), diethyl pyrocarbonate (DEPC) were purchased from Sigma-Aldrich; phenol:chloroform:isoamyl alcohol (25:24:1) solution was purchased from Thermo Fisher Scientific; (Z)-4-(3,5-difluoro-4-hydroxybenzylidene)-1,2-dimethyl-1H-imidazol-5(4H)-one (DFHBI) was purchased from Lucerna technologies; One Shot® TOP10 Chemically Competent *E. coli* cells were purchased from Invitrogen; nickel- nitrilotriacetic acid agarose resin (Ni-NTA) and 0.1 ml tubes were purchased from Qiagen; E.Z.N.A.® MicroElute RNA Clean Up Kit was purchased from Omega Biotek.

All the material used for the preparation of the home made S30 extract was purchased according to Noireaux's indications[39] except for the following: Yeast extract and Tryptone to prepare 2xYT medium, polyethylene glycol (PEG) 8000 Da, all of the 20 amino acids, Tris base were purchased from Sigma-Aldrich; SnakeSkin™ Dialysis Tubing was purchased from Thermo Fisher Scientific; 10X TBE buffer was purchased from Euroclone; *E. coli* BL21 Rosetta 2 (DE3) from Novagen was used in place of *E. coli* BL21 Rosetta 2 and was received from Prof. Friedrich C. Simmel

2.1.2 Instruments

FastPrep®-24 from MP biomedicals was used for the bead-beating step in S30 extract; Rotor-Gene Q from Qiagen was used for the measurements of protein and RNA levels in S30 and PURE system reactions; Infinite M200 plate reader from Tecan was used to measure luminescence.

2.1.3 Water treatment with DEPC for nuclease inhibition:

1 L milliQ water (18 Ω) was mixed with 1 ml of DEPC under stirring overnight, then unbound DEPC was degraded by autoclave at 121 °C for 20'.

2.1.4 Plasmids and cloning

All of the plasmids were assembled with the Gibson method[88] and DNA primers were purchased at Eurofins MWG. Linear fragments of double-stranded DNA were created by PCR using Phusion polymerase. A 50 µl mix contained 1X HF buffer, 0.2 mM each dNTPs, 0.5 µM of each primer, 0.02 U/µl Phusion, ~ 0.2 ng/µl plasmid DNA template. The reaction ran on a thermal cycler with the following protocol: 98°C for 2 min for initial denaturation, then 29 cycles of 98 °C for 5 s, annealing temperature for 10 s and 72 °C for 15 s/kb and a final extension step at 72 °C for 10 min. The annealing temperature was calculated for each primer by the online tool IDT oligo analyzer.

PCR fragments were treated with 0.4 U/µl of DpnI for at least 1 h at 37 °C, then mixed together with a premixed Gibson assembly stored at -80 °C. The final volume of reaction was 10 µl and the mixture contained: 100 mM Tris pH 7.4, 10 mM MgCl$_2$, 0.2 mM each dNTP, 10 mM DTT, 6.25 mM PEG 8000, 1 mM NAD. The DNA was added according to the following table:

Table 2 - 1. Volumes for Gibson assembly.

PCR amplicon size	Volume (µl)
< 1kb	0.5
1-4 kb	0.5-1
4-8 kb	2
8-12 kb	3.5

PCR products were always loaded on a 1X TBE 1% agarose gel for analysis. When the PCR product contained some extra bands, the correct band was extracted from the agarose gel with Promega Wizard gel and PCR clean-up kit and added to the mix to fill the final volume. Mixes were then incubated at 50 °C and then added to *E. coli* TOP10 competent cells for transformation. Details of DNA sequences and assembly are described in the following tables.

Table 2 - 2. Primers used for PCR amplification.

Primer ID	Sequence (5'-3')
DC010	GCGGATCCGAATTCAATTAGTTTGAACTTATAAGGAGAATAATCT ATGGCTTCCTCCGAAGACG
DC046	AAGTGGCGAGCCCGATCTTCCCCAT
DC057	AATTCGAGCTCCGTCGACAAG
DC058	CATGCTAGCCATATGTATATCTCC

DC079	CGTACTAGTTAACTAGTACGCCCTATAGTGAGTCGTATTAATTTCGC
DC083	GGAGATATACATATGGCTAGCATGATGGCCCAACAATCACCCTATTCA
DC084	CTTGTCGACGGAGCTCGAATT TCAATCGCTTTTCAGCAACACCTCTT
DC094	GGAGATATACATATGGCTAGCATGATGTCCAGATTAGATAAAAGTAAAGTG
DC096	CTTGTCGACGGAGCTCGAATTTCAGGACCCACTTTCACATTTAAGTTG
DC098	CTTGTCGACGGAGCTCGAATTCTGCCCGCTTTCCAGTC
DC099	CTTGTCGACGGAGCTCGAATTTCACTGCCCGCTTTCCAGTC
DC100	ATGTATATCTCCTTCTTAAAGTTAAACA
DC101	GCGCAACGCAATTAATGTAAGTTAG
DC103	TCTCTATCACTGATAGGGACCCTATAGTGAGTCGTATTAATTTC
DC104	GGATATAGTTCCTCCTTTCAGCAAA
DC111	ATGGCTTCCTCCGAAGAC
DC133	CCCTATAGTGAGTCGTATTA
DC145	TGTTTAACTTTAAGAAGGAGATATACATATGAAACCAGTAACGTTATACG
DC146	TAATACGACTCACTATAGGGCCCCTCTAGAAATAATTTTGTTTAACTTTAAGA AGGAG
DC147	CTAACTTACATTAATTGCGTTGCGCAAGTGGCGAGCCCGATCTTCCCCAT
DC157	GAGTCGTATTAACCGGCTGCAGATCTCGATCCTCTACGCCG
DC159	CAGTCGA AAGACTGGGCCTTTCGTTTTAT GTGATGTCGGCGATATAGGC
DC160	CCAGTCTTTCGACTGAGCCTTTCGTTTTAT GGATATAGTTCCTCCTTTCAGCAAA
DC161	ATGGGGAAGATCGGGCTCGCCACTT
DC162	AAGTGGCGAGCCCGATCTTCCCCATTAATACGACTCACTATAGGGGAATT
DC163	AGTTCAAACTAATTGAATTCGGATCCGCTCACTGCCCGCTTTCCAGTC
DC199	ATTAATTGCGTTGCGCAGCAGCCCAGTAGTAGGTTG
DC200	AAGTGGCGAGCCCGATCTTCCCCATTAATACGACTCACTATAGGG
DC257	CTAACTTACATTAATTGCGTTGCGC
DC258	CTAACTTACATTAATTGCGTTGCGC
DC275	ATATCCGGATTGGCGAAT
DC276	TTAAGCACCGGTGGAGTG
DC277	AGCTCAGTCCTAGGTACAGTGCTAGCTACTAGAGTCACACAGGAAA
DC278	TACCTAGGACTGAGCTAGCCGTCAATGAGCGCAACGCAA
DC299	AGTACTTTCCTGTGTTACTCTAGTA
DC300	TGCCTGGCTCTAGTATTATTACCTTGCTGCTGACGC
DC302	TAACACAGGAAAGTACTATGTTCTCTTTCTTCCTTGAAAACC
DC306	ATTCGCCAATCCGGATAT
DC307	CCACCGGTGCTTAAGGGAGAATGCGGCC

FC267	CCGGTTAATACGACTCACTATAGCCTGTACTATAGTGCAGGTGGAAGATTGT GAGCGGATAACAATTCC
JF001A fw	TAACTCGAGCACCACCACCACCAC
RL007	CCCCTCTAGAAATAATTTTGTTTA
T9002g FW	TAATAATACTAGAGCCAGGCATC

Table 2 - 3. Relevant sequences used in this section. The table reports the insert of relevant plasmids whose information is not available elsewhere.

ID	Sequence (5'-3')
pT7_lacO_mRFP1_ spinach	TAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCT CTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCTT CCTCCGAAGACGTTATCAAAGAGTTCATGCGTTTCAAAGTTCGTATG GAAGGTTCCGTTAACGGTCACGAGTTCGAAATCGAAGGTGAAGGTG AAGGTCGTCCGTACGAAGGTACCCAGACCGCTAAACTGAAAGTTAC CAAAGGTGGTCCGCTGCCGTTCGCTTGGGACATCCTGTCCCCGCAG TTCCAGTACGGTTCCAAAGCTTACGTTAAACACCCGGCTGACATCCC GGACTACCTGAAACTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTG TTATGAACTTCGAAGACGGTGGTGTTGTTACCGTTACCCAGGACTCC TCCCTGCAAGACGGTGAGTTCATCTACAAAGTTAAACTGCGTGGTAC CAACTTCCCGTCCGACGGTCCGGTTATGCAGAAAAAAACCATGGGTT GGGAAGCTTCCACCGAACGTATGTACCCGGAAGACGGTGCTCTGAA AGGTGAAATCAAAATGCGTCTGAAACTGAAAGACGGTGGTCACTACG ACGCTGAAGTTAAAACCACCTACATGGCTAAAAAACCGGTTCAGCTG CCGGGTGCTTACAAAACCGACATCAAACTGGACATCACCTCCCACAA CGAAGACTACACCATCGTTGAACAGTACGAACGTGCTGAAGGTCGT CACTCCACCGGTGCTTAAGCCCGGATAGCTCAGTCGGTAGAGCAGC GGCCGGACGCAACTGAATGAAATGGTGAAGGACGGGTCCAGGTGT GGCTGCTTCGGCAGTGCAGCTTGTTGAGTAGAGTGTGAGCTCCGTA ACTAGTCGCGTCCGGCCGCGGGTCCAGGGTTCAAGTCCCTGTTCGG GCGCCA
His::MBP::EsaR	ATGAAAATCCATCACCATCACCATCACGAAGAAGGTAAACTGGTAAT CTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTA AGAAATTCGAGAAGATACCGGAATTAAAGTCACCGTTGAGCATCCG GATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATG GCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCT CAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGG ACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAG CTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAAC AAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGG CGCTGGATAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTT

28

	CAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGAC
	GGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACGACATTAAAGA
	CGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCT
	GGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACT
	CCATCGCAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGACCATC
	AACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATT
	ATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACC
	GTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAAC
	AAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGA
	AGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCG
	CTGAAGTCTTACGAGGAAGAGTTGGTGAAAGATCCGCGGATTGCCG
	CCACCATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCC
	GCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAAC
	GCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCG
	CAGACTAATTCGATCGAGAACCTGTACTTCCAGGGTGGTGGTGGTG
	GTTTCTCTTTCTTCCTTGAAAACCAAACAATAACGGATACGCTTCAGA
	CTTACATACAGAGAAAGTTATCTCCGCTGGGTAGTCCGGATTACGCT
	TACACTGTTGTGAGCAAAAAAATCCTTCAAATGTTCTGATTATTTCC
	AGTTATCCTGACGAATGGATTAGGTTATACCGCGCTAACAACTTTCA
	GCTGACCGATCCCGTTATTCTCACGGCCTTTAAACGCACCTCGCCGT
	TTGCCTGGGATGAGAATATTACGCTGATGTCCGGCCTGCGGTTCAC
	CAAAATTTTCTCTTTATCCAAGCAATACAACATCGTTAACGGCTTTAC
	CTATGTCCTGCATGACCACATGAACAACCTTGCTCTGTTGTCCGTGA
	TCATTAAAGGCAACGATCAGACTGCGCTGGAGCAACGCCTTGCTGC
	CGAACAGGGCACGATGCAGATGCTGCTGATTGATTTTAACGAGCAG
	ATGTACCGACTGGCAGGCACCGAAGGCGAGCGAGCCCCGGCGTTA
	AATCAGAGCGCGGACAAAACGATATTTTCCTCGCGTGAAAATGAGGT
	GTTGTACTGGGCGAGTATGGGCAAAACCTATGCTGAGATTGCCGCT
	ATTACGGGCATTTCTGTGAGTACCGTGAAGTTTCACATCAAGAATGT
	GGTCGTGAAACTGGGCGTCAGTAACGCCCGACAGGCTATCAGACTG
	GGTGTAGAACTGGATCTTATCAGACCGGCAGCGTCAGCAGCAAGGT
	AA

Table 2 - 4. Plasmids used in this section and relative cloning strategies.

Plasmid ID	Backbone	Insert	Source	Cloning strategy	
				Primers	**Template**
pET21b			Novagen		
BBa_K731500	pSB1C3	pLacIq_LacI_ pTac_lacO	Registry of standard biological parts		
BBa_C0040	pSB1C3	TetR	Registry of standard biological parts		
BBa_T9002	pSB1A3	pTet_LuxR_ pLux_sfGFP	Registry of standard biological parts		
DC024A	pET21b (LacI removed from the backbone)	pT7_LacI::His		DC057/DC101	pET21B
				DC147/DC133	pET21B
				DC145/DC098	BBa_ K731500
				DC146/DC100	-
DC013A	pET21b	pT7_lacO_TrpR		DC083/DC084	*E. coli* genome
				DC057/DC058	pET21b
DC019A	pET21b	pT7_lacO_TetR		DC094/DC096	BBa_C0040
				DC057/DC058	pET21b
DC049A	DC024A	pT7_LacI		DC099/JF001 A fw	DC024A
DC021A	pET21b	pT7_tetO_ mRFP1_spinach		DC103/RL007	FC013A
DC032A	pET21b	pT7_trpO_ mRFP1_spinach		DC079/RL007	FC013A
FC013A	pET21b	pT7_lacO_ mRFP1_spinach	Mansy Lab		
DC076A	pET21b	pT7_trpO_ mRFP1_spinach _pT7_TrpR		DC199/DC161	DC013A
				DC159/DC101	DC013A
				DC160/DC200	DC032A

FC043A	pET21b	pT7_lacO_His:: MBP::EsaR	Mansy Lab		
DC053A	DC024A	pT7_esaO_ mRFP1_spinach		DC147/DC157	FC013A
				FC267/DC101	FC013A
DC035A	DC024A	pT7_lacO_ mRFP1_spinach _ pT7_lacI		DC161/DC159	DC049A
				DC160/DC162	FC013A
DC052A	DC024A	pT7_lacO_lacI_ mRFP1		DC010/DC101	FC013A
				DC100/DC147	FC013A
				DC145/DC163	DC049A

2.1.5 DNA purification with phenol:chloroform mix

When DNA quality was not high enough (A260/A280 ≥ 1.8 and A260/230 ≥ 2), an extra step of purification was required. 50 µl of DNA were diluted in 500 µl of DEPC-treated water and mixed together with 1 volume (vol) of phenol:chloroform:isoamyl alcohol (25:24:1). The tube was vortexed and spun at max speed for 10 min to separate the aqueous from the organic phase. The upper aqueous phase was transferred to a new tube containing 1 vol of chloroform and spun at max speed for 5 min. The upper part was then mixed together with 2.5 vol of ethanol and 1/10 vol of 3 M sodium acetate pH 5.5 and incubated at -20 °C for at least 1 h. Then DNA was precipitated by centrifugation at max speed for 30 min at 4 °C. The supernatant was discarded and 200 µl of 70% ethanol were added to the pellet without resuspending. The pellet was spun again at max speed for 10 min at 4 °C. The pellet was dried at 65 °C on a thermoblock and resuspended in 20 µl of DEPC-treated water.

2.1.6 In vitro *gene expression with PURE system*

DNA templates were prepared by PCR amplification using Phusion polymerase and primers DC046 and DC104. The reaction was assembled as described above and the thermal protocol included changes in the annealing temperature (touchdown PCR): 72 °C for 15 cycles, 68 °C for 10 cycles and 65 °C for the last 10 cycles. The extension time was set at 15 sec.

The reactions were prepared with some modifications from the manufacturer's instructions. The final volume was increased by 8%, DNA was added to 12.5 nM and the reaction mixture was supplemented with 0.75 U/µl RNase inhibitor and 60 µM DFHBI. When needed, inducer molecules were added to the following concentrations: 10 µM

3OC6, 0.5 mM IPTG, 1 mM aTc, 0.5 mM IPA. Fluorescence kinetics of both transcription and translation were monitored with real-time PCR cycler Rotor-Gene Q. Transcription levels were monitored through spinach aptamer bound to DFHBI on channel Green (excitation: 470 ± 10 nm; emission: 510 ± 5 nm), while translation levels were monitored by the expression of the fluorescent protein mRFP1 on channel Orange (excitation: 585 ± 5 nm; emission: 610 ± 5 nm). Fluorescence levels were recorded every 5 min at 37 °C for ~13 h.

2.1.7 In vitro *gene expression with home-made S30* E. coli *extract*

A reaction mix for *in vitro* transcription and translation based on a crude *E. coli* extract was prepared according to the protocol of Noireaux and colleagues[39] with the modifications indicated in § 2.1.1. The final volume of the reaction was 10 µl with 20 nM of plasmid DNA (RL082A). When needed 3OC6 HSL was supplemented to 10 µM. Gene expression was verified by luminescence.

2.1.8 Fluorescence standard curve and data normalization

In order to convert fluorescence raw data into molar units, a standard curve was made both for mRFP1 and for spinach aptamer together with DFHBI. mRFP1 protein was expressed in *E. coli* BL21 (DE3) pLysS cells transformed with plasmid FC011A and purified with Ni-NTA column according to manufacturer's instructions. The protein concentration was measured by nanodrop spectrophotometer and converted according to its extinction coefficient ($\varepsilon^{584\,nm}$ = 44 000 M^{-1} cm^{-1}).[89] Spinach RNA aptamer fused at 3'-UTR of mRFP1 mRNA was transcribed *in vitro* using a home-made T7 RNA polymerase and plasmid DC032A in the following reaction mix prepared according to the indications from Seelig:[90] 35 mM $MgCl_2$, 2 mM spermidine, 200 mM HEPES adjusted to pH 7.5 with KOH, 1 mg/ml BSA, 4 mM DTT, 5 mM each NTP, 1 mU/µl yeast inorganic pyrophosphatase, 0.4 U/µl RNase inhibitor, 3 U/µl T7 RNA polymerase, 0.2 µM DNA template. The reaction was incubated at 37 °C for 4 h, then DNase I was added to a final concentration of 20 mU/µl together with its buffer provided by the supplier and incubated at 37°C for 1 h. The RNA was purified with E.Z.N.A.® MicroElute RNA Clean Up Kit and quantified at the nanodrop spectrophotometer by measuring the absorbance at 260 nm. The concentration was converted in molarity with a dedicated online tool (biotools.nubic.northwestern.edu/OligoCalc.html). Different concentrations of mRFP1 (0.2, 0.5, 1 and 2 µM) were prepared in HEPES buffer (50 mM HEPES, 10 mM $MgCl_2$, 100 mM

KCl, pH 7.6) and different concentrations of spinach mRNA (0.06, 0.15, 0.3 and 0.6 µM) were prepared with HEPES buffer together with 60 µM DFHBI.

2.1.9 Luciferase assay

To 10 µl of S30 reaction incubated for 4 h at 30 °C, 10 µl of a 2X luciferase mix from Promega, containing luciferin and ATP was added. Luminescence was measured on a 384-wells plate with a Tecan Infinite M-200 with an integration time of 1 s.

2.1.10 Chemically competent E. coli cells

E. coli TOP10 stored in glycerol stocks were grown overnight in 5 ml of LB, then reinoculated in 50 ml of LB to a starting dilution of 1:100. Cells were grown up to OD_{600} = 0.5, then chilled on ice for 10 min and harvested by centrifugation at 5000 g for 10 min at 4 °C. The pellet was resuspended in 15 ml of transformation buffer (10 mM Tris-HCl, pH 7.0, 50 mM $CaCl_2$), chilled on ice for 15 min and spun down again at 5000 g for 10 min at 4°C. The pellet was resuspended in 4 ml of transformation buffer and 20% glycerol. Cells were flash frozen in liquid nitrogen and stored at -80°C.

2.1.11 Transformation of E. coli cells

One aliquot of chemically competent *E. coli* was thawed on ice and incubated for 30 min with the DNA to be transformed. Heat shock was for 1 min at 42 °C, placed on ice for 2 min, then 800 µl of LB were added to the cells. The culture was incubated at 37 °C with shaking at 220 rpm for 1 h.

2.1.12 Purification of the repressor EsaR

The coding sequence of EsaR was fused both to a maltose binding protein (MBP) and a His tag and cloned in plasmid pET21b to create plasmid FC043A, according to the indications of Schu *et al.*[91] *E. coli* BL21 (DE3) pLysS cells were transformed with FC043A and bacteria were grown at 37 °C shaking at 220 rpm until they reached OD_{600} = 0.5. Protein expression was induced by the addition of 1 mM IPTG and cells were kept growing at 37 °C shaking at 220 rpm for 4 h. Protein was then purified with Ni-NTA column according to manufacturer's instructions.

2.2 Results

2.2.1 Design of genetic circuits and tests of the transcriptional repressors in PURE system

For all of the repressors to be tested, two plasmids were designed. One plasmid contained the repressor coding sequence under the constitutive promoter T7, and another plasmid containing a reporter gene under the control of a T7 promoter fused to the specific operator close to the transcription start site: in position +3 from transcription start site for LacI and +4 for TetR and TrpR. A similar design was previously shown to be functional for LacI and TetR repression activity in *in vitro* transcription and translation reactions,[92,93] while for TrpR there was no such data available. The reporter gene coded for the fluorescent protein mRFP1[94] to monitor translation levels and the spinach aptamer[95] that was placed at the 3'-UTR of the mRNA to monitor transcription levels. This aptamer was developed to bind a ligand designed to resemble the GFP chromophore: (Z)-4-(3,5-difluoro-4-hydroxybenzylidene)-1,2-dimethyl-1H-imidazol-5(4H)-one (DFHBI). This molecule free in solution has a very low fluorescence quantum yield that is strongly increased by the binding of the aptamer. Such a construct was previously demonstrated to be a good tool for the real-time monitoring of both transcription and translation levels in the PURE system.[36]

All of the genetic elements were amplified by polymerase chain reaction (PCR) in order to obtain linear fragments of double stranded DNA spanning from the promoter to the transcriptional terminators. The DNA fragments encoding the repressors were all tested at different concentrations in the PURE system in the presence and absence of the respective inducer molecule (Figure 2 - 2). TrpR showed a good transcriptional repression with no addition of the ligand tryptophan, possibly because what was contained in the reaction mix was enough to induce the activity of the repressor (Figure 2 - 2 A). The analog IPA though was not capable of derepression at the concentrations reported *in vivo*.[87] Reporter genes under the control of LacI and TetR (Figure 2 - 2 C, D) both showed a lower level of gene expression when compared with the reporter gene exploited in TrpR circuit, but while LacI could be derepressed, TetR could not in contrast to previously reported data.[93] Lac circuit showed the best level of switch between on and off states and was chosen for further characterization.
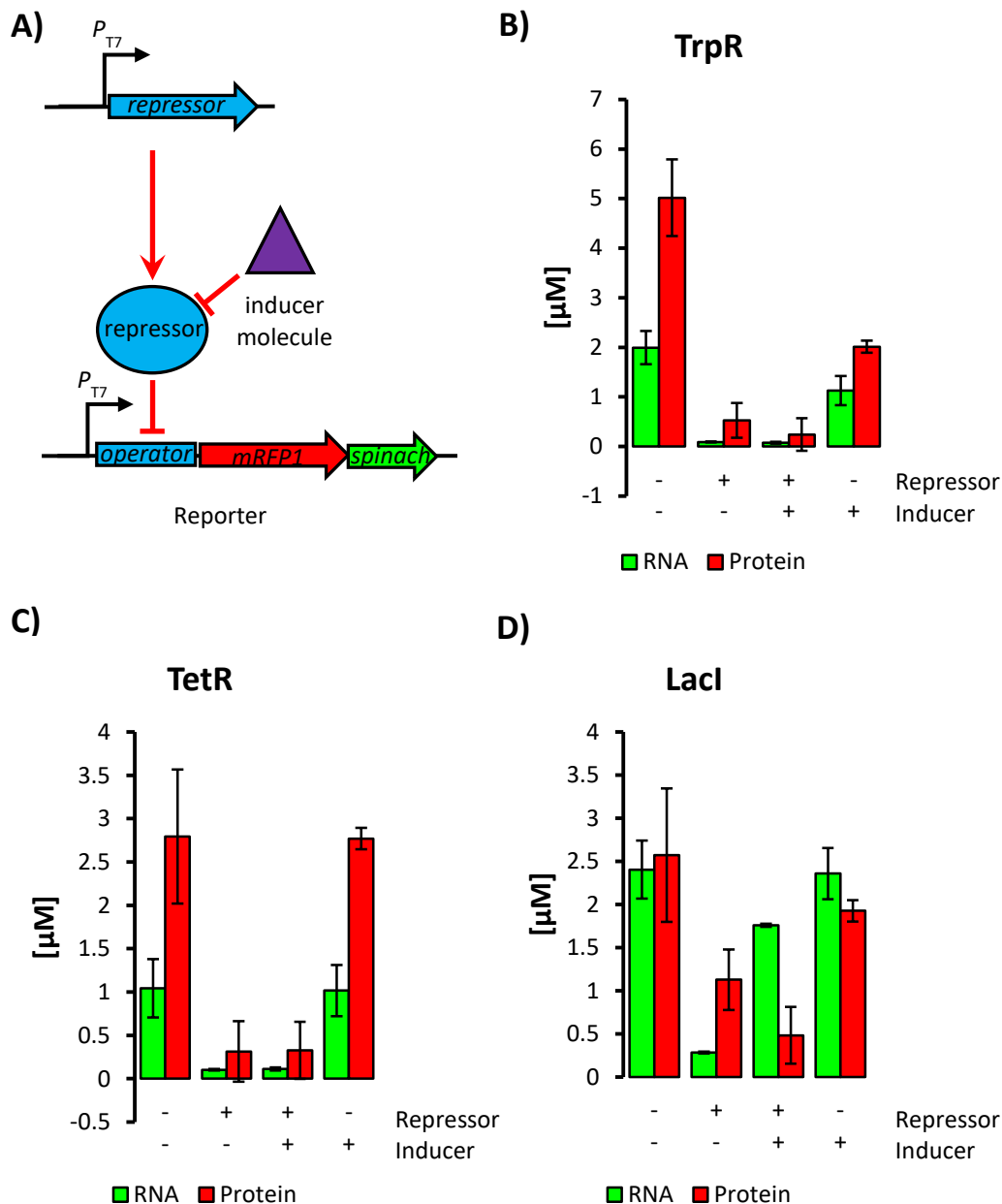
Figure 2 - 2. Test on transcriptional repressors LacI, TetR and TrpR. A) Repressors and reporter genes were encoded in two separate PCR fragments under the control of a T7 promoter. The constitutive expression of the repressor gene keeps the reporter gene in an off state, that is activated by the specific inducer molecule. B) - D) RNA (green bars) and protein (red bars) concentrations reached after 6 hours of incubations at 37 °C. The fluorescence values measured for spinach aptamer bound to DFHBI and mRFP1 were converted into molar concentration as described in § 2.1.8. The charts compare the expression levels of the reporter genes in the presence ("+") or absence ("-") of the DNA fragments encoding for the repressor and in presence or absence of inducer molecules, that were added to final concentrations indicated in § 2.1.6. All of the values reported are technical triplicates of each sample and the error bars indicates the standard deviation. Among the three repressors tested, LacI showed the best gene induction, although the difference between induced (repressor +, inducer +) and repressed (repressor +, inducer -) state was clear in transcription but not in translation. This result was thought to be caused by the ratio between repressor and reporter DNA, that slightly favors repressor gene expression (the exact ratio is 1:1.15).

2.2.2 Deep characterization of LacI circuit in PURE system and some attempts at improvement

In order to improve the difference between on and off states, the molar ratio between the two plasmids was gradually changed to favor the repressor rather than the reporter gene. As it can be observed in Figure 2 - 3, a 1:1 molar ratio was enough for proper gene derepression, and increasing amounts of repressor's DNA reduced the final yield of protein after induction. It is possible that most of the elements required for gene expression are directed towards the synthesis of the repressor more than the reporter gene.

As the 1:1 molar ratio showed a good difference between on and off states, the two genes were joined together in the same plasmid. Such a construct, rather than two separate linear fragments, would allow for an equal encapsulation efficiency of the two genes when used in artificial cells. Two plasmid designs were compared: one containing the two coding sequences under the control of two separate T7 promoters, while the other
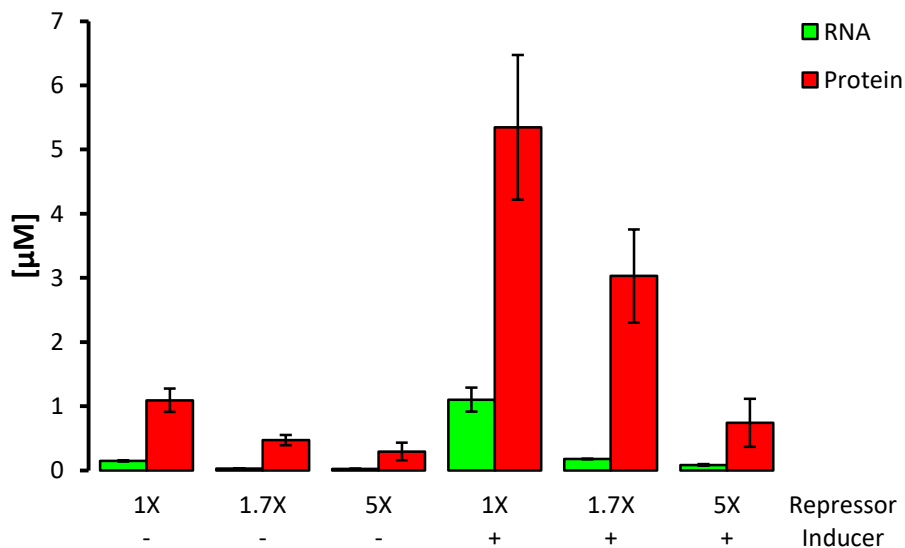
Figure 2 - 3. Effect of different molar ratios between the DNA fragment encoding the LacI repressor and the DNA fragment encoding the reporter gene under the control of lac operator. To validate whether the induction of gene expression could be improved by decreasing the amount of repressor, three different molar ratios of the DNA fragments carrying the *lacI* gene and the reporter gene were tested in the same genetic circuit described in Figure 2 – 2 A. The exact molar ratio of 1:1 between reporter and repressor DNA is enough to show a good difference between on and off state, both in RNA (green bars) and protein (red bars) levels. Increasing the molar ratio still provided activation but an excessive amount of repressor's DNA resulted detrimental to the reaction, possibly because of a sequestration of all the elements required for gene expression that are directed towards the synthesis of the repressor gene more than the reporter gene. Fluorescence data were analyzed as in Figure 2 – 2.
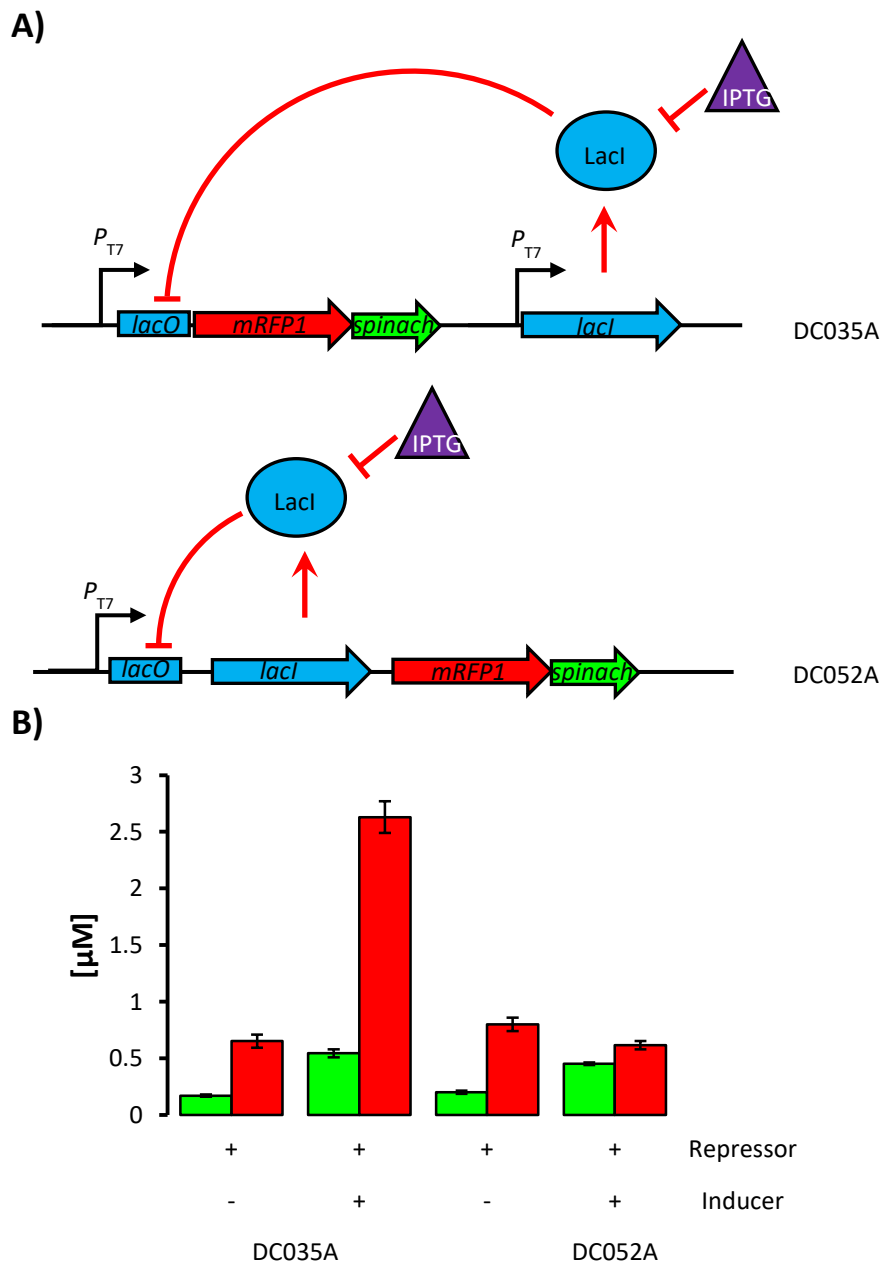
Figure 2 - 4. In single plasmid constructs gene induction occurred when reporter and repressor genes are under the control of two separate promoters. A) DC035A carried two genes, the repressor *lacI* and the reporter gene *mRFP1-spinach* encoded by two different promoters. Both genes were under the control of T7 RNA polymerase but the reporter gene contained also the operator sequence for LacI (*lacO*). DC052A instead carried both reporter and repressor gene under the control of the same T7 promoter and *lacO*; B) gene induction was observed both in transcription (green bars) and in translation levels (red bars) in the plasmid containing two separate promoters but not in the bicistronic construct where only transcription rates increased after the induction. These results correlate with the observed influences of gene position on its expression.[36] Fluorescence data were analyzed as in Figure 2 – 2.

was a bicistronic construct containing both repressor and reporter genes under the control of the same T7 promoter fused to a lac operator (Figure 2 - 4). While the bicistronic construct did not report any induction at the translational level, the plasmid designed with two promoters showed a good switch between on and off states. The behavior of the bicistronic levels can be explained with a recent characterization of gene expression in the PURE system.[36] From this recent work aimed at characterizing the efficacy of gene expression within an operon, it appeared clear how the position within the operon affects the final gene expression: coding sequences further from the transcription start site resulted in less efficient translation than coding sequences closer to the transcription start site. In the bicistronic design, the RNA levels monitored by spinach aptamer and DFHBI refer to the mRNA coding for both LacI repressor, in the first position within the operon, and mRFP1, in the second position. The increased signal observed for transcription levels in the presence of the inducer molecule did not correlate with translation levels of mRFP1 because of the effect observed in the PURE system related to gene position within an operon.

