UNIVERSITAT POLITÈCNICA DE VALÈNCIA

ETSIAMN

BACHELOR´S DEGREE IN BIOTECHNOLOGY

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Design and optimization of a genetic circuit for the internal control of the population density of an *Escherichia coli* culture

Course 2018/2019

Author: Adrián Requena Gutiérrez

Thesis director: Nuria Andrés Colas

Thesis co-director: Alejandro Vignoni

Valencia, July 2019

Title: Design and optimization of a genetic circuit for the internal control of the population density of an *Escherichia coli* culture

Abstract

Synthetic biology allows the construction of genetic circuits with a desired behavior taking as an example, modifying and combining different genetic elements and specific molecular mechanisms that exist in nature. This capability gives synthetic biology a great potential in the use of all the tools that biology offers. As part of synthetic biology, mathematical modeling allows to make predictions of the behavior of a genetic circuit, from known data of the different biochemical reactions that occur in the organism and its kinetic parameters. In addition, it allows to know which parameters of these reactions must be modified to obtain the desired behavior and, with it, to make a rational design of the genetic constructs.

In this work, synthetic biology methods have been used to design a genetic circuit that limits the population density of an *Escherichia coli* culture within a range. For this, the genetic circuit has been based on the use of quorum sensing, discovered in the bacterium *Vibrio fischeri*, and consists of a lethal gene whose expression is dependent on population density. In this way, when the population density reaches a certain value, the expression of the lethal gene is activated and produces the death of a large part of the cell population. Those individuals that survive will continue multiplying until reaching again the sufficient population density for the activation of the lethal gene. Thus, an oscillating behaviour is achieved between two points of population density. In addition, a transcriptional unit of green fluorescent protein (GFP) regulated in the same way as the lethal gene has been incorporated into the construct, so that it can serve as a marker to monitor the behaviour of the circuit. This type of genetic circuits can be applied in the design of synthetic organisms for periodic drug delivery in situ.

In particular, for the rational design of this genetic circuit, a mathematical model that can represent the desired dynamic behavior has been developed . In order to obtain predictions of this behavior, the different parts that compose the genetic circuit have been characterized and these results have been combined with the developed model. The circuit has been constructed using high-efficiency assembly methods characteristic of synthetic biology, which allow to construct the circuit from its different basic parts of DNA. The behavior of the bacteria with the incorporated genetic circuit has been followed by measuring the optical density of the culture and the fluorescence emitted by the GFP. These measurements have been used to compare and adjust the model to the actual behavior. Finally, this process has also made it possible to determine which biological parameters must be adjusted to achieve the desired behavior.

Keywords: *Escherichia coli*; Synthetic Biology; Rational design; Quorum sensing; Population control; Oscillator

Author: Adrián Requena Gutiérrez

Thesis director: Nuria Andrés Colás

Thesis co-director: Alejandro Vignoni

Valencia, July 2019

Título: Diseño y optimización de un circuito genético para el control interno de la densidad poblacional de un cultivo de Escherichia coli.

Resumen

La biología sintética permite la construcción de circuitos genéticos con un comportamiento deseado, tomando como ejemplo, modificando y combinando diferentes elementos genéticos y mecanismos moleculares concretos que existen en la naturaleza. Esta capacidad le ha dado a la biología sintética un gran potencial en el aprovechamiento de todas las herramientas que ofrece la biología. Como parte de la biología sintética, la modelación matemática permite hacer predicciones del comportamiento de un circuito genético, a partir de los datos conocidos de las diferentes reacciones bioquímicas que ocurren en el organismo y de sus parámetros cinéticos. Además, permite conocer qué parámetros de estas reacciones deben ser modificados para obtener el comportamiento deseado y, con ello, realizar un diseño racional de las construcciones genéticas.

En este trabajo se han utilizado métodos de biología sintética para el diseño de un circuito genético para limitar la densidad poblacional de un cultivo de *Escherichia coli* dentro de un rango. Para ello, el circuito genético se ha basado en el uso del quorum sensing, descubierto en la bacteria *Vibrio fischeri*, y consiste en un gen letal cuya expresión es dependiente de la densidad poblacional. De esta manera, cuando la densidad poblacional alcanza un determinado valor se activa la expresión del gen letal y produce la muerte de gran parte de la población celular. Aquellos individuos que sobreviven continuarán multiplicándose hasta alcanzar nuevamente la densidad poblacional suficiente para la activación del gen letal. De esta forma se consigue un comportamiento oscilante entre dos puntos de densidad poblacional. Además, se ha incorporado a la construcción una unidad transcripcional de la proteína fluorescente verde (GFP) regulada de la misma manera que el gen letal, de modo que sirva de marcador para poder realizar un seguimiento del comportamiento del circuito. Este tipo de circuitos genéticos puede tener aplicación en el diseño de organismos sintéticos para la liberación periódica de fármacos in situ.

En particular, para el diseño racional de este circuito genético se ha desarrollado un modelo matemático que permite representar el comportamiento dinámico deseado. Para obtener predicciones de dicho comportamiento, se han caracterizado las diferentes partes que componen al circuito genético y estos resultados se han combinado con el modelo desarrollado. El circuito se ha construido utilizando métodos de ensamblaje de alta eficiencia propios de la biología sintética, que permiten construir el circuito a partir de sus diferentes partes básicas de DNA. El comportamiento de las bacterias con el circuito genético incorporado se ha seguido mediante la medición de la densidad óptica del cultivo y la fluorescencia emitida por la GFP. Dichas medidas han sido utilizadas para comparar y ajustar el modelo al comportamiento real. Por último, este proceso también ha permitido determinar qué parámetros biológicos deben ajustarse para alcanzar el comportamiento deseado.

Palabras clave: *Escherichia coli*; Biología sintética; Diseño racional; *Quorum sensing*; Control poblacional; Oscilador

Autor: Adrián Requena Gutiérrez

Tutora: Nuria Andrés Colás

Cotutor: Alejandro Vignoni

Valencia, julio 2019

Agradecimientos

En primer lugar, me gustaría agradecer la dedicación de mis supervisores, Alejandro y Nuria. También a Yadira y Jesús por el buen ambiente que siempre hay en el DISA.

Además, me gustaría agradecer a mis amigos de toda la vida, que han sido un gran apoyo durante estos cuatro años: Jorge, David, Javi, Omar, Roberto, Laura, Álvaro, Christian, Arturo y, en especial, a Lucas. También me gustaría acordarme de mis compañeros Josep, Paula y Almudena, con quienes he estado aprendiendo tanto todo este tiempo, y de mis socios, Roger y Carlos, con los que estoy creciendo mucho como profesional y como persona.

Por último, quiero agradecer todo lo que mis padres han hecho por mi, que siempre han estado, en las buenas y en las malas, y que sé que siempre estarán.

Contents

1 Introduction 1

1.1 Synthetic Biology . . . 1

1.1.1 Design: Modeling . . . 1

1.1.2 Build: GoldenBraid DNA assembly method . . . 3

1.2 Synthetic biological devices . . . 7

1.2.1 Genetic oscillators . . . 7

2 Objectives 10

3 Materials and methods 11

3.1 Design . . . 11

3.1.1 Deterministic mathematical models . . . 11

3.2 Build . . . 12

3.2.1 Creation of new GoldenBraid destination plasmids . . . 12

3.2.2 GoldenBraid assembly . . . 14

3.2.3 Transformation . . . 15

3.2.4 Colony PCR . . . 16

3.2.5 Agarose gel electrophoresis . . . 16

3.2.6 Cell cultures . . . 17

3.2.7 Bacterial glycerol stock . . . 18

3.2.8 Plasmid isolation: miniprep . . . 18

3.3 Test . . . 18

3.3.1 Measurements of fluorescence and OD_{600} *in vivo* . . . 18

3.3.2 Model parameter values estimation . . . 19

4 Results and discussion 21

4.1 Design: Genetic circuit and its mathematical model . . . 21

4.1.1 Designed genetic circuit . . . 21

4.1.2 Obtained mathematical model . . . 22

4.2 Build: Assembled plasmids . . . 25

4.2.1 Construction of new GoldenBraid Level 1 destination plasmids . . . 25

4.2.2 Constructions of transcriptional units in the new Level 1 destination plasmids . . . 26

4.2.3 Combination of 2 transcriptional units in Level 2 destination plasmids . . . 26

4.2.4 Construction of the genetic circuit by combination of multiple transciptional units in a Level 1 destination plasmid 28
4.3 Test: Characterization of DNA basic parts and test of the designed genetic circuit 28
4.3.1 Characterization of the interaction of LuxR with the pLux promoter 28
4.3.2 Characterization of the lethality of the E-lysis protein 31
4.3.3 Test of the genetic oscillator circuit 33
4.4 Redesign of the genetic circuit 34

5 Conclusions 36

Bibliography 37

Appendix I: DNA sequences 41

Appendix II: Reactions of the genetic circuit model 60

Appendix III: Models used in DNA basic parts characterization 63

List of Figures

1.1 The synthetic biology cycle with the three principal phases: design, build and test (Boada Acosta, 2018). 2

1.2 Schematic representation of the DNA basic part domestication process using PCR or DNA synthesis products of the DNA basic parts and the pUPD2 plasmid. . . . 4

1.3 Schematic representation of an assembly of a transcriptional unit using the DNA basic parts in pUPD2 plasmids and the Level 1 destination plasmids both digested with BsaI. 5

1.4 Schematic representation of a combination of 2 transcriptional units in Level 2 destination plasmids using transcriptional units in the two Level 1 plasmids and the Level 2 destination plasmids both digested with BsmBI. 6

1.5 Schematic representation of a combination of multiple transcriptional units in a Level 1 plasmid using constructions in the two Level 2 plasmids and the Level 1 destination plasmids both digested with BsaI. 6

1.6 Genetic oscillator used by Din et al. (2016) for producing synchronized cycles of bacterial lysis for *in vivo* drug delivery. 8

4.1 Representation of the genetic circuit designed in this work implemented in *E. coli*. 22

4.2 Simulation of the temporal evolution of the population density of *E. coli* population carrying the designed genetic circuit with the constructed mathematical model and the provisional values for the kinetic parameters. 25

4.3 Representation of the different transcriptional units assembled in the new GoldenBraid Level 1 destination plasmids. 27

4.4 Representation of the different transcriptional units combined in GoldenBraid Level 2 destination plasmids. 27

4.5 Representation of the genetic circuit designed in this work into a GoldenBraid Level 1 destination plasmid. 28

4.6 Temporal evolution of the population density (number of cells) and the expression of the GFP per cell of *E. coli* carrying the pARSO11 plasmid at different AHL concentrations. 29

4.7 Comparison between the experimental and the simulated results for the population density (number of cells) over time of *E. coli* carrying the pARSO11 plasmid at different AHL concentrations. 30

4.8 Comparison between the experimental and the simulated results of the GFP expression per cell of *E. coli* carrying the pARSO11 plasmid in steady state at different AHL concentrations. 31

4.9 Temporal evolution of the *E. coli* population density with the pARSO11 plasmid incorporated over time at different AHL conentrations.. 32

4.10 Comparison between the experimental and the simulated population density (number of cells) over time of *E. coli* carrying the pARSO13 plasmid at different AHL concentrations. 33

4.11 population density (number of cells) over time of *E. coli* carrying the pARSO13 plasmid at different AHL concentrations. 34

List of Tables

3.1 Composition of a PCR reaction using the Q5® High-Fidelty DNA polymerase. . 12

3.2 Conditions used for a PCR reaction using the Q5® High-Fidelty DNA polymerase to amplify the backbone from pDGB1 α1. 12

3.3 Conditions used for a PCR reaction using the Q5® High-Fidelty DNA polymerase to amplify the mRFP1 transcriptional unit from BBa_K2656109. 13

3.4 Composition of a restriction reaction with EcoRI and PstI. 13

3.5 Composition of a DNA ligation reaction with T4 DNA ligase. 14

3.6 DNA basic parts used in this work. 14

3.7 GoldenBraid destination plasmids used in this work. 14

3.8 Composition of a Golden Gate reaction. 15

3.9 Conditions needed to perform a Golden Gate reaction. 15

3.10 Composition of a colony PCR. 16

3.11 Conditions used for the colony PCR. 16

3.12 Conditions used for the measurement of OD_{600} and GFP fluorescence. 18

4.1 Parameter values used in the model simulation. 24

4.2 Estimated values of the parameters of the growth equation for the different AHL amounts used in the characterization of the interaction of LuxR with the pLux promoter. 30

Abbreviations

AHL : Acyl homoserine lactone

CamR : Cloramphenicol resistance

CDS : Coding sequence

cPCR : colony PCR

KanR : Kanamycin resistance

MEFL : Molecules of equivalent fluorescein

OD : Optical density

ODE : Ordinary differential equation

RBS : Ribosome binding site

SpecR : Spectinomycin resistance

TF : Transcription factor

1 Introduction

1.1 Synthetic Biology

Synthetic biology is a research field that combines the investigative nature of biology with the constructive nature of engineering (Purnick and Weiss, 2009). It is becoming an inclusive theoretical and technical framework in which to approach biological systems with the conceptual tools and language imported from electrical circuitry and mechanical manufacturing (de Lorenzo and Danchin, 2008).

The goal of synthetic biology is to modify the behavior of organisms and engineer them to perform new tasks that they are unable to do naturally (Adrianatoandro et al., 2006). This goal is pursued by the creation of new organisms through the rational combination of standardized biological parts decoupled from their natural context (de Lorenzo and Danchin, 2008).

The aspect which differentiates synthetic biology from genetic engineering and other biotechnology approaches, is the application of techniques which are normally used in engineering design and development (The Royal Academy of Engineering, 2009). This methodology can be illustrated in a 3 phases iterative cycle (Figure 1.1), that correspond with the design of the solution to a problem, the construction of that solution and the test of it.

1.1.1 Design: Modeling

As part of the design phase, it is very useful to have a tool that allows to predict the behavior of a new biologic system. It enables to make a rational design of the new system without having to perform large numbers of trial-and-error experiments. This tool is a model.

A model is a simplified representation of a system in some form useful for a given purpose (Boada Acosta, 2018). It is possible to use qualitative models or even simple graphical models (Chandran et al., 2008). However, qualitative models have the limitation of being based on interpretations that cannot be used in mathematical analysis or in computer algorithms and the amount of data necessary to fit these models often prohibits their use (Arkin, 2001). This is why quantitative models are being more widely used.

In the context of the biological systems, the quantitative models are usually dynamic models (concretely, kinetic models). These models typically consist of systems of differential equations that represent the temporal evolution of the different components of a biological system

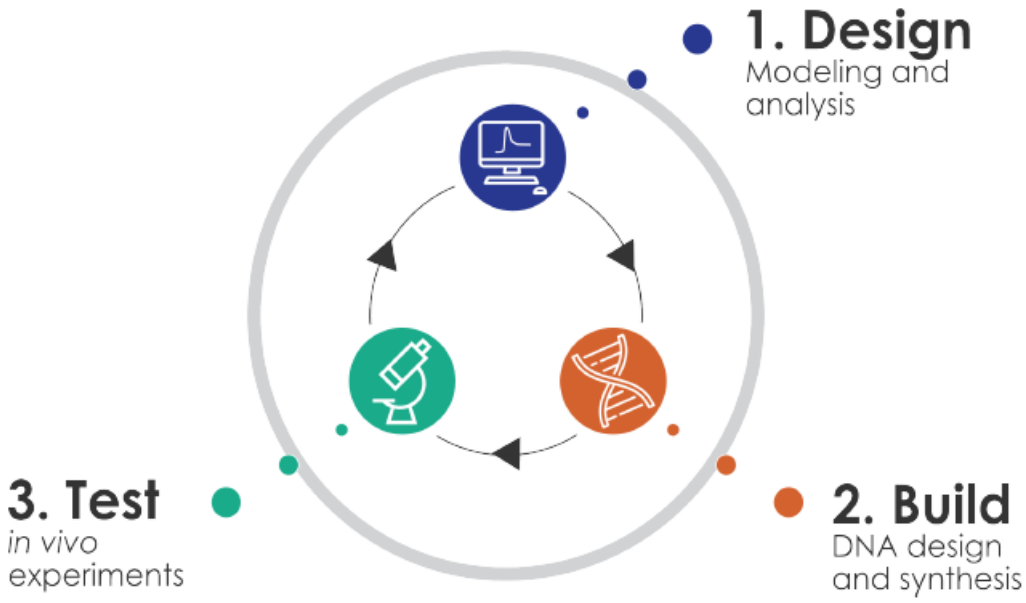

Figure 1.1. The synthetic biology cycle with the three principal phases: design, build and test (Boada Acosta, 2018).

(Villaverde and Banga, 2014). To construct a kinetic model, a series of simplifications are made, such as assuming that enzymatic kinetics are restricted to an approximation by means of the law of mass action or that the distribution of the reagents is represented as if it were continuous (Chen et al., 2010). Despite this, kinetic modeling of biological circuits has a long tradition and its usefulness for the design and characterization of biological systems is proven (Villaverde and Banga, 2014)(Chen et al., 2010).

A widely used way to construct a dynamic model of a biological system is to start from a representation of the biochemical reactions that occur in the system and apply the law of mass action to obtain a system of ordinary differential equations (ODEs) (Picó et al., 2015). The law of mass action establishes that the reaction rate is proportional to a kinetic parameter and to the product of the concentrations of the reaction substrates elevated to their stoichiometric coefficient (Chellaboina et al., 2009). In this way a deterministic model is obtained, which does not take into account the probabilistic factors of the complex biological reality but that can be taken as a starting point to obtain more complex models (Picó et al., 2015). In addition, it is important to emphasize that it is needed to simplify the networks of interactions that occur in biological systems without losing biological relevance. In this way, the data obtained from the models can be interpreted (Borhnholdt, 2005).

Once the model has been obtained and there are experimental data from the biological system, it is possible to improve the approach that the model can make. This is done by an adjustment process to obtain indirectly those unknown parameters of the model that cannot be obtained from experimental data (Lillacci and Khammash, 2010). Obtaining these parameters, in addition to improving the model, serves to know in greater depth the system and characterize its components with respect to the model. These data can be used to improve the implementation of the system (Boada et al., 2016) or even to design new systems (Decoene et al., 2018)(Canton et al., 2008).

1.1.2 Build: GoldenBraid DNA assembly method

One of the engineering features that are taken by synthetic biologists is the use of abstraction to design new biological systems (Adrianatoandro et al., 2006). Essentially, it means considering each component coded in DNA, such as promoters, ribosome biding sites (RBS), coding sequences (CDS) or transcriptional terminators as isolated parts. Once these parts have been characterized, which allows the rational design of the desired genetic circuit, they can be combined to produce new routes and biological devices whose behavior is predictable under defined conditions. Based on this, the goal envisioned by synthetic biology is to build increasingly complex systems from these parts of DNA: combine basic parts to produce genes, join these genes to build pathways and devices, and finally combine pathways to build synthetic chromosomes and genomes (Ellis et al., 2011).

However, in order to achieve these objectives, it has been necessary to develop highly efficient systems that allow DNA parts to be assembled and tested (Ellis et al., 2011). Those are the DNA assembly methods. Today there is a wide variety of quick and reliable assembly methods that work in different ways. The choice of the assembly method can influence the final structure of the genetic constructs (Ellis et al., 2011), so the choice of the method used is an important part of a synthetic biology work.

The DNA assembly methods can be classified according to whether they are based on the use of restriction enzymes or whether they are based on overlapping sequences (Weninger et al., 2016).

It is very important to know two concepts. The first one is the concept of assembly standard. A standard defines how the DNA parts are flanked taking into account the sites of recognition and cutting of the restriction enzymes to be used. In this way, parts that are standardized with the same standard can be combined in a directional way. The other important concept is the concept of scar. Scars are sets of non-functional nucleotides resulting from the ligation of two DNA fragments that share the same standard and have been digested in order to join them (Chan et al., 2015)(KNIGHT, 2008).

The DNA assembly methods based on restriction enzymes are also classified into two groups: those based on Type II restriction enzymes and those based on Type IIS restriction enzymes (Weninger et al., 2016). Methods based on Type IIS restriction enzymes employ the advantage that, with these enzymes, the cleavage site does not coincide with the recognition site as it occurs with Type II restriction enzymes (Loenen et al., 2014). This allows to design different overhangs for the same restriction enzyme allowing in turn, to make multiple parts assemblies in a single reaction (Engler et al., 2008).