A final attempt was made to correct the problems associated with the leaky gene expression in the off state of the LacI genetic circuit with the PURE system. A second operator was thought to reduce the interaction of the promoter with the RNA polymerase and a mutant of LacI was reported to show higher differences between on and off state. Neither a second operator upstream of the T7 promoter nor the described mutation in the repressor[96] sequence served the intended purpose (Figure 2 - 5).

2.2.3 Genetic circuits based on T7 promoter showed a high background activation

Among the genetic circuits tested, the system based on LacI demonstrated to be tunable although the off state was never satisfactory. An explanation for this can be found again in the limitation of the PURE system highlighted from the abovementioned work from Chizzolini *et al.* A set of mutants of T7 promoters was tested for different abilities of gene expression, both at the transcriptional and at the translational level. From the data, it was clear how the regulation of a strong RNA polymerase such as T7 could not offer good tunability. In fact, the different strength of the T7 promoters tested, while showing a gradient in RNA levels, did not control well the protein levels, i.e. protein expression was either high or low without the possibility of achieving intermediate levels.[36] It is likely that the levels of LacI expressed were not enough to keep a proper off state in the absence of the respective inducer molecule. To verify this hypothesis, a repressor similar to the previously described was purified and tested for the effect on the expression of a reporter

gene similarly designed. The repressor EsaR involved in the quorum sensing mechanism of the plant pathogen *Pantoea stewartii* was taken into consideration because a detailed protocol for the purification with current technologies was available.[91] The activity of the
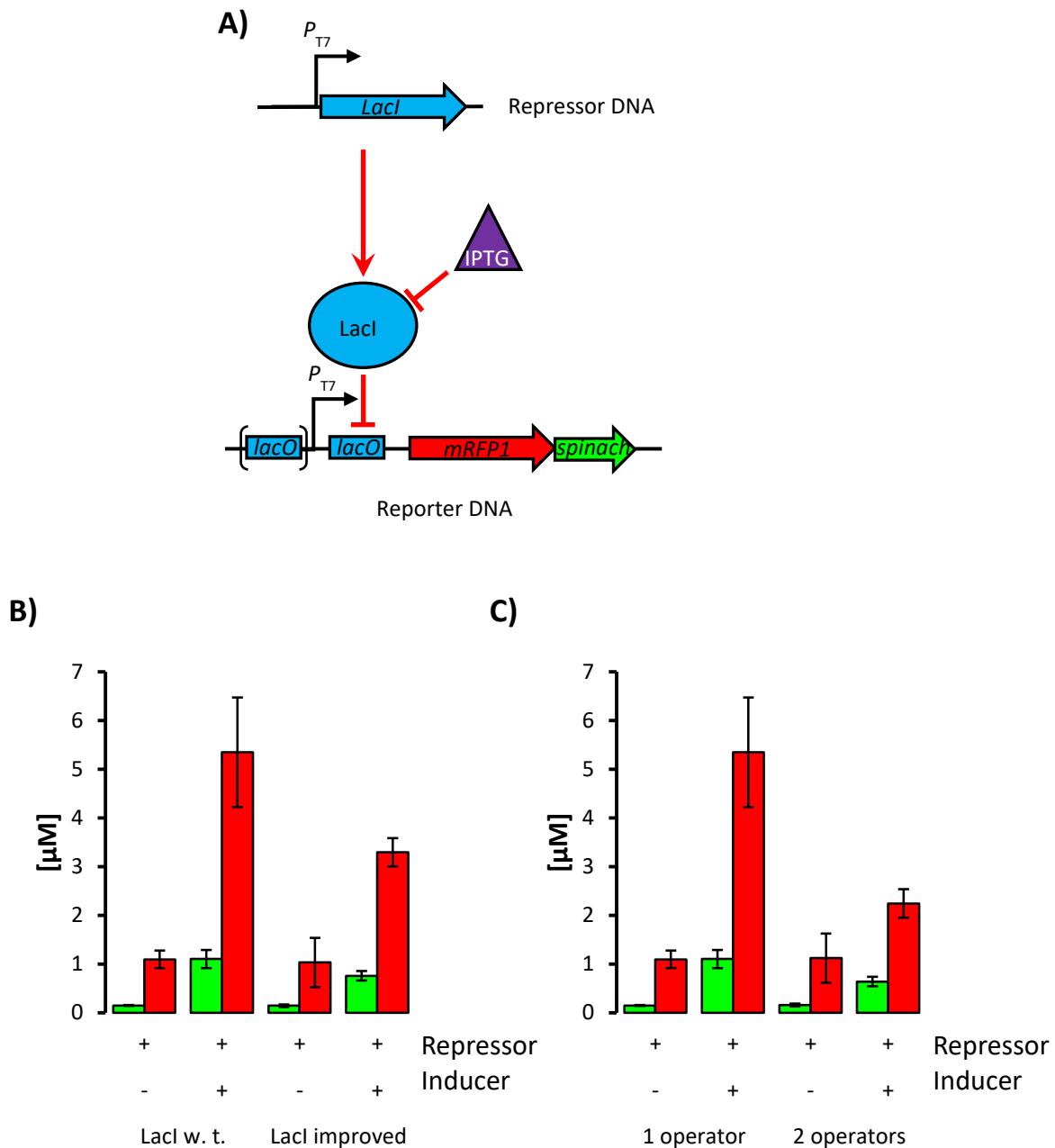
Figure 2 - 5. Trials for LacI circuit improvements. A) Similarly to the systems described earlier, a DNA fragment carried the gene coding for LacI or one improved version described in the literature under the constitutive expression of the T7 promoter, and the other fragment carried the reporter DNA under the control of T7 promoter and 1 or 2 lac operator sequences. The repressor constitutively expressed binds to the operator sequences preventing the reporter gene to be expressed. When the inducer molecule IPTG is present in solution, gene expression is activated. B) The addition of a second operator sequence close to the promoter showed a decrease in gene expression activation for LacI. C) A previously described mutant of the repressor LacI reported for improved activation *in vivo* was not showing the same behavior *in vitro*. Fluorescence data were analyzed as in Figure 2 – 2.
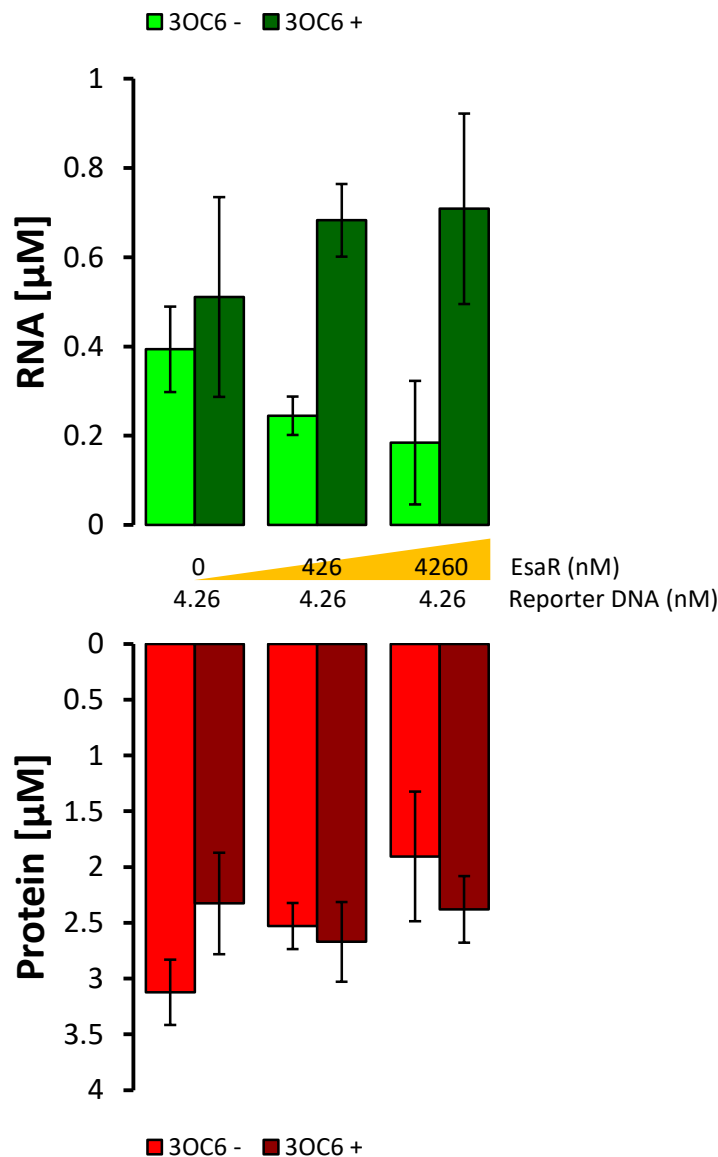
Figure 2 - 6. The purified EsaR repressor decrease transcription rates but not translation rates in the PURE system. Several concentrations of purified repressor were subjected to a PCR fragment carrying a reporter gene under the control of the promoter T7 fused to the operator sequence of EsaR. Increasing levels of repressor decreased RNA levels but not protein levels, confirming that T7 RNA polymerase cannot provide a fine regulation of gene expression *in vitro*. Fluorescence data were analyzed as in Figure 2 – 2.

repressor is analogous to TetR and LacI, therefore a reporter gene was designed under the control of T7 promoter fused to EsaR operator sequence and several concentrations of the purified repressor were tested for gene repression, in presence or absence of the inducer molecule *N*-(-3-oxohexanoyl)-L-homoserine lactone (3OC6 HSL). The purified repressor showed a good activity at the transcriptional level and the effect correlated with

increasing ratios between repressor protein and reporter DNA. Nonetheless, the levels of protein expression did not correlate accordingly, therefore confirming the strong limitations in modulating gene expression with the T7 promoter. Decreasing the strength of the promoter either by mutations or by the action of a transcriptional repressor is not sufficient for a fine regulation of gene expression (Figure 2 - 6).

2.2.4 A genetic circuit based on E. coli *promoters and the regulator LuxR was a valid alternative to T7-based genetic circuits*

To overcome the limitations of the T7 promoter, alternative circuits were considered for the regulation of gene expression. The commercially available PURE system lacks *E. coli* RNA polymerase so all of the constructs exploited were forced to use T7 promoters. S30 extracts, instead, not only provide *E. coli* RNA polymerases but also allows for the possibility to exploit different regulatory circuits. Being more similar to the cytoplasm of *E. coli*, the reaction can more reliably exploit many systems that are functional *in vivo* as it contains almost all of the factors involved.
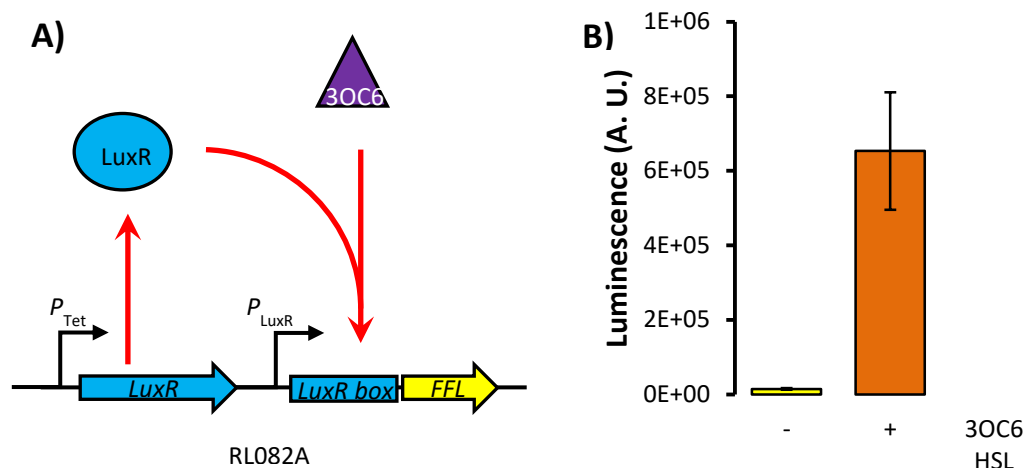
Figure 2 - 7. Gene regulation in LuxR/3OC6 HSL system tested with the reporter gene coding for the Firefly luciferase. A) The reporter gene Firefly Luciferase was put under the control of the promoter P_{LuxR} containing a LuxR box for the binding of the transcriptional activator LuxR in the plasmid RL082A. In this genetic circuit, the activator is constitutively expressed in an inactive form under the control of the promoter P_{tet} and is not able to bind the LuxR box until the inducer molecule 3OC6 HSL is present in solution. When bound to the cofactor, LuxR drives the binding of *E. coli* RNA polymerase and successive gene transcription. B) An S30 reaction was run with plasmid RL082A in the presence ("+") or absence ("-") of the inducer molecule 3OC6 HSL. Luciferase activity was measured after 4 h of incubation at 30 °C by the addition of the enzyme's substrate luciferin. Luminescence measurements (integration time: 1 s) showed no background activation in absence of 3OC6 HSL. The error bars represent the standard deviation between two technical replicates.

A valid alternative was then found in a genetic circuit based on the quorum sensing signaling pathway of *Vibrio fischeri*, that exploits the transcription factor LuxR. A genetic circuit was designed from a modification of the plasmid BBa_T9002 available from the registry of standard biological parts. This plasmid carries the LuxR coding sequence under the constitutive promoter pTet and a reporter gene under the control of the inducible promoter PLuxR. In the presence of the same inducer molecule interacting with EsaR (3OC6 HSL), the transcription factor binds to a sequence enclosed in PLuxR (LuxR box) and recruits RNA polymerase to start transcription (Figure 2 - 1 A). Conversely to the repressors described above, this protein behaves as a transcriptional activator, functional in the presence of the inducer molecule. The final regulation is the same described for EsaR, but the molecular mechanism is different: both genes regulated by EsaR and genes regulated by LuxR are expressed in presence of the inducer molecule 3OC6 but while EsaR is a transcriptional repressor that blocks the expression of the gene of interest in absence of 3OC6, LuxR is instead a transcriptional activator unable to start the transcription in absence of the inducer molecule.

This system was previously shown to be active both *in vivo* and *in vitro*,[67] and since no further characterization was required, a reporter gene alternative to mRFP1/spinach was chosen to control gene expression also within lipid vesicles. As discussed in the introduction, lipid vesicles are not extremely efficient in the encapsulation of an active TX/TL reaction. Therefore, the reporter gene should allow a great discrimination between the "on" and the "off" state of the circuit. Firefly luciferase seemed the best option, because of the high luminescent signal catalyzed by the enzyme compared to background activity (Figure 2 - 7). Given the good regulation observed and the fact that the inducer molecule could easily permeate the lipid membrane,[97] this genetic circuit was chosen for the building of artificial cells described in the following chapter.

Chapter 3

Towards artificial cells consortia

The first chapter discusses some of the technical limitations related to the interactions between artificial and natural cells showing how the need for the creation of communities of artificial cells is mainly linked to the necessity of an efficient use of available resources. This chapter summarizes the central aim of the thesis, that is to evaluate the feasibility of the creation of communities of artificial cells linked together in a communication network. To this end, the possibility to set communication pathways between different kinds of artificial cells was tested. To the end of complementing reciprocal defects, not only lipid vesicles were taken into consideration but also other kinds of artificial cells. A collaboration with Prof. Stephen Mann at the University of Bristol was aimed at evaluating the possibility of creating communication pathways between lipid vesicles carrying TX/TL reactions and proteinosomes with catalytic activity. These structures based on protein-polymer conjugates offer a semi-permeable membrane, whose pore size can vary according to the modified protein used and to all those conditions affecting protein folding, such as temperature, pH and ionic strength. Enzymes can perform catalytic reactions both on the membrane and in the lumen of these compartments.[6]
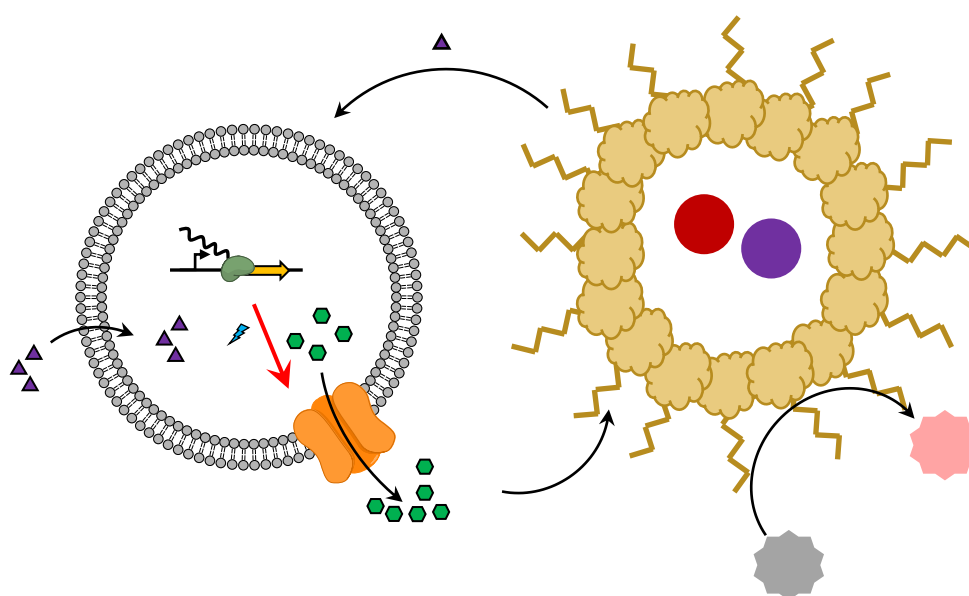
Figure 3 - 1. Schematic representation of the communication pathway between proteinosomes and liposomes. Glucose oxidase proteinosomes (on the right) contain two enzymes: horseradish peroxidase and LuxI (purple and red circles). LuxI catalyzes the synthesis of the inducer molecule 3OC6 HSL (purple triangles) that activates the expression of a pore-forming protein in a lipid vesicle (on the left). The pores will localize at the membrane of liposomes allowing the release of glucose (green hexagons) that will be subjected to an enzymatic reaction catalyzed by glucose oxidase and horseradish peroxidase (the conversion of the non-fluorescent molecule Amplex Red, grey shape, into the fluorescent molecule Resorufin, pink shape, described in details in Figure 3 - 2). The two artificial cell will communicate through the conditional exchange of two molecules, 3OC6 HSL and glucose.
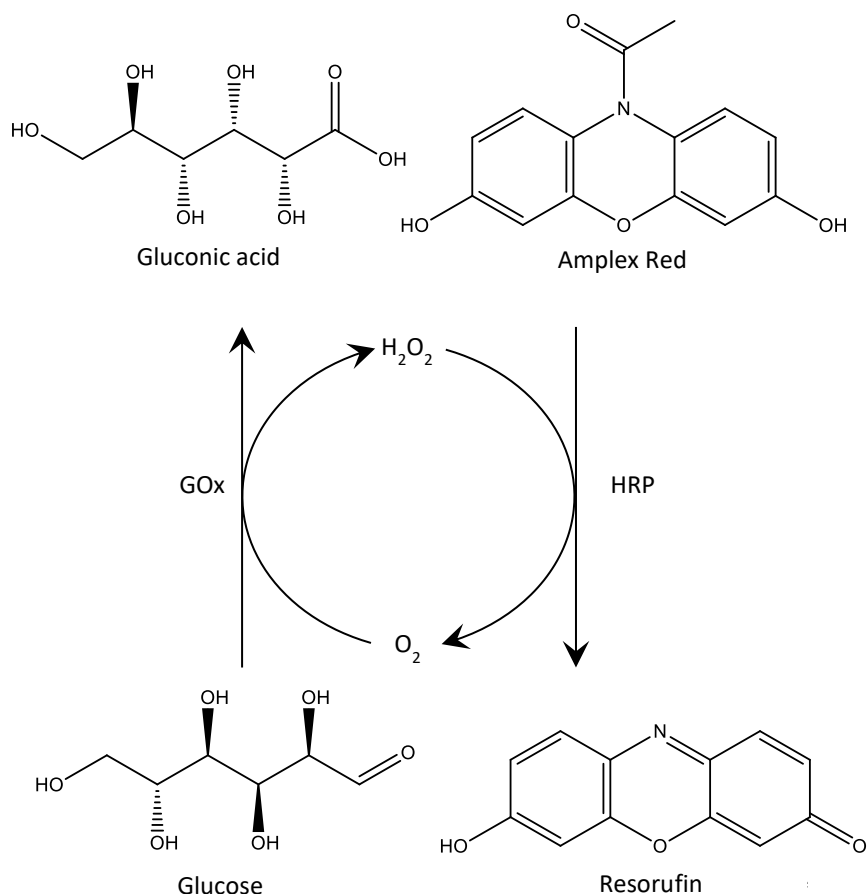
Figure 3 - 2. Enzymatic reactions catalyzed by the proteinosome. The hydrogen peroxide produced by glucose oxidase (GOx) serves as a substrate for horseradish peroxidase (HRP) to convert amplex red into resorufin. Conversely to amplex red, resorufin is fluorescent, therefore the kinetics of the reactions can be monitored by measuring the fluorescence signal of resorufin (excitation: 563 nm; emission: 587 nm).

Two molecules were mainly involved in this communication pathway: the quorum sensing molecule 3OC6 HSL described in the previous chapter for the ability to induce the expression of genes under the control of LuxR promoters and glucose that can serve as a substrate for the enzyme glucose oxidase, modified to create proteinosomes as described by Huang *et al.*[6] (Figure 3 - 1). Gene induction led by 3OC6 HSL would allow the formation of pores inside of liposomes thus releasing encapsulated glucose. Proteinosomes made of glucose oxidase and encapsulating horseradish peroxidase (HRP) would catalyze a reaction that converts Amplex (or Ampliflu) Red into the fluorescent molecule resorufin (Figure 3 - 2). The communication pathway could then be easily monitored through fluorescence. Further, the enzyme responsible for 3OC6 HSL synthesis could be trapped inside of glucose oxidase proteinosomes and used to trigger gene expression, thus creating a two ways communication pathway where proteinosomes send a signal to liposomes (3OC6 HSL) which will reply with another message (glucose).

In order to create proteinosomes, glucose oxidase was modified in Stephen Mann's laboratory to carry the hydrophobic polymer poly(*N*-isopropylacrylamide) (PNIPAAm). The resulting amphipathic macromolecules form Pickering emulsions around water-in-oil droplets. These structures are stabilized by covalent crosslinking for a stable transfer into water phase where they can be in contact with liposomes.

3.1 Materials and methods

3.1.1 Reagents and general supplies

Glucose oxidase from *Aspergillus niger*, Horseradish peroxidase, cholesterol, repel silane, EDTA, sepharose 4b were purchased from SIGMA; 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) was purchased from Avanti Polar Lipids; Amylose column was purchased from New England Biolabs; *E. coli* JM109 competent cells were purchased from Invitrogen; a protease inhibitor cocktail was purchased from Serva; ultrafiltration membrane filters made of regenerated cellulose with a 10 kDa cut-off were purchased from Millipore.

3.1.2 Instruments

T10 basic ULTRA-TURRAX disperser from IKA was used for the homogenization of lipid vesicles prepared with the FDEL method; Osmometer typM from Löser was used for osmolarity measurements; Amicon Stirred Ultrafiltration Cells from Millipore was used to concentrate protein solutions.

3.1.3 Plasmids and cloning

All the plasmids were assembled with the Gibson method[88] as described in the previous chapter. DNA primers and λ-holin sequence were synthesized at Eurofins MWG

Table 3 - 1. Primers used for PCR amplification.

Primer ID	Sequence (5'-3')
DC238	TAGAGTCACACAGGAAAGTACTAG ATGGATTCTGATATCAATATCAAA
DC239	GATGCCTGGCTCTAGTATTATTA TCAGTTGGTCATTTCTTCTT
DC263	AACAACAACCTCGGGATCGAGGGAAGGATTTCAGAATTCATGACTATAATGAT AAAAAAATCGGATT
DC265	CCCGAGGTTGTTGTTATTGTTATTGTTGTTGTTGTTCGAGCTCGAATTAGTCTG CGCG
DC271	ACTCTAGAGGATCCTCATTAATTTAAGACTGCTTTTTTAAACTG
DC272	TGAGGATCCTCTAGAGTCGACCTGCAGGC
T9002g fw	TAATAATACTAGAGCCAGGCATC
T9002g rev	CTAGTACTTTCCTGTGTGACTCTA

Table 3 - 2. Relevant sequences used in this section. The table reports synthetic genes or inserts of relevant plasmids whose information are not available on online catalogs.

Sequence ID	Sequence (5'-3')
αHL	ATGGATTCTGATATCAATATCAAAACCGGCACCACCGATATCGGCTCCAAT ACCACCGTTAAAACCGGTGATCTGGTGACCTATGATAAAGAAAACGGTAT GCATAAAAAAGTGTTTTACTCGTTTATTGACGATAAAAACCATAACAAAAAA CTGCTGGTCATCCGCACCAAAGGCACCATTGCGGGTCAATACCGTGTGTA CTCCGAAGAAGGTGCGAACAAAAGCGGTCTGGCTTGGCCGTCTGCCTTT AAAGTGCAGCTGCAACTGCCGGATAATGAAGTGGCGCAGATTTCAGATTA TTATCCGCGTAATAGCATCGATACCAAAGAATATATGAGTACCCTGACCTA TGGTTTTAATGGCAATGTTACCGGTGATGATACGGGTAAAATTGGCGGTC TGATTGGCGCCAATGTGTCCATTGGTCATACGCTGAAATACGTGCAACCG GATTTCAAAACCATTCTGGAAAGTCCGACCGATAAAAAAGTGGGGTTGGAA AGTTATCTTCAACAACATGGTGAATCAGAACTGGGGTCCGTACGATCGCG ATTCCTGGAATCCGGTTTATGGCAATCAGCTGTTTATGAAAACCCGCAAC GGTAGTATGAAAGCGGCGGATAATTTTCTGGACCCGAACAAAGCCTCAAG CCTGCTGTCCAGCGGTTTTAGCCCCGGATTTTGCCACGGTTATTACCATGG ATCGCAAAGCCAGCAAACAGCAGACCAACATTGATGTGATCTACGAACGT GTGCGTGATGATTATCAACTGCATTGGACCTCAACCAATTGGAAAGGCAC CAATACCAAAGATAAATGGACGGATCGCAGTTCAGAACGCTACAAAATTG ATTGGGAAAAAGAAGAAATGACCAACTGA
FFL	ATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTATCCTCT AGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATAC GCCCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAA CATCACGTACGCGGAATACTTCGAAATGTCCGTTCGGTTGGCAGAAGCTA TGAAACGATATGGGCTGAATACAAATCACAGAATCGTCGTATGCAGTGAA AACTCTCTTCAATTCTTTATGCCGGTGTTGGGCGCGTTATTTATCGGAGTT GCAGTTGCGCCCGCGAACGACATTTATAATGAACGTGAATTGCTCAACAG TATGAACATTTCGCAGCCTACCGTAGTGTTTGTTTCCAAAAAGGGGTTGCA AAAAATTTTGAACGTGCAAAAAAAATTACCAATAATCCAGAAAATTATTATC ATGGATTCTAAAACGGATTACCAGGGATTTCAGTCGATGTACACGTTCGT CACATCTCATCTACCTCCCGGTTTTAATGAATACGATTTTGTACCAGAGTC CTTTGATCGTGACAAAACAATTGCACTGATAATGAATTCCTCTGGATCTAC TGGGTTACCTAAGGGTGTGGCCCTTCCGCATAGAACTGCCTGCGTCAGAT TCTCGCATGCCAGAGATCCTATTTTTGGCAATCAAATCATTCCGGATACTG CGATTTTAAGTGTTGTTCCATTCCATCACGGTTTTGGAATGTTTACTACACT CGGATATTTGATATGTGGATTTCGAGTCGTCTTAATGTATAGATTTGAAGA AGAGCTGTTTTTACGATCCCTTCAGGATTACAAAATTCAAAGTGCGTTGCT AGTACCAACCCTATTTTCATTCTTCGCCAAAAGCACTCTGATTGACAAATA CGATTTATCTAATTTACACGAAATTGCTTCTGGGGGCGCACCTCTTTCGAA AGAAGTCGGGGAAGCGGTTGCAAAACGCTTCCATCTTCCAGGGATACGA

	CAAGGATATGGGCTCACTGAGACTACATCAGCTATTCTGATTACACCCGA
	GGGGGATGATAAACCGGGCGCGGTCGGTAAAGTTGTTCCATTTTTTGAAG
	CGAAGGTTGTGGATCTGGATACCGGGAAAACGCTGGGCGTTAATCAGAG
	AGGCGAATTATGTGTCAGAGGACCTATGATTATGTCCGGTTATGTAAACAA
	TCCGGAAGCGACCAACGCCTTGATTGACAAGGATGGATGGCTACATTCTG
	GAGACATAGCTTACTGGGACGAAGACGAACACTTCTTCATAGTTGACCGC
	TTGAAGTCTTTAATTAAATACAAAGGATATCAGGTGGCCCCCGCTGAATTG
	GAATCGATATTGTTACAACACCCCAACATCTTCGACGCGGGCGTGGCAGG
	TCTTCCCGACGATGACGCCGGTGAACTTCCCGCCGCCGTTGTTGTTTTGG
	AGCACGGAAAGACGATGACGGAAAAAGAGATCGTGGATTACGTCGCCAG
	TCAAGTAACAACCGCGAAAAAGTTGCGCGGAGGAGTTGTGTTTGTGGAC
	GAAGTACCGAAAGGTCTTACCGGAAAACTCGACGCAAGAAAAATCAGAGA
	GATCCTCATAAAGGCCAAGAAGGGCGGAAAGTCCAAATTGTAA

Table 3 - 3. Plasmids used in this section

Plasmid ID	Backbone	Insert	Source	Cloning strategy	
				Primers	Template
pET21b			Novagen		
pET21b_αHL C-ter	pET21b	T7_lacO_αHL::His	Mansy Lab		
pMAL-c4x		pTac_lacO_MBP	New England Biolabs		
RL078A	pSB1A3	pTet_LuxR_pLux_LuxI	Mansy Lab		
DC128A	pMAL-c4x	pTac_lacO_MBP::LuxI		DC263/ DC271	RL078A
				DC265/ DC272	pMAL-c4x
DC119A	pSB1A3	pTet_LuxR_pLux_αHL (α-hemolysin)		T9002g fw/T9002 g rev	RL078A
				DC238/ DC239	pET21b_ αHL C-ter
RL082A	pSB1A3	pTet_LuxR_pLux_FFL (firefly luciferase)	Mansy Lab		

3.1.4 LuxI purification

The enzyme LuxI was purified following the indications by Schaefer *et al.*[98] with some modifications. The gene was cloned in pMAL-c4x and *E. coli* JM109 were transformed with the plasmid for expression and purification. A protease inhibitor cocktail was added according to manufacturer's instructions and glycerol was removed from the column wash buffer to allow the liquid to pass through the column faster. The protein was dialyzed against a buffer containing 20 mM sodium phosphate pH 7.4, 0.1 mM EDTA, 100 mM NaCl, 1 µM DTT. The solution was concentrated with an Amicon stirred cell with cellulose filter with a pore size of 10 kDa, then supplemented with 2% DMSO to allow the use in S30 reaction as suggested by Sun *et al.*[37] and stored at -80 °C in 100 µl aliquots.