The GoldenBraid standard is based on the Golden Gate methodology (Sarrion-Perdigones et al., 2011), which employs Type IIS restriction enzymes to make assemblies of multiple parts of DNA, directionally and in a single pot reaction (Engler et al., 2008). The standards based on this method of assembly leave only 4nt of scar between every two parts (Weber et al., 2011), being able even, to avoid that no scar appears if the design is done in an intelligent way.

The GoldenBraid assembly is a modular strategy that allows the binary assembly of multigenic constructions starting from standard DNA basic parts in pUPD2 plasmid (Level 0) using two levels of destination plasmids. These two levels are named α (Level 1) and Ω (Level 2) (Sarrion-Perdigones et al., 2011) and for each level there are two different destination plasmids.

Level 0: DNA basic parts

As mentioned before, the DNA basic parts are promoters, RBSs, CDSs and transcriptional terminators. They must follow the standard adopted by any Type IIS standard, the one defined in the original MoClo paper (Weber et al., 2011) and they must be cloned into the pUPD2 plasmid (Figure 1.2)(Vazquez-Vilar et al., 2017). To adapt the parts to the standard or domesticate the parts, they must be flanked with specific sequences as represented in Figure 1.2. The DNA basic parts flanked can be obtained either by PCR or by DNA synthesis. Once the parts are flanked by the correct sequences, they can be inserted into pUPD2. A schematic representation of the result is found in Figure 1.2.

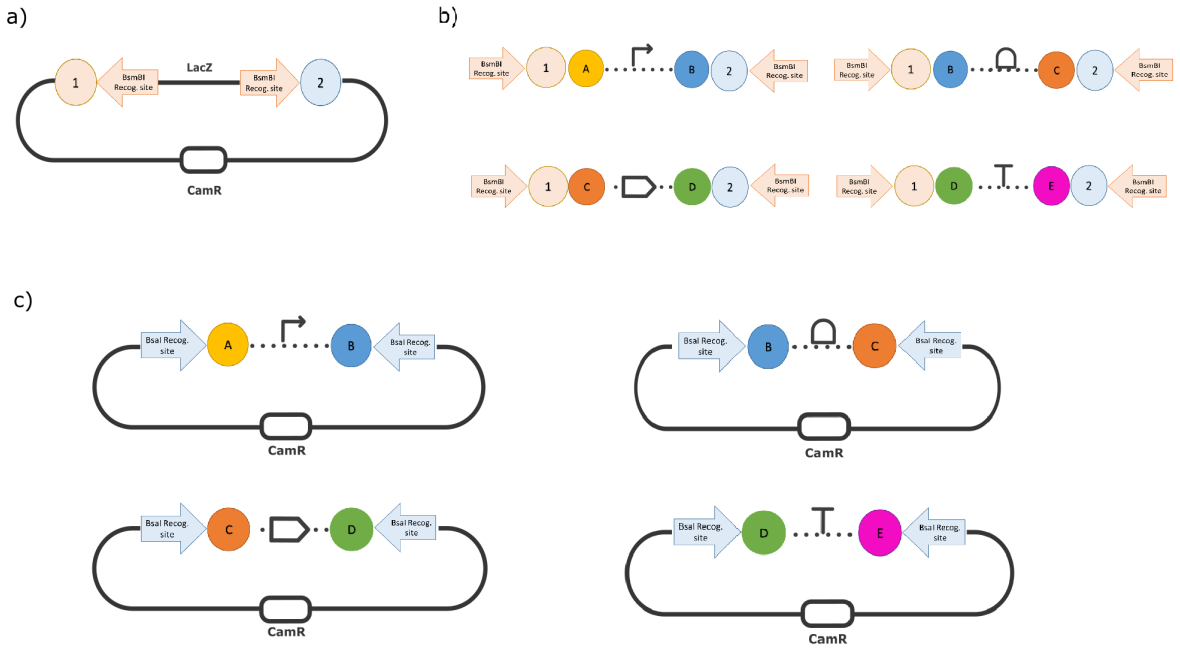

Figure 1.2. Schematic representation of the DNA basic part domestication process using PCR or DNA synthesis products of the DNA basic parts and the pUPD2 plasmid. The sequence represented by each number is 1: CTCG and 2: TGAG. The sequence represented by each letter is A:GGAG, B: TACT, C: AATG, D: GCTT and E: CGCT. CamR is cloramphenicol resistance. a) The pUPD2 plasmid. b) The different DNA basic parts flanked by the specific sequences necessary for their domestication. The basic parts from left to right and from top to bottom are: promoter, RBS, CDS and transcriptional terminator. c) The DNA basic parts domesticated and cloned into the pUPD2 plasmid.

Level 1: Transcriptional units

With the DNA basic parts adapted to the standard and into a pUPD2 plasmid, it is possible to combine them into a Level 1 destination plasmid in order to create a transcriptional unit (Sarrion-Perdigones et al., 2011) which is a sequence of nucleotides in DNA that codes for a single RNA molecule, along with the sequences necessary for its transcription (Pierce, 2016). This process is illustrated in Figure 1.3.

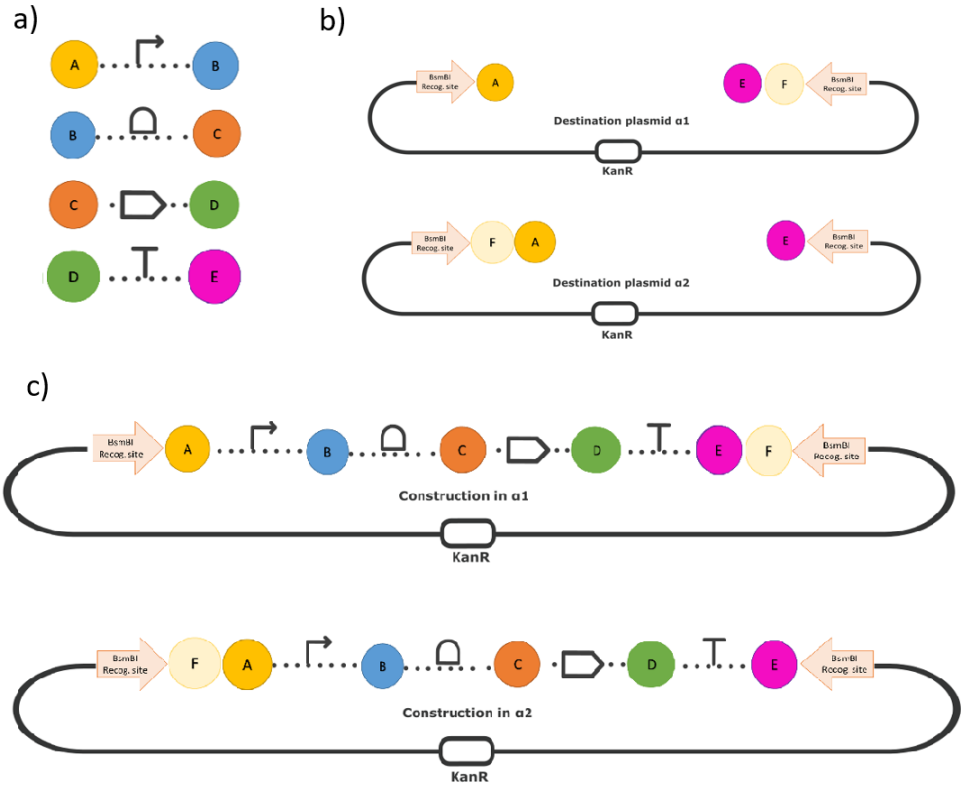

Figure 1.3. Schematic representation of an assembly of a transcriptional unit using the DNA basic parts and the Level 1 destination plasmids both digested with BsaI. The sequence represented by each letter is A: GGAG, B: TACT, C: AATG, D: GCTT, E: CGCT and F: GTCA. KanR is kanamycin resistance. a) The DNA basic parts cleaved from pUPD2 using BsaI. b) Level 1 destination plasmids digested with BsaI. c) The result of the assembly being B the scar between the promoter and the RBS, C the scar between the RBS and the CDS and D, the scar between the CDS and the terminator.

Level 2: combination of 2 transcriptional units

Once a construction has been made on an α1 destination plasmid and another on an α2 destination plasmid, it is possible to combine these constructions into a Level 2 destination plasmid (Sarrion-Perdigones et al., 2011). This process is illustrated in Figure 1.4.

Loop: Combination of multiple transcriptional units constructions

Finally, when constructions have been made in Ω1 and Ω2 destination plasmids, it is possible to combine them into a Level 1 destination plasmid (Sarrion-Perdigones et al., 2011). This process is illustrated in Figure 1.5.

Making this process iterative, the assembly of multigenic constructions can be achieved with less reactions and steps than in other assembly methods. This is the main advantage of this method. However, this method requires more planning than others, since once the constructions are started, the order and with which constructions it will be possible to combine them are defined.