3.1.5 Preparation of vesicles with the FDEL method

POPC and cholesterol were dissolved in chloroform to 10 mg/ml and the needed amount was transferred into a glass round bottom flask to allow chloroform evaporation through rotary evaporation. The lipid mixtures were resuspended in DEPC-treated water to a final concentration of 24.5 mM with a molar ratio between POPC and cholesterol of 1:2. Liposome mixtures were successively homogenized for 1 min at power 4 with T10 basic ULTRA-TURRAX disperser. The solution was split into 100 µl aliquots in 2 ml tubes, flash-frozen in liquid nitrogen and dehydrated overnight at rotavapor. Aliquots were stored at -20°C and each of them was thawed on ice prior to use.

3.1.6 Permeability test for glucose

POPC:cholesterol 1:2 (24.5 mM) vesicles were prepared with the FDEL method and resuspended with 10 mM calcein dissolved in HEPES buffer (50mM HEPES, 10mM MgCl$_2$, 100mM KCl, pH 7.6) in the round bottom flask after evaporation with a rotary evaporator. Vesicles were then extruded to 100 nm and fractionated with sepharose 4b column. To 50 µl of vesicles, 1 vol of 1 M glucose was added and kinetics were followed at the fluorimeter with the following parameters: excitation 495 nm, emission 515 nm, 37 °C, acquisition every 5 min for 13 h. Triton X-100 was later added to a final concentration of 0.3% (vol/vol) to verify the integrity of vesicles over time: the increase in fluorescence observed correlated to the break of vesicles, indicating that before the addition of the detergent, vesicles were intact and the data collected are reliable.

3.1.7 In vitro gene expression with home-made S30 *E. coli* extract

In vitro transcription/translation reactions based on S30 crude extract were prepared as described in the previous chapter with the following modifications: DNA was added to a final concentration of 25 nM, 100 mM glucose was added when needed and 10 µM 3OC6 HSL was added both from the outside or together in a non-encapsulated S30 reaction. The reactions were incubated at 30 °C for 4 h.

3.1.8 Pore formation tests with calcein

POPC:cholesterol 1:2 (24.5 mM) vesicles were prepared with the FDEL method and resuspended in 50 µl of 80 mM calcein dissolved in HEPES buffer and vortexed for 1 min. Vesicles were transferred to a 2 ml tube filled with HEPES buffer and spun at 6000 g for 5 min, the buffer was replaced and this step was repeated 3 additional times until the buffer appeared clear and the pellet of vesicles was finally resuspended in 500 µl of HEPES buffer. Pore formation was tested on a 384-wells plate by incubating 18 µl of vesicles with 2 µl of S30 reaction or controls as indicated in the results section. The S30 reaction was assembled as in § 3.1.7 and incubated at 30°C for 4 h prior to use in combination with calcein-loaded vesicles. Fluorescence levels were measured at Tecan Infinite M-200 with the following parameters: excitation 485 ± 9 nm, emission 515 ± 20 nm, gain 50, room temperature.

3.1.9 Pore formation tests with Amplex red

Vesicles were prepared with 100 mM glucose with the same procedure described in the previous paragraph. A 25X enzymatic mixture was prepared to contain 250 µM amplex red, 5 U/ml HRP and 50 U/ml GOx. Pore formation was tested in a 384-wells plate by incubating 21.5 µl of vesicles with 2.5 µl of 6 mg/ml α-hemolysin or HEPES buffer and 1 µl of the 25X enzymatic mixture. Fluorescence levels were measured at Tecan Infinite M-200 with the following parameters: excitation 560 ± 9 nm, emission 595 ± 20 nm, gain 50, room temperature. Raw fluorescent data were converted by means of a standard curve created with 1X enzymatic mixture and different concentrations of glucose (0, 1E-08, 1E-07, 1E-06, 1E-05 M)

3.1.10 Gene expression inside lipid vesicles prepared with FDEL method

Three aliquots of vesicles prepared with FDEL method were resuspended with a total of 150 µl of S30 reaction supplemented with 20 µM GFPmut3B, previously expressed and purified via Ni-NTA column as described for EsaR repressor. Vesicles were extruded 11 times with three membranes cut-offs (0.2, 1 and 3 µm) and then loaded on a sepharose 4b column equilibrated with HEPES buffer, used also for the elution and fractionation of vesicles. GFPmut3B fluorescence of the fractions was measured at Tecan Infinite M-200 with the following parameters: excitation 480 nm, emission 511 nm, gain 50, room temperature. The first fluorescent fractions, that contained vesicles, were separated from the others. Vesicles were then incubated at 30°C and 50 µl aliquots were collected every hour to measure glucose release with amplex red enzymatic assay as described earlier.

3.1.11 Gene expression inside lipid vesicles prepared with Pautot's method

A modification of the protocol developed by Hadorn[47,99,100] was used for the formation of vesicles. All the glassware and plasticware were incubated with repel silane for 10 min at room temperature to avoid nonspecific interactions with vesicles. For the same reason, 96-well plates with U-shaped bottom were additionally treated with salmon sperm DNA and dry milk for 10 min. A solution of POPC and a solution of cholesterol both dispersed in chloroform were added to a round bottom glass flask. Chloroform was evaporated under a stream of nitrogen and lipid mixtures were resuspended in mineral oil to a final concentration of 200 µM, then sonicated in a warm bath at 50 °C and incubated overnight at room temperature in the dark.

In a 96-well plate, 100 µl of a solution made with 0.88 M alanine pH 8 ("hosting solution") was pipetted at the bottom of the well, and 50 µl of a lipid mixture was added on top and phospholipids were allowed to sediment at the interphase for 2 h at room temperature. An S30 reaction was assembled with the addition of 180 mM sucrose and combined with lipid mixture to 2% (vol/vol) to create a water-in-oil emulsion by mechanical agitation. The emulsion was poured on top of the well and the vesicles were formed by spinning the plate at 1500 g for 3 min at 4 °C. The oil phase was removed by aspiration and the vesicles were washed once with the hosting solution and spun with the same conditions. 10 µl of vesicles were resuspended and loaded on a microscope glass slide, sealed with nail polish and incubated at 37 °C for 3 h before observation by fluorescence microscopy.

3.2 Results

S- adenosylmethionine

3-Oxohexanoyl-ACP

ACP

S-methyl-5'-thioadenosine

N-(3-Oxohexanoyl)-L-homoserine lactone
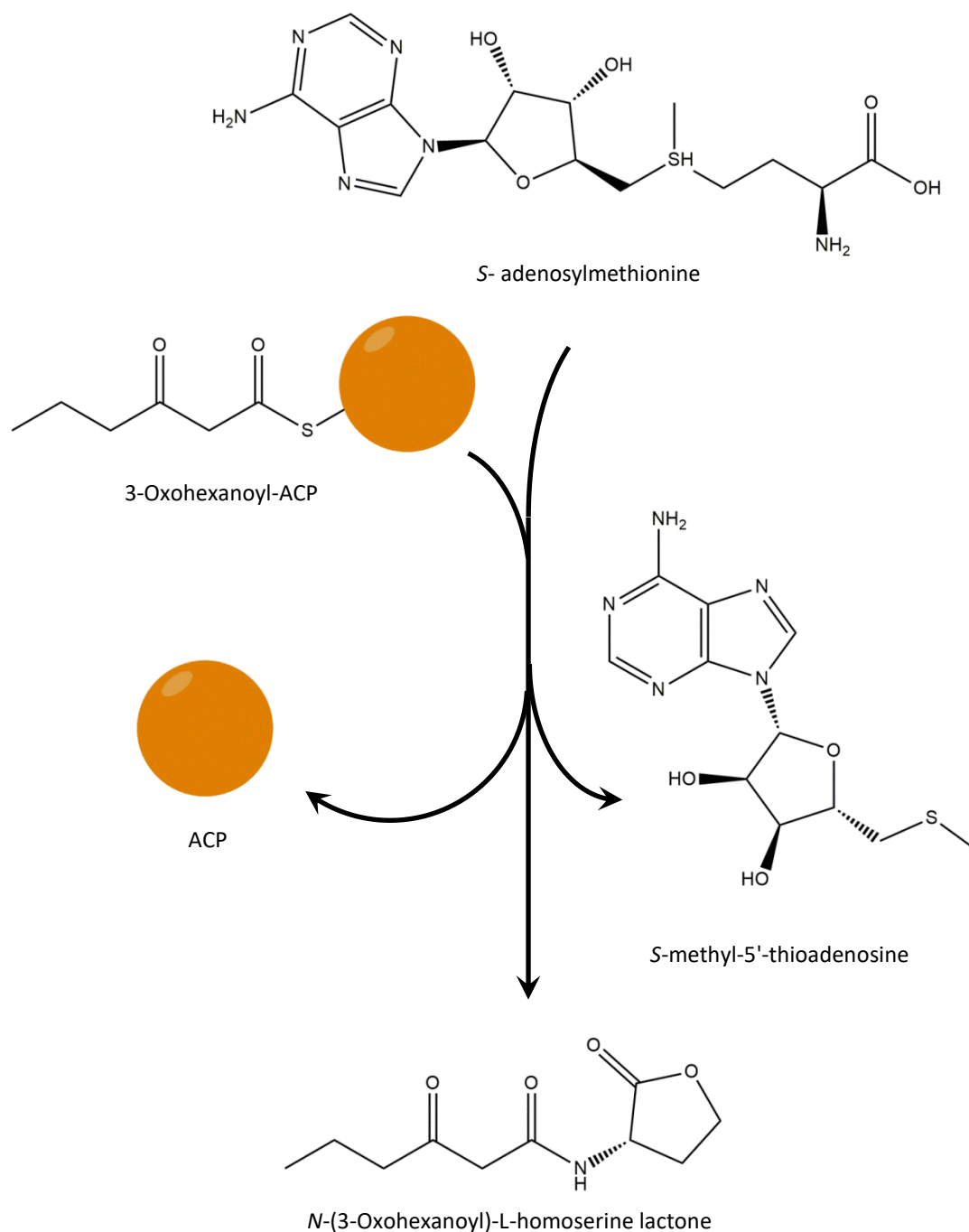
Figure 3 - 3. Enzymatic reactions catalyzed by LuxI. The quorum signaling molecules acyl-homoserine lactones are produced by specific synthases that form a covalent bond between a lactone ring derived from *S*-adenosylmethionine and an acyl chain transferred from acyl carrier protein (ACP). LuxI from *V. fischeri* catalyzes the synthesis of *N*-(3-Oxohexanoyl)-L-homoserine lactone.

In order to assess the feasibility of the designed communication pathway, the two directions were analyzed separately. The first analysis focused on proteinosomes sending a message to liposomes, in a mechanism where proteinosomes trigger gene expression

in liposomes through the production of 3OC6 HSL. The second mechanism foresees the delivery of glucose from liposomes to proteinosomes to induce catalytic activity.

3.2.1 LuxI purification and test for activity

The enzyme responsible for the synthesis of the quorum sensing molecule, LuxI, was purified and assayed *in vitro* with an S30 reaction for the induction of gene expression. The enzyme is classified as an acyl-homoserine lactone synthase and its substrates are S-adenosylmethionine (SAM) and an acyl carrier protein (ACP) loaded with a six carbon hydrocarbon (3-oxohexanoyl-ACP) (Figure 3 - 3).[98,101] Acyl carrier proteins are involved in the fatty acid synthetic pathway and the loading of the proper acyl group requires the precursor acetyl-CoA.[102] In order to test the activity of the purified enzyme, an S30 reaction was supplemented with acetyl-CoA and S-adenosylmethionine. The reporter gene circuit exploited was the previously described modification of BBa_T9002 shown to express firefly luciferase in the presence of the inducer molecule 3OC6 HSL.

0	0	0.01	0	[3OC6 HSL]
4.62	4.62	0	0	[MBP-LuxI]
0.01	0	0	0	[Acetyl-CoA]
0.3	0	0	0	[SAM]
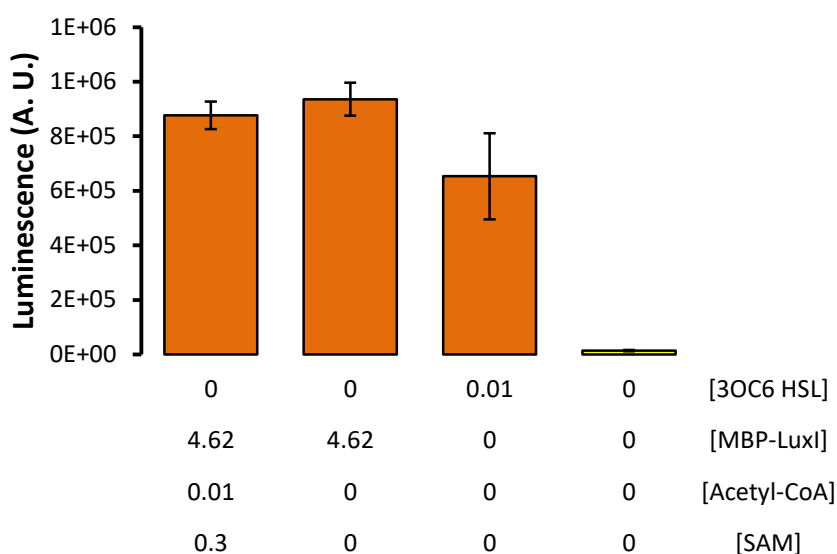
Figure 3 - 4. Luciferase expression in the S30 reaction under the control of LuxR promoter induced by a purified enzyme MBP-LuxI. The induction of the Firefly luciferase gene was tested in an S30 reaction assembled with plasmid RL082A and the elements indicated in the chart (values are reported in mM). After 4 hours of incubation at 30°C, the reaction was supplemented with the substrate of luciferase and the luminescent signal was measured with an integration time of 1 s. The error bars represent the standard deviation between technical duplicates. The enzyme is the only factor required for the production of the inducer molecule 3OC6 HSL and shows good gene induction, comparable to what was observed when pure 3OC6 HSL is added to the solution. The addition of the precursor molecules acetyl-CoA and *S*-adenosylmethionine (SAM) is not necessary for the induction of gene expression, despite previous reports.

Consistently to the observations of Schaefer *et al.,*[98] the enzyme was very active even if the fused maltose binding protein (MBP) that was used for purification was not cleaved off. The fact that the enzyme retained the catalytic activity despite the presence of the fusion tag may result helpful for the stability of the system as the dimensions of the protein will forbid the passage across the porous membrane of proteinosomes. Conversely to what was reported for the LuxI homolog in *Pseudomonas aeruginosa,*[103] the enzyme was active in *E. coli* extract also in absence of the precursors of the acyl-homoserine lactone. The *E. coli* extract seemed to retain enough enzymatic activity to both synthesize S-adenosylmethionine and to load an acyl carrier protein with a six carbons acyl group (Figure 3 - 4).

3.2.2 Glucose permeability
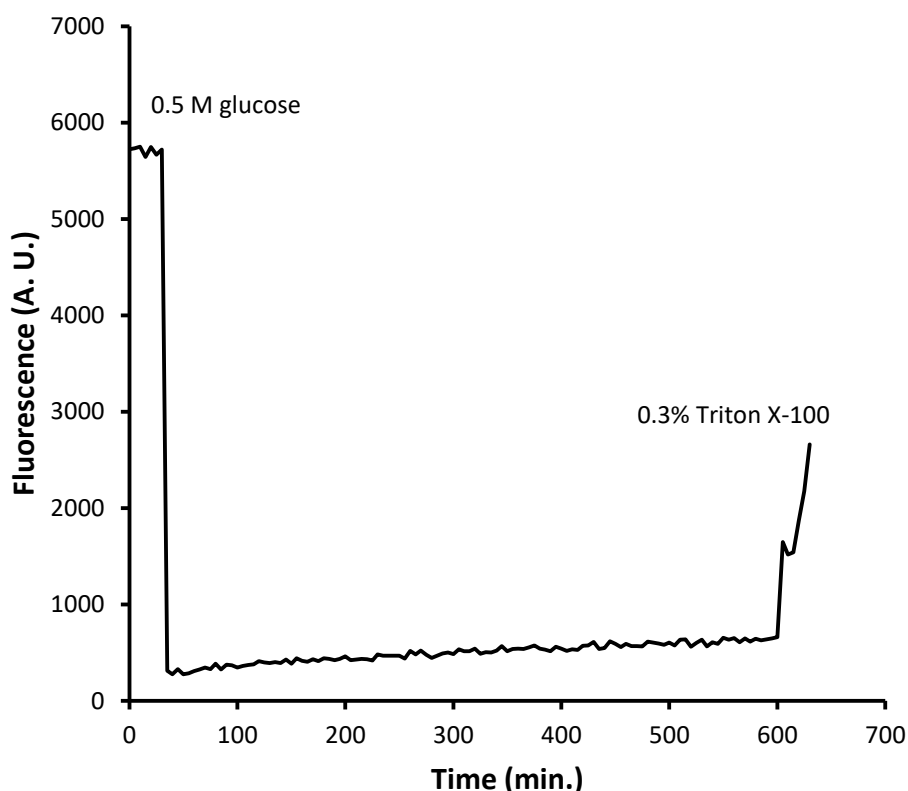
Figure 3 - 5. Permeability test of POPC:cholesterol vesicles to glucose. The addition of glucose induces the shrinkage of vesicles and a consequent drop in fluorescence of encapsulated calcein. The chart reports the data on the fluorescence values of calcein (excitation: 485 ± 9 nm; emission: 515 ± 20 nm) and the error bars represent the standard deviation between technical duplicates. No fluorescence restoration due to glucose permeation and balance of osmotic pressure was observed after 13 h. The detergent Triton X-100 was finally added to verify the integrity of vesicles over time.

Together with other sugars, glucose is well-known to have a very low permeability to lipid membranes containing POPC, to the point that it can be used to evaluate the change in membrane permeability induced by pores formed by synthetic polymers[104,105] or
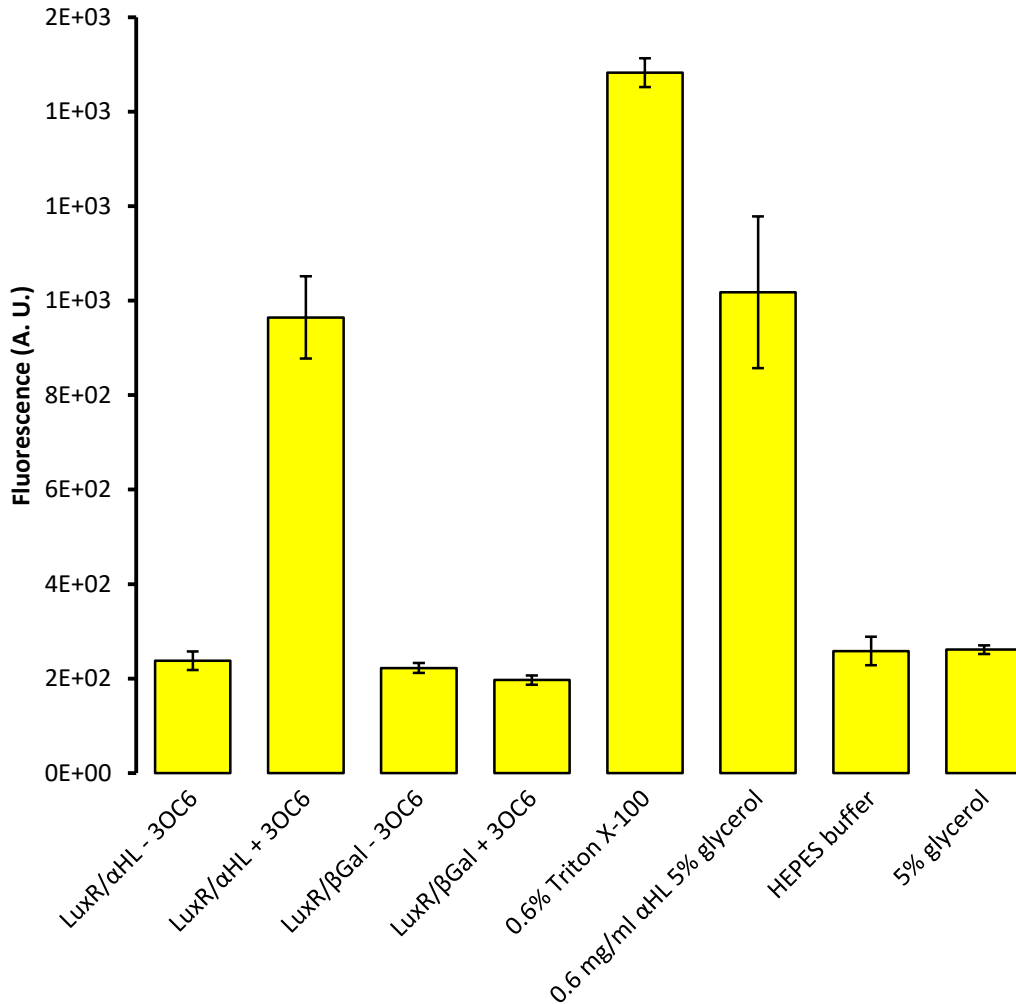
Figure 3 - 6. Calcein release by means of αHL gene induction. Vesicles loaded with 80 mM calcein were incubated together with S30 reactions or Triton X-100, purified α-hemolysin (αHL), HEPES buffer (50 mM HEPES, 10 mM MgCl$_2$, 100 mM KCl, pH 7.6), 5% glycerol. S30 reactions contained plasmid DNA carrying LuxR gene constitutively expressed and α-hemolysin (α HL) or β-galactosidase (βGal) genes under the control of LuxR promoter. Prior to incubation with vesicles, S30 reactions were incubated at 30°C for 4 h. The chart reports the data on the fluorescence of calcein (excitation: 485 ± 9 nm; emission: 515 ± 20 nm) and the error bars represent the standard deviation between technical duplicates. Only when α-hemolysin was expressed, calcein was released outside of the vesicles and an increase of fluorescence was observed because of a dilution effect.

small peptides.[106,107] Nonetheless a quick control on the permeability across liposomes was performed prior to proceeding with the experiments aimed at establishing communication pathways between liposomes and proteinosomes. To assess the retention of glucose inside of the lipid vesicles, a shrink-swell assay was performed with liposomes

loaded with 10 mM calcein.[108] The assay is based on the ability of calcein to self-quench at increasing concentrations so that the shrinkage caused by a hypertonic solution decreased fluorescence levels. A permeable molecule would re-equilibrate the osmotic pressure resulting in an increase of fluorescence. As reported in Figure 3 - 5, POPC:cholesterol vesicles did not allow the passage of glucose, because no fluorescence increase was observed after 13 h of incubation.

3.2.3 Pore formation tests on liposomes

In the conceived pathway, the LuxR/3OC6 HSL circuit would be responsible for the activation of expression of a pore-forming protein that should induce the release of glucose from the vesicle lumen. Releasing a small molecule such as glucose would not require a large pore, so alpha-hemolysin (αHL) from *Staphylococcus aureus,* with a pore size of ca. 1.5 nm,[109] can serve this purpose, as it has been previously shown for the release of a molecule of similar dimensions such as IPTG.[1]
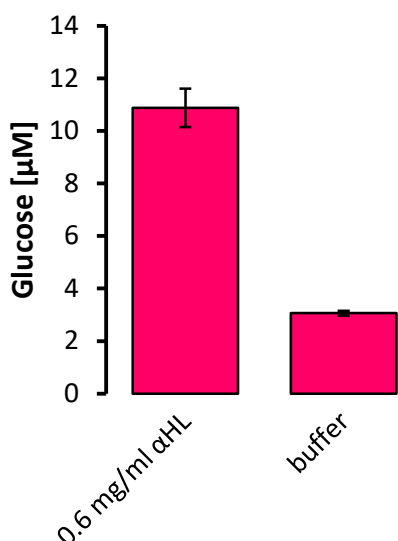
Figure 3 - 7. Glucose release by passive diffusion through the αHL pore. Vesicles loaded with 100 mM glucose were purified through sepharose column and mixed with α-hemolysin. The passage of glucose outside of lipid vesicles through α-hemolysin pores was verified with the enzymatic reaction described in Figure 3 - 2 (raw fluorescence data are converted to the concentration of glucose detected by means of the standard curve described in § 3.1.9). The error bars referred to the standard deviation of technical duplicates. The results reported here show how the pores created by the protein allows the free diffusion of glucose and the consecutive enzymatic activity.

In order to test the ability of S30 reactions to synthesize a functional pore-forming protein, the coding sequence of α-hemolysin was cloned in the genetic circuit based on LuxR/3OC6 HSL system. The test was performed by incubating the proteins produced in S30 reactions with calcein-loaded liposomes for 5 min. Liposomes loaded with 80 mM

calcein are largely self-quenched, but if the membrane is destabilized by the action of a detergent or pore-forming proteins, calcein would be diluted in the external solution thus leading to increased fluorescence.[110] As reported in Figure 3 - 6, the S30 reaction was able to produce a functional pore-forming protein and cause calcein release only in the on state. To further verify the passage of glucose through α-hemolysin pores, POPC:cholesterol vesicles loaded with 100 mM glucose were incubated with purified α-hemolysin and the levels of glucose were verified with the enzymatic reactions catalyzed by glucose oxidase and horseradish peroxidase in the presence of amplex red, with the same reaction performed by proteinosome (Figure 3 - 7).

3.2.4 Glucose release from liposomes

The major issue faced with the project was related to the removal of unencapsulated material. The FDEL method to create lipid vesicles does not allow for a complete encapsulation of the desired solution. The encapsulation of the TX/TL reaction relies on the resuspension of a previously dried lipid film, therefore part of the reaction mix may not be included within liposomes. Although gene expression can be silenced externally by the addition of DNase and RNase, glucose cannot be removed similarly. It was previously demonstrated how it is possible to remove unencapsulated PURE system by dialysis,[1] but we also observed a dramatic decrease in gene expression for the S30 reaction during the time required for dialysis. Therefore, a faster alternative was investigated. Size-exclusion chromatography was chosen as a first trial: because of their large size, vesicles are eluted earlier than the non-encapsulated reaction mixture, including glucose, that is instead retained in the column for a longer time. To help the collection of the proper fraction and exclude all of the unencapsulated solution, GFP was added to the S30 reaction to be used as a fluorescent marker, and vesicles were extruded before loading on a sepharose column. The extrusion step ensures a homogeneous size distribution to allow a good separation of vesicles on sepharose column. The induction of gene expression of α-hemolysin did not show any specific glucose release, monitored through the reaction catalyzed by glucose oxidase and horseradish peroxidase (Figure 3 - 8). It is possible that the procedure was affecting the efficacy of the S30 reaction so that the number of active vesicles was not enough to induce a proper level of glucose release and overcome the fluorescence background signal.

An alternative to the FDEL method, also the method developed by Pautot was taken into consideration. Vesicles formed in this way are theoretically free of the non-encapsulated solution and do not require any purification step. Nonetheless, some of the

Figure 3 - 8. Induction of glucose release through αHL pores with lipid vesicles prepared with FDEL method. An S30 reaction assembled with glucose and the plasmid carrying the genetic circuit LuxR/αHL was encapsulated in POPC:cholesterol vesicles prepared with FDEL method. A control encapsulating glucose in HEPES buffer was added. Vesicles were extruded to 0.2 (A), 1 (B) and 3 μm (C). Half of the vesicles were incubated with 10 μM of the inducer molecule 3OC6 and glucose release through αHL pores was measured through the enzymatic reaction described in Figure 3 - 2 over time. In none of the three cases, a specific glucose release was detected.

vesicles may pop during the formation process and some oil has to be removed from the top of the solution. To allow vesicle formation by centrifugation, the density between internal and external solution had to be increased while keeping an osmotic equilibrium. To address this purpose, sucrose was supplemented to the S30 reaction and 0.88 M alanine pH 8.0 was used externally. Alanine was chosen as a good alternative to glucose that was previously described to serve this purpose.[47,100] Being the molecule responsible for signaling, glucose could not be used as the external solution for vesicles formation. Alanine is mostly zwitterionic at pH conditions used in the S30 reaction so less likely to interfere with the polar heads of phospholipids, and being smaller than glucose it increases more the difference in density between inner and outer solution.
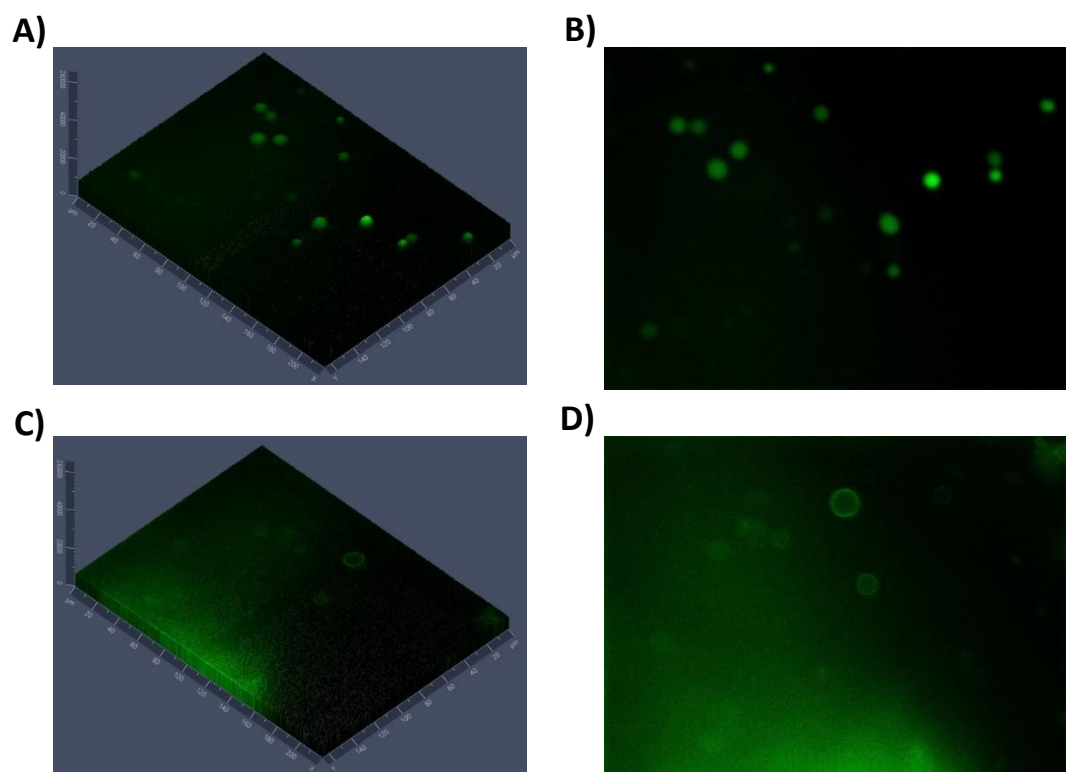
Figure 3 - 9. Gene expression in POPC vesicles prepared with Pautot's method. Ypet fluorescent protein was expressed under the control of a T7 promoter in an S30 reaction (A and B) but no fluorescence was observed in absence of template DNA (C and D). A) and C) histogram charts of the fluorescence observed respectively in B) and D).

To assess the efficacy of the S30 reaction in vesicles prepared with this method, the production of a Ypet fluorescent protein was tested inside of liposomes encapsulating a reaction mixture assembled with a plasmid DNA encoding the protein under T7 promoter and the purified T7 RNA polymerase. As observed under the microscope, Ypet could be expressed very well in this system (Figure 3 - 9 B). Nonetheless, measuring the release of glucose from such a system was not possible due to the background detected

(Figure 3 - 10). It can be presumed that liposomes are not stable enough and may pop, thus releasing their content.
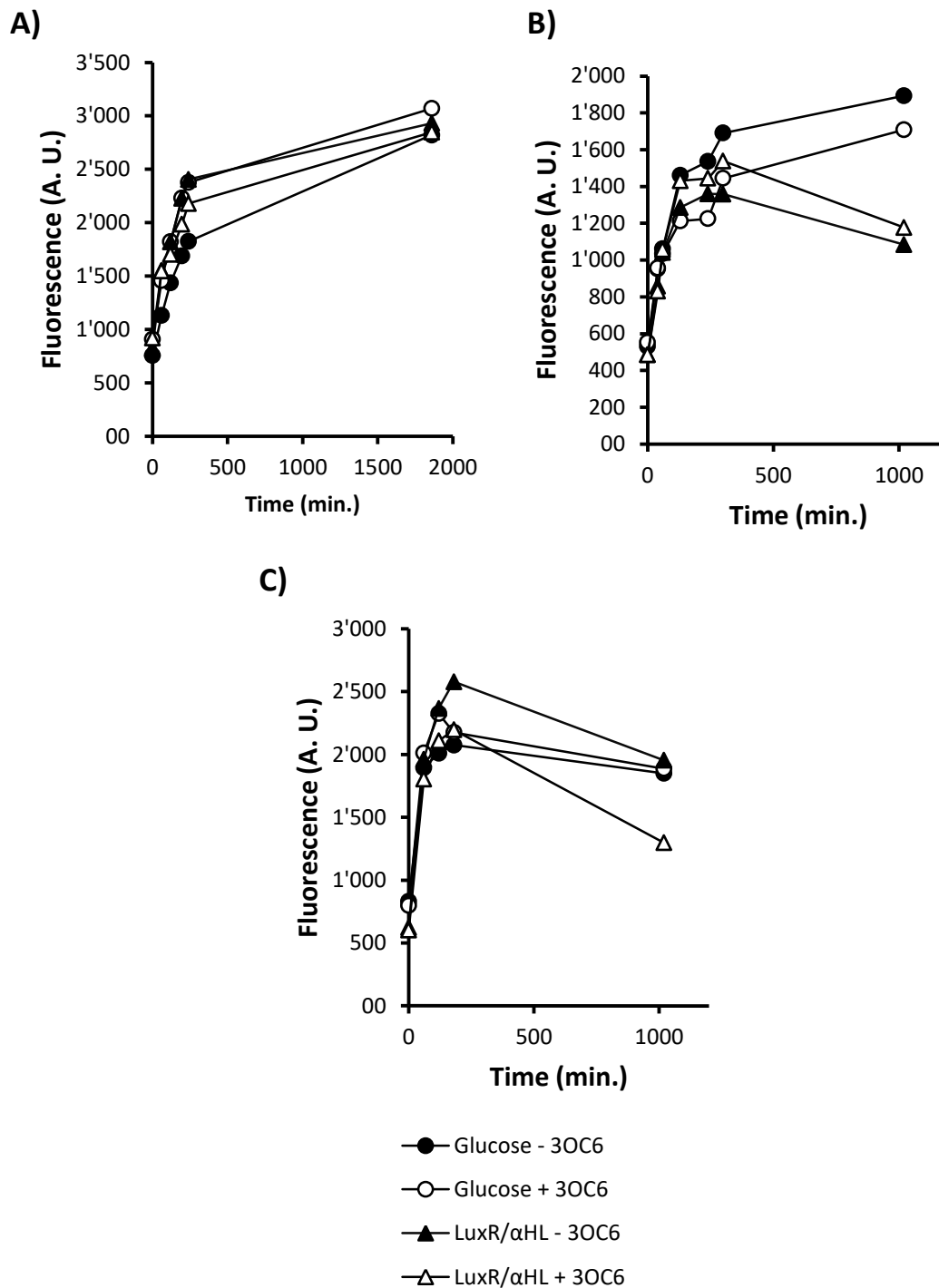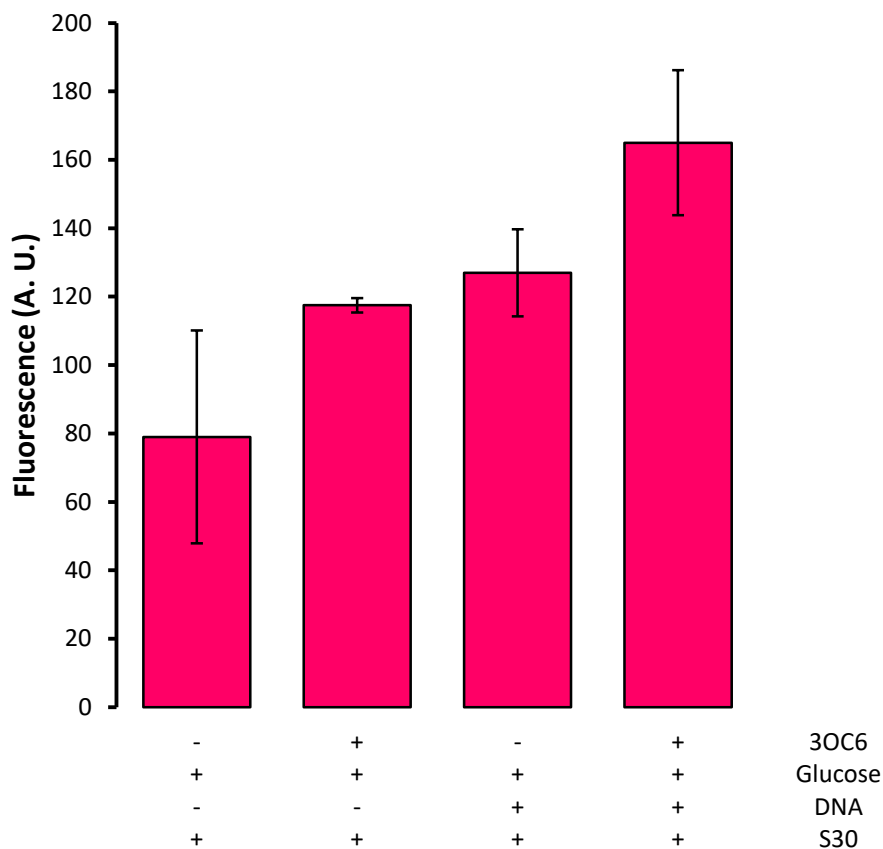

Figure 3 - 10. Induction of glucose release through αHL pores with lipid vesicles prepared with Pautot's method. An S30 reaction containing glucose and the plasmid carrying the genetic circuit LuxR/αHL was used to create a w/o emulsion in a mixture of POPC dispersed in mineral oil and create vesicles. A control without DNA was included. The inducer molecule was added from the outside and glucose release was measured after 3 h of incubation at 30 °C by means of resorufin fluorescence (excitation: 560 ± 9 nm; emission: 595 ± 20 nm) produced in the enzymatic reaction described in Figure 3 - 2. The error bars refer to standard deviation between technical duplicates. Again, no particular difference in fluorescence was observed for all the samples tested.

3.2.5 A consortium involving bacteria

Considering all the problems related to the stability of liposomes, a new communication pathway was taken into consideration (Figure 3 - 11). In this new design, liposomes containing S30 reaction should not undergo any purification step that may destabilize either the reaction or the membrane. The release of the communication molecule would be entrusted to bacteria rather than artificial cells and two different kinds of liposomes would be exploited. One liposome would contain the S30 reaction responsible for the production of a trigger molecule that activates the release of the chemical message contained in another population of liposomes. The trigger molecule

would activate the production of a pore-forming protein in bacteria by means of a genetic switch similar to LuxR/3OC6 HSL, and the protein would bind the membrane of glucose-loaded liposomes causing the release of the cargo by passive diffusion. Keeping the chemical message separated from the S30 reaction means that no purification step would be required for S30-loaded vesicles with higher chances of retaining the activity of the S30 solution.
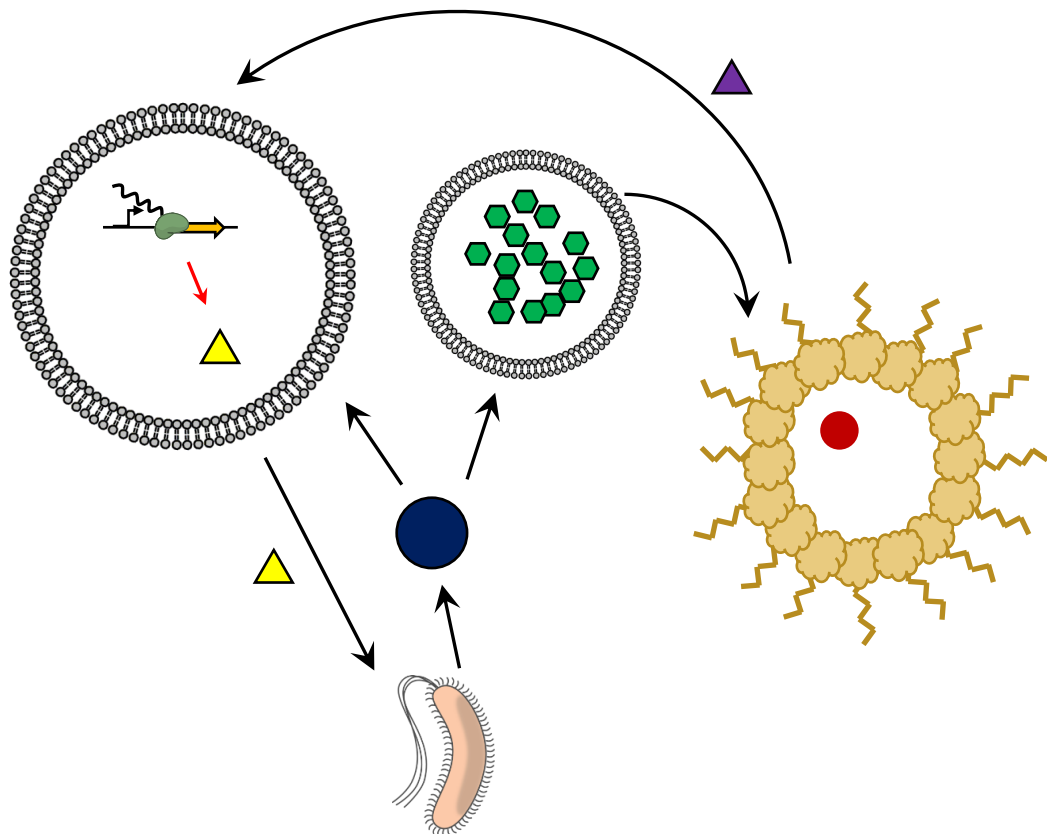
Figure 3 - 11. Schematic representation of a community of artificial and natural cells. To avoid all the problems related to the purification of vesicles from glucose, an alternative communication pathway was designed to keep the chemical message separated from the S30 reaction. One population of lipid vesicles would encapsulate the S30 reaction and another population would carry glucose. The expression of a pore-forming protein (blue circle) would be induced in bacteria by lipid vesicles synthesizing a freely diffusible molecule active for gene induction (yellow triangle). The protein would disrupt bacterial membranes, then diffuse outside and attack all the lipid vesicles inducing glucose release. As described earlier, proteinosomes would be responsible for the beginning of the communication pathway by producing a trigger molecule (purple triangle) to induce gene expression in lipid vesicles.

The protein responsible for the explosion of vesicles was chosen to disrupt any kind of lipid membranes so that once produced they will be also released because of their action on the stability of the bacterial membrane. *E. coli* BL21 Rosetta 2 (DE3) cells were tested for the production of known pore-forming proteins such as holin from bacteriophage λ and ColE7, that should create bigger pores than α-hemolysin. First of all, it was tested

whether the membrane of bacteria was destabilized by the production of the protein. The turbidity of bacterial culture was monitored over time and ColE7 demonstrated to be more active in exploding bacteria than λ-holin as shown by a drop in optical density values in
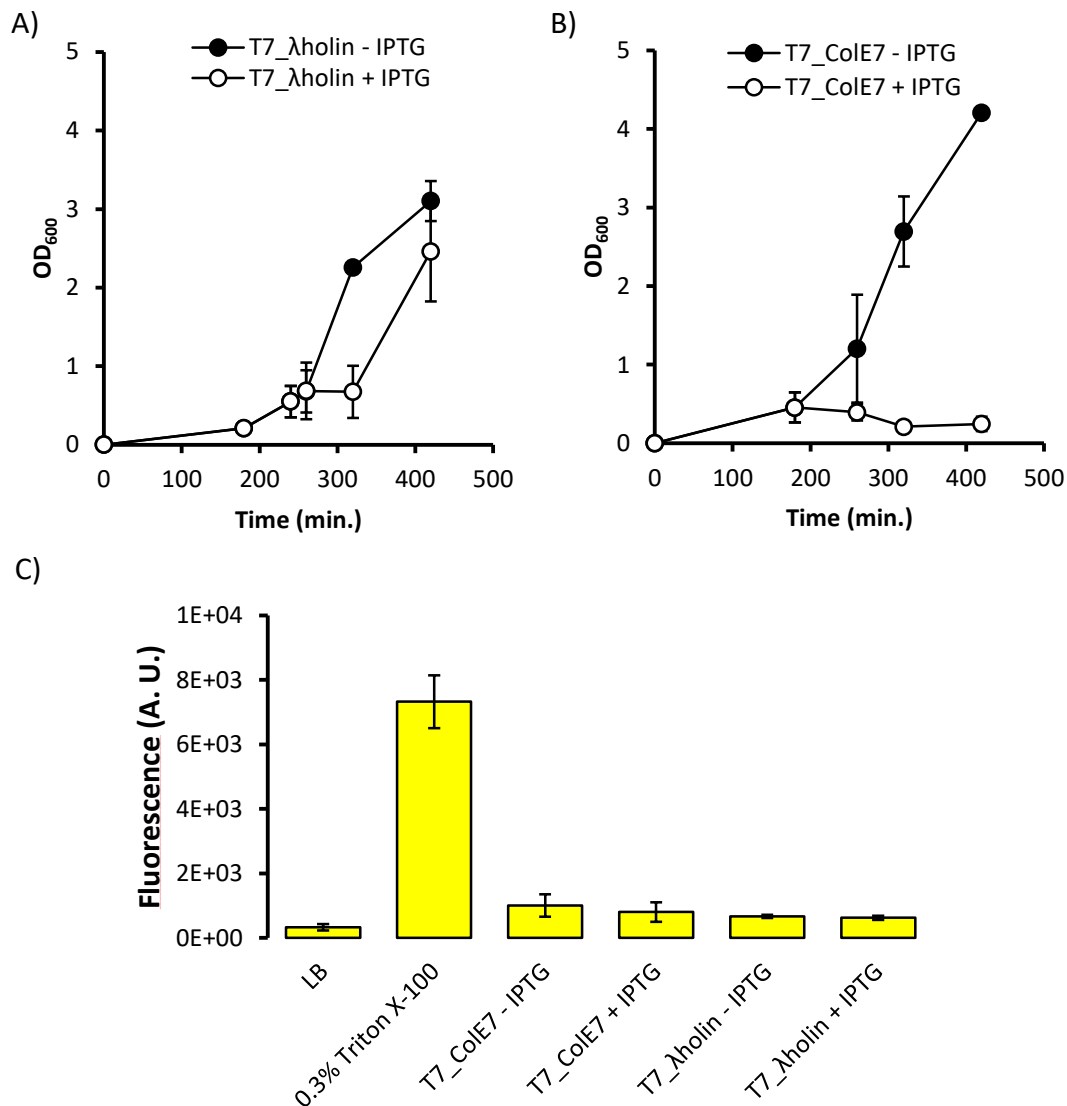
Figure 3 - 12. Pore-forming proteins expressed in bacteria. *E. coli* BL21 Rosetta 2 (DE3) cells were transformed with plasmids carrying the genes of the pore-forming proteins λ-holin and ColE7 under the control of an inducible promoter controlled by LacI. The ability to form pores in bacterial membranes was measured as the decrease in bacteria concentration by optical density changes after the induction with IPTG. The chart reports the absorbance values measured at 600 nm every 60 minutes. The error bars represent the standard deviation between biological duplicates. ColE7 (B) showed a higher effect than λ-holin (A). C) After 7 hours of induction, bacteria were incubated with vesicles loaded with 80 mM calcein to verify pore-formation on membranes. Fluorescence values of calcein were measured as indicated in Figure 3 - 6 and the error bars refer to the standard deviation between biological duplicates. Despite the proteins were able to reduce the number of bacteria, it was not possible to observe any release of calcein from liposomes when incubated with the cultures of bacteria producing the proteins. It is likely that all the polypeptides synthesized were sequestered by the membrane of the producer bacteria.

Figure 3 – 12 A and B. The cultures were successively incubated with calcein-loaded vesicles to verify the formation of pores as it was described in § 3.2.3 for αHL produced in S30 extract. Fluorescence of calcein inside vesicles was quenched because of the high concentration of the fluorophore and it was expected to increase in case of a destabilization of the membrane and the consecutive release. After the disruption of bacterial membranes, the pore-forming protein was thought to be present in the liquid medium so that a simple addition of the medium to calcein-loaded vesicles could result in calcein release due to a destabilization of lipid membranes. Conversely to the expectations, no effect on lipid membranes was observed when bacteria were incubated with vesicles after the protein was produced. It is likely that after the production, the pore-forming proteins attached to bacterial membranes so that no interaction with lipid vesicles was possible (Figure 3 - 12 C).

A new construct is now under investigation inspired by the work of Saeidi *et al.*[9] In this new design not only ColE7 would be under the control of a genetic switch but also the protein perfringolysin-O (PFO). This protein undergoes a conformational change that leads to the formation of pores in lipid membranes only in presence of cholesterol.[111] The human pathogen *Clostridium perfringens* exploits this property to infect human cells containing cholesterol while preserving its membrane that instead lacks cholesterol. *E. coli* membranes do not contain cholesterol and a water-soluble version of the protein can be efficiently produced by the removal of an N-terminal signal peptide.[112] In this new design, *E. coli* cells would explode by the action of ColE7 and release PFO that would cause the break of cholesterol-containing lipid vesicles.

Part B

Development of biosensors for possible integration with artificial cell

Chapter 4

Selection for a malachite green DNA aptamer for use in a sensor molecule

The concept of artificial cells described in the previous chapter relies heavily on biosensing. Simple biosensors, that are not based on complex structures such as artificial cells, can be developed by engineering biological molecules. These elements could be used alone or possibly integrated with artificial cells. Several biological elements can be engineered to detect molecular signals and among the described sensors,[113,114] aptamers offer a versatile and easy to engineer module.

Nucleic acids sequences can be engineered *in vitro* through methods of directed evolution, with a technique described in the 1990s named SELEX (Systematic evolution of ligands by exponential enrichment). Although it was L. Gold to coin the term, both G. Joyce and J. Szostak together with A. Ellington set up similar protocols independently.[115–118] This procedure is based on a pool of randomized single-stranded nucleic acids, referred to as "library", that is added to the molecule for which an aptamer has to be created. Through iterative steps of binding, elution and amplification, the desired oligonucleotide is evolved in a test tube. The combination of aptamers in modular sequences can be a useful tool for the creation of biosensors. The fusion of an aptamer with ribozymes, for example, can induce a nucleolytic cleavage in presence of the ligand.[119] Two aptamers can be combined together to create a sequence that can undergo a conformational change induced by the simultaneous presence of the two ligands. The change can be easily detected through aptamers that are able to increase the quantum yield of barely fluorescent fluorophores.[95,120–123] A biosensor for a small biologically relevant molecule can be based on a single-stranded nucleic acid sequence containing an aptamer developed for the binding of the small molecule (the analyte) fused together with an aptamer capable of binding a fluorophore. In a previously described work,[124] two aptamer sequences, one able to bind a fluorophore and one able to bind an analyte, were joined together with an intermediate sequence that induced a conditional conformational change in presence of the analyte so that the nucleic acid switched from a conformation unable to bind a fluorophore to a conformation able to do so. The final result was a sensor that shows fluorescence increase only in presence of the molecule to be detected.

Because of the small dimensions of this type of sensors, it is possible to consider applications such as live imaging in narrow spaces where other molecules could not fit. Neurotransmitters, for example, are molecules hard to visualize because of the small gap where they localize extracellularly. Among the known neurotransmitter, an RNA aptamer for the binding of dopamine was developed through *in vitro* evolution[125] and it was shown to retain its affinity also in the DNA version.[126] DNA is more resistant to hydrolysis than RNA and therefore it is much more recommended for the use as a biosensor. The project aimed at building a single-stranded DNA molecule from the combination of a dopamine aptamer and a fluorescent aptamer, to create a dopamine biosensor for live imaging of the

neurotransmitter. To our knowledge, there was only one example of DNA aptamer able to increase the quantum yield of a poorly fluorescent molecule.[121] The ligand of this aptamer was a modified version of the Hoechst dye and it was not commercially available thus posing a limit to the possible range of users: the project was aimed at developing a universal tool exploited by the largest number of scientist and to this end it requires all the components to be easily accessible.

A more versatile DNA fluorescent aptamer had then to be engineered and malachite green (MG) seemed a good candidate ligand for a couple of practical reasons: it can be purchased at a very low price and there is an RNA aptamer shown to increase its fluorescence quantum yield of ~2000 fold.[120,127] The aptamer was selected by affinity to the ligand and discovered later to increase its fluorescence quantum yield, so a DNA aptamer could potentially do the same. Conversely to what was described for the dopamine aptamer, the DNA version of the MG aptamer was not able to increase the fluorescence of malachite green, therefore a SELEX procedure was set to find a binder among three different pools of DNA molecules.

4.1 Materials and methods

4.1.1 DNA sequences

DNA libraries for strategies 1 and 3 were purchased from IDT technologies; DNA library for strategies 2, DNA oligonucleotides for PCR were purchased from Eurofins Genomics. All the DNA libraries contained a core variable region flanked by two constant portions for PCR amplification after every round of selection. Details of the library design are discussed in the results.

Table 4 - 1. Single-stranded DNA sequences. DNA sequences were all sent lyophilized and primers were purified by standard desalting while all the libraries were purified by HPLC.

ID	Sequence (5'-3')	Modification	Purpose
DC026	GCGGATAACAATTCCCCTCT	5' biotin	PCR amplification of the first library
DC027	GCGGATAACAATTCCCCTCT		qPCR of the first library
DC028	GCTGTCCACCAGTCATGCTA		PCR amplification and qPCR of the first library
DC047	GTAAAACGACGGCCAG		Sequencing
DC031	GCGGATAACAATTCCCCTCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTAGCATGACTGGTGGACAGC		DNA library 1
MGApt(T)	GGATCGCGACTGGCGAGAGCCAGGTAACGAATCGATCC		DNA version of the RNA aptamer
MGApt(U)	GGAUCGCGACUGGCGAGAGCCAGGUAACGAAUCGAUCC		DNA version of the RNA aptamer with deoxyuridine
DC070	GGAACACTATCCGACTGGCACC		PCR amplification and qPCR of libraries 2 and 3
DC071	CGGGATCCTAATGACCAAGG		qPCR of libraries 2 and 3
DC072	GGAACACTATCCGACTGGCACCGGATCGNNACTGGCGAGAGCCAGNTAANNNATCGATCCACCTTGGTCATTAGGATCCCG		DNA library 2 with 6 randomized nucleotides
DC073	CGGGATCCTAATGACCAAGG	5' biotin	PCR amplification of libraries 2 and 3

Degenerated MG aptamer	GGAACACTATCCGACTGGCACC <u>GGATCGCGACTGGCGAGAGCC AGGTAACGAATCGATCC</u> CCTTGGTCATTAGGATCCCG		DNA library 3. The underlined part was synthesized in order to have 85 % of the indicated nucleotide and 15% random nucleotides
AJM.RBS.ON			Control DNA used for fluorescence test
FC001A			Plasmid DNA encoding MG RNA aptamer under the control of T7 promoter

4.1.2 RNA sequences

The MG RNA aptamer was synthesized from plasmid FC001A (5'-GGAUCGCGACUGGCGAGAGCCAGGUAACGAAUCGAUCC-3') as described in § 4.1.8 and purified as in § 4.1.9. Total mRNA was extracted from *E. coli* DH5α with GeneJET RNA purification kit according to manufacturer's instructions.

4.1.3 Reagents and general supplies

Ultrafree-MC microcentrifuge filters with pore size of 0.45 µm, Malachite green (MG), adipic acid dihydrazide agarose, N,N-dimethylformamide (DMF), N,N,N′,N′-Tetramethylethylenediamine (TEMED), sodium bicarbonate (NaHCO3), liquid chromatography columns, LB and LB agar media, magnesium chloride (MgCl$_2$), spermidine, 1,4-dithiothreitol (DTT), HEPES, 99% ethanol, sucrose, sodium dodecyl sulfate (SDS), bromophenol blue, xylene cyanol, 0.5 M EDTA pH 8.0, ammonium persulfate (APS) were purchased from Sigma-Aldrich; 50 U/µl T7 RNA polymerase, 100 mM ATP, 100 mM GTP, 100 mM CTP, 100 mM UTP, yeast inorganic pyrophosphatase were purchased from New England Biolabs; Malachite green isothiocyanate (MGI), dynabeads C1 streptavidin beads, were purchased from Life technologies; TOPO-TA cloning kit was purchased from Invitrogen; 2X SSO advanced mix for qPCR and Micro Bio-Spin Chromatography Columns were purchased from Biorad; Phusion polymerase and mRNA extraction kit (GeneJET RNA purification kit) were purchased from Thermo Fisher Scientific; Wizard® Plus SV Minipreps DNA Purification System was purchased from Promega; ALUGRAM® aluminium sheets SIL G/UV254 (cat.# 818133) was purchased from Macherey Nagel.

4.1.4 Buffers and solutions

Buffers for selection included: 10X MG selection buffer 1 (50 mM $MgCl_2$, 50 mM KCl, 1450 mM NaCl, 200 mM Tris-HCl pH 7.4), 1X MG elution buffer (1X MG selection buffer, 1 mM MG), 10X MG selection buffer 2 (50 mM $MgCl_2$, 50 mM KCl, 1033.3 mM NaCl, 200 mM Tris-HCl pH 7.4) used from 2nd cycle onwards to compensate the NaCl concentration after the dynabeads purification. Buffers for dynabeads purification included: 2X DB B&W (10 mM Tris-HCl pH7.5, 1 mM EDTA pH 8.0, 2 M NaCl, 0.01% w/v Tween-20), 1X DB elution buffer (100 mM NaOH), 1X DB neutralization buffer (100 mM HCl, 5 mM Tris-HCl pH 7.5).

4.1.5 Affinity resin for malachite green

The column was created according to the protocol described by Grate and Wilson.[120] Briefly, 3 mg of malachite green isothiocyanate were solubilized in 300 ml of DMF and incubated with 10 ml of adipic acid dihydrazide agarose previously equilibrated with 0.1 M $NaHCO_3$ (pH 8.3). The reaction proceeded overnight at room temperature in the dark. Then the column was stored at +4 °C wrapped in aluminum foil. Prior to each use, the resin was washed on a column by gravity adding 2 vol of DMF, 5 vol of DEPC-treated water, 2 vol of 1X MG selection buffer 1. It was finally resuspended with 1 vol of 1X MG selection buffer and kept for storage at +4 °C.

4.1.6 Selection cycles

The basic protocol was established on a cyclic repetition of the following steps: 1) a negative selection to exclude nonspecific binders by incubation with adipic acid dihydrazide; 2) positive selection by incubation with MG column, wash and elution; 3) quantification of bound molecules through qPCR; 4) PCR amplification to get a number of molecules similar to the start; 5) purification of ssDNA through dynabeads C1 streptavidin magnetic beads 6) quantification through qPCR.

1) The starting number of molecules for each cycle was ~10^{14}. The DNA was heated at 95 °C for 10 min, chilled on ice for 5 min and incubated with adipic acid dihydrazide agarose for 30 min. The solution was pipetted on Microbiospin columns and DNA was spun down by gentle centrifugation at 1000 g for 30 s. In step 2) 1 volume of previously washed MG column was added to eluted DNA and incubated in darkness at room temperature for 1 h; unbound DNA was removed by centrifugation on Micro Bio-Spin columns at 1000 g for 30 s; the column was washed with 15 vol of MG selection buffer

with the same centrifugation speed and time; bound DNA was then eluted with 5 vol of MG elution buffer and concentrated by ethanol precipitation in 20 µl of DEPC-treated water as described in § 2.1.5.

qPCR was performed with Biorad mix SSO advanced in steps 3 and 6). A 10 µl reaction contained 1X SSO advanced mix, 2.5 µM of each primer and 1 µl of DNA solution. DNA amplification was monitored in 0.1 ml tubes with real-time PCR cycler Rotor-Gene Q in channel green with the following thermal protocol: 95°C for 30 s for initial denaturation and then 40 cycles of 95 °C for 5 s and 63 °C for 5 s. The absolute number of molecules was estimated through a standard curve prepared with several dilutions of the library DC031 in a range between 1 and 1000 pmol. The fraction of bound molecules was calculated on the ratio between the number of molecules bound to MG column and those collected after dynabeads purification.

PCR amplification was performed with Phusion polymerase in step 4). A 1000 µl mix contained 1X HF buffer, 0.2 mM each dNTPs, 0.5 µM of each primer, 0.02 U/µl Phusion, ~ 10^7 molecules/µl DNA. The reaction ran on a thermal cycler with the following protocol: 98°C for 2 s for initial denaturation, then 20 cycles of 98 °C for 5 s, 63 °C for 5 s and 72 °C for 5 s and a final extension step at 72 °C for 10 min.

ssDNAs were separated with dynabeads C1 in step 5). 120 µl of dynabeads were pipetted into a 1.5 ml tube placed on a magnetic rack. The buffer was removed and beads were washed 3 times with DB B&W buffer; the PCR product was resuspended with dynabeads and incubated on rotation at room temperature for 15 min. The tube was placed back on the magnetic rack and the supernatant was discarded. Beads were washed 3 times with B&W buffer and the non-biotinylated strands were detached with 250 µl of 1X DB elution buffer by shaking the beads with NaOH for 5 min. Beads were placed back on the magnet rack and the supernatant was transferred to a new tube. The solution was neutralized with 250 µl of 1X DB neutralization buffer and 1X selection buffer was added to a final volume of 600 µl.

4.1.7 Cloning in TOPO-TA and sequencing

The cloning was performed as described by the supplier. To a PCR reaction run with Phusion polymerase, 1 µl of home-made Taq polymerase was added and incubated at 72 °C for 10 min to allow the incorporation of few adenosines in 3' position. Then 4 µl of the PCR were mixed with 1 µl of TOPO vector and 1 µl of a salt solution provided by the supplier. The mixture was then incubated at room temperature for 5 min and added to *E. coli* TOP10 for transformation. 20 sequences from cycles 1, 2, 5, 9 and 17 were sent for sequencing with primer DC047.

4.1.8 In vitro *transcription with T7 RNA polymerase*

In vitro transcription was performed with T7 RNA polymerase following the indications from Seelig.[90] A reaction mixture was set to a final volume of 100 µl with the following components: 35 mM MgCl$_2$, 2 mM spermidine, 200 mM HEPES adjusted to pH 7.5 with KOH, 1 mg/ml BSA, 4 mM DTT, 5 mM each NTP, 1 mU/µl yeast inorganic pyrophosphatase, 0.4 U/µl RNase inhibitor, 3 U/µl T7 RNA polymerase, 0.2 µM DNA template. The reaction was incubated at 37 °C for 4 h, then DNase I was added to a final concentration of 20 mU/µl together with its buffer provided by the supplier and incubated at 37°C for 1 h. 5 mM EDTA was finally added.

4.1.9 RNA purification through polyacrylamide gel

A solution with 8 M urea, 8% acrylamide:bisacrylamide (37.5:1) was mixed. APS and TEMED were both added to a final concentration of 0.1% and the solution was poured on a gel apparatus and waited for polymerization. While the gel polymerized, RNA was precipitated by the addition of 800 µl of 99% ethanol. RNA was centrifuged at maximum speed at 4°C for 30 min. The supernatant was removed and the pellet dried at 94 °C. RNA was resuspended in 50 µl of a solution containing 14% sucrose, 0.07% SDS, 0.035% bromophenol blue, 0.035% xylene cyanol FF, 63 mM Tris, 63 mM boric acid, 7 mM EDTA pH 8. The solution was loaded on the polyacrylamide gel submerged in 1X TBE buffer and a constant voltage of 140 V was applied for 4 h.

The RNA was later cut from the gel through UV shadowing and extracted by the crush-and-soak method.[90] Briefly, the cut gel band was ground with a pipette tip in a 1.5 ml tube, then 500 µL of TE buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA) and 120 U of RNase inhibitor were added and the sample was incubated tumbling at 37°C overnight. The supernatant was isolated by centrifugation at maximum speed for 5 min, then 250 µl of TE buffer were added to extract more sample by tumbling at 37 °C for 2 h. The supernatant was again collected by centrifugation and any residual of polyacrylamide gel were removed through 0.45 µm microcentrifuge filters.