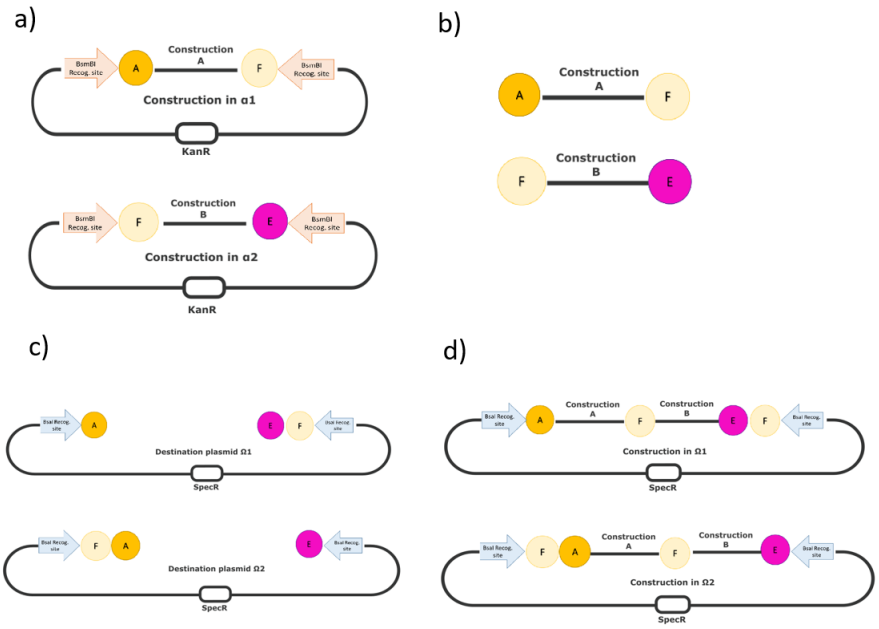

Figure 1.4. Schematic representation of a combination of 2 transcriptional units in Level 2 destination plasmids using transcriptional units in the two Level 1 plasmids and the Level 2 destination plasmids both digested with BsmBI. The sequence represented by each letter is A: GGAG, E: CGCT and F: GTCA. KanR is kanamycin resitance and SpecR is spectinomycin resistance. a) Constructions in Level 1 plasmids. b) The constructions cleaved from the Level 1 plasmids using BsmBI. c) Level 2 destination plasmids digested with BsmBI. d) The result of the assembly being F the scar between the two constructions.

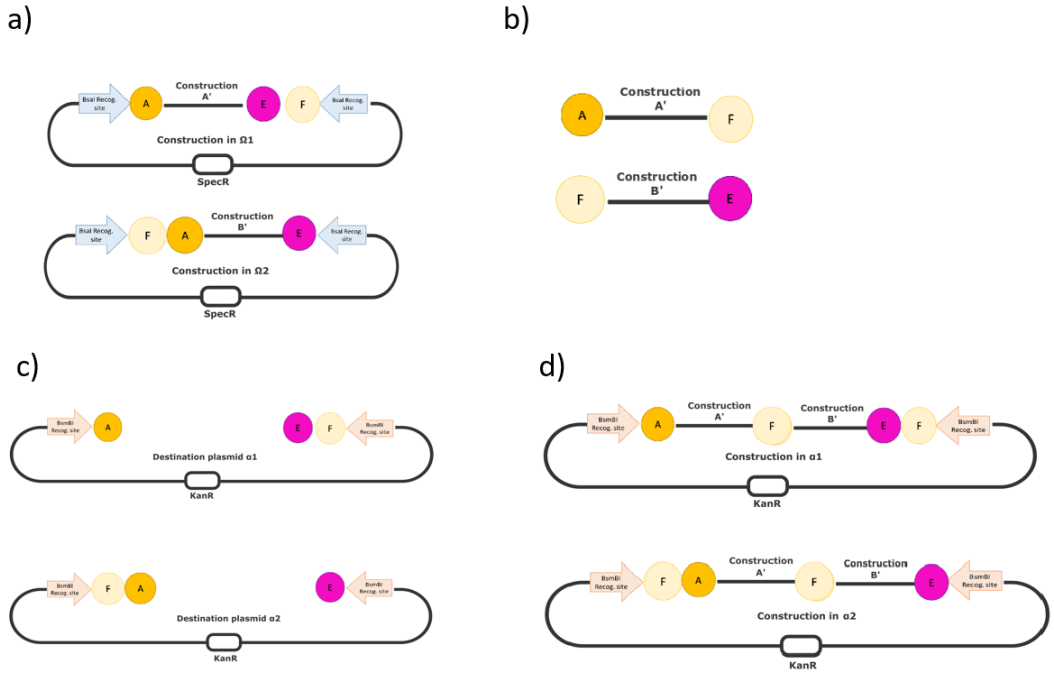

Figure 1.5. Schematic representation of an assembly of a construction in a Level 1 plasmid using constructions in the two Level 2 plasmids and the Level 1 destination plasmids both digested with BsaI. The sequence represented by each letter is A: GGAG, E: CGCT and F: GTCA. KanR is kanamycin resitance and SpecR is spectinomycin resistance. a) Constructions in Level 2 plasmids. b) The constructions cleaved from the Level 2 plasmids using BsaI. c) Level 1 destination plasmids digested with BsaI. d) The result of the assembly being F the scar between the two constructions.

1.2 Synthetic biological devices

Following the idea of combining parts to build increasingly complex systems, the level reached by combining different transcriptional units or genes is the level of synthetic biological device (Adrianatoandro et al., 2006). To define a synthetic biological device, it is possible to start from the definition of a device used in other fields of engineering: a thing made or adapted for a particular purpose. Using this definition and the idea of abstraction, explained above, synthetic biological devices are defined as that combination of biological parts designed by humans for a specific purpose (Endy, 2005). There are different types of engineering devices that find their biological equivalent such as sensors (Liu et al., 2015)(Slomovic et al., 2015) or logic gates which are devices that emit their output conditionally (Anderson et al., 2007). And as in the case of engineering devices, a particular biological device is defined by its input and its output (Canton et al., 2008).

Since the beginning of this century, genetic circuits have been designed with the intention of creating devices that are useful in the different fields where biotechnology has an impact (Cameron et al., 2014). The first genetic circuits to be implemented were simple systems such as the toggle switch or the repressilator. The toggle switch is a genetic circuit with two transcriptional states, which can change from one to the other with the addition of an external inducer, and its transcriptional state is maintained over time and generations without the need to add more inducer (Gardner et al., 2000). And the repressilator is a system composed of three transcriptional units made up of three different repressible promoters, each one with the CDS of the repressor of one of the other two promoters, thus achieving an oscillating clock (Elowitz and Leibler, 2000). These simple circuits did not need an exhaustive rational design.

However, since the beginning of this decade, large collections of parts have emerged, and synthetic biologists have begun to implement increasingly more complex circuits that require a rational design approach with quantitative data (Mukherji and Oudenaarden, 2009)(Ceroni et al., 2010).

1.2.1 Genetic oscillators

A genetic oscillator is a genetic network in which there are feedback loops. The simplest example is a gene whose expression represses its expression, so that when high values of the protein produced by the gene are reached, the production of that protein is reduced and when the levels of the protein are low again, the expression of the protein is increased again (Uriu, 2016). In nature there are different examples of genetic oscillators, such as the cell cycle (Pomerening et al., 2005) or the circadian clocks (Goldbeter et al., 2012).

Since the beginning of the design of biological devices, oscillators have been along with switches, the main types of devices designed (Tyson et al., 2008). In fact, the previously commented repressilator is one of these devices (Elowitz and Leibler, 2000). There are different types of genetic oscillators being designed such as Goodwin oscillators, based on proteins whose expression represses whose expression, or amplified negative feedback oscillators, based on two genes: the first one activates the expression of the second one and the second one represses the expression of the first one (Purcell et al., 2010).

It has been proven on numerous occasions in the literature that genetic circuits with oscillating behavior are tunable in different ways: changing the environmental conditions to which the oscil-

lator responds (Stricker et al., 2008) or changing the regulation and degradation of the proteins used (Tomazou et al., 2018). Another thing that has been proven is that using computational tools it is possible to design a genetic circuit with oscillatory behavior from its expected behavior and not from its parts. Thus, it is selected, in an automated way, which components must be combined to obtain that behavior (Otero-Muras and Banga, 2016).

A particularly interesting type of genetic oscillator are those based on the activation of a lethal gene due to quorum sensing. In this way it is achieved an oscillation in which, when the population density is low, there is no expression of the lethal gene, but when the population density is high, the lethal gene is expressed and part of the population dies, returning to low levels of population density. In this case, the feedback loop is based on the death of part of the population. In 2004 the first of these devices appeared: a population control designed for *Escherichia coli* (*E. coli*), whose purpose was to maintain a constant population density using the ccdB lethal gene and the *Vibrio fischeri* (*V. fischeri)* cell comunication (You et al., 2004). In 2010, the phenomena that occur behind the oscillations are explained using a circuit that causes the population density to oscillate if it is activated by IPTG, using the *Vibrio fischeri* cell comunication and the φX174 lysis gene (Marguet et al., 2010). The *V. fischeri* cell comunication is based on the lux operon. This operon has a gene (luxI) that codifies for the protein that produces the messenger molecule (a homoserine lactone) and a regulator gene (luxR) that codifies for the protein which binds to the messenger molecule and then, interacts with the operator, activating the transcription. Finally, in 2016, a genetic circuit is designed using an oscillator of this type, with, again, the cell comunication of *V. fischeri* and the φX174 lysis gene, built to be implemented in *Salmonella enterica* subsp. *enterica* serovar Typhimurium to deliver periodically drugs *in vivo* by bacterial lysis (Din et al., 2016). The genetic circuit used for the drug delivery *in vivo* is represented in Figure 1.6.

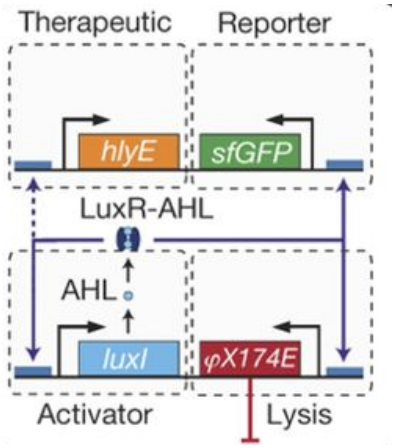

Figure 1.6. Genetic oscillator used by Din et al. (2016) for producing synchronized cycles of bacterial lysis for *in vivo* drug delivery. luxI produces acyl homoserine lactone (AHL) that joins LuxR so that it can activate the expression of luxI, sfGFP, the therapeutic gene (hlyE) and the lysis gene (φX174E). Since AHL can difuse through the membrane, the medium will have more AHL as there are more cells producing it. And as the more AHL, the more luxI expression there is, a positive feedback is produced in the formation of AHL. When an AHL threshold is reached, the expression of the lethal gene is sufficient to lyse the cells, reducing the number of cells producing AHL and producing the feedback loop that defines this oscillator. The expression of sfGFP and the therapeutic gene occurs at the moment when the lethal gene expression occurs (Din et al., 2016).

This circuit has a real applicability in the field of biomedicine but in spite of this, the design of the circuit has not been done in a rational way and that does not allow the tuneability of the system. For this reason, it is important to develop a methodology that allows the creation and optimization of these biological devices that can be applied in the different fields of biotechnology.