4.1.10 Fluorescence tests on DNA and RNA aptamers

To assess the increase in fluorescence of malachite green aptamer in presence of nucleic acids, a 10 µl mix was assembled in a 0.1 ml tube containing the following: 1X MG selection buffer, 10 µM MG RNA aptamer, 50 µm malachite green. To this mix, DNA

sequences arose from the selection and control single-stranded DNAs were added to different concentrations: 1, 10 and 100 µM. Fluorescence was monitored with the real-time PCR cycler Rotor-Gene Q on red channel (excitation: 625 ± 5 nm; emission: 660 ± 10 nm) with gain 10 at 37 °C. The fluorescence values were normalized by calculating the ratio between fluorescence values measured for the molecule mixed together with nucleic acids sequences over the values measured for the free molecule.

4.2 Results

The design of the selection procedure was based on few considerations. First of all, the complexity of the library had to be statistically big enough to ensure the presence of the desired molecule. Among the *in vitro* selection protocols, numbers vary between 10^{13} and 10^{15}.[128] The malachite green RNA aptamer is 38 nt in length, so a library of 38 random nucleotides is more than enough to obtain a number of molecules at least equal to what is reported in the majority of protocols ($4^{38} \approx 10^{23}$). The amount of sequences synthesized corresponded to 10^{16} molecules, that are guaranteed to be all different one to another by the complexity of the library. 1/10 of this library was used for the selection, in order to ensure multiple trials. The core random region was also flanked by two constant sequences for quantification and amplification (Figure 4 - 1). The separation of molecules able to bind the analyte from molecules unable to bind the analyte was carried out with a home made resin obtained by the conjugation of agarose beads with malachite green isothiocyanate.[120] The need for a physical support for the molecule of interest could lead to the amplification of non-specific binders, more likely to interact with the support rather than with malachite green. To exclude these molecules, the DNA pool was always incubated with agarose beads free from malachite green prior to the malachite green resin. These procedures were respectively called negative and positive selection.

Figure 4 - 1. Design of the DNA library. A central core random region of 38 nucleotides is flanked by two constant regions, to allow primer's matches for PCR amplification and quantification.

The amplification of DNA after each round of selection posed some issues. While RNA was synthesized by the polymerases as a single strand, DNA aptamers were required to be separated from the complementary sequences after the amplification steps. For such a purpose, a 5'-biotinylated primer was included in the PCR reactions and the products were caught through streptavidin beads. The biotinylated primer matched the aptamer and was in 5' position of the complementary sequence, while the primer matching the complementary sequence was not modified. An increase in pH was enough to break hydrogen bonds and separate the two strands, allowing the collection of the lone strand selected for MG binding. This strand was released in solution while leaving the complementary strand bound to streptavidin beads.

From all these considerations the following cyclic protocol was designed: 1) negative selection to exclude non-specific binders by incubation with adipic acid

75

dihydrazide agarose beads; 2) positive selection by incubation with MG column, wash and elution; 3) quantification of bound molecules through qPCR; 4) PCR amplification to get a number of molecules similar to the start; 5) purification of single-stranded DNA (ssDNA) through dynabeads C1 streptavidin magnetic beads 6) quantification through qPCR (Figure 4 - 2). To monitor the proceeding of the selection, some of the DNA sequences picked from several cycles were cloned and sent for sequencing.
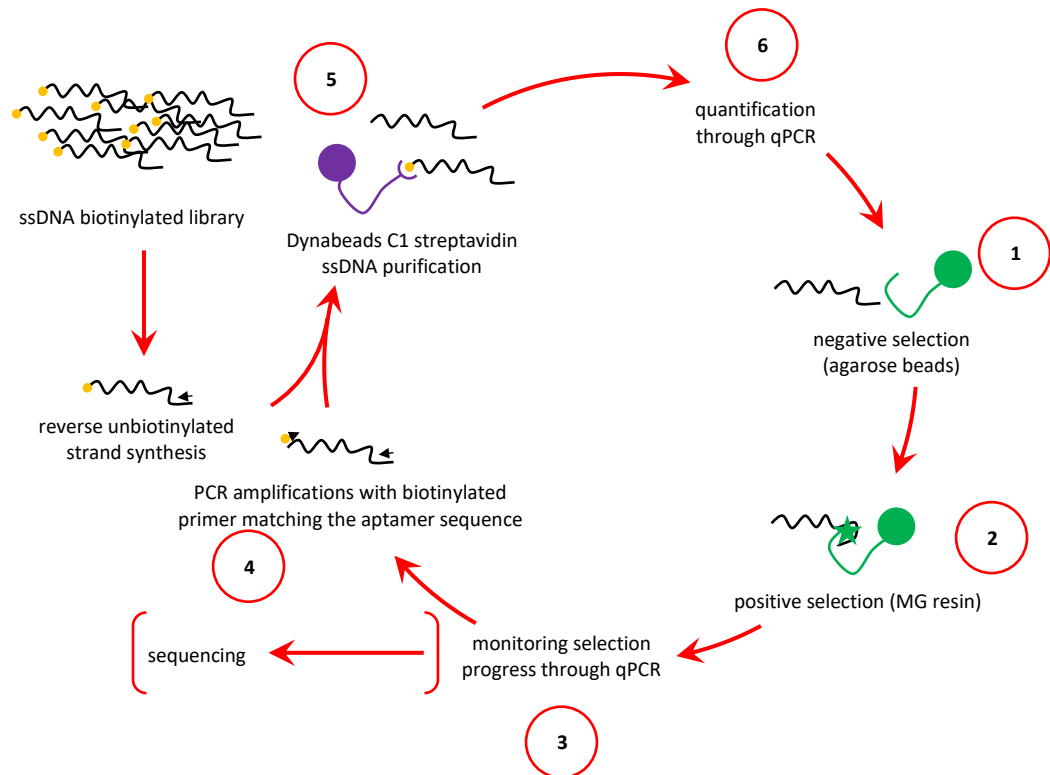

Figure 4 - 2. Diagram flow of the selection cycle. The selection protocol is based on the repetition of 6 steps. The non-specific binders are first removed by incubating the sequences to agarose resin free from malachite green (1), then they are incubated with malachite green resin (2) and the number of bound molecules is quantified by qPCR (3). The bound molecules are amplified by PCR using a 5'-biotinylated primer to obtain a linear double-stranded DNA fragment where only the sequences complementary to the selected sequences are biotinylated (4). The amplified aptamers are separated from the complementary sequences by means of magnetic streptavidin beads (5) and successively quantified by qPCR (6). By mistake, the library was synthesized with a biotinylation in 5' position and it could not be used for the selection because the first purification step with streptavidin beads would ensure the collection of the complementary strand, rather than the potential aptamers selected. Therefore the complementary strand of the library was synthesized with a non-modified primer and separated by streptavidin magnetic beads (5) prior to beginning the selection process.

4.2.1 Preliminary steps to verify the validity of the protocol

Prior to beginning the SELEX protocol, two preliminary experiments had to be performed to ensure the validity and the necessity of the procedure. In order to demonstrate that the *in vitro* evolution of a DNA aptamer was a necessary step for the development of a DNA-based biosensor, the possibility that a DNA version of the RNA aptamer was able to bind malachite green had to be excluded. For this purpose, a DNA oligonucleotide containing the sequence of the malachite green RNA aptamer was synthesized. The second experiment was aimed at evaluating the binding ability of the home-made malachite green column.
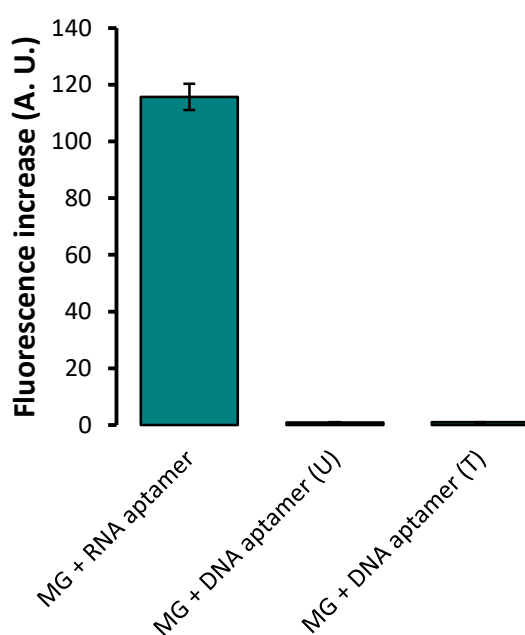

Figure 4 - 3. Test on the ability of DNA versions of MG RNA aptamer to increase malachite green fluorescence. Two single-stranded DNA oligonucleotides carrying the sequence of malachite green RNA aptamer were compared for the ability to increase the fluorescence of malachite green. In order to have a version more similar to the RNA aptamer, one of the two oligonucleotides contained uridine in place of thymidine. The fluorescence values were measured with real-time PCR cycler Rotor-Gene Q on channel Red (excitation: 625 ± 5 nm; emission: 660 ± 10 nm). The chart reports the increase in fluorescence of malachite green in presence of the sequences tested, expressed as the ratio between fluorescence values of the molecule together with nucleic acids sequences over the free molecule. The error bars represent the standard deviations between technical triplicates of the same sample. None of the two sequences showed fluorescence levels similar to the RNA aptamer.

In order to compare the ability of the two versions of the aptamer for malachite green binding, both DNA and RNA sequences were mixed with the ligand and tested for the ability to increase the fluorescence of malachite green. DNA was synthesized in two versions, one with regular nitrogenous bases, while the other containing deoxyuridine in place of deoxythymidine, to make the DNA more similar to RNA. Figure 4 - 3 reports the

increase in fluorescence of the fluorophore in presence of the RNA aptamer and in presence of the two DNA versions of the aptamer. Both DNA sequences were poor in promoting a fluorescence increase, indicating that no binding on the fluorophore occurred or, at least, no binding capable to significantly increase the fluorescence quantum yield. Similarly, the malachite green column was tested for its RNA aptamer retention ability. As observed in Figure 4 - 4, although lower, the increase in fluorescence was still higher with MG RNA aptamer than what was observed for random RNA, obtained from a total mRNA extraction of *E. coli* cells.
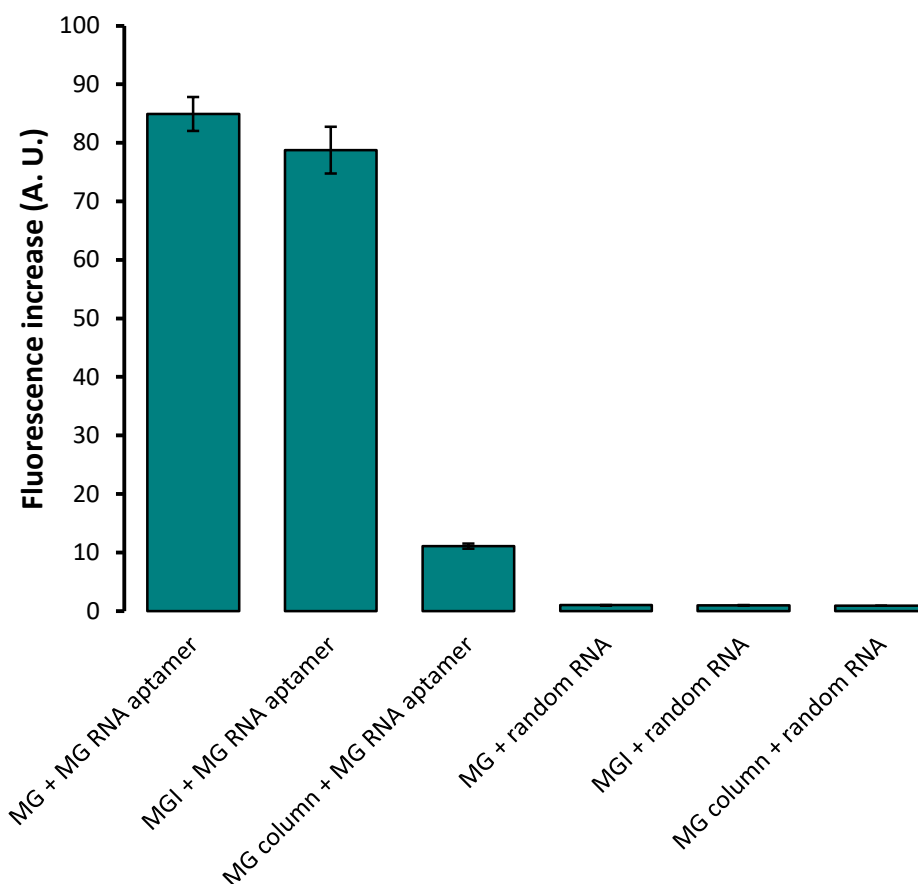

Figure 4 - 4. Verification of the ability of the RNA aptamer to bind MG column by fluorescence increase of the fluorophore. In order to verify the integrity of the malachite green resin, the MG RNA aptamer was tested for the ability to increase the fluorescence of the fluorophore. Fluorescence data were analyzed as reported in Figure 4 – 3. In the presence of the RNA aptamer, the fluorescence signal of malachite green column increased, although to a lower extent than what was observed in the presence of pure molecules MG or MGI (malachite green isothiocyanate, a modified version of the fluorophore used for the creation of the resin). This behavior may be explained either as the result of a reduced stability of malachite green after the synthesis of the resin or as a lower than expected number of molecules bound to the agarose beads. As the column seems to carry functional MG molecules it was not considered worth getting deeper to know the exact number of active molecules.
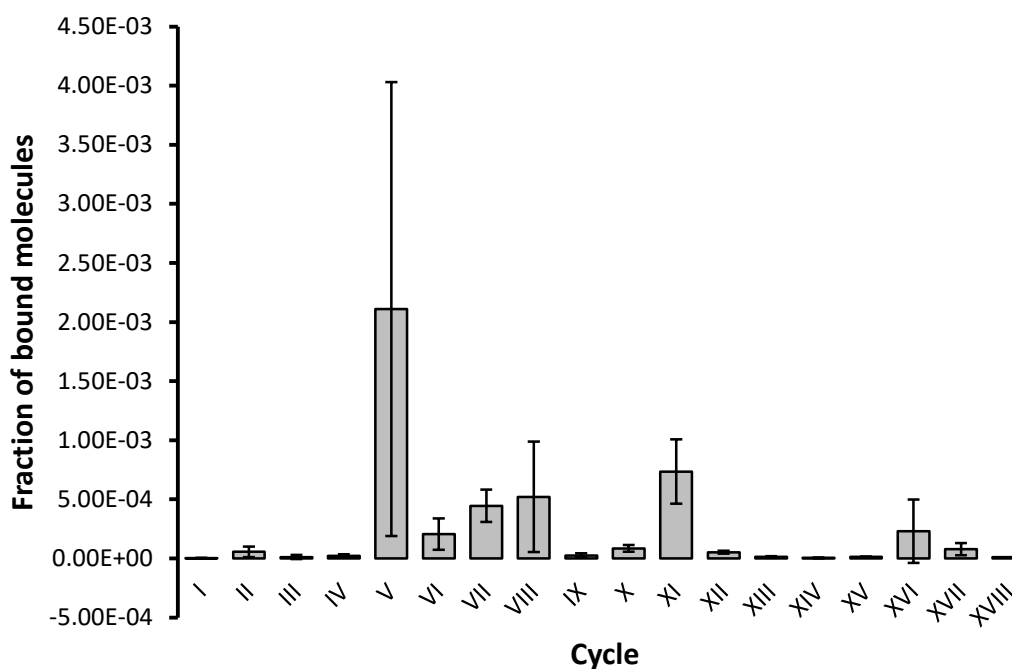
4.2.2 Selection from a randomized DNA library

Figure 4 - 5. Selection progress. The chart reports the number of DNA molecules obtained after each round of selection calculated by qPCR. The values are reported as the fraction of DNA strands obtained after the selection, over the number of strands obtained after the separation with streptavidin magnetic beads. The error bars represent the standard deviation of technical triplicates. The huge error bars observed for some cycles may be related to technical errors occurred during the setting of the qPCR reaction. Despite the number was expected to increase and maintain a Plateau level, it was observed an unconventional behavior in which the number of bound molecules increased and dropped down more than once: excluding the value associated to cycle V that may be resulted from a mistake, the number of bound molecules increased up to cycle VIII and dropped in cycle IX, it increased again up to cycle XI and dropped again the next cycle; a small increase of bound molecules seemed to be linked to cycle XVI, although it could have been related to technical mistakes.

Having verified the functionality of the column, the selection could start. By mistake, the library was designed to carry a 5' biotinylation, but the protocol for the isolation of single-stranded DNA fragments with streptavidin beads allows the collection of the non-biotinylated strand, as described earlier. Therefore the use of a biotinylated library would forbid the collection of the sequences bound to malachite green. For this reason, the complementary strand was first synthesized with Phusion polymerase. Then the non-biotinylated strands were collected by streptavidin beads separation and subjected to the selection procedure. The progress of the selection was monitored through qPCR (Figure 4 - 5). If the selection worked properly, the number of bound molecules should have increased after every round until finally leveling off. Instead, an odd behavior was

observed. The number of bound molecules increased, then dropped down and then increased again. As the results from qPCR seemed inconsistent, some of the collected DNA molecules were sent for sequencing from cycles 1, 2, 5, 9 and 17. The final results are summarized in Figure 4 - 5. From a representative number of samples sequenced for each cycle, it is clear that some of the sequences accumulated over the selection procedure.
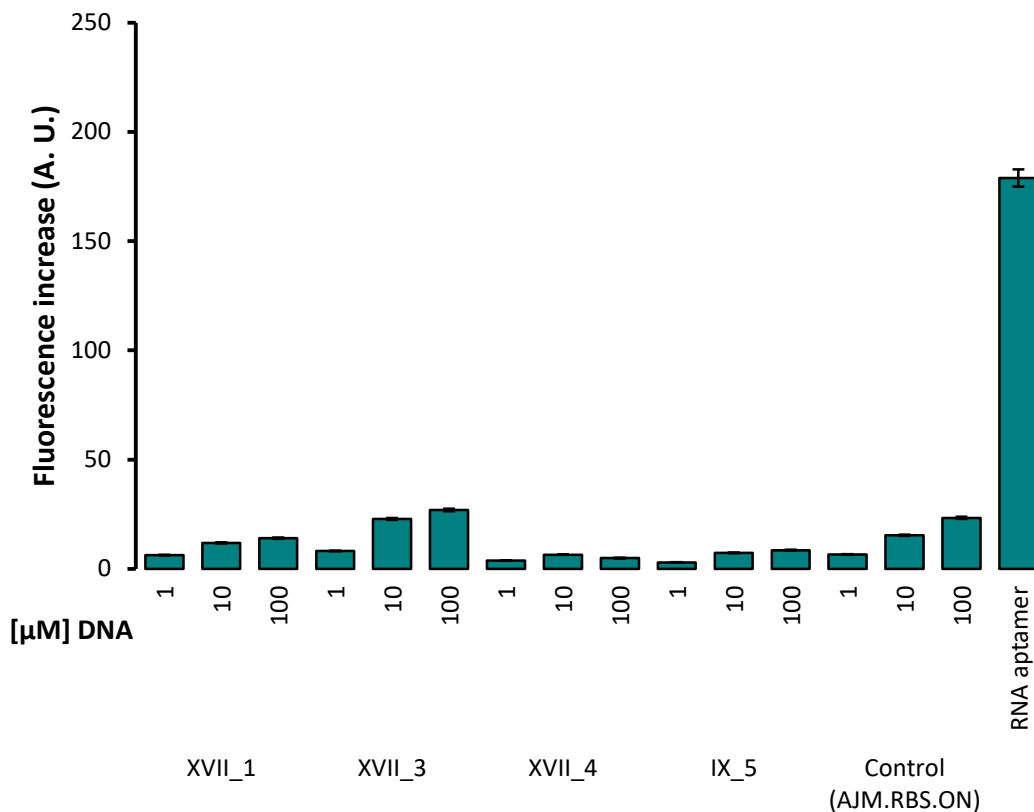

Figure 4 - 6. Fluorescence test on selected sequences. The most frequent sequences as resulting from Table 4 – 2 were tested for their ability to increase the fluorescence of malachite green. Fluorescence values were analyzed as in Figure 4 – 3. Sequences XVII_1 and XVII_3 show a slight increase in fluorescence correlating to DNA concentration. Nonetheless, the signals are not significantly different from what was observed for control DNA. It is likely that the DNA sequences had some non-specific interactions with malachite green that resulted in slight increase in fluorescence.

The most represented sequences were tested for fluorescence in the presence of the ligand malachite green. Because of a possible active involvement in the binding of malachite green, these sequences contained intentionally also the constant regions used for the amplification. As shown in Figure 4 - 6, a relevant increase in fluorescence was observed for none of the tested sequences. The poor fluorescent signal was not higher than what could be observed for control DNA sequences, that are some single-stranded DNA sequences with no homology neither for RNA aptamer nor for the constant regions of the DNA library used for the selection. These results suggest that the sequences

obtained by the selection process were not able to stabilize malachite green and increase its fluorescence quantum yield.

Table 4 - 2. Sequencing results. The table shows the sequencing results obtained from the selection and the relative frequency for each of the analyzed cycle. The table reports only the variable regions, excluding the two constant extreme sequences. Almost all of the sequences are composed of 38 nucleotides, except for few. Sequences with different length from what was expected might have aroused because of some mistakes occurred during the PCR amplification steps: it is likely that some secondary structure may have negatively affected the amplification of some DNA sequences. Counts are highlighted in red when no repetition was observed and in green when present more than once. The color code helps the visualization of the accumulation of repeated sequences while the selection proceeds. The most abundant sequences were named with an ID code and tested in presence of malachite green.

		Cycle				
ID	**Sequence (5'-3')**	**I**	**II**	**V**	**IX**	**XVII**
	CCAAAACTAGCCCGTAATTAATTAATCCAGTGCGATTAAA	**1**	0	0	0	0
	CTAAATAGTTTAATTCCCTACGAGCCCGGCAAAAAGAGCT	**1**	0	0	0	0
	TATTAAAGTATCAAAAAGCCAAAACCTTCACATACCTTCT	**1**	0	0	0	0
	CGGCCTGAGCGAAACTACACCGGACACAGTCAACTAGAAA	**1**	0	0	0	0
	CCGAATCTAGAACACAAGGATAAGCGCGTAGCCTTTACGC	**1**	0	0	0	0
	CATAGCACTTGTCATTGTTGAATAAAGCCGCGTTTGCAGA	**1**	0	0	0	0
	ATTCCTGTTTCCCTGTCCAACACGATGTAATCAACCATTC	**1**	0	0	0	0
	GTCTCAGCGGGCACCCTGTAATAATAACACGCGCGACTTA	**1**	0	0	0	0
	CCGTGAAACAGTGAGATAGTAGTAACGTGTTGCAAACAAA	**1**	0	0	0	0
	ACCGAATCTGTAAGCGAGTTACTCTCATACTTGTACGAAA	**1**	0	0	0	0
	ATATACAGATAGTTGCGTCTGTCGCGCCCGCGGAACTAAT	**1**	0	0	0	0
	TACTTTCCATGCGTCCAATTAAAACTAAAACGGCCGGACC	**1**	0	0	0	0
	AAACAGTTTTCGTGACGATGGAAGCTCAAGAAGGACAACT	**1**	0	0	0	0
	TGGACAATGTCTAATTCGAATGGGTTGAAATTTAGAAGTG	**1**	0	0	0	0
	TAAGCGGAACGCCTGCACAATATCAGCGCGATCAACTAGT	**1**	0	0	0	0
	CTCACCGAATTCCCGCGTTATACCGGATATCATCCCGTAT	**1**	0	0	0	0
	TGACCACAGCAGGCAAAACCTGTGCTGTAGTGGGTCGACT	**1**	0	0	0	0
	AGGGGAATAAGCTATCTAGAACAGTGTAGACACTCGTCAG	**1**	0	0	0	0
	CACGGGGCCAACGGCAAACTCTGTCTAAGACAGTTCTGAC	**1**	0	0	0	0
	ACACTACGACCTGGGAGAACCACCTCTGCATAGAGATAAA	**1**	0	0	0	0
	AATTATATCTAATTCTAGCCGTAATTGGCAGTTCCTGATC	0	**1**	0	0	0
	TACTAAGATCGATGCCTGAGGGACCCCCTAGAAAGCATTG	0	**1**	0	0	0
	TTCATAAGGAGTATCCTCCCATGAGCACATGCACGTAAGG	0	**1**	0	0	0
	GGCAGCGAAGCTACATACCTTCCCAGCACCATCGACTAAA	0	**1**	0	0	0
	TGCTAGATAAGACCATCTGATAGAAAAAATGATAGATGAC	0	**1**	0	0	0
	TCCGTGACCTCGGTCTCTCCCTTAACGCACTACTCCCTTG	0	**1**	0	0	0
	TAACATATTTAGTCAGGACAGCGTCGAGAGGGCAGGATCA	0	**1**	0	0	0

|						
	CACTTAACAATAATCTGACACAAATCACCTATATATGTGC	0	**1**	0	0	0
	AGAAGAACAAACACGTGCTGCAAATTTCGCACCTATGTGGA	0	**1**	0	0	0
	AAAACCCAATATAAGTACGATAAGCGCCATGTTCAAGACC	0	**1**	0	0	0
	ACATATGCTGGAACATAACGTGGGTTTCAGCTGTCGAAAG	0	**1**	0	0	0
	TAAGTATCAAGAGAATGCTAACTAGGATGGTCGCCCGACA	0	**1**	0	0	0
	CCCCACTCCCGACCGCCCTCGAAGTCTGTTTCTCAGAAGC	0	**1**	0	0	0
	GAGAGGGATATGAGCCGATCGCCGATAAACGTGTAGCTTT	0	**1**	0	0	0
	AGCGCAAAATAGTTAGACCGTCCGGCTCTCATGTATGAAA	0	**1**	0	0	0
	GTCCAATAAGATCGACTGCGAACAGCCGGCGTGGCGTCTT	0	**1**	0	0	0
	CCTCTAGCAGATCGAACATGCCATAGTATCTCTTTAGGTC	0	**1**	0	0	0
	TTATTCAACCTTTACTTGCGATCAAATAGTCTTCTTCATT	0	**1**	0	0	0
	CCCCCCCGACCTTTAATGTCACGACCCCAGTAAGTCGACC	0	**1**	0	0	0
	CGCACAGTACGTGGGTTAACATGGAAACACACGGCATGTC	0	0	**2**	0	0
	AGGCGAAACTGGCTCATGCACCTATGAGCGTGGACGAGCG	0	0	**6**	**4**	0
	CTGTGACTGGGCGGTTCCTCCGCCGTGGGCCGGCT	0	0	**1**	0	0
	CGTACAGGAAAATGTCGGTGCGCGCGTTCAGTATATCAGG	0	0	**1**	0	0
	GGATCGTGACAAACGAACAATTGTACGAGGCATGCCCTGC	0	0	**1**	0	0
	ATGGCGGTACGTCTCGTACAAGCACTGGGCGACATGTAAG	0	0	**1**	0	0
	ATGAGCGAACCAACGCCTGGTGTTGACTGAACGCGTAGTC	0	0	**1**	0	0
	GCTCATAGCATCTGTGAAAGGTCATAGGACGCTGGACTGT	0	0	**1**	0	0
	GCTCATTTCACGTGCAATGACTGATCGGGTTGCGATGTGG	0	0	**1**	0	0
	CCTCTTCAAGGACTTGTCAGCCGTCAACACACCATCGCAC	0	0	**2**	0	0
	CGTGACTGGGATCCGGTACGGTCCGTGGGCAACATAGCAG	0	0	**2**	0	0
MGAptIX_5	AGGTTTTCCAAGGTACGAAGTAGCTTCCTCGTGT	0	0	0	**7**	**12**
	CACGCGGTTCAAATTCCTCAACGTCTAGTTGACCGATCCT	0	0	0	**2**	**2**
	CAGAGGAGTAACTTTGGGTGTATAGCACAAAAATTCAGCC	0	0	0	**5**	**2**
	GACAAAGTGCGAAACCGCTGGAAGCGAATTCTCTTCGTAT	0	0	0	**2**	**3**
	GTACGACTGAAACCTGTTGTCATAAGGTTTGTGGCTGCCT	0	0	0	**1**	**1**
	GACCAAGAGCATGGCAACGTTCAGTTGCCGACTGGATCGG	0	0	0	**7**	0
	ATCGACAGGTGTACGTTCAGTACGGAGAGCCGGTATAACC	0	0	0	**6**	0
	TTGGAAAACATTCCGCACTTAATCTAAGTGCCCTCACAAT	0	0	0	**1**	0
	GCCTAACTGGTCGGAAAACTCGTCTAGAGTTTCAGCACGA	0	0	0	**2**	**1**
	CGGAGGCAAAGACTGAATTAACGAGACTTAATTGTGAGAC	0	0	0	**1**	0
	ATGCTTTAGGTGGTGCCGCAACGGAGGTACTAGGCACGAC	0	0	0	**1**	0
	GAAGCGGAGCACTTCGAAGCATCCGAAGGGACGAGTGGAT	0	0	0	**5**	0
	GCGCCCTCATAACCACGACCGTGAATTCCGCCCACCCAGT	0	0	0	**1**	0
MGAptXVII_1	TCACTTCTTCGTACAGTTGAAAGATTGCTGTCGGAGGAA	0	0	0	0	**11**
	CATTTGCACGTACAGGAATTTGTGCTATGACGACAGTAG	0	0	0	0	**5**
MGAptXVII_3	ACCGGAAACTCATCCACGCATTTGCTTTTGCGCCTCGTGT	0	0	0	0	**4**
MGAptXVII_4	GTGTGAAATTCCGCCGTTCAGGCCCTCGTTCTCCGCAAT	0	0	0	0	**6**

.
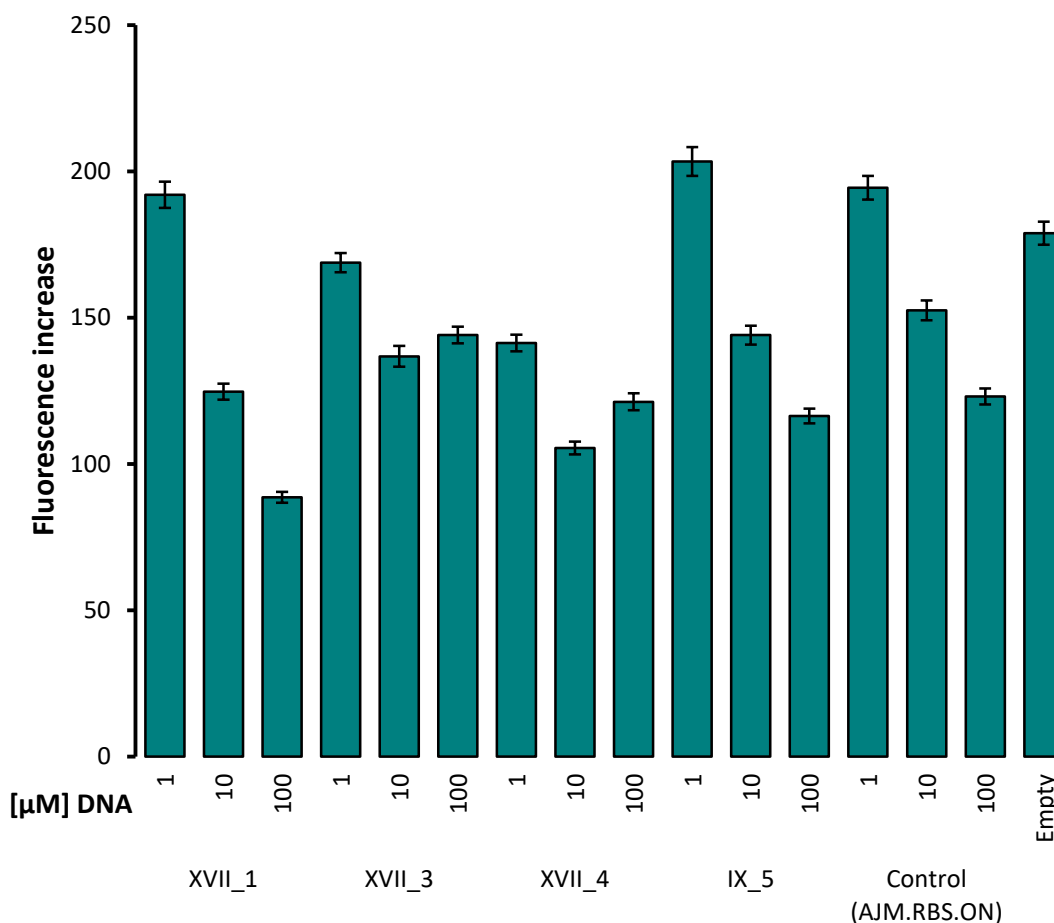
Figure 4 - 7. Test of the competition with RNA aptamer. Despite the accumulation observed in sequencing results during the selection process, none of the sequences tested showed the ability to increase malachite green fluorescence. This behavior may be explained by the fact that these DNA strands set interactions with malachite green different from what the RNA aptamer can do. To verify whether the potential aptamers are able to bind the fluorophore, the most frequent DNA sequences as resulting from Table 4 – 2 were incubated together with the RNA aptamer and tested for their ability to interfere with the binding of RNA by monitoring the fluorescence signal. Fluorescence data were analyzed as in Figure 4 – 3. No clear competition was observed and it was not possible to correlate the fluorescence values with a specific DNA sequence obtained from the selection process.

To finally check the effective ability in binding, the selected sequences were evaluated for their ability in disturbing the binding of the RNA aptamer. Mixed together with malachite green and the RNA aptamer, a good binder should decrease the level of fluorescence due to the interaction between MG and the RNA sequence. Again, several concentrations of each aptamer were not able to induce a specific decrease in the fluorescence signal, resulting in a similar signal from what was observed with control DNAs (Figure 4 - 7).

4.2.3 Degenerated libraries

A strategy based on a completely random library did not give the expected results, therefore two new libraries were designed with a different approach. Both of them were not completely randomized but were derived from the RNA aptamer sequence. Besides the constant extreme sequences required for quantification and amplification, one library (library 2) contained a core region with the malachite green RNA aptamer sequence and 6 randomized nucleotides in the positions more involved in binding the ligand, while the other (library 3) contained a degenerated sequence of the RNA aptamer. Library 2 was designed according to information found from the crystal structure analysis of the RNA aptamer bound to malachite green,[129] and library 3 was designed to have each position randomly mutated at a frequency of 15%. Both the strategies were intended to keep the sequence as close as possible to the RNA aptamer and increase the chances to find slight differences able to result in a novel DNA aptamer.

After few cycles of selection, some samples were sent for sequencing to monitor the progress of the selection but no sequence was repeated more than once for library 2

Table 4 - 3. Sequencing results from library 2. The bold characters in the library highlight the variable nucleotides. The differences in the sequences obtained from the selection procedure are indicated by colored characters.

ID	Sequence (5'-3')	Counts
Library	GGATCG**NN**ACTGGCGAGAGCCAG**N**TAA**NNN**ATCGATCC	
1	GGATCG**TT**ACTGGCGAGAGCCAG**A**TAA**TTG**ATCGATCC	1
2	GGATCG**GG**ACTGGCGAGAGCCAG**T**TAA**GCT**ATCGATCC	1
3	GGATCG**CT**ACTGGCGAGAGCCAG**A**TAA**GGT**ATCGATCC	1
4	GGATCG**CG**ACTGGCGAGAGCCAG**C**TAA**CTC**ATCGATCC	1
5	GGATCG**AT**ACTGGCGAGAGCCAG**C**TAA**GGA**ATCGATCC	1
6	GGATCG**TA**ACTGGCGAGAGCCAG**A**TAA**GTT**ATCGATCC	1
7	GGATCG**GG**ACTGGCGAGAGCCAG**G**TAA**CGG**ATCGATCC	1
8	GGATCG**AG**ACTGGCGAGAGCCAG**T**TAA**TTT**ATCGATCC	1
9	GGATCG**TA**ACTGGCGAGAGCCAG**G**TAA**GTG**ATCGATCC	1
10	GGATCG**GC**ACTGGCGAGAGCCAG**G**TAA**GAG**ATCGATCC	1
11	GGATCG**GT**ACTGGCGAGAGCCAG**T**TAA**GTT**ATCGATCC	1
12	GGATCG**TA**ACTGGCGAGAGCCAG**G**TAA**GTC**ATCGATCC	1
13	GGATCG**CA**ACTGGCGAGAGCCAG**C**TAA**TGG**ATCGATCC	1
14	GGATCG**TT**ACTGGCGAGAGCCAG**C**TAA**TAG**ATCGATCC	1
15	GGATCG**CT**ACTGGCGAGAGCCAG**G**TAA**CAG**ATCGATCC	1
16	GGATCG**AG**ACTGGCGAGAGCCAG**T**TAA**CAT**ATCGATCC	1

Table 4 - 4. Sequencing results from library 3.

| ID	Sequence (5'-3')	Counts
1	GGATAACGACTGGCTAGATTCAGGTAACGAGTCGTTTC	1
2	CGATGGCGACTTATGAGAGCTAGGGAAAGAATCGAGCC	2
3	GGACCGCGACTGGCGACAGCCAGGTAACTAATGGATCG	1
4	GGATCGCGACTGGCGGGAGGTAGCTAACGAATCGATCC	1
5	GGATCACGACTGGCCAGATTCAGGGAGCGAATTGATGG	1
6	GGCTAGCGAAGGGCTAGACCCAGGTCACGATTCAACCT	2
7	GAATCACGCGTGGCGAGAGCCAGGTCGCGAGTCGGTAC	1
8	TGATTGCAAATGACGCGAGACAGGTGTCGACTCGATCT	1
10	GGAAGGCGACTGCCTAAAGTCAGCTAACTAATCGGTGC	1
11	GCATCGCGAGTGGCGATAGCCCGGTAACGGATAGTTTC	1
12	GCGTCGTTACTGACGAAAGCTAGGTAACGAAGCGATCC	1
14	GGTTCACCACTGGCGAGAGCCAGGTGACGATTCGATCC	1
15	GGACCGCGATTTTCGAGTGCCAGGTAACAGCCTGATCC	1
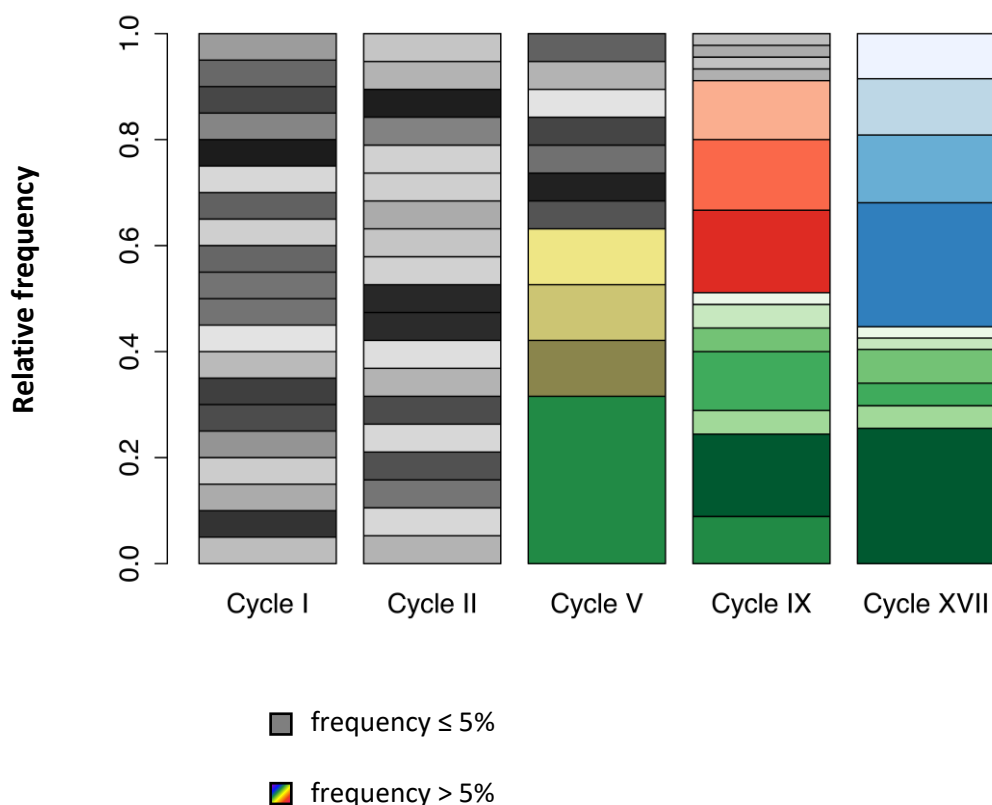
Figure 4 - 8. Sequence accumulation on selection carried with the first strategy. The relative frequency of each sequence listed in Table 4 - 2 was plotted in this color-coded chart. Grey-scale sequences showed a frequency below 5%, while each colored bar represents a sequence with a statistically more relevant frequency. An accumulation of repeated sequences is observed from cycle V onwards and some of the sequences were found in more than one of the cycle analyzed.

(Table 4 – 3) and just a couple were repeated twice for library 3 (Table 4 – 4).

Among the library tested and the relative selection strategy, the strategy based on a completely random library showed an accumulation of some sequences from a random pool of single-stranded DNA oligonucleotides (Figure 4 - 8) although none of the most frequent sequences showed a clear binding effect, nor a fluorescence increase. The results collected suggest that a different strategy of selection would be more appropriate for such an aptamer. The procedure exposed here relied on the fact that the sole ability for binding would be a sufficient prerequisite for the increase of the fluorescence quantum yield, as it occurred for the RNA aptamer selection.[120] It is likely that the selection procedure was not sufficiently stringent, thus resulting in the accumulation of weak malachite green binders that were not able nor to increase the fluorescence of the fluorophore nor to compete with the binding of the RNA aptamer.

Chapter 5

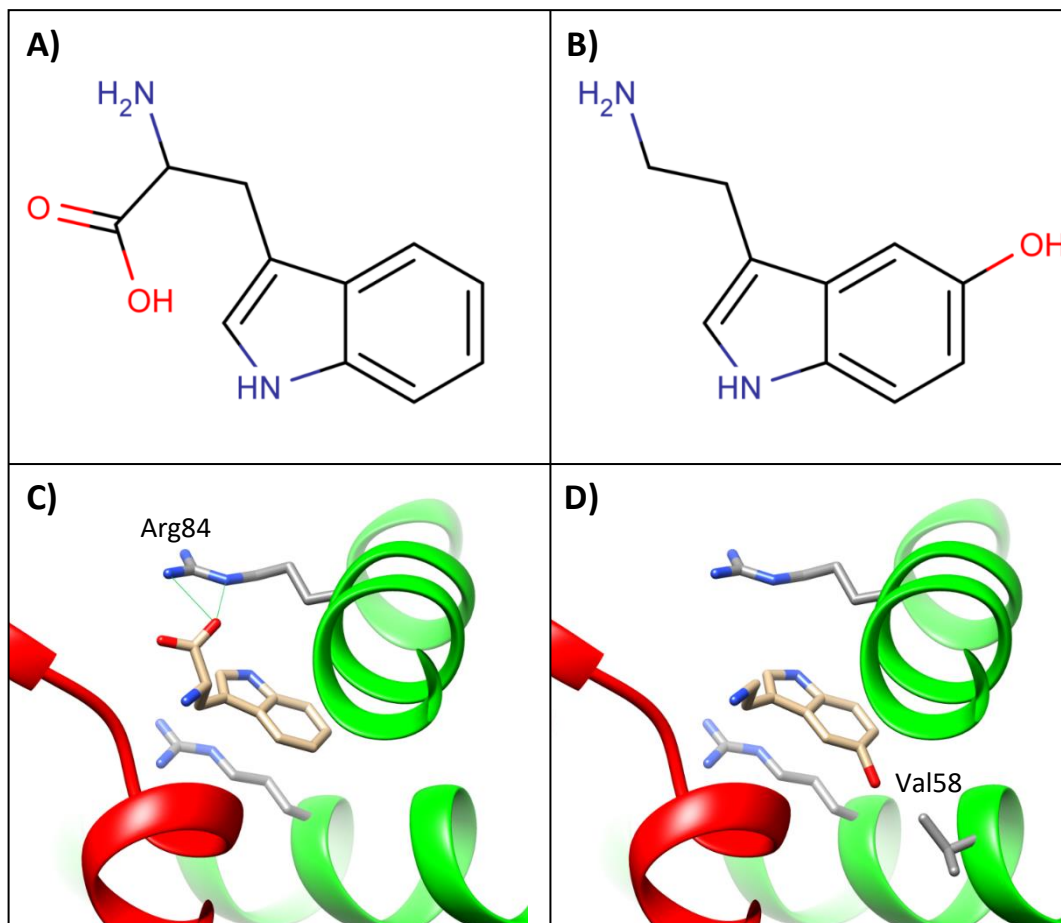**Engineering TrpR to sense the neurotransmitter serotonin**

Figure 5 - 1. Comparison of the two ligands. Chemical structures of A) tryptophan and B) serotonin. C) Interactions within the binding pocket of tryptophan and D) putative interactions of serotonin. Arg84 interacts with -COOH group of tryptophan, while Val58 is very likely to be close to the -OH group joint to carbon ζ3 of serotonin.

All of the genetic circuits described in chapter 2 were intended at establishing a solid genetic circuit to possibly activate a communication pathway. As discussed in the introducing chapter, the induction of a genetic switch in response to a chemical signal from a natural cell adds extra value to the system. An artificial cell based on *E. coli* TX/TL system can be used to sense the presence of prokaryotic chemical messages because it is possible to reproduce the exact molecular mechanisms occurring *in vivo*: all of the elements exploited in live cells are most likely present also in the *E. coli* TX/TL reaction. Conversely, the same kind of artificial cell may not be appropriate to reproduce the mechanisms occurring in eukaryotes. A good solution to this problem could rely on the possibility to engineer some of the macromolecules of the TX/TL machinery to actively interact with molecules released by eukaryotic cells.

The work described in this chapter began with the observation that the neurotransmitter serotonin has a very similar structure to the amino acid tryptophan (Figure 5 - 1 A, B). As discussed in chapter 2, tryptophan is involved in a very well described regulatory mechanism, demonstrated to be functional also *in vitro*, therefore it

can be exploited to construct a genetic circuit regulated by the neurotransmitter. The structure of the repressor TrpR was extensively studied and many details of the interactions with the ligand are available.[130,131] What follows in this chapter is a summary of some strategies to engineer TrpR on an attempt to change the affinity from tryptophan to serotonin.

5.1 Materials and methods

5.1.1 Reagents and general supplies

Serotonin hydrochloride was purchased from Sigma

5.1.2 Plasmids and cloning

All of the plasmids were assembled with the Gibson method[88] as described in chapter 2. DNA primers and the synthetic gene TrpR_S30 were purchased at Eurofins MWG.

Table 5 - 1. Relevant sequences used in this section. TrpR_S30 was synthesized by Eurofins Genomics MWG; pRNA sequence, containing malachite green aptamer, was sent by Prof. Friedrich C. Simmel from the Technical University of Munich; promoter pDC135A arose as a spontaneous mutation while cloning DNA plasmids listed below.

ID	Sequence (5'-3')
TrpR_S30	CTAACTTACATTAATTGCGTTGCGCTCATTGACAGCTAGCTCAGTC CTAGGGATTGTGCTAGCTACTAGAGTCACACAGGAAAGTACTATGG CTAGCATGATGGCCCAACAATCACCCTATTCAGCAGCGATGGCAG AACAGCGTCACCAGGAGTGGTTACGTTTTGTCGACCTGCTTAAGAA TGCCTACCAAAACGATCTCCATTTACCGTTGTTAAACCTGATGCTG ACGCCAGATGAGCGCGAAGCGTTGGGGACTCGCGTGCGTATTGTC GAAGAGCTGTTGCGCGGCGAAATGAGCCAGCGTGAGTTAAAAAAT GAACTCGGCGCAGGCATCGCGACGATTACGCGTGGATCTAACAGC CTGAAAGCCGCGCCCGTCGAGCTGCGCCAGTGGCTGGAAGAGGT GTTGCTGAAAAGCGATTGATAATACTAGAGCCAGGCATCAAATAAA ACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGT TTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGG GTGGGCCTTTCTGCGTTTATATACTAGAGGCTGTTGACAATTAATC ATCGAACTAGTTAACTAGTACGCAAGTTCACGTACTAGAGAAAGAG GAGAAATACTAGATGGCTTCCTCCGAAGAC
pTac_mRFP1_pRNA	GTTGACAATTAATCATCGGCTCGTATAATGTGTGGCCCCTCTAGAA ATAATTTTGTTTAAAAGAGGAGAAATACTAGATGGCTTCCTCCGAAG ACGTTATCAAAGAGTTCATGCGTTTCAAAGTTCGTATGGAAGGTTC CGTTAACGGTCACGAGTTCGAAATCGAAGGTGAAGGTGAAGGTCG TCCGTACGAAGGTACCCAGACCGCTAAACTGAAAGTTACCAAAGGT GGTCCGCTGCCGTTCGCTTGGGACATCCTGTCCCCGCAGTTCCAG TACGGTTCCAAAGCTTACGTTAAACACCCGGCTGACATCCCGGACT

	ACCTGAAACTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGTTAT
	GAACTTCGAAGACGGTGGTGTTGTTACCGTTACCCAGGACTCCTC
	CCTGCAAGACGGTGAGTTCATCTACAAAGTTAAACTGCGTGGTACC
	AACTTCCCGTCCGACGGTCCGGTTATGCAGAAAAAACCATGGGTT
	GGGAAGCTTCCACCGAACGTATGTACCCGGAAGACGGTGCTCTGA
	AAGGTGAAATCAAAATGCGTCTGAAACTGAAAGACGGTGGTCACTA
	CGACGCTGAAGTTAAAACCACCTACATGGCTAAAAAACCGGTTCAG
	CTGCCGGGTGCTTACAAAACCGACATCAAACTGGACATCACCTCC
	CACAACGAAGACTACACCATCGTTGAACAGTACGAACGTGCTGAA
	GGTCGTCACTCCACCGGTGCTTAAGGGAGAATGCGGCCGCCGAC
	CAGAATCATGCAAGTGCGTAAGATAGTCGCGGGTCGGCGGCCGCA
	TAAAAATTGTCATGTGTATGTTGGGCGCAGGACTCGGCTCGTGTAG
	CTCATTAGCTCCGAGCCGAGTCCTCGAATACGAGCTGGGCACAGA
	AGATATGGCTTCGTGCCCAGGAAGTGTTCGCACTTCTCTCGTATTC
	GATTGCGCCCACATACTTTGTTGAGGATCCCGACTGGCGAGAGCC
	AGGTAACGAATGGATCCTCAATCATGGCAA
pDC135A+5'UTR	TTGACAGCTTGCTCAGTCCTAGGGATTGTGCTAGCTACTAGAGTAA
	CACAGGAAAGTACT

Table 5 - 2. Plasmids used in this section and relative cloning strategies.

Plasmid ID	Backbone	Insert	Source	Cloning strategy	
				Primers	Template
DC124A	DC024A	pJ23117_TrpR_pTet _mRFP1_ spinach		DC101/ DC111	DC032A
				DC257/ DC258	TrpR_S30
DC135A	DC024A	pDC135A_TrpR_ pTet_mRFP1_ spinach		DC277/ DC278	DC124A
GB008A	DC024A	pTac_mRFP1_ pRNA	Mansy Lab		
LG013A	DC024A	pDC135A_TrpR_ pTet_mRFP1_pRNA		DC275/ DC276	DC135A
				DC307/ DC306	GB008A
LG011A	DC024A	pDC135A_EsaR_ pTet_mRFP1_pRNA		DC300/ DC302	FC033A
				DC299/ T9002g FW	LG013A

Table 5 - 3. TrpR mutants and plasmids encoding the sequences. Two type of constructs were created to be used in S30 or in PURE system reactions. The difference is indicated in brackets and refers to the promoter used for the expression of the mutants and the *in vitro* TX/TL reaction.

Mutations	Mutant ID	Mutant ID (DC135A, S30)	Mutant ID (T7, PURE)
V58N, R84L	TrpR*001		DC091A
V58N, I82A, R84L	TrpR*002		DC093A
V58N, R84L, S86A	TrpR*003		DC077A
V58N, I82A, R84L, S86A	TrpR*004		DC094A
T44A, V58N, R84L	TrpR*005		DC096A
T44A, V58N, I82A, R84L	TrpR*006		DC098A
T44A, V58N, R84L, S86A	TrpR*007		DC099A
T44A, V58N, I82A, R84L, S86A	TrpR*008		DC100A
R84L	TrpR*009		DC084A
V58N	TrpR*010		DC082A
T44A	TrpR*011		DC081A
I82A, R84L	TrpR*012		DC086A
R84L, S86A	TrpR*013		DC087A
I82A, R84L, S86A	TrpR*014		DC088A
I82A	TrpR*015		DC083A
S86A	TrpR*016		DC085A
T44A, V58N	TrpR*017		DC089A
V58N, I82A	TrpR*018		DC090A
V58N, S86A	TrpR*019		DC092A
T44A, V58N, I82A	TrpR*020		DC095A
T44A, V58N, S86A	TrpR*021		DC097A
V58D	TrpR*022		DC101A
V58Q	TrpR*023		DC102A
V58E	TrpR*024		DC103A
R84I	TrpR*025		DC104A
R84V	TrpR*026		DC105A
R84F	TrpR*027		DC106A
R84M	TrpR*028		DC107A
L41A	TrpR*029		DC108A
M42A	TrpR*030		DC109A
L43A	TrpR*031		DC110A
P45A	TrpR*032		DC111A
R54A	TrpR*033		DC112A

I57A	TrpR*034		DC113A
T81A	TrpR*035		DC114A
T83A	TrpR*036		DC115A
N87A	TrpR*037		DC116A
S88A	TrpR*038		DC117A
V58S, R84W	TrpR*043	LG001A	
V58N, R84V	TrpR*044	LG002A	
V58T, R84I	TrpR*045	LG003A	
V58T, R84V	TrpR*046	LG004A	
V58Q, R84I	TrpR*047	LG005A	
V58Q, R84F	TrpR*048	LG006A	
w.t.		LG013A	DC076A
EsaR		LG011A	DC125A

5.1.3 In silico *analysis of TrpR*

The files 1TRR.pdb and 2QEH.pdb were downloaded at www.rcsb.org and used for structure analysis with the software UCSF Chimera, downloaded at www.cgl.ucsf.edu/chimera. The file 1TRR contains the atomic coordinates of the *E. coli* holorepressor TrpR bound to the operator sequence on DNA while 2QEH contains the atomic coordinates of D7r4, a salivary biogenic amine-binding protein from the malaria mosquito *Anopheles gambiae* bound to serotonin. 1TRR was used as the basis for the analysis of the binding pocket and the characterization of the mutants and 2QEH was used for the atomic coordinates of serotonin. The closest residues to the ligand were identified by restricting the visualized amino acids in a radius within 5 Å from the ligand.

The analysis of the mutants was carried out by Dr. Luca Belmonte, a postdoctoral fellow in our research group. The wild type protein structure was first desolvated and deprived of the ligand, then mutants were generated with the plugin VMD that introduces point mutations in a user-defined position and to find the best rotamer for the lateral chain. Then water molecules were added again together with sodium and chloride ions to neutralize the charge of the protein, energy minimization was performed and a short molecular dynamics was run. All these steps were done with Yasara and these "relaxed" structures were subjected to docking with the software Vina to evaluate the interaction between the two ligands, tryptophan and serotonin, with the binding pocket area.

5.1.4 In vitro *gene expression and fluorescence measurements*

PURE system reactions were assembled in a final volume of 3 µl with the components indicated in § 2.1.6 mixed together at the same proportion. After 4 h incubation at 37 °C, the mixtures were diluted to 9 µl with HEPES buffer (50 mM HEPES, 10 mM $MgCl_2$, 100 mM KCl, pH 7.6). S30 reactions were assembled as described in § 2.1.7, in a final volume of 10.5 µl. Each of the mutants was tested in absence or in presence of serotonin that was added to a final concentration of 10 mM and the same condition was tested in duplicate in PURE system reactions and in triplicates in S30. Transcription levels were monitored by different fluorophores with the real-time PCR cycler Rotor-Gene Q. Translation levels were monitored by the expression of the fluorescent protein mRFP1 on channel Orange (excitation: 585 ± 5 nm; emission: 610 ± 5 nm). Transcription levels were monitored through spinach aptamer bound to DFHBI on channel Green (excitation: 470 ± 10 nm; emission: 510 ± 5 nm) in PURE system and through malachite green aptamer bound to malachite green on Red Channel (excitation: 625 ± 5 nm; emission: 660 ± 10 nm) in S30 reactions. For PURE system reactions, only the end point was measured, while for S30 reactions the whole kinetic was followed for ~20 h with a measurement every 20 s for the first 20 min and then every 5 min. Raw fluorescence data were converted in molar concentrations as described in § 2.1.8. In addition to the standard curves created for mRFP1 and spinach aptamer with DFHBI, a standard curve for the malachite green signal was created using malachite green aptamer synthesized as described in § 4.1.2 added to different concentrations (0, 0.61225, 1.2245, 2.449 µM) to a solution containing 10 µM malachite green and HEPES buffer.

5.1.5 In vivo *gene expression and fluorescence measurements*

E. coli TOP10 cells were transformed as indicated in chapter 2 with plasmids carrying TrpR mutations. One colony for each was picked and grown overnight in 5 ml of LB medium supplemented with 100 µg/ml ampicillin. Then bacteria were reinoculated with a starting dilution of 1:1000 until they reached an absorbance at 600 nm = 0.1-0.2. 100 µl of culture were transferred to 4 wells of a 96 plate, to 2 of them serotonin was added to 4 mM and the plate was incubated at 37 °C for 24 hours. Fluorescence levels were measured on a Tecan Plate reader with the following parameters: excitation 577 nm, emission 607 nm, gain 75.

5.2 Results

TrpR is produced as an inactive aporepressor that dimerizes in a conformation able only for weak and aspecific interactions with DNA. Tryptophan forces a conformational change in the repressor by the simultaneous interactions with two monomers. Two symmetric binding pockets form specular interactions between two molecules of tryptophan and two monomers of TrpR (Figure 5 - 2).[130–133] One of the first characterization of the binding sites was conducted by affinity tests on tryptophan analogs.[130] From those data it was deduced that the indole group of tryptophan is trapped in a hydrophobic pocket composed of glycine 85 and two arginine residues, 54 and 84 (Figure 5 - 3 A). The interaction is further stabilized by hydrogen bonds both on carboxylic and aminic groups of the molecule. Arg84 plays a double role by interacting with the carboxylic group together with Thr44, while Ser88 and backbone chains of Leu41 and 43 stabilize the aminic group. Many of these interactions can be easily checked on the crystal structure of the holorepressor deposited on pdb under the name 1TRR.[131] Some of them are not visualized, like the hydrogen bonds set by Ser88 and Thr44, but many amino acids are enclosed in a 5 Å radius from the ligand even if no active role in binding was described for them.
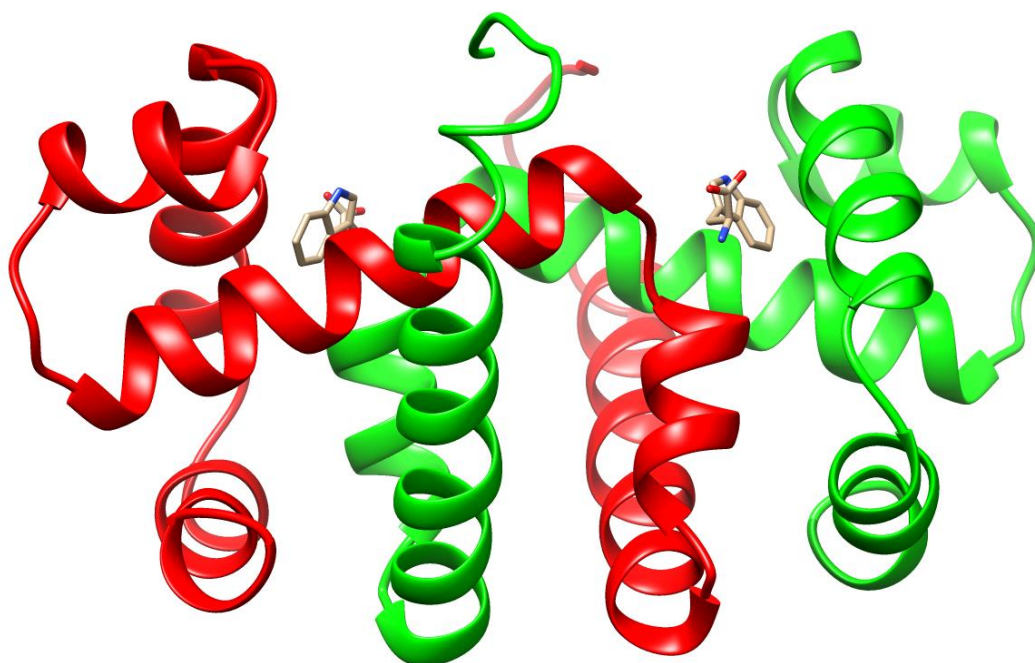

Figure 5 - 2. 3D structure of the holorepressor TrpR. Two monomers (indicated by red and green colors) interact with each other through the ligand. Two symmetric binding pockets are formed by the cooperation of the two monomers, some amino acids involved in the binding belong to one chain, while some to the other chain.
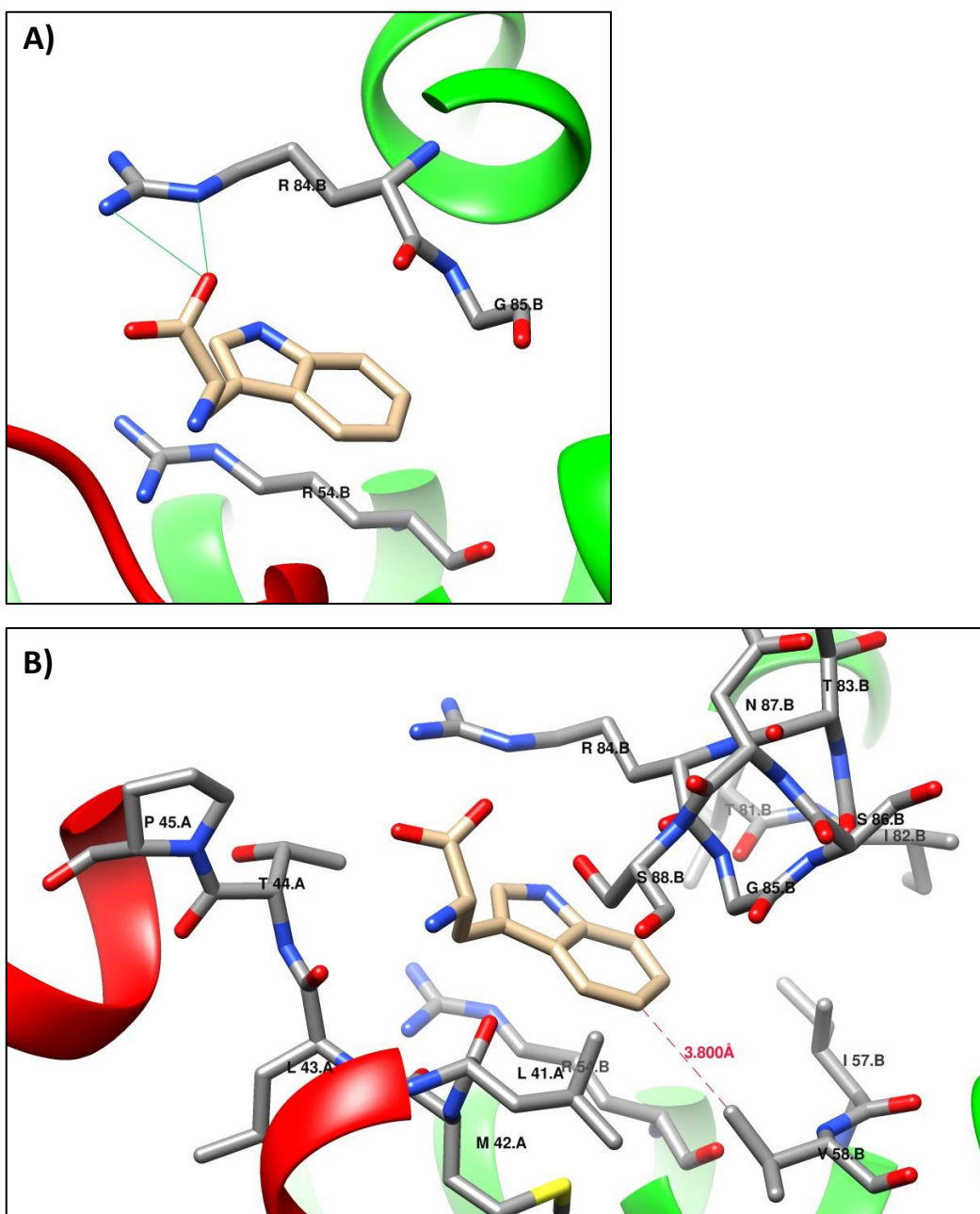
Figure 5 - 3. Details on interactions of tryptophan within the binding pocket. A) The indole group is stabilized by hydrophobic interactions with Arg54, Arg84 and Gly85. The carboxylic group sets hydrogen bonds with Arg84. B) The distance between carbon from Val58 and carbon from tryptophan is highlighted in the figure and makes position 58 a good candidate for a stabilization with the -OH group joint to carbon ζ3 of serotonin. Amine group of tryptophan interacts with backbone chain of Leu41 and 43.