2 Objectives

The main objective of this work is to create a framework for the rational design and optimization of genetic oscillators based on the use of quorum sensing and a lethal gene to establish an internal control of population density in *E. coli* cultures. To do this the following specific goals were set:

1. Design a genetic circuit that gives *E. coli* an oscillatory population density.
2. Create a mathematical model that simulates the behavior of *E. coli* with a genetic circuit that gives *E. coli* an oscillatory population density.
3. Build the genetic circuit that gives *E. coli* an oscillatory population density.
4. Characterize the different DNA basic parts used for the construction of the genetic circuit that gives *E. coli* an oscillatory population density.
5. Test the built genetic circuit in *E. coli*.
6. Redesign the genetic circuit using the characterization and test data and combining them with the mathematical model to optimize the resulting behavior.

3 Materials and methods

3.1 Design

3.1.1 Deterministic mathematical models

To build a model of a system, it is necessary to visualize which components form this system and how they are correlated. The first step was making graphical representations with the different compartments of the system (Endler et al., 2009) in which the different components of the system, their products and their interactions are represented.

The next step was writing the reactions that occur in the system with their corresponding stoichiometry and associating each one with a kinetic constant. These reactions include both chemical transformations and the transport of molecules through different compartments. The kinetic constants were obtained from previous scientific studies (Endler et al., 2009) or in some cases they were considered as parameters to be estimated from the experimental data.

Then, applying the law of mass action, the ODEs were obtained (Picó et al., 2015). To apply the law of mass action, the tool sencillo.py, a version of Facile (Siso-Nadal et al., 2007) adapted to Python, was used. This tool uses a text file with the different reactions and provides a file with the ODEs and the kinetic constants that describe the system ready for its use in MATLAB®. These ODEs were used together with custom made scripts to perform the simulations in the software tool MATLAB ®.

The following assumptions were made:

- In each compartment, the concentration of the species is uniform, so diffusion ratios are not considered (Chandran et al., 2008).
- Cell volume is considered to be constant, so its variation along the cell cycle is not taken into account for the calculations of the ratios of transport through the membrane.
- The metabolic reactions that do not intervene in the behavior given by the genetic circuit are not taken into account. It is assumed that they all are part of the growth equation.
- The number of RNA polymerases (RNAp) and the number of copies of each plasmid are considered to remain constant. In addition, the number of nucleotides added by RNAp per unit of time is considered to be always the same. Therefore, the constitutive transcription processes that without these assumptions would consist of two reactions (binding of the

RNAp to DNA, and production of mRNA) and three species (RNAp, DNA and mRNA) (Picó et al., 2015), consist of a single reaction with a single specie (mRNA) (Reaction 3.1), and DNA and RNAp are included in the kinetic constant.

$$\emptyset \xrightarrow{k_1} \text{mRNA} \tag{3.1}$$

- In the same way, the number of ribosomes, aminoacids and tRNAs are considered to be constant, so these species are not taken into consideration in the translation reactions, leaving this process as a single reaction with two species (Reaction 3.2).

$$\text{mRNA} \xrightarrow{k_2} \text{mRNA} + \text{protein} \tag{3.2}$$

- Protein maturation is not taken into account but is considered in the translations rates.

3.2 Build

3.2.1 Creation of new GoldenBraid destination plasmids

The first step to create new Level 1 destination plasmids was to add recognition sites for EcoRI and PstI to the ends of the pDGB1 α1 backbone (Sarrion-Perdigones et al., 2011). To do this, the primers bbfw and bbrev (sequences in Appendix I) were designed. Then a PCR using the Q5® High-Fidelty DNA polymerase (New England Biolabs) was carried out. The composition of the reaction is found in Table 3.1 and the conditions used are found in Table 3.2.

Table 3.1. Composition of a PCR reaction using the Q5® High-Fidelty DNA polymerase.

Component	**Volume (µl)**
5x Q5® reaction buffer	5
10 mM dNTPs	2
Primer Forward 10 µM	1.25
Primer Reverse 10 µM	1.25
Q5® DNA polymerase	0.25
Template at 1 ng/µl	2
ddH_2O	12.5
Total Volume	25

Table 3.2. Conditions used for a PCR reaction using the Q5® High-Fidelty DNA polymerase to amplify the backbone from pDGB1 α1.

Name	**Temperature**	**Time**	**Number of cycles**
Initial Denaturation	98°C	5 minutes	1
Denaturation	98°C	10 seconds	30
Annealing	63°C	30 seconds	
Extension	72°C	1:30 minutes	
Final Extension	72°C	2 min	1

Then, the selector genes for the destination plasmids α1 and α2 were created. To do this, 2 pairs of primers to amplify a constitutive monomeric red fluorescent protein (mRFP1) transcriptional unit were designed for a PCR reaction using the Q5® High-Fidelty DNA. This transcriptional unit was amplified from BBa_K2656109 that belongs to the Valencia UPV iGEM 2018 part collection. The first primer pair (mrfpalpha1fw and mrfp1alpha1rev (sequences in Appendix I)) flanks the transcpritional unit with the same recognition and cutting sites as the lacZ in pDGB1 α1. The second one (mrfpalpha2fw and mrfp1alpha2rev (sequences in Appendix I)) flanks the transcpritional unit with the same recognition and cutting sites as the lacz in pDGB1 α2. Both pairs of primers add EcoRI and PstI restriction sites to the ends of the sequence. The composition of the reactions is in Table 3.1 and its conditions are in Table 3.3.

Table 3.3. Conditions used for a PCR reaction using the Q5® High-Fidelty DNA polymerase to amplify the mRFP1 transcriptional unit from BBa_K2656109.

Name	Temperature	Time	Number of cycles
Initial Denaturation	98°C	2 minutes	1
Denaturation	98°C	10 seconds	30
Annealing	63°C	30 seconds	
Extension	72°C	30 seconds	
Final Extension	72°C	2 min	1

The next step was to digest the products of the different PCRs with EcoRI and PstI, preparing the reaction as indicated in Table 3.4 and incubating it at 37°C for 1 hour. The enzymes were inactivated by incubating the reactions at 80°C for 20 minutes. And finally, two ligation reactions (one per each type of insert) were made with the products of the digestions. The components of a ligation reaction are in Table 3.5. The reaction was incubated at room temperature for 15 minutes. Then, the products of the ligations were cloned into *E. coli* (see transformation section).

Table 3.4. Composition of a restriction reaction with EcoRI and PstI.

Component	Amount
10x Cutsmart® Buffer	5 µl
PCR product	2000 ng
EcoRI	1 µl
PstI	1 µl
ddH_2O	up to 50 µl
Total	50 µl

Table 3.5. Composition of a DNA ligation reaction with T4 DNA ligase.

Component	Volume (µl)
T4 DNA ligase Buffer	2
Digested insert	1.68
Digested backbone	1.25
T4 DNA ligase	1
ddH2O	up to 20
Total	20

3.2.2 GoldenBraid assembly

The DNA basic parts cloned into pUPD2 used in this work are listed in Table 3.6 and were taken from the part collection designed for the iGEM competition in 2018 by the Valencia UPV team (VALENCIA UPV IGEM, 2018b). They were assembleb to build constructions into the GoldenBraid destination plasmids that are listed in Table 3.7.

Table 3.6. DNA basic parts used in this work.

Name	Type	Description	Source
J23106	Promoter	Constitutive promoter	BBa_K2656004
pLux	Promoter	Inducible promoter: positively regulated by LuxR in the presence of acyl homoserine lactone	BBa_K2656003
B0030	RBS	Strong RBS	BBa_K2656009
B0032	RBS	Weak RBS	BBa_K2656010
GFPmut3b	CDS	Green fluorescent protein coding sequence	BBa_K2656022
LuxR	CDS	Lux transcription factor coding sequence	BBa_K2656016
LuxI	CDS	AHL synthase coding sequence	BBa_K2656019
E-lysis protein	CDS	Lysis gene from φX174 phage	BBa_K2656015
B0015	Terminator	Double transcriptional terminator	BBa_K2656026

Table 3.7. GoldenBraid destination plasmids used in this work.

Name	Level	Antibiotic resistance	Selector Gene	Source
pARK α1	α1	Kanamycin	mRFP1	Section Construction of new GoldenBraid Level 1 destination plasmids of this work
pARK α2	α2	Kanamycin	mRFP1	Section Construction of new GoldenBraid Level 1 destination plasmids of this work
pDGB1 Ω1	Ω1	Spectinomycin	LacZ	Sarrion-Perdigones et al., 2011
pDGB1 Ω2	Ω2	Spectinomycin	LacZ	Sarrion-Perdigones et al., 2011

To assemble a transcriptional unit into a Level 1 destination plasmid (pARK α1 or pARK α2), 75ng of pUPD2 plasmid containing each of the basic parts and 75ng of Level 1 destination plasmid were used in a Golden Gate reaction (Table 3.8, 3.9) with the endonuclease BsaI.

To combine two different Level 1 constructions (one must be in an α1 plasmid and the other one in an α2 plasmid) into a Level 2 destination plasmid (pDGB1 Ω1 or pDGB1 Ω2), 75ng of each

Level 1 plasmid and 75ng of the Level 2 destination plasmid were used in a Golden Gate reaction (Table 3.8, 3.9) with the endonuclease BsmBI.

To combine two different Level 2 constructions (one must be in an Ω1 plasmid and the other one in an Ω2 plasmid) into a Level 1 destination plasmid, 75ng of each Level 2 plasmid and 75ng of the Level 1 destination plasmid were used in a Golden Gate reaction (Table 3.8, 3.9) with the endonuclease BsaI.

Table 3.8. Composition of a Golden Gate reaction.

Component	**Volume** (µl)
Type IIS restriction enzyme (BsaI or BsmBI)	1
T4 DNA ligase	1
T4 DNA ligase buffer 10x	1.5
BSA 10x	1.5
DNA	75 ng of each plasmid used
ddH_2O	up to 15 µl
Total	15

Table 3.9. Conditions needed to perform a Golden Gate reaction.

Temperature (°C)	**Time** (min)	**Number of cycles**
37	10	1
37 16	3 4	25
50	10	1
80	10	1

The assembled plasmids were cloned in *E. coli* (see section Transformation) and verified by colony PCR (cPCR) and by sequencing.

3.2.3 Transformation

All transformations were done by electroporation using *E. cloni* 10G® electrocompetent cells (Lucigen). To transform bacteria by electroporation, chill an aliquot of 50 µl electrocompetent cells on ice and add from 1 to 5µl of DNA resulting from a DNA assembly. Gentile homogenize and place the cells into a cold electroporation cuvette. Place the cuvette into the electroporator and electroporate the cells with 1700V. Pull out the cuvette from the electroporator and add 700µl SOC to the cuvette. Place the content of the cuvette in a culture tube and incubate at 37°C and 250 rpm for 1 hour. Finally, plate the cells with the appropiate antibiotic using a sterile glass rod and incubate overnight at 37°C.

3.2.4 Colony PCR

To make a cPCR, the first step is to select an isolated colony and inoculate 50µl of sterile deionized water with the colony. All cPCRs performed for this work were made with the Taq DNA polymerase kit with Standard Taq buffer (New England Biolabs). The composition of a cPCR is found in Table 3.10. The conditions used to perform the cPCR are found in Table 3.11. The primers used for each one of the cPCR performed are VerTUfw and VerTUrv whose sequence is found in the Appendix I. With these primers the region in which the genetic constructs in the GoldenBraid destination vectors are found is amplified.

Table 3.10. Composition of a colony PCR.

Component	**Volume (µl)**
10x Standard Taq buffer	2
10 mM dNTPs	2
Primer Forward 10µM	0.4
Primer Reverse 10µM	0.4
Taq DNA polymerase	0.1
Colony inoculated in water	5
ddH_2O	10.1
Total Volume	20

Table 3.11. Conditions used for the colony PCR.

Name	**Temperature**	**Time**	**Number of cycles**
Initial Denaturation	95°C	2 min	1
Denaturation	95°C	20 sec	30
Annealing	56°C	1 min	
Extension	68°C	1 min/kb	
Final Extension	68°C	5 min	1

3.2.5 Agarose gel electrophoresis

All gels used in this work were done using agarose at 1% and RedSafe^{TM} Nucleic Acid Staining (20000x) (Intron).

The samples were charged into the gel using Gel Loading Dye Purple (6X) (New England Biolabs) and O'GeneRuler 1 kb DNA Ladder (Fisher Scientific).

3.2.6 Cell cultures

Liquid cultures

LB Broth

All liquid cultures, except for the use of SOC in transformations and M9 in experimental measurements, were carried out using LB Broth (Lennox) as the culture medium and using 1000x stock solutions of the different antibiotics used.

To prepare this culture medium, add 10g tryptone, 5g yeast extract, 5g NaCl and 1L distilled water to a bottle. Homogenize and sterilize by autoclave. Let it chill and add 1mL 1000x antibiotic solution stock.

To make a culture from a single colony or from a glycerol stock, in aseptic conditions, inoculate 4mL LB Broth with the bacteria and incubate overnight at 37°C and 250 rmp.

SOC

To prepare this culture medium the first step is to prepare SOB. To prepare SOB, add 20g tryptone, 5g yeast extract, 2mL of 5M NaCl, 2.5mL of 1M KCl, 10mL of 1M $MgCl_2$, 10mL of 1M $MgSO_4$ to a bottle. Adjust to 1L with distilled water, homogenize and sterilize by autoclave. To make SOC with SOB, let SOB chill and, in aseptic conditions add 20mL of 1M filter-sterilized glucose.

To make a culture from a single colony or from a glycerol stock, in aseptic conditions, inoculate 4mL SOC with the bacteria and incubate overnight at 37°C and 250 rmp.

M9 Minimal Media

To prepare M9 Minimal Media, disolve 12.5g of M9 Medium Broth Powder (VWR Amresco Life Science) in 1L of distilled, deionized water. Homogenize and sterilize by autoclave. Let it chill and, in aseptic, condition add 2mL of filter-sterilized 1M $MgSO_4$, 10mL of filter-sterilized 20% glucose and 0.1mL of filter-sterilized 1M $CaCl_2$. If it will be used with bacteria with a plasmid that gives resistance to an antibiotic, add 1mL of 1000x stock solution of that antibiotic.

To make a culture from a single colony or from a glycerol stock, in aseptic conditions, inoculate 4mL M9 minimal media with the bacteria and incubate overnight at 37°C and 250 rmp.

Petri dish cultures

All Petri dish cultures were done using LB Broth (Lennox) with agar as culture medium and using 1000x stock solutions of the different antibiotics used.

To make this culture medium, add 10g tryptone, 5g yeast extract, 5g NaCl, 20g Agar-agar and 1L distilled water to a bottle. Homogenize and sterilize by autoclave. Let it chill and add 1mL 1000x antibiotic sotck solution. Finally, in aseptic conditions, extend 25mL of the contents of the bottle in each Petri dish.

To make a culture in a Petri dish from a liquid culture, in aseptic conditions, spread the bacteria on the plate and incubate overnight at 37°C.

3.2.7 Bacterial glycerol stock

To make a bacterial glycerol stock, add 1440µl overnight LB cell culture and 360µl of 80% glycerol to a cryovial in aseptic conditions. Gentile homogenize and store at -80°C.

3.2.8 Plasmid isolation: miniprep

All plasmid purifications were done using the QIAprep Spin Miniprep Kit (Qiagen). To use this kit, it is started from an overnight culture in LB of *E. coli* containing the plasmid that is wanted to purify. The cultures are placed in microcentrifuge tubes, the cells are lysed and the lysate is clarified by centrifugation. The clarified is placed in an affinity column and the DNA is bound to the column. After this, the column is washed and then the DNA is eluted in elution buffer.

3.3 Test

3.3.1 Measurements of fluorescence and OD_{600} in vivo

All experimental measurements were taken with Biotek Cytation^{TM} 3, a multi-mode reader that allows to take measures of fluorescence and optical density (OD) while cells are being incubated at the desired conditions in a 96 well plate. The conditions used in this work are in Table 3.12. The cells were *E. cloni*® cells (Lucigen) transformed with the corresponding plasmid.

For all experiments, it was started from two different isolated colonies of bacteria that incorporated the corresponding plasmid. With these two colonies, two cultures in LB with the corresponding antibiotic were prepared and incubated overnight at 37°C and 250 rpm. Then, 3 replicas of 3 mL cultures in M9 with the corresponding antibiotic at initial OD_{600}=0.04 with each of the overnight cultures were made and incubated at 37°C and 250 rpm for 4 hours. Dilutions of 3 ml culture at OD_{600}=0.1 were made for each replica. The next step was loading a 96 well black plate with clear and flat bottom with 10 samples of 200 µl of each culture and filling the wells in the edges of the plate with M9 to use it as blank.

Table 3.12. Conditions used for the measurement of OD_{600} and GFP fluorescence.

Time between measures	5 min
Temperature	37°C
Shake	Double orbital (Continiously)
Absorbance wavelength	600 nm
Excitacion wavelength	485 nm
Emission wavelength	528 nm

For each particular characterization that was done, a different experiment was designed:

Characterization of the interaction between LuxR and the pLux promoter by the addition of external acyl homoserine lactone (AHL)

For this experiment, the measurement time was set to 9 hours and the time in which the induction took place was set to 1 hour, having 8 hours of measurement after induction. Ten different concentrations of AHL (N-3-Oxohexanoyl-L-homoserine lactone, Santa Cruz Biotechnology) were used: 0 nM, 0.1 nM, 5 nM, 8 nM, 10 nM, 25 nM, 50 nM, 100 nM and 1000 nM.

Characterization of the lethality of the E-lysis protein by the addition of external AHL

In this case, the measurement time was set to 5 hours and the time in which the induction took place was set to 1 hour, having 4 hours of measurement after induction. The ten different amounts of AHL used in this case were: 0 nM, 0,1 nM, 0.2 nM, 0.3 nM, 0.4 nM, 0.5 nM, 0.6 nM, 0.7 nM, 0.8 nM and 0.9 nM,

Test of the designed genetic oscillator circuit

For this test, the measurement time was set to 24 hours and no induction was needed.

3.3.2 Model parameter values estimation

Once the experimental data were obtained, an estimation of the value of the model parameters, whose value cannot be obtained from the literature was performed. To do this, the chosen strategy was the novel methodology used by Boada et al. (2019). It consists on a multi-objective optimization based methodology to simultaneously estimate the different unknown parameters values of a model. This is achieved by minimizing the error between experimental data and simulations predictions for each of the variables that can be measured, such as OD_{600} or fluorescence. The methodology has four main steps:

1. Multi-objective problem definition or Cost function definition: specifies the objectives (at least two) to be minimized as a function of the model parameters that are intended to be estimated using the available experimental data. This step defines both the number of objectives $J(\Theta)$, and the decision variables (i.e. parameters) Θ to be estimated, with their corresponding initial value ranges.

2. The Multi-objective optimization process: runs the spMODE algorithm (Reynozo-Meza et al., 2013) to obtain the solution of the optimization. The solution obtained is not a single value, but a Pareto Front ($J_i^*(\Theta)$). That is a set of equally optimal solutions, those that minimize the objectives $J_i(\Theta)$. The corresponding Pareto set includes the values of the parameters of the model associated to the Pareto Front. The spMODE is a multi-objective optimization algorithm based on differential evolution, which actively searches for all the solutions in the parameter space along the Pareto front.

3. The Multi-criteria decision making process: consists of the selection of the preferred solution according to the designer's criteria. To this end, tools that simplify the analysis and visualization of the trade-offs among the competing objectives are used. Such visualization and analysis is not a trivial task when the number of objectives is larger than three or

the number of decision variables in the Pareto set is large. To perform the multi-criteria decision making process Level diagrams (Blasco et al., 2017) and clustering of the solutions obtained are used.

4. Validation of the obtained model: implies comparing the model predictions using the optimized parameters with experimental validation data not used during the optimization process.