A putative binding pocket was then defined with the following residues: Leu41, Met42, Leu43, Thr44, Pro45 from one monomer and Arg54, Ile57, Val 58, Thr81, Ile82, Thr83, Arg84, Gly85, Ser86, Asn87, Ser88 from the other (Figure 5 - 4). The hypothetical binding pocket had to be defined better to include only the residues that are strictly necessary for ligand binding. To this end, a set of mutants was designed to carry

individual substitutions of the residues of the putative binding pocket and was tested for a loss in repressor activity. Mutants were tested in a genetic circuit in the PURE system, similarly to what was described for the repressors in chapter 2 (Figure 5 - 5 A). A plasmid carried the sequence of TrpR mutants under the control of a constitutive T7 promoter and the sequence of the reporter gene *mRFP1-spinach* under the control of a T7 promoter fused to TrpR operator sequence. Gene expression was monitored both at transcription and translation levels in order to visualize any possible loss in repressor activity. In presence of an active repressor, the reporter gene cannot be expressed, while inactive repressors are associated with a good expression of the reporter gene. Together with the mutants the wild type repressor and an unrelated transcriptional repressor were added to the experiment as controls. The control repressor EsaR used for this experiment belong to the plant pathogen *Pantoea stewartii*[134] and should not interact with TrpR operator sequence. Most of the amino acids were substituted with alanine but a couple of residues were substituted with some other that may facilitate the interaction with serotonin. A deeper analysis of the interactions between the ligand and the amino acids of the repressor was carried in order to define some mutations that could help the switch for ligand affinity, from tryptophan to serotonin.
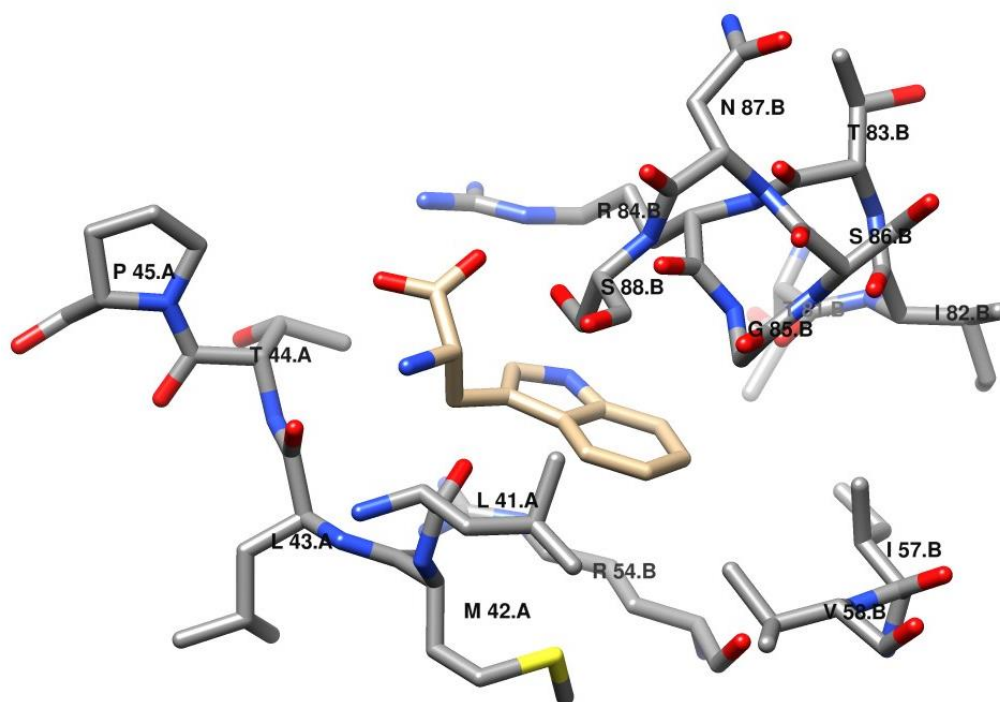

Figure 5 - 4. Residues within a 5 Å radius from the ligand. A putative binding pocket for tryptophan was defined by analyzing the residues enclosed in 5 Å radius from the ligand: Leu41, Met42, Leu43, Thr44, Pro45 from one monomer (A) and Arg54, Ile57, Val 58, Thr81, Ile82, Thr83, Arg84, Gly85, Ser86, Asn87, Ser88 from the other monomer (B).

The differences between the tryptophan and serotonin consist of a carboxyl group of tryptophan and a hydroxyl group joined to carbon ζ3 of the indole group of serotonin (Figure 5 - 1). As discussed above, Arg84 interacts with the carboxylic group of tryptophan, therefore hydrophobic residues such as leucine, isoleucine and valine were taken into consideration as well as tryptophan, more likely to forbid interactions with the carboxylic group by steric hindrance. On the other side of the molecule, the closest amino acid to the carbon ζ3 is Val58, therefore this residue is a possible interactor to the hydroxyl group of serotonin. The neurotransmitter has a very high pK_a,[135] so the hydroxyl group is protonated at neutral pH, making position 58 suitable for amino acids with hydrogen bonds acceptor atoms in their side chain.

The data collected on single mutants tested in genetic circuits in PURE system reactions (Figure 5 - 5) showed that the amino acids actively involved in the binding can be restricted to Arg54, Ile57, Thr81, Ile82, Arg84, and to a less extent also Val58 and Thr83 that reported some effects when mutated. All of these mutants lost their ability to repress the reporter gene. The fact that mutations on Leu41 and 43 do not induce a great effect on ligand binding indicates that these single mutations are not affecting the overall folding of the repressor so that the interactions described to occur with the backbone chain are not affected. The addition of serotonin showed some reduction in gene expression only for some of the mutants, while other did not show any effect (Figure 5 - 6). Some mutants – such as the one carrying T44A, V58N and I82A or the one carrying T44A, V58N, S86A – showed a good difference of gene expression in presence or absence of serotonin. Nonetheless, it is not possible to find a clear pattern among these data that can help understanding which mutants favor the binding of serotonin. As it can be clearly observed from translation data, in fact, the levels of repression observed for the mutants in presence of serotonin are never close to what is shown by the wild type repressor in presence of the natural ligand, tryptophan. As discussed in chapter 2, one of the main reason why this is occurring involves the strength of the promoter. T7 RNA polymerase revealed to be a highly processive enzyme, although TrpR is very strong in its action. It may not be so easy to obtain a mutant as strong in gene repression as the original one, so the screening method had to be changed with S30 extract for the same considerations reported in chapter 2, for the genetic circuits based on T7 promoters. The substitution with a genetic circuit based on *E. coli* promoters resulted in a higher degree of control in gene expression.
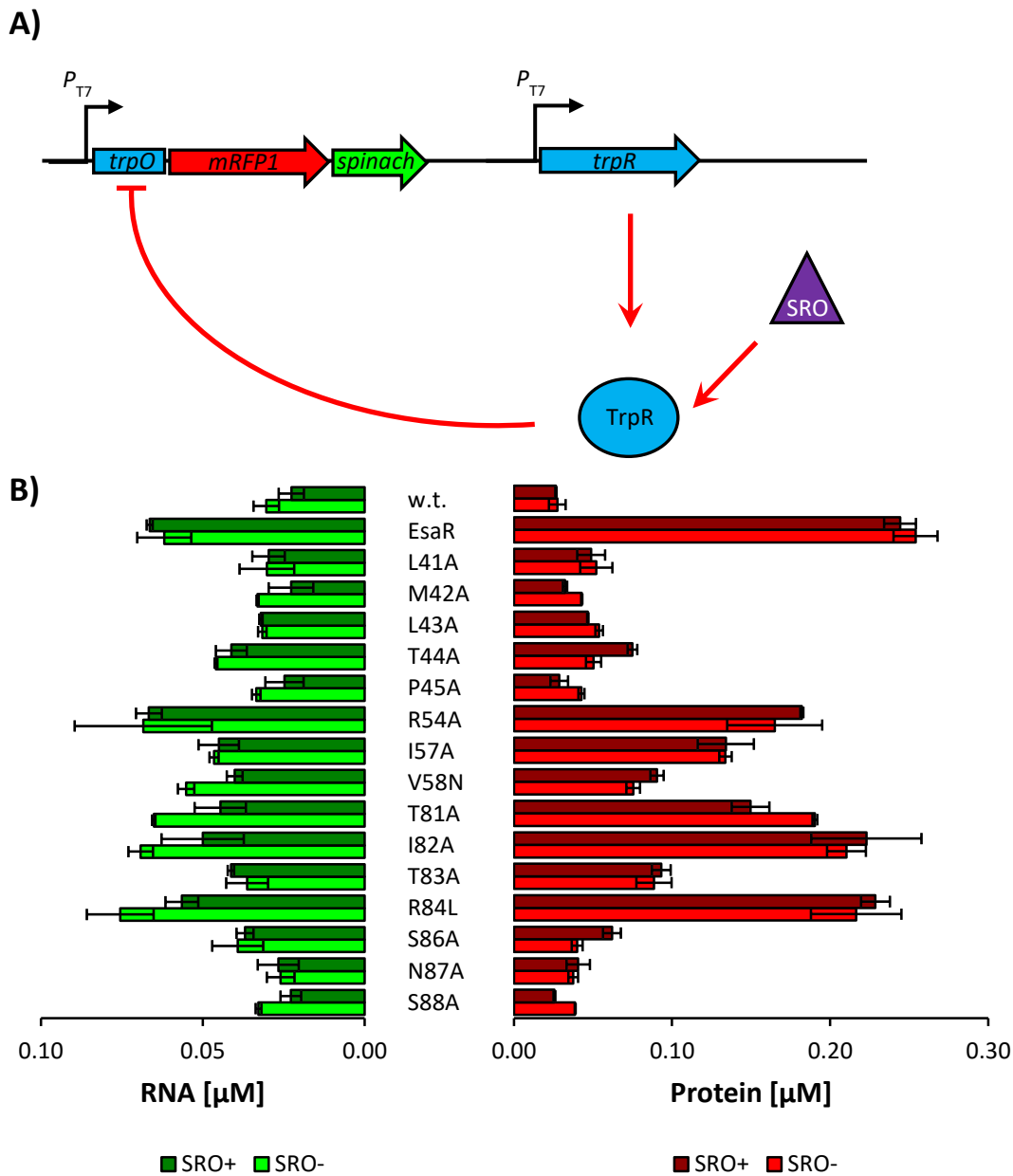
Figure 5 - 5. Single mutations on TrpR tested in the PURE system. A) Two genes were encoded in the same plasmid under the control of two separate T7 promoters. One gene coded for the TrpR mutants tested, while the other coded for the reporter gene *mRFP1* with spinach aptamer fused at 3'-UTR of the mRNA as described in chapter 2. The repressor and the mutants are constitutively expressed, while the promoter upstream of the reporter gene is fused to the operator sequence for TrpR repressor. B) Transcription (green bars) and translation (red bars) levels of the reporter gene regulated by TrpR single amino acid mutants are reported, both in the presence (dark bars) or in the absence (light bars) of serotonin. Tryptophan concentration inside the PURE system is enough to induce repression as observed for the wild type repressor. Raw fluorescence values of both spinach aptamer bound to DFHBI and mRFP1 were converted in molarity using the standard curves described in § 2.1.8. The chart reports technical duplicates for each mutant.
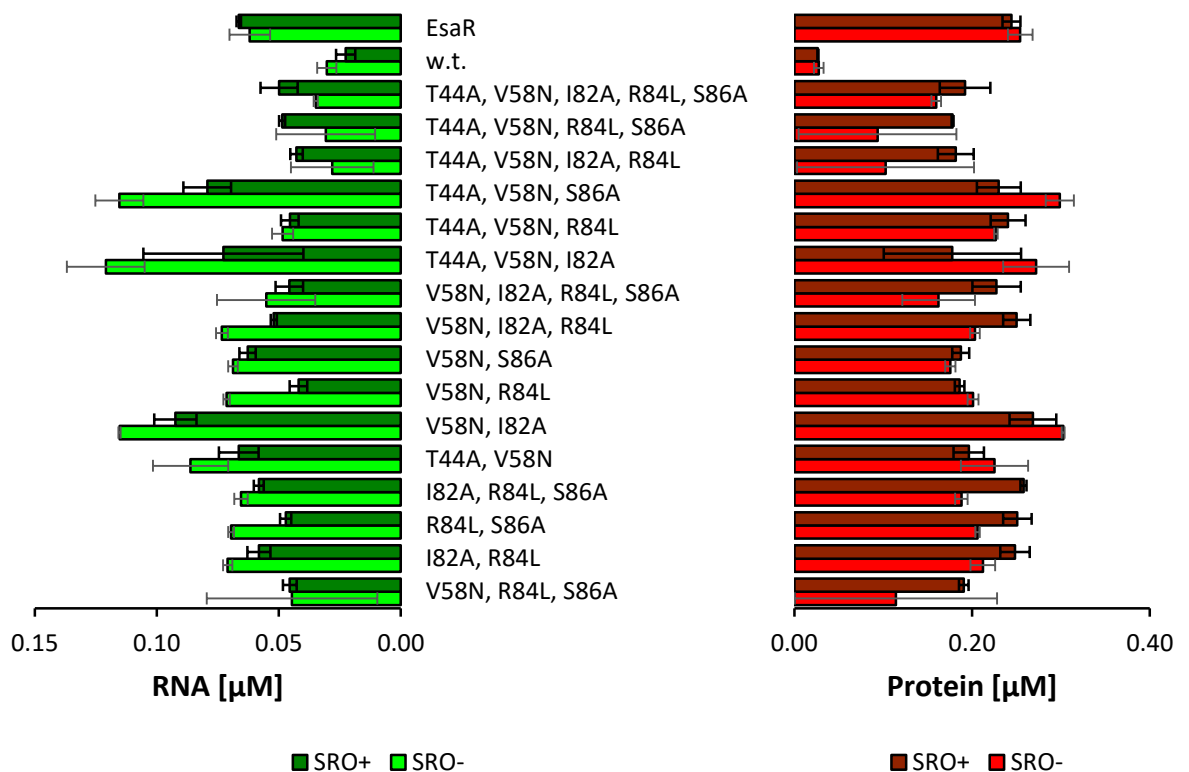
Figure 5 - 6 Some mutants screened in the PURE system for the ability to bind serotonin showed no specific interactions with serotonin. Some TrpR mutants were designed to carry mutations on amino acids close to tryptophan, according to the analysis carried on the crystal structure of TrpR described in Figure 5 - 4. The mutants were cloned in the construct depicted in Figure 5 - 5 A and tested in the PURE system in the presence (dark bars) or absence (light bars) of serotonin. Both transcription (green bars) and translation levels (red bars) were monitored and raw fluorescence data were analyzed as in Figure 5 – 5. Some slight reduction in gene expression was observed for few genes, but the levels are not close to what was observed for the wild type.

As discussed earlier, *E. coli* promoters can be more easily tuned *in vivo* or *in vitro* in a crude cell extract. A new genetic circuit was designed taking inspiration from plasmid BBa_K560000 available from the registry of standard biological parts. This part encodes for a hybrid protein derived by the fusion of TrpR DNA binding domain and the photoactive domain of a plant protein to create an optogenetic repressor.[136] The promoter sequence and 5' UTR of both the reporter gene and repressor were kept as described in the registry of standard biological parts. The *E. coli* promoter controlling the expression of the repressor belongs to a series of mutants described by Prof. J. Anderson in the registry of standard biological parts. Considering that the sequence used in this construct (J23117) is reported to have poor gene expression, a stronger promoter sequence was tested in parallel (J23106). The strength of the promoter was chosen to correctly balance the resources for the synthesis of the two genes, the repressor and the reporter. A too strong

promoter would sequester all the resources for the expression of the repressor gene, thus leading to a low expression of the reporter gene also in presence of a weak repressor. During the mutagenesis of the promoter, a random mutation appeared both in the promoter and in the RBS. As these sequences are not described by Anderson, the mutation was assigned with a new name as the plasmid containing it, DC135A. The wild type (w. t.) repressor showed a good activity so the mutants were successively cloned under the same promoter and tested for gene repression.

Table 5 - 4. Docking results of the 6 mutants most capable of switching ligand affinity from tryptophan to serotonin. The table reports the minimum energy values for the binding of 6 mutants coming from the combinatorial analysis of mutants in positions 58 and 84. Among the 56 mutants screened, these 6 reported a lower or equal energy for the binding of serotonin when compared to tryptophan.

ID	Mutant	Energy serotonin (kcal/mol)	Energy tryptophan (kcal/mol)
TrpR*043	V58S, R84W	-6.0	-5.7
TrpR*044	V58N, R84V	-6.2	-5.9
TrpR*045	V58T, R84I	-6.4	-6.4
TrpR*046	V58T, R84V	-6.2	-6.2
TrpR*047	V58Q, R84I	-6.9	-5.5
TrpR*048	V58Q, R84F	-5.9	-5.7

In S30 reaction, spinach aptamer could not be used to monitor transcription levels because of some background fluorescence deriving from the reaction mix. So the 3' UTR of the reporter gene was replaced to contain a malachite green aptamer, shown by Noireaux as a good alternative for monitoring transcription levels in S30 reactions.[137] The behavior observed in RNA kinetics indicated a high degradation rate within 2 h, that is expected for the reasons discussed in chapter 1 regarding the presence of nucleases inside of the cell extract. Conversely to the data reported for the PURE system, transcription levels do not refer to the end-point reactions but indicate the maximum amount of RNA obtained before the degradation began.
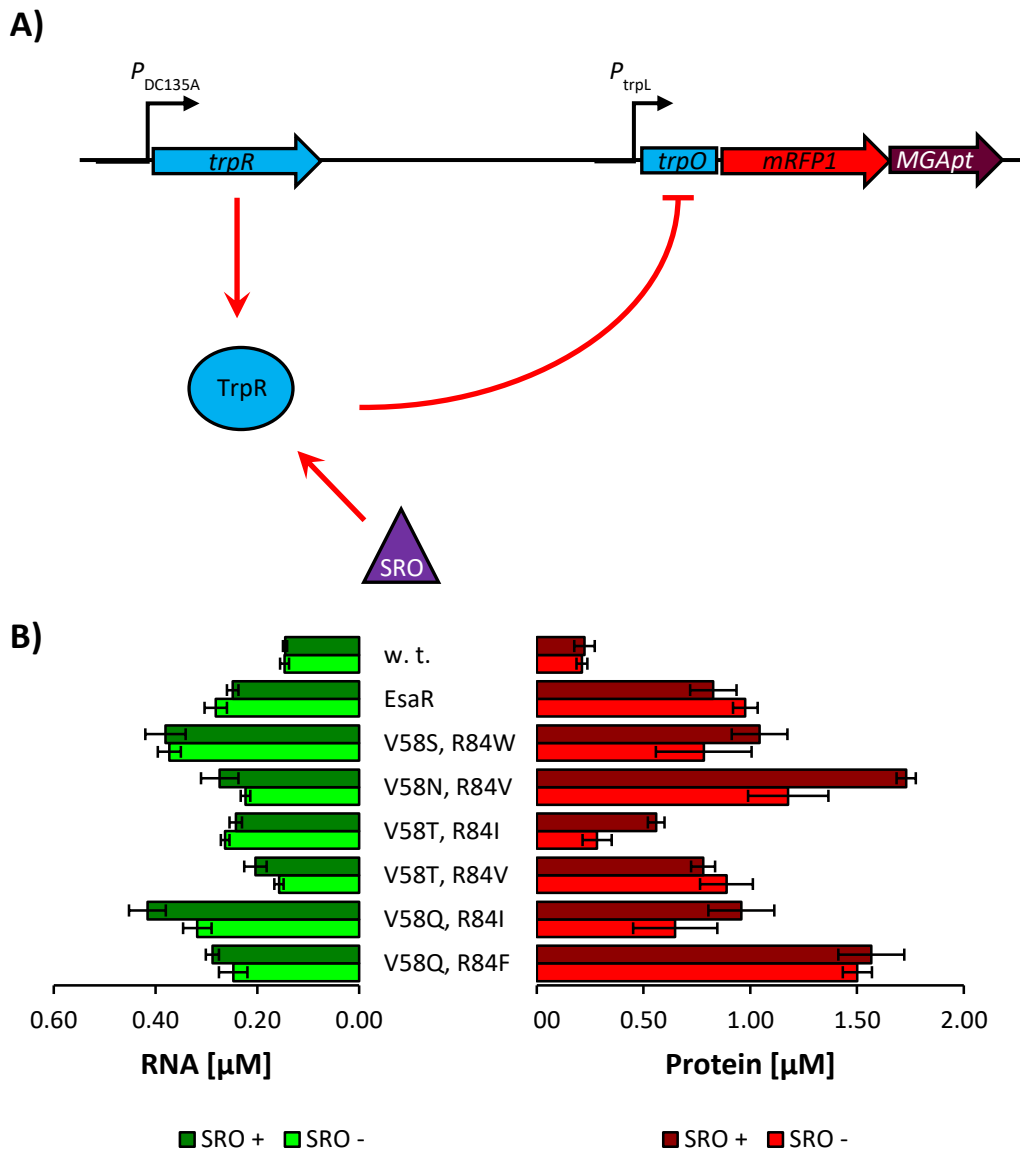
Figure 5 - 7. Test of the 6 TrpR mutants predicted to change affinity from tryptophan to serotonin in the S30 extract. 56 possible combinations of TrpR mutants carrying mutations only in position 58 and 84 were tested *in silico* for a change in affinity from tryptophan to serotonin. The 6 mutants with the best affinity for serotonin were cloned in the genetic construct schematized in A) and tested for the ability to repress gene expression. A) A constitutive *E. coli* promoter drives the expression of TrpR mutants and the activity of the repressor is measured with a reporter gene under the control of the promoter P_{trpL}. As the detection of the signal from spinach aptamer and DFHBI was not possible in S30 reaction, spinach aptamer at 3'-UTR of the reporter gene was replaced by the malachite green aptamer. B) The plasmids carrying the 6 mutants were tested in S30 reaction and transcription (green bars) and translation (red bars) levels of the reporter gene were measured both in the presence (dark bars) and in the absence (light bars) of serotonin. Fluorescence values were analyzed as in Figure 5 – 5. These data confirmed that some of the mutants lost affinity for tryptophan, as the transcription and translation values were considerably higher than the control wild type repressor. Nonetheless, the addition of serotonin did not affect gene expression.

In this step, the attention was then focused on the two residues described earlier for a possible major role in ligand binding: Arg84 and Val58. A combination of the following amino acids was tested *in silico* for the ability of binding both the molecules, serotonin or tryptophan. For position 58 asparagine, glutamic acid, glutamine, aspartic acid, histidine, serine, threonine, tyrosine while for position 84 alanine, valine, isoleucine, leucine, phenylalanine, tryptophan and methionine. The analysis was based on the docking of the mutated proteins together with the two ligands, serotonin and tryptophan.

The solutions with the lowest energy values were taken into consideration and analyzed for the location of the ligand with respect to the binding pocket of TrpR. Among the 56 possible combinations, 6 showed promising behaviors: 1) Val58Ser, Arg84Trp; 2) Val58Asn, Arg84Val; 3) Val58Thr, Arg84Ile; 4) Val58Thr, Arg84Val; 5) Val58Gln, Arg84Ile; 6) Val58Gln, Arg84Phe. All of them reported a lower or equal energy when bound to serotonin rather than tryptophan (Table 5 - 4). These six mutants were tested for gene repression in presence or absence of serotonin in S30. All of the six mutants lost their affinity for the original ligand, but none of the mutants showed any specific decrease in gene expression in presence of serotonin (Figure 5 - 7).

Taken together all these data suggest that two positions cannot be enough to find a proper mutant able to switch the affinity for a new ligand. It was found that eight positions are strictly necessary for the binding of tryptophan but testing all the possible combinations *in silico* (20^8 = 2.56E10) would require too much time. A good solution to this problem could be given by a directed evolution of the protein, carried *in vivo* or *in vitro*. A strategy could be designed to have all the possible combinations of mutants mixed together and subjected to a selection procedure where only the desired mutant(s) would survive.

Chapter 6

Conclusions

The work presented in this thesis highlights the major advantages and limitations faced by some kinds of artificial cells when used for communication with natural cells. One of the big interest in biotechnology is the engineering of natural cells for the interaction with other cells with the aim of developing treatments against pathogens or tumors. Despite the promising results obtained, these strategies rely on the use of living organisms that can always be subjected to possible side effects not completely controllable. Alternatively, cellular mimics have the potential to reduce the number of non-specific interactions because of a reduced complexity. A system assembled from simple components of defined properties guarantees a higher degree of control but lacks some fundamental mechanisms that can affect the efficiency. While decreasing the power of control, the multiplicity of the elements composing natural cells allows for self-sustainment and energy regeneration resulting in higher efficiency compared to artificial cells. As a consequence, engineered cells can host multiple functions, while cellular mimics are strongly limited.

Rather than improving energy regeneration mechanisms, this thesis proposed an alternative strategy to overcome efficiency-related problems through the combination of different kinds of artificial cells. To date, several types of compartments can host enzymatic reactions or perform gene expression and the diverse setups offer different ranges of advantages and disadvantages. If combined together these artificial cells could potentially compensate reciprocal defects and result more efficient in communication with natural cells. A mixed population of artificial cells may find an application when, for example, a target message delivered from a natural cell cannot permeate the membrane of a lipid vesicle and activate gene expression as a response to that. Rather than changing the permeability of the compartment that could possibly have negative effects on the encapsulated TX/TL reaction, it is possible to include a third player in the communication pathway to convert the chemical message into a molecule permeable to lipid membranes. Porous membranes, such as proteinosomes or colloidosomes, could host a cascade of enzymatic reactions to allow this conversion and bridging the gap between the target natural cell and the cellular mimic responsible for message delivery.

As a proof of concept, this thesis investigated the possibility to create communication pathways between two different kinds of cellular mimics: lipid vesicles carrying *in vitro* gene expression and proteinosomes able for catalysis. The designed communication pathway involved two chemical messages by means of which the two cellular mimics integrated both a sensing and a sending mechanism. Proteinosomes would be able to catalyze the synthesis of the quorum sensing molecule 3OC6 HSL from *V. fischeri*. The molecule would subsequently permeate the lipid vesicles and activate the expression of a pore-forming protein to release the encapsulated glucose that would serve

as a substrate for the catalytic activity of proteinosomes. The coupling with the dye amplex red would allow the monitoring of the communication pathway, through the enzymatic reaction catalyzed by glucose oxidase and horseradish peroxidase carried by proteinosomes.

The communication pathway was investigated for each step and the related problems were handled separately. The design of the whole pathway required an initial step dedicated to the improvement of the genetic switch responsible for the delivery of a chemical message. The combination of prokaryotic transcriptional repressor together with the constitutive promoter of the bacteriophage T7 resulted in a poor regulation because of a background gene expression in the repressed state. Consistent with a previous characterization of T7 RNA polymerase activity in the *in vitro* TX/TL reaction mixture PURE system,[36] this behavior was explained by the high processivity of the polymerase, that is it is difficult to eliminate completely transcription and only a small amount of transcript is needed to produce much protein. These factors allowed for a good decrease in transcription rates but not to the extent of completely decreasing protein expression. A better genetic regulation was found in *V. fischeri*[67] and was further investigated both for the regulation of a reporter gene and the activation of the pore-forming protein α-hemolysin from *S. aureus*. Gene regulation revealed to be tight enough to be used for the communication pathway.

Unfortunately, when this genetic circuit was integrated into lipid vesicles to regulate the release of glucose, it was difficult to monitor a specific signal related to the presence of the inducer molecule. The most plausible explanation for this behavior can relate to a reduced stability of lipid vesicles carrying TX/TL reactions and to the efficacy of encapsulation of a complete and active gene expression system. It has been recently reported how the stability of lipid-based compartments is affected by the presence of a high concentration of proteins, typical of *in vitro* TX/TL systems.[138] The heterogeneous mixture composing the S30 reaction may not be ideal for vesicles integrity that could not survive the purification procedures aimed at avoiding background catalytic activity in proteinosomes. More efforts should be put to find the proper conditions for the stable formation of liposomes able for gene synthesis. Some alternative lipid compositions are reported to overcome stability problems and, in particular, PEGylated lipids are described to shield the bilayer between lipid emulsion droplets from interactions with proteins.[138]

For time-related reasons, a deep characterization of the proper lipid compositions was not possible, therefore the communication pathway was changed in order to include engineered bacteria. The function carried by bacteria is based on genetic elements previously established,[9] making this third player very likely to fill the gap in the communication pathway between liposomes and proteinosomes. The decision to include

engineered cells in the communication pathway was forced by limitations due to the early stage of this research field. The number of established cellular mimics able to sense and/or respond to external stimuli is still limited, therefore it is necessary to integrate some missing functions with engineered cells. Nonetheless, a communication pathway involving both artificial and natural cells has the potential to show how it is possible to create synthetic communication pathways that can be used to establish efficient interactions with a target natural cell.

6.1 Future perspective

As discussed earlier, the improvement brought by artificial cells communities in communication with natural cells may have possible therapeutic applications in the fight against dangerous cells. The catalytic activity of glucose oxidase is used by several species of fungi to produce hydrogen peroxide and kill bacteria. The cooperation between artificial cells to activate the specific release of glucose could be used to kill bacteria. Cellular mimics could be further engineered to respond to chemical messages released by pathogens. Bacteria like *Pseudomonas aeruginosa* exploit quorum sensing pathways similar to those described for *V. fischeri*,[101] so it is possible to assemble an artificial cells community to activate a death pathway only in presence of the pathogen. One of the major problems given by *P. aeruginosa* infections is related to the biofilm formation, a process that protects the pathogen from the immune system and antibiotics. The early detection of quorum sensing molecules may prevent the formation of the biofilm by the activation of a killer communication pathway.

Although the number of possible applications is very likely to rapidly increase, the purpose of this PhD project was to describe a method rather than looking for possible applications. The creation of artificial cells consortia offers a modular approach where different cellular mimics cooperate for efficient communication pathways with target natural cells. Coherently to what is proposed here for artificial cells, it has been shown how engineered cells can be integrated into logic networks to increase the control of the response to an external stimulus.[139,140] Artificial cells networks could similarly combine multiple input signals and give a tight and efficient control of the output message. Artificial cells with different properties have the potential to expand the possible communication networks beyond what is possible through genetic engineering. The use of synthetic amphipathic polymers is helping the creation of novel cellular mimics able to respond to regulatory mechanisms different from gene expression or enzymatic activity.[141] The assembly of polymersomes with specific chemical properties can be used to create

compartments able to undergo specific cargo release induced by hydrolysis,[142] change in pH,[143,144] redox reactions,[145–148] small molecules,[149] light[150–153] and temperature.[154–156]

The integration of multiple communication pathways requires the absence of crosstalk to ensure a tight control on the response to stimuli coming from target cells. Orthogonality can be achieved by genetic modifications as reported for engineered cells modified to reduce the interference between natural occurring quorum sensing mechanisms.[157] As discussed earlier, the design of artificial cells networks is not constrained to genetic engineering and orthogonal communication pathways can be set by the use of different components. A recent example of orthogonal communication pathways between artificial cells was described between liposomes encapsulating two different kinds of *in vitro* gene expression systems: one based on *E. coli* extract and the other based on HeLa cells extract.[158] In this communication pathway the crosstalk was avoided by the use of different TX/TL machineries.

In order to perform a desired function, a consortium of artificial cell needs the cells to be always colocalized. The technologies under investigation for this purpose are based on microfluidics devices that allow a high degree of control on several features such as size distribution of the compartments of cellular mimics.[159–163] For these reasons, microfluidics offers a good platform for the assembly of multiple cellular mimics in more complex structures. The integration of a 3D printer with a microfluidic device allowed the connection of lipid droplets in a structure that was described as a synthetic tissue.[138] This tissue consisted of lipid droplets carrying gene expression systems connected with each other by lipid bilayers. The inducible activation of α-hemolysin pores connects the lumen of the droplets and allow for molecular communication between compartments. It is also possible to colocalize different cellular mimics within bigger compartments, thus creating a structure analogous to eukaryotic cells.[164,165] The examples described are applied to lipid droplets but the versatility of these technologies allows the use with different kinds of artificial cells.

As the number of cellular mimics is gradually increasing over the years, it is possible that a collection of standardized parts will be soon available. Similar to the genetic elements collected in the registry of standard biological parts, a dataset of artificial cells with defined properties and functions could be available in the future. The integration of multiple functions could be easily designed by means of this "catalog" to create communities of artificial cells interacting with each other to address a problem that ultimately requires the exchange of chemical messages with natural cells. Once the data set will be big enough it could also be possible to design algorithms for the automation of the design of new pathways. A web-based platform was recently developed for the rational design of metabolic pathways through the interrogation of enzyme databases. An

enzymatic cascade can be designed for the engineering of bacteria to synthesize a desired product.[166] Similarly, we can envision a platform developed with the purpose of scanning a library of defined artificial cells and designing the best communication pathway for an efficient interaction with natural cells.

Bibliography

1. Lentini, R. *et al.* Integrating artificial with natural cells to translate chemical messages that direct E. coli behaviour. *Nat. Commun.* **5,** 4012 (2014).

2. Martini, L. & Mansy, S. S. Cell-like systems with riboswitch controlled gene expression. *Chem. Commun. (Camb).* **47,** 10734–6 (2011).