4 Results and discussion

4.1 Design: Genetic circuit and its mathematical model

4.1.1 Designed genetic circuit

For the design of a genetic oscillator based on the activation of the expression of a lethal gene by quorum sensing, the following transcriptional units were selected:

- Expression of GFP activated by LuxR in the pressence of AHL.
- Constitutive expression of LuxR.
- Constitutive expression of LuxI.
- Expression of the lysis gene (E-lysis protein) from phage φ174 activated by LuxR in the pressence of AHL.

These transcriptional units were assembled, as indicated in Materials and Methods, forming the genetic circuit represented in Figure 4.1.

With this circuit an oscillating behavior is achieved as follows:

1. At the start of an experiment, there is no AHL in the medium. AHL can difuse through the cellular membrane and bacteria starts to produce AHL since they express LuxI. Because of this, it exists a concentration gradient of AHL, causing the produced AHL to exit from the cells.
2. As population density increases and the number of cells producing AHL increases, the concentration gradient decreases and the AHL begins to accumulate inside the cells.
3. The AHL threshold value is reached inside the cell so that the expression of the lethal gene is sufficient to cause the death of part of the population. At this time, the expression of GFP also reaches its maximum.
4. As the number of cells producing AHL has been reduced, the levels of this molecule drop and the cells that remain after the activation of the lethal gene in part of the population grow again and begin to produce more AHL, achieving the oscillating behavior.

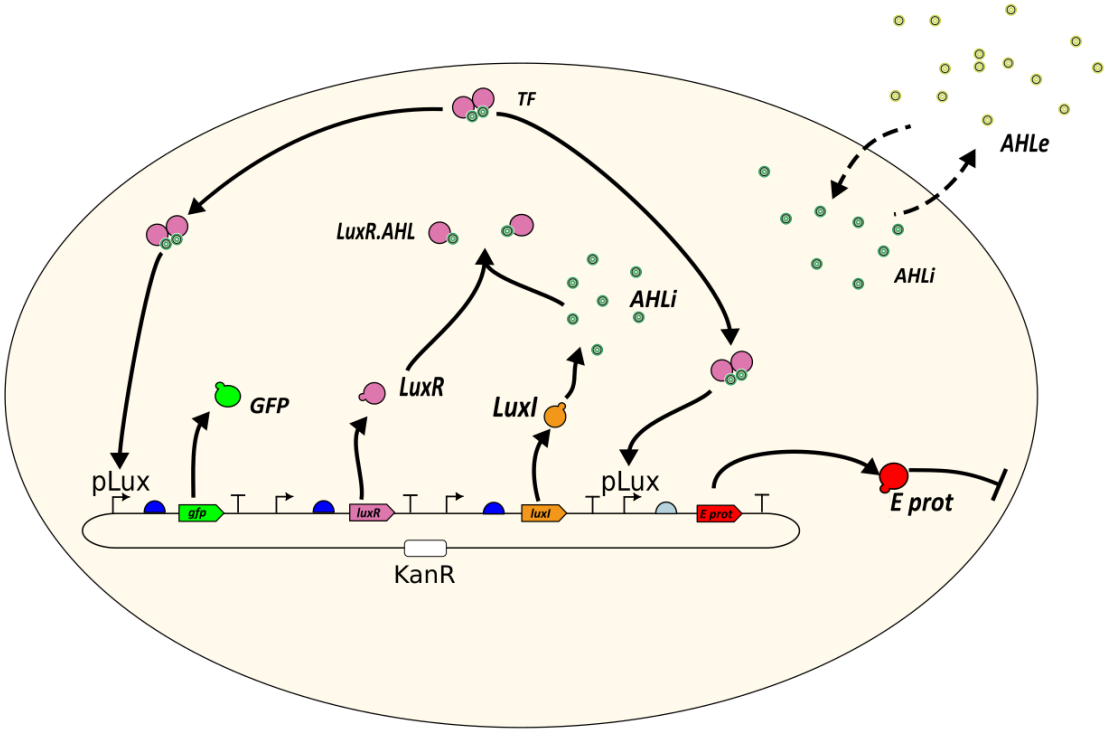

Figure 4.1. Representation of the genetic circuit designed in this work implemented in *E. coli*. The promoters of the luxI and the luxR genes are the constitutive promoter J23106. The RBS of the E-lysis protein gene is the weak RBS B0032, while for all other genes is the strong RBS B0030. The transcriptional terminator of all genes is B0015.

As represented in Figure 4.1, LuxR acts as a transcription factor (TF) if a homodimer is formed from two LuxRAHL heterodimers. The formation of both dimers is reversible (Qin et al., 2007).

4.1.2 Obtained mathematical model

Prior to the construction of the designed genetic circuit, a mathematical model, with provisional values for the kinetic parameters, was used to carry out a simulation of its resulting behaviour, in order to test if the designed genetic circuit could work as expected. With the designed genetic circuit to give *E. coli* an oscillatory population density behavior, and using the cellular schema of Figure 4.1, a mathematical model was constructed as explained in Materials and Methods. In this way, the reactions that occur in a cell carrying the designed genetic circuit were written (Appendix II), and using these reactions, a system of 13 ODEs was obtained.

$$\frac{dX}{dt} = \mu X - \frac{\mu X^2}{K_{\mathrm{mx}}} - \frac{k_{\mathrm{me}}[E]^n}{\theta^n + [E]^n} X \tag{4.1}$$

$$\frac{d[mLuxR]}{dt} = k_{\mathrm{R1}} - (k_{\mathrm{R3}} + \mu)[mLuxR] \tag{4.2}$$

$$\frac{d[LuxR]}{dt} = k_{\mathrm{R2}}[mLuxR] - k_{\mathrm{T1}}[LuxR][AHL_i] + k_{\mathrm{T\text{-}1}}[LuxRAHL] - (k_{\mathrm{R4}} + \mu)[LuxR] \tag{4.3}$$

$$\frac{d[mLuxI]}{dt} = k_{\mathrm{I1}} - (k_{\mathrm{I3}} + \mu)[mLuxI] \tag{4.4}$$

$$\frac{d[LuxI]}{dt} = k_{\mathrm{I2}}[mLuxI] - (k_{\mathrm{I4}} + \mu)[LuxI] \tag{4.5}$$

$$\frac{d[AHL_i]}{dt} = k_{\mathrm{A1}}[LuxI] + k_{\mathrm{A2}}Vc[AHL_e] - k_{\mathrm{T1}}[LuxR][AHLi] + k_{\mathrm{T\text{-}1}}[LuxRAHL] - (k_{\mathrm{A2}} + k_{\mathrm{A3}} + \mu)[AHL_i] \tag{4.6}$$

$$\frac{d[AHL_e]}{dt} = k_{\mathrm{A2}}[AHL_i] - X(k_{\mathrm{A2}}Vc[AHL_e]) - k_{\mathrm{A4}}[AHL_e] \tag{4.7}$$

$$\frac{d[LuxRAHL]}{dt} = k_{\mathrm{T1}}[LuxR][AHL_i] - k_{\mathrm{T\text{-}1}}[LuxRAHL] - 2k_{\mathrm{T2}}[LuxRAHL]^2 + 2k_{\mathrm{T\text{-}2}}[TF] - (k_{\mathrm{T3}} + \mu)[LuxRAHL] \tag{4.8}$$

$$\frac{d[TF]}{dt} = k_{\mathrm{T2}}[LuxRAHL]^2 - k_{\mathrm{T\text{-}2}}[TF] - (k_{\mathrm{T4}} + \mu)[TF] \tag{4.9}$$

$$\frac{d[mGFP]}{dt} = k_{\mathrm{G1}}CN_g\frac{[TF]}{[TF] + k_{\mathrm{G2}}} + \alpha_g\frac{k_{\mathrm{G2}}}{[TF] + k_{\mathrm{G2}}} - (k_{\mathrm{G4}} + \mu)[mGFP] \tag{4.10}$$

$$\frac{d[GFP]}{dt} = k_{\mathrm{G3}}[mGFP] - (k_{\mathrm{G5}} + \mu)[GFP] \tag{4.11}$$

$$\frac{d[mE]}{dt} = k_{\mathrm{E1}}CNe\frac{[TF]}{[TF] + k_{\mathrm{E2}}} + \alpha\frac{k_{\mathrm{E2}}}{[TF] + k_{\mathrm{E2}}}(k_{\mathrm{E4}} + \mu)[mE] \tag{4.12}$$

$$\frac{d[E]}{dt} = k_{\mathrm{E3}}[mE] - (k_{\mathrm{E5}} + \mu)[E] \tag{4.13}$$

Where (4.1) is the temporal evolution of the biomass (number of cells) taking into account the effect of the E-lysis protein, (4.2) is the temporal evolution of the amount of the LuxR mRNA, (4.3) is the temporal evolution of the amount of the LuxR protein, (4.4) is the temporal evolution of the amount of the LuxI mRNA, (4.5) is the temporal evolution of the amount of the LuxI protein, (4.6) is the temporal evolution of the intracellular AHL, (4.7) is the temporal evolution of extracellular AHL, (4.8) is the temporal evolution of the heterodimer LuxR-AHL, (4.9) is the temporal evolution of the transcription factor, (4.10) is the temporal evolution of the GFP mRNA, (4.11) is the temporal evolution of the GFP, (4.12) is the temporal evolution of the E-protein mRNA and (4.13) is the temporal evolution of the E-lysis protein. The concentration of all species is expressed in nM, which inside the cell is equivalent to molecules/cell (Boada Acosta, 2018).

The parameters used in the equations and their values are found in Table 4.1. The values of the parameters were obtained from previous works that used the same type of mathematical models (Boada Acosta, 2018)(Boada et al., 2019)(VALENCIA UPV IGEM, 2018a). The unknown parameters were given a provisional value.

This model, with provisional values for the kinetic parameters was used to carry out a simulation of the resulting behavior. This simulation was made assuming that in a cell population, there are different subpopulations with different phenotypes. For this, an algorithm was created that first gives a random value to each kinetic constant for each subpopulation following a normal distribution with mean in the given provisional value. After this, the number of cells is assigned to each subpopulation, also following a normal distribution. Finally, the differential equations are solved taking into account that at each time interval in which the equations are solved, the individuals are again distributed among the different subpopulations. The result of this simulation is in Figure 4.2.

Table 4.1. Parameter values used in the model simulation.