3. Li, M. *et al.* In vitro gene expression and enzyme catalysis in bio-inorganic protocells. *Chem. Sci.* **2,** 1739 (2011).

4. Gupta, A. *et al.* Encapsulated fusion protein confers 'sense and respond' activity to chitosan-alginate capsules to manipulate bacterial quorum sensing. *Biotechnol. Bioeng.* **110,** 552–562 (2013).

5. Huang, X. *et al.* Interfacial assembly of protein-polymer nano-conjugates into stimulus-responsive biomimetic protocells. *Nat. Commun.* **4,** 2239 (2013).

6. Huang, X., Li, M. & Mann, S. Membrane-mediated cascade reactions by enzyme-polymer proteinosomes. *Chem. Commun. (Camb).* **50,** 6278–80 (2014).

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

8. Caschera, F. & Noireaux, V. Synthesis of 2.3 mg/ml of protein with an all Escherichia coli cell-free transcription-translation system. *Biochimie* **99,** 162–168 (2014).

9. Saeidi, N. *et al.* Engineering microbes to sense and eradicate Pseudomonas aeruginosa, a human pathogen. *Mol. Syst. Biol.* **7,** 521 (2011).

10. Xiang, S., Fruehauf, J. & Li, C. J. Short hairpin RNA-expressing bacteria elicit RNA interference in mammals. *Nat. Biotechnol.* **24,** 697–702 (2006).

11. Anderson, J. C., Clarke, E. J., Arkin, A. P. & Voigt, C. A. Environmentally Controlled Invasion of Cancer Cells by Engineered Bacteria. *J. Mol. Biol.* **355,** 619–627 (2006).

12. Gupta, S., Bram, E. E. & Weiss, R. Genetically programmable pathogen sense and destroy. *ACS Synth. Biol.* **2,** 715–723 (2013).

13. Kotula, J. W. *et al.* Programmable bacteria detect and record an environmental signal in the mammalian gut. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 4838–43 (2014).

14. Hwang, I. Y. *et al.* Reprogramming microbes to be pathogen-Seeking killers. *ACS Synth. Biol.* **3,** 228–237 (2014).

15. Pinheiro, V. B. *et al.* Synthetic Genetic Polymers Capable of Heredity and Evolution. *Science* **336,** 341–344 (2012).

16. Malyshev, D. A. *et al.* A semi-synthetic organism with an expanded genetic alphabet. *Nature* **509,** 1–17 (2014).

17. Rackham, O. & Chin, J. W. A network of orthogonal ribosome·mRNA pairs. *Nat. Chem. Biol.* **1,** 159–166 (2005).

18. Terasaka, N., Hayashi, G., Katoh, T. & Suga, H. An orthogonal ribosome-tRNA pair via engineering of the peptidyl transferase center. *Nat. Chem. Biol.* **10,** 555–557 (2014).

19. Wang, K., Neumann, H., Peak-Chew, S. Y. & Chin, J. W. Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nat. Biotechnol.* **25,** 770–777 (2007).

20. Neumann, H., Wang, K., Davis, L., Garcia-Alai, M. & Chin, J. W. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **464,** 441–444 (2010).

21. Folcher, M. *et al.* Mind-controlled transgene expression by a wireless-powered optogenetic designer cell implant. *Nat. Commun.* **5,** 5392 (2014).

22. Nirenberg, M. W. & Matthaei, J. H. H. The dependence of cell-free protein synthesis in E. coli upon naturally occurring or synthetic polyribonucleotides. *Proc. Natl. Acad. Sci. U. S. A.* **47,** 1588–602 (1961).

23. Hoagland, M. B., Stephenson, M. L., Scott, J. F., Hecht, L. I. & Zamecnik, P. C. A soluble ribonucleic acid intermediate in protein synthesis. *J. Biol. Chem.* **231,** 241–57 (1958).

24. Chiao, A. C., Murray, R. M. & Sun, Z. Development of prokaryotic cell-free systems for synthetic biology. *bioRxiv* 1–38 (2016).

25. Mandel, M. & Higa, A. Calcium-dependent bacteriophage DNA infection. *J. Mol. Biol.* **53,** 159–162 (1970).

26. Smith, H. O. & Welcox, K. W. A Restriction enzyme from Hemophilus influenzae: I. Purification and general properties. *J. Mol. Biol.* **51,** 379–391 (1970).

27. Johnston, W. A. & Walker, J. M. *Cell-Free Protein Synthesis. Methods in Molecular Biology* **1118,** (2014).

28. Zemella, A., Thoring, L., Hoffmeister, C. & Kubick, S. Cell-Free Protein Synthesis: Pros and Cons of Prokaryotic and Eukaryotic Systems. *Chembiochem* **16,** 2420–31 (2015).

29. Lian, Q., Cao, H. & Wang, F. The Cost-Efficiency Realization in the Escherichia coli-Based Cell-Free Protein Synthesis Systems. *Appl. Biochem. Biotechnol.* **174,** 2351–2367 (2014).

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

31. Wang, H. H. *et al.* Multiplexed *in Vivo* His-Tagging of Enzyme Pathways for *in Vitro* Single-Pot Multienzyme Catalysis. *ACS Synth. Biol.* **1,** 43–52 (2012).

32. Garamella, J., Marshall, R., Rustad, M. & Noireaux, V. The all E. coli TX-TL Toolbox 2.0: a platform for cell-free synthetic biology. *ACS Synth. Biol.* acssynbio.5b00296 (2016). doi:10.1021/acssynbio.5b00296

33. Stögbauer, T., Windhager, L., Zimmer, R. & Rädler, J. O. Experiment and mathematical modeling of gene expression dynamics in a cell-free system. *Integr. Biol.* **4,** 494 (2012).

34. Calviello, L., Stano, P., Mavelli, F., Luisi, P. L. & Marangoni, R. Quasi-cellular systems: stochastic simulation analysis at nanoscale range. *BMC Bioinformatics* **14 Suppl 7,** S7 (2013).

35. Lentini, R. *et al.* Fluorescent Proteins and *in Vitro* Genetic Organization for Cell-Free Synthetic Biology. *ACS Synth. Biol.* **2,** 482–489 (2013).

36. Chizzolini, F., Forlin, M., Cecchi, D. & Mansy, S. S. Gene position more strongly influences cell-free protein expression from operons than T7 transcriptional promoter strength. *ACS Synth. Biol.* **3,** 363–371 (2014).

37. Sun, Z. Z., Yeung, E., Hayes, C. A., Noireaux, V. & Murray, R. M. Linear DNA for rapid prototyping of synthetic biological circuits in an escherichia coli based TX-TL cell-free system. *ACS Synth. Biol.* **3,** 387–397 (2014).

38. Chappell, J., Jensen, K. & Freemont, P. S. Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic Acids Res.* **41,** 3471–3481 (2013).

39. Sun, Z. Z. *et al.* Protocols for Implementing an Escherichia coli Based TX-TL Cell-Free Expression System for Synthetic Biology. *J. Vis. Exp.* 50762 (2013). doi:10.3791/50762

40. Singer, S. J. & Nicolson, G. L. The Fluid Mosaic Model of the Structure of Cell Membranes. *Science* **175,** (1972).

41. Makino, K. & Shibata, A. Chapter 2: Surface Properties of Liposomes Depending on Their Composition. *Adv. Planar Lipid Bilayers Liposomes* **4,** 49–77 (2006).

42. Israelachvili, J. N. *et al.* Physical principles of membrane organization. *Q. Rev. Biophys.* **13,** 121 (1980).

43. Kikuchi, H. *et al.* Gene delivery using liposome technology. *J. Control. Release* **62,** 269–277 (1999).

44. Spencer, A. C., Torre, P. & Mansy, S. S. The Encapsulation of Cell-free Transcription and Translation Machinery in Vesicles for the Construction of Cellular Mimics. *J. Vis. Exp.* e51304–e51304 (2013). doi:10.3791/51304

45. Sunami, T., Matsuura, T., Suzuki, H. & Yomo, T. in *Methods in molecular biology (Clifton, N.J.)* (eds. Endo, Y., Takai, K. & Ueda, T.) **607,** 243–256 (Humana Press, 2010).

46. Pautot, S., Frisken, B. J. & Weitz, D. A. Production of unilamellar vesicles using an inverted emulsion. *Langmuir* **19,** 2870–2879 (2003).

47. Hadorn, M. *et al.* Defined dna-mediated assemblies of gene-expressing giant unilamellar vesicles. *Langmuir* **29,** 15309–15319 (2013).

48. Elani, Y., Law, R. V & Ces, O. Protein synthesis in artificial cells: using compartmentalisation for spatial organisation in vesicle bioreactors. *Phys. Chem. Chem. Phys.* **17,** 15534–7 (2015).

49. Fujii, S., Matsuura, T. & Yomo, T. Membrane Curvature Affects the Formation of alpha-Hemolysin Nanopores. *ACS Chem Biol* **10,** 1694–1701 (2015).

50. Fujii, S. *et al.* Liposome display for in vitro selection and evolution of membrane proteins. *Nat. Protoc.* **9,** 1578–91 (2014).

51. Caschera, F. & Noireaux, V. Compartmentalization of an all- E. coli Cell-Free Expression System for the Construction of a Minimal Cell. *Artif. Life* **22,** 185–195 (2016).

52. Pickering, S. U. CXCVI.—Emulsions. *J. Chem. Soc., Trans.* **91,** 2001–2021 (1907).

53. Dinsmore, A. D. *et al.* Colloidosomes: Selectively permeable capsules composed of colloidal particles. *Science* **298,** 1006–1009 (2002).

54. Kwon, Y.-C., Hahn, G.-H., Huh, K. M. & Kim, D.-M. Synthesis of functional proteins using Escherichia coli extract entrapped in calcium alginate microbeads. *Anal. Biochem.* **373,** 192–196 (2008).

55. Gåserød, O., Sannes, A. & Skjåk-Bræk, G. Microcapsules of alginate–chitosan. II. A study of capsule stability and permeability. *Biomaterials* **20,** 773–783 (1999).

56. Oparin, A. I. *The origin of life on the Earth*. (Moscow worker publisher, 1924).

57. Dora Tang, T., van Swaay, D., deMello, A., Ross Anderson, J. L. & Mann, S. In vitro gene expression within membrane-free coacervate protocells. *Chem. Commun. Chem. Commun. Chem. Commun* **51,** 11429–11432 (2015).

58. Qiao, Y., Li, M., Booth, R. & Mann, S. Predatory behaviour in synthetic protocell communities. *Nat. Chem.* 1–23 (2016). doi:10.1038/nchem.2617

59. Tang, T.-Y. D., Antognozzi, M., Vicary, J. A., Perriman, A. W. & Mann, S. Small-molecule uptake in membrane-free peptide/nucleotide protocells. *Soft Matter* **9,** 7647 (2013).

60. Crosby, J. *et al.* Stabilization and enhanced reactivity of actinorhodin polyketide synthase minimal complex in polymer–nucleotide coacervate droplets. *Chem. Commun.* **48,** 11832 (2012).

61. Long, M. S., Cans, A. S. & Keating, C. D. Budding and asymmetric protein microcompartmentation in giant vesicles containing two aqueous phases. *J. Am. Chem. Soc.* **130,** 756–762 (2008).

62. Andes-Koback, M. & Keating, C. D. Complete budding and asymmetric division of primitive model cells to produce daughter vesicles with different interior and membrane compositions. *J. Am. Chem. Soc.* **133,** 9545–9555 (2011).

63. Torre, P., Keating, C. D. & Mansy, S. S. Multiphase water-in-oil emulsion droplets for cell-free transcription-translation. *Langmuir* **30,** 5695–5699 (2014).

64. Elani, Y., Law, R. V & Ces, O. Vesicle-based artificial cells as chemical microreactors with spatially segregated reaction pathways. *Nat. Commun.* **5,** 5305 (2014).

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

66. Kobori, S., Ichihashi, N., Kazuta, Y. & Yomo, T. A controllable gene expression system in liposomes that includes a positive feedback loop. *Mol. Biosyst.* **9,** 1282–5 (2013).

67. Schwarz-Schilling, M., Aufinger, L., Mückl, A. & Simmel, F. C. Chemical communication between bacteria and cell-free gene expression systems within linear chains of emulsion droplets. *Integr. Biol.* **8,** 564–570 (2016).

68. Zhu, T. F. & Szostak, J. W. Coupled growth and division of model protocell membranes. *J. Am. Chem. Soc.* **131,** 5705–5713 (2009).

69. Zhu, T. F., Adamala, K., Zhang, N. & Szostak, J. W. Photochemically driven redox chemistry induces protocell membrane pearling and division. *Proc. Natl. Acad. Sci.* **109,** 9828–9832 (2012).

70. Osawa, M., Anderson, D. E. & Erickson, H. P. Reconstitution of Contractile FtsZ Rings in Liposomes. *Science* **320,** (2008).

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

72. Merkle, D., Kahya, N. & Schwille, P. Reconstitution and anchoring of cytoskeleton inside giant unilamellar vesicles. *Chembiochem* **9,** 2673–2681 (2008).

73. Pontani, L. L. *et al.* Reconstitution of an actin cortex inside a liposome. *Biophys. J.* **96,** 192–198 (2009).

74. Carvalho, K. *et al.* Cell-sized liposomes reveal how actomyosin cortical tension drives shape change. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 16456–61 (2013).

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

76. Gardner, L. P., Mookhtiar, K. A. & Coleman, J. E. Initiation, elongation, and processivity of carboxyl-terminal mutants of T7 RNA polymerase. *Biochemistry* **36,** 2908–2918 (1997).

77. Gilbert, Walter and Müller-Hill, B. The Isolation of the Lac repressor. *Bioessays* **12,**

41–43 (1966).

78.   Yang, H. L., Zubay, G. & Levy, S. B. Synthesis of an R plasmid protein associated with tetracycline resistance is negatively regulated. *Proc. Natl. Acad. Sci.* **73,** 1509–1512 (1976).

79.   Cohen, G. & Jacob, F. Inhibition of the synthesis of the enzymes participating in the formation of tryptophan in Escherichia coli. *Comptes rendus Hebd. des séances l'Académie des Sci.* **248,** 3490–2 (1959).

80.   Jobe, A., Sadler, J. E. & Bourgeois, S. lac Repressor-operator interaction. *J. Mol. Biol.* **85,** 231–248 (1974).

81.   Oxender, D. L., Zurawskit, G. & Yanofsky, C. Attenuation in the Escherichia coli tryptophan operon: Role of RNA secondary structure involving the tryptophan codon region. *Proc. Natl. Acad. Sci. U. S. A.* **76,** 5524–5528 (1979).

82.   Lee, F. & Yanofsky, C. Transcription termination at the trp operon attenuators of Escherichia coli and Salmonella typhimurium: RNA secondary structure and regulation of termination. *Proc. Natl. Acad. Sci. U. S. A.* **74,** 4365–9 (1977).

83.   Zurawski, G. *et al.* Translational control of transcription termination at the attenuator of the *Escherichia coli* tryptophan operon. *Proc. Natl. Acad. Sci. U. S. A.* **75,** 5988–5992 (1978).

84.   Sadler, J. R. & Novick, A. The properties of repressor and the kinetics of its action. *J. Mol. Biol.* **12,** 305–327 (1965).

85.   Lederer, T. *et al.* Tetracycline analogs affecting binding to Tn10-encoded Tet repressor trigger the same mechanism of induction. *Biochemistry* **35,** 7439–7446 (1996).

86.   Silverman, S. K. Aptamers, Ribozymes, and Deoxyribozymes Identified by In Vitro Selection. *Funct. Nucleic Acids Anal. Appl.* 47–108 (2007). doi:10.1007/978-0-387-73711-9_3

87.   Pauley, R. J., Fredricks, W. W. & Smith, O. H. Effect of tryptophan analogs on derepression of the Escherichia coli tryptophan operon by indole-3-propionic acid. *J. Bacteriol.* **136,** 219–226 (1978).

88.   Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 343–5 (2009).

89.   Vrzheshch, E. P. *et al.* Optical properties of the monomeric red fluorescent protein mRFP1. *Moscow Univ. Biol. Sci. Bull.* **63,** 109–112 (2008).

90.   Seelig, B. mRNA display for the selection and evolution of enzymes from in vitro-translated protein libraries. *Nat. Protoc.* **6,** 540–552 (2011).

91.   Schu, D. J., Ramachandran, R., Geissinger, J. S. & Stevens, A. M. Probing the impact of ligand binding on the Acyl-homoserine lactone-hindered transcription

factor EsaR of Pantoea stewartii subsp. stewartii. *J. Bacteriol.* **193,** 6315–6322 (2011).

92.   Iyer, S., Karig, D. K., Norred, S. E., Simpson, M. L. & Doktycz, M. J. Multi-Input Regulation and Logic with T7 Promoters in Cells and Cell-Free Systems. *PLoS One* **8,** 1–12 (2013).

93.   Karig, D. K., Iyer, S., Simpson, M. L. & Doktycz, M. J. Expression optimization and synthetic gene networks in cell-free systems. *Nucleic Acids Res.* **40,** 3763–3774 (2012).

94.   Campbell, R. E. *et al.* A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 7877–82 (2002).

95.   Paige, J. S., Wu, K. Y. & Jaffrey, S. R. RNA mimics of green fluorescent protein. *Science* **333,** 642–646 (2011).

96.   Gatti-Lafranconi, P., Dijkman, W. P., Devenish, S. R. A. & Hollfelder, F. A single mutation in the core domain of the lac repressor reduces leakiness. *Microb. Cell Fact.* **12,** 67 (2013).

97.   Kaplan, H. B. & Greenberg, E. P. Diffusion of autoinducer is involved in regulation of the *Vibrio fischeri* luminescence system. *J. Bacteriol.* **163,** 1210–1214 (1985).

98.   Schaefer, a L., Val, D. L., Hanzelka, B. L., Cronan, J. E. & Greenberg, E. P. Generation of cell-to-cell signals in quorum sensing: acyl homoserine lactone synthase activity of a purified Vibrio fischeri LuxI protein. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 9505–9509 (1996).

99.   Hadorn, M. & Eggenberger Hotz, P. DNA-Mediated Self-Assembly of Artificial Vesicles. *PLoS One* **5,** e9886 (2010).

100.   Hadorn, M., Boenzli, E. & Hotz, P. E. A Quantitative Analytical Method to Test for Salt Effects on Giant Unilamellar Vesicles. *Sci. Rep.* **1,** 1–8 (2011).

101.   Fuqua, C. & Greenberg, E. P. Listening in on bacteria: acyl-homoserine lactone signalling. *Nat. Rev. Mol. cell Biol.* **3,** 685–695 (2002).

102.   Voet, D., Voet, J. G. & Pratt, C. W. *Fundamentals of biochemistry : life at the molecular level.* (Wiley, 2006).

103.   Jiang, Y. *et al. In vitro* biosynthesis of the *Pseudomonas aeruginosa* quorum-sensing signal molecule N-butanoyl-L-homoserine lactone. *Mol. Microbiol.* **28,** 193–203 (1998).

104.   Widanapathirana, L., Li, X. & Zhao, Y. Hydrogen bond-assisted macrocyclic oligocholate transporters in lipid membranes. *Org. Biomol. Chem.* **10,** 5077 (2012).

105.   Yoshimoto, M. *et al.* Preparation and characterization of reactive and stable glucose oxidase-containing liposomes modulated with detergent. *Biotechnol. Bioeng.* **81,** 695–704 (2003).

106. Mally, M., Majhenc, J., Svetina, S. & Žekš, B. The response of giant phospholipid vesicles to pore-forming peptide melittin. *Biochim. Biophys. Acta - Biomembr.* **1768,** 1179–1189 (2007).

107. Mally, M., Majhenc, J., Svetina, S. & Ekš, B. Ž. Mechanisms of Equinatoxin II-Induced Transport through the Membrane of a Giant Phospholipid Vesicle.

108. Chen, P. Y., Pearce, D. & Verkman, a S. Membrane water and solute permeability determined quantitatively by self-quenching of an entrapped fluorophore. *Biochemistry* **27,** 5713–5718 (1988).

109. Song, L. *et al.* Structure of Staphylococcal α-Hemolysin, a Heptameric Transmembrane Pore. *Science* **274,** (1996).

110. Patel, H., Tscheka, C. & Heerklotz, H. Characterizing vesicle leakage by fluorescence lifetime measurements. *Soft Matter* **5,** 2849 (2009).

111. Ramachandran, R., Heuck, A. P., Tweten, R. K. & Johnson, A. E. Structural insights into the membrane-anchoring mechanism of a cholesterol-dependent cytolysin. *Nat. Struct. Biol.* **9,** 823–7 (2002).

112. Wade, K. R. *et al.* An intermolecular electrostatic interaction controls the prepore-to-pore transition in a cholesterol-dependent cytolysin. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 2204–9 (2015).

113. Turner, A. P. F. Biosensors: sense and sensibility. *Chem. Soc. Rev.* **42,** 3184 (2013).

114. Goode, J. A., Rushworth, J. V. H. & Millner, P. A. *Biosensor Regeneration: A Review of Common Techniques and Outcomes. Langmuir* **31,** (2015).

115. Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **346,** 818–822 (1990).

116. Robertson, D. L. & Joyce, G. F. Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature* **344,** 467–468 (1990).

117. Tuerk, C. & Gold, L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249,** (1990).

118. Jijakli, K. *et al.* The in vitro selection world. *Methods* **106,** 3–13 (2016).

119. Levy, M. & Ellington, A. D. ATP-dependent allosteric DNA enzymes. *Chem. Biol.* **9,** 417–26 (2002).

120. Grate, D. & Wilson, C. Laser-mediated, site-specific inactivation of RNA transcripts. *Proc. Natl. Acad. Sci.* **96,** 6131–6136 (1999).

121. Sando, S., Narita, A. & Aoyama, Y. Light-up Hoechst-DNA aptamer pair: Generation of an aptamer-selective fluorophore from a conventional DNA-staining dye. *ChemBioChem* **8,** 1795–1803 (2007).

122. Dolgosheina, E. V. *et al.* RNA Mango Aptamer-Fluorophore: A Bright, High-Affinity

Complex for RNA Labeling and Tracking. (2014). doi:10.1021/CB500499X

123. Filonov, G. S., Moon, J. D., Svensen, N. & Jaffrey, S. R. Broccoli: Rapid Selection of an RNA Mimic of Green Fluorescent Protein by Fluorescence-Based Selection and Directed Evolution. *J. Am. Chem. Soc.* **136,** 16299–16308 (2014).

124. Stojanovic, M. N. & Kolpashchikov, D. M. Modular aptameric sensors. *J. Am. Chem. Soc.* **126,** 9266–9270 (2004).

125. Mannironi, C., Di Nardo, A., Fruscoloni, P. & Tocchini-Valentini, G. P. *In Vitro* Selection of Dopamine RNA Ligands [†]. *Biochemistry* **36,** 9726–9734 (1997).

126. Walsh, R. & Derosa, M. C. Retention of function in the DNA homolog of the RNA dopamine aptamer. (2009). doi:10.1016/j.bbrc.2009.08.084

127. Babendure, J. R., Adams, S. R. & Tsien, R. Y. Aptamers Switch on Fluorescence of Triphenylmethane Dyes. *J. Am. Chem. Soc.* **125,** 14716–14717 (2003).

128. Stoltenburg, R., Reinemann, C. & Strehlitz, B. SELEX-A (r)evolutionary method to generate high-affinity nucleic acid ligands. *Biomol. Eng.* **24,** 381–403 (2007).

129. Baugh, C., Grate, D. & Wilson, C. 2.8 Å Crystal Structure of the Malachite Green Aptamer. *J. Mol. Biol.* **301,** 117–128 (2000).

130. Marmorstein, R. Q., Joachimiak, A., Sprinzl, M. & Sigler, P. B. The structural basis for the interaction between L-Tryptophan and the Escherichia coli trp aporepressor. *J. Biol. Chem.* **262,** 4922–4927 (1987).

131. Lawson, C. L. & Carey, J. Tandem binding in crystals of a trp represser/operator half-site complex. *Nature* **366,** 178–182 (1993).

132. Schevitz, R. W., Otwinowski, Z., Joachimiak, A., Lawson, C. L. & Sigler, P. B. The three-dimensional structure of trp repressor. *Nature* **317,** 782–6 (1985).

133. Otwinowski, Z. *et al.* Crystal structure of trp repressor/operator complex at atomic resolution. *Nature.* **335,** 321–329 (1988).

134. von Bodman, B. S., Farrand, S. K., Beck von Bodman, S. & Farrand, S. K. Capsular polysaccharide biosynthesis and pathogenicity in Erwinia stewartii required induction by an N-acylhomoserinelactone autoinducer. *J. bacteriol.* **177,** 5000–5008. (1995).

135. Mazák, K., Dóczy, V., Kökösi, J. & Noszál, B. Proton Speciation and Microspeciation of Serotonin and 5-Hydroxytryptophan. *Chem. Biodivers.* **6,** 578–590 (2009).

136. Strickland, D., Moffat, K. & Sosnick, T. R. Light-activated DNA binding in a designed allosteric protein. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 10709–10714 (2008).

137. Siegal-Gaskins, D., Tuza, Z. A., Kim, J., Noireaux, V. & Murray, R. M. Gene circuit performance characterization and resource usage in a cell- free 'breadboard'. *ACS*

Synth. Biol. **3,** 1–15 (2013).

138. Booth, M. J., Schild, V. R., Graham, A. D., Olof, S. N. & Bayley, H. Light-activated communication in synthetic tissues. *Sci. Adv.* **2,** 1–12 (2016).

139. Müller, M. *et al.* Designed cell consortia as fragrance-programmable analog-to-digital converters. *Nat. Chem. Biol.* (2017). doi:10.1038/nchembio.2281

140. Regot, S. *et al.* Distributed biological computation with multicellular engineered networks. *Nature* **469,** 207–211 (2011).

141. Palivan, C. G. *et al.* Bioinspired polymer vesicles and membranes for biological and medical applications. *Chem. Soc. Rev.* **45,** 377–411 (2016).

142. Ahmed, F. & Discher, D. E. Self-porating polymersomes of PEG–PLA and PEG–PCL: hydrolysis-triggered controlled release vesicles. *J. Control. Release* **96,** 37–53 (2004).

143. Shen, L., Du, J., Armes, S. P. & Liu, S. Kinetics of pH-Induced Formation and Dissociation of Polymeric Vesicles Assembled from a Water-Soluble Zwitterionic Diblock Copolymer. *Langmuir* **24,** 10019–10025 (2008).

144. Lomas, H. *et al.* Efficient Encapsulation of Plasmid DNA in pH-Sensitive PMPC-PDPA Polymersomes: Study of the Effect of PDPA Block Length on Copolymer-DNA Binding Affinity. *Macromol. Biosci.* **10,** 513–530 (2010).

145. Jia, L. *et al.* Reduction-Responsive Cholesterol-Based Block Copolymer Vesicles for Drug Delivery. *Biomacromolecules* **15,** 2206–2217 (2014).

146. Liu, G., Wang, X., Hu, J., Zhang, G. & Liu, S. Self-Immolative Polymersomes for High-Efficiency Triggered Release and Programmed Enzymatic Reactions. *J. Am. Chem. Soc.* **136,** 7492–7497 (2014).

147. Nahire, R. *et al.* Multifunctional polymersomes for cytosolic delivery of gemcitabine and doxorubicin to cancer cells. *Biomaterials* **35,** 6482–6497 (2014).

148. Thambi, T. *et al.* Bioreducible polymersomes for intracellular dual-drug delivery. *J. Mater. Chem.* **22,** 22028 (2012).

149. Kim, H., Kang, Y. J., Kang, S. & Kim, K. T. Monosaccharide-Responsive Release of Insulin from Polymersomes of Polyboroxole Block Copolymers at Neutral pH. *J. Am. Chem. Soc.* **134,** 4030–4033 (2012).

150. Nazemi, A. *et al.* Dendrimersomes with photodegradable membranes for triggered release of hydrophilic and hydrophobic cargo. *Chem. Commun.* **50,** 11122 (2014).

151. Griepenburg, J. C. *et al.* Caging Metal Ions with Visible Light-Responsive Nanopolymersomes. *Langmuir* **31,** 799–807 (2015).

152. Cabane, E. *et al.* Photoresponsive polymersomes as smart, triggerable nanocarriers. *Soft Matter* **7,** 9167 (2011).

153. Blasco, E., Serrano, J. L., Piñol, M. & Oriol, L. Light Responsive Vesicles Based on

Linear–Dendritic Block Copolymers Using Azobenzene–Aliphatic Codendrons. *Macromolecules* **46,** 5951–5960 (2013).

154. Wong, C. K. *et al.* Polymersomes Prepared from Thermoresponsive Fluorescent Protein-Polymer Bioconjugates: Capture of and Report on Drug and Protein Payloads. *Angew. Chemie Int. Ed.* **54,** 5317–5322 (2015).

155. Qin, S., Geng, Y., Discher, D. E. & Yang, S. Temperature-Controlled Assembly and Release from Polymer Vesicles of Poly(ethylene oxide)-block- poly(N-isopropylacrylamide). *Adv. Mater.* **18,** 2905–2909 (2006).

156. Cheng, R. *et al.* Biodegradable chimaeric polymersomes mediate highly efficient delivery of exogenous proteins into cells. *J. Control. Release* **152,** e136–e137 (2011).

157. Grant, P. K. *et al.* Orthogonal intercellular signaling for programmed spatial behavior. *Mol. Syst. Biol.* **12,** 849 (2016).

158. Adamala, K. P., Martin-Alarcon, D. A., Guthrie-Honea, K. R. & Boyden, E. S. Engineering genetic circuit interactions within and between synthetic minimal cells. *Nat. Chem.* 1–9 (2016). doi:10.1038/NCHEM.2644

159. Carugo, D. *et al.* Liposome production by microfluidics: potential and limiting factors. *Sci. Rep.* **6,** 25876 (2016).

160. Kim, S.-H., Kim, J. W., Kim, D.-H., Han, S.-H. & Weitz, D. A. Enhanced-throughput production of polymersomes using a parallelized capillary microfluidic device. *Microfluid. Nanofluidics* **14,** 509–514 (2013).

161. Vanswaay, D., Tang, T. Y. D., Mann, S. & DeMello, A. Microfluidic Formation of Membrane-Free Aqueous Coacervate Droplets in Water. *Angew. Chemie - Int. Ed.* **54,** 8398–8401 (2015).

162. Thiele, J. *et al.* Fabrication of Polymersomes using Double-Emulsion Templates in Glass-Coated Stamped Microfluidic Devices. *Small* **6,** 1723–1727 (2010).

163. Shum, H. C., Kim, J.-W. & Weitz, D. A. Microfluidic Fabrication of Monodisperse Biocompatible and Biodegradable Polymersomes with Controlled Permeability. *J. Am. Chem. Soc.* **130,** 9543–9549 (2008).

164. Deng, N.-N., Yelleswarapu, M., Zheng, L. & Huck, W. T. S. Microfluidic Assembly of Monodisperse Vesosomes as Artificial Cell Models. *J. Am. Chem. Soc.* **139,** 587–590 (2017).

165. Baxani, D. K. *et al.* Bilayer Networks within a Hydrogel Shell : A Robust Chassis for Artificial Cells and a Platform for Membrane Studies. 1–6 (2016). doi:10.1002/anie.201607571

166. Carbonell, P., Parutto, P., Herisson, J., Pandit, S. B. & Faulon, J.-L. XTMS: pathway design in an eXTended metabolic space. *Nucleic Acids Res.* **42,** W389-94

(2014).

Appendix

The appendix contains the list of scientific publications to which I contributed during the development of my PhD project.

Xenobiotic Life

in: Glieder A, Kubicek CP, Mattanovich D, Wiltschi B, Sauer M, eds. "Synthetic Biology"
SE - 10.

Dario Cecchi, & Sheref S. Mansy.

I equally contributed to the writing of this chapter together with the corresponding author.

Integrating artificial with natural cells to translate chemical messages that direct *E. coli* behavior.

Roberta Lentini, Silvia Perez Santero, Fabio Chizzolini, **Dario Cecchi**, Jason Fontana, Marta Marchioretto, Cristina Del Bianco, Jessica L. Terrell, Amy C. Spencer, Laura Martini, Michele Forlin, Michael Assfalg, Mauro Dalla Serra, William E. Bentley and Sheref S. Mansy

For the development of this project, I contributed to the flow cytometry experiments. I was also actively involved in the discussion of the results.

Gene position more strongly influences cell-free protein expression from operons than T7 transcriptional promoter strength

Fabio Chizzolini, Michele Forlin, **Dario Cecchi**, & Sheref S. Mansy.

For the development of this project, I contributed to the design and the assembly of DNA sequences. I was also actively involved in the discussion of the results.

Cell-free translation is more variable than transcription

Fabio Chizzolini, Michele Forlin, Noël Yeh Martín, Giuliano Berloffa, **Dario Cecchi**, & Sheref S. Mansy.

For the development of this project, I contributed to the design and the assembly of some of the DNA sequences and to the optimization of the protocol for the production of a home-made *E. coli* S30 extract.