Parameter	Description	Value	Unit	Source
k_{I1}	Constitutive transcription rate with J23106	171.15	min^{-1}	VALENCIA UPV IGEM, 2018
k_{I2}	Translation rate of LuxI with B0030	7.8541	min^{-1}	Boada et al., 2019
k_{I3}	Degradation rate of mLuxI	0.247	min^{-1}	Boada Acosta, 2018
k_{I4}	Degradation rate of LuxI	0.027	min^{-1}	Boada Acosta, 2018
k_{R1}	Constitutive transcription rate with J23106	171.15	min^{-1}	VALENCIA UPV IGEM, 2018
k_{R2}	Translation rate of LuxR with B0030	7.8541	min^{-1}	Boada et al., 2019
k_{R3}	Degradation rate of mLuxR	0.247	min^{-1}	Boada Acosta, 2018
k_{R4}	Degradation rate of LuxR	0.027	min^{-1}	Boada Acosta, 2018
k_{A1}	Synthesis rate of AHL by LuxI	0.04	min^{-1}	Boada Acosta, 2018
k_{A2}	Diffusion rate of AHL through the cell membrane	2	min^{-1}	Boada Acosta, 2018
Vc	Typical volume of *E. coli* divided by the volume of culture used in the experiments	1.1E-11		Boada Acosta, 2018
k_{A3}	Degradation rate of intracellular AHL	0.057	min^{-1}	Boada Acosta, 2018
k_{A4}	Degradation rate of AHL in the culture medium	0.04	min^{-1}	Boada Acosta, 2018
k_{T1}	Association rate of AHL with LuxR	40	min^{-1}	Boada Acosta, 2018
k_{T-1}	Dissociation rate of LuxR.AHL	10	min^{-1}	Boada Acosta, 2018
k_{T2}	Dimerization rate of LuxR.AHL	20	min^{-1}	Boada Acosta, 2018
k_{T-2}	Dissociation rate of TF	1	min^{-1}	Boada Acosta, 2018
k_{T3}	Degradation rate of LuxR.AHL	0.156	min^{-1}	Boada Acosta, 2018
k_{T4}	Degradation rate of TF	0.32	min^{-1}	Boada Acosta, 2018
CNg	Copy number of the GFP transcriptional unit in a cell	500	units/cell	It is in a pUC derivative plasmid (MORGAN, 2014)
α_g	Basal transcription rate with pLux	0.01	min^{-1}	Provisional value
k_{G1}	Maximun rate of transcription with the pLux promoter	1	min^{-1}	Provisional value
k_{G2}	Number of TF molecules resulting in half maximum transcription rate	1	molecules	Provisional value
k_{G3}	Translation rate of GFP with B0030	7.8541	min^{-1}	Boada et al., 2019
k_{G4}	Degradation rate of mGFP	0.247	min^{-1}	Boada Acosta, 2018
k_{G5}	Degradation rate of GFP	0.0058	min^{-1}	Boada et al., 2019
CNe	Copy number of the E-lysis protein transcriptional unit in a cell	500	units/cell	It is in a pUC derivative plasmid (MORGAN, 2014)
α_e	Basal transcription rate with pLux	0.01	min^{-1}	Provisional value
k_{E1}	Maximun rate of transcription with the pLux promoter	1	min^{-1}	Provisional value
k_{E2}	Number of TF molecules resulting in half maximum transcription rate	1	molecules	Provisional value
k_{E3}	Translation rate of E with B0032	0.034	min^{-1}	Boada et al., 2019
k_{E4}	Degradation rate of mE	0.247	min^{-1}	Boada Acosta, 2018
k_{E5}	Degradation rate of E	2	min^{-1}	Provisional value
μ	Specific growth rate	0.0174	min^{-1}	Provisional value
K_{mx}	Maximum growth capacity (OD_{600})	1.174		Provisional value
K_{me}	Maximum rate of cell lysis due to E-lysis protein	0.02	molecules	Provisional value
n	Hill coefficient of lysis function	1		Provisional value
Θ	Number of E-lysis protein resulting in half maximum lysis	1	min^{-1}	Provisional value

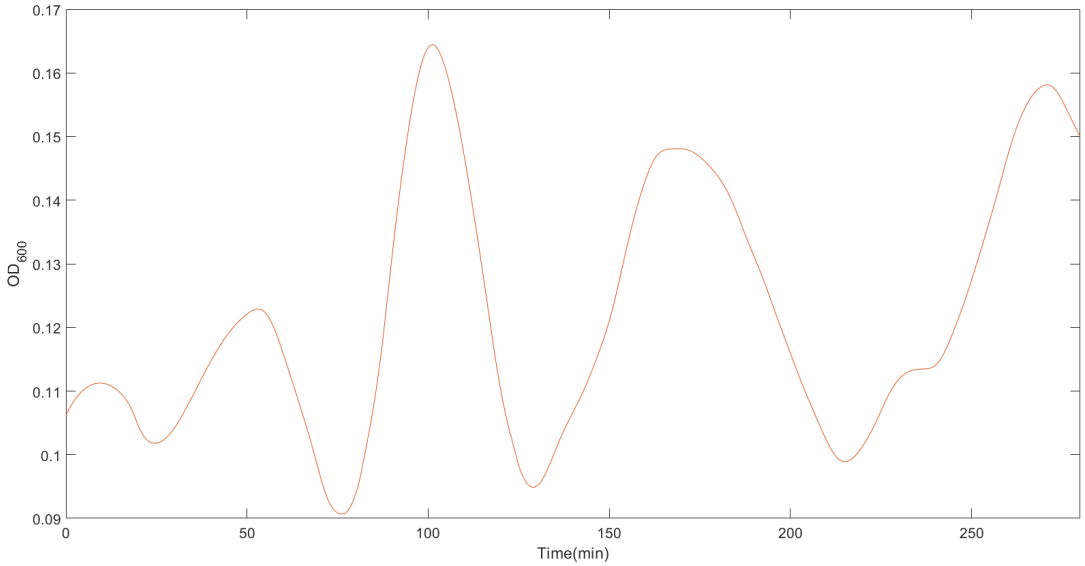

Figure 4.2. Simulation of the temporal evolution of the population density of *E. coli* population carrying the designed genetic circuit with the constructed mathematical model and the provisional values for the kinetic parameters.

From this obtained simulation results, it can be observed that the population density oscillates. Therefore, it was concluded that this genetic circuit design could give *E. coli* the oscillatory population density behavior. Following this result, the genetic circuit was built.

4.2 Build: Assembled plasmids

In order to construct the designed genetic circuit and all the plasmids necessary for the different characterizations, the GoldenBraid method was used as described in Materials and Methods chapter. Additionally, two new GoldenBraid Level 1 destination plasmids were designed and built.

4.2.1 Construction of new GoldenBraid Level 1 destination plasmids

A new pair of GoldenBraid Level 1 vectors (α1 and α2) were designed and constructed as indicated in Materials and Methods. These new two plasmids are named pARK α1 and pARK α2 and their sequences are in Appendix I.

These plasmids use a red fluorescent protein transcriptional unit as selector gene instead of lacZ. Using red fluorescence instead of white/blue selection has the advantage of avoiding the adding of X-gal and IPTG to the culture medium, reducing the complexity and cost of the transformats selection processes. However it has a disadvantage: it cannot be used for the assembly of constructions that contains a transcriptional unit with a red fluorescent protein. They were used in the following steps of this work.

4.2.2 Constructions of transcriptional units in the new Level 1 destination plasmids

Five transcriptional units were built into the new GoldenBraid Level 1 destination plasmids using the DNA basic parts as indicated in Materials and Methods:

- pARKA11: GFP transcriptional unit into pARK α1 plasmid. The expression of GFP is controlled by the pLux promoter and a strong RBS (B0030). Using only this plasmid, no effect would be observed on the bacteria except basal expression of GFP, nor with the addition of external AHL, as there is no transcriptional unit that produces the LuxR protein.
- pARKA12: LuxI transcriptional unit into pARK α1 plasmid. The expression of LuxI is controlled by a constitutive promoter and a strong RBS (B0030). With this plasmid, LuxI which is the AHL synthase, is produced in a constitutive way. Bacteria containing this plasmid produce AHL.
- pARKA13: E-lysis protein transcriptional unit into pARK α1 plasmid. The expression of the E-lysis protein is controlled by the pLux promoter and a weak RBS (B0032). As in the case of pARKA11 no effect would be observed on the bacteria, nor with the addition of external AHL, as there is no transcriptional unit that produces the LuxR protein.
- pARKA21: LuxR transcriptional unit into pARK α2 plasmid. The expression of LuxR is controlled by a constitutive promoter and a strong RBS (B0030). With this plasmid, bacteria would produce the LuxR protein in a constitutive way.
- pARKA22: The same construction as pARKA13 but into a pARK α2 destination plasmid. This plasmid was assembled to combine it with pARKA12

All these plasmids are represented in Figure 4.3 and their sequences are in Appendix I.

4.2.3 Combination of 2 transcriptional units in Level 2 destination plasmids

Three constructions with two transcriptional units each one were built into GoldenBraid Level 2 destination plasmids using the plasmids described in the previous section as indicated in Materials an Methods:

- pARSO11: Result of joining the transcriptional units of pARKA11 and pARKA21 into GoldenBraid Ω1 plasmid. With this plasmid, LuxR protein is constitutively produced and GFP is inducibly produced. When AHL is added externally, it binds to LuxR and activates the production of GFP. This plasmid was used to characterize the interaction of LuxR with the pLux promoter at different AHL concentrations.
- pARSO12: Result of joining the transcriptional units of pARKA13 and pARKA21 into GoldenBraid Ω1 plasmid. With this plasmid, LuxI is constitutively produced, and therefore bacteria produce AHL. But, as there is no transcriptional unit of LuxR, there will be no observable effect on the behavior of the bacteria except for the basal expression of E-lysis protein.

- pARSO21: Result of joining the transcriptional units of pARKA12 and pARKA22 into GoldenBraid Ω2 plasmid. With this plasmid, LuxR protein is constitutively produced and GFP is inducibly produced. When AHL is added externally, it binds to LuxR and activates the production of the E-lysis protein, producing the death of the bacteria. This plasmid was used to characterize the lethality of the E-lysis protein.

These three plasmids are represented in Figure 4.4 and their sequences are in Appendix I.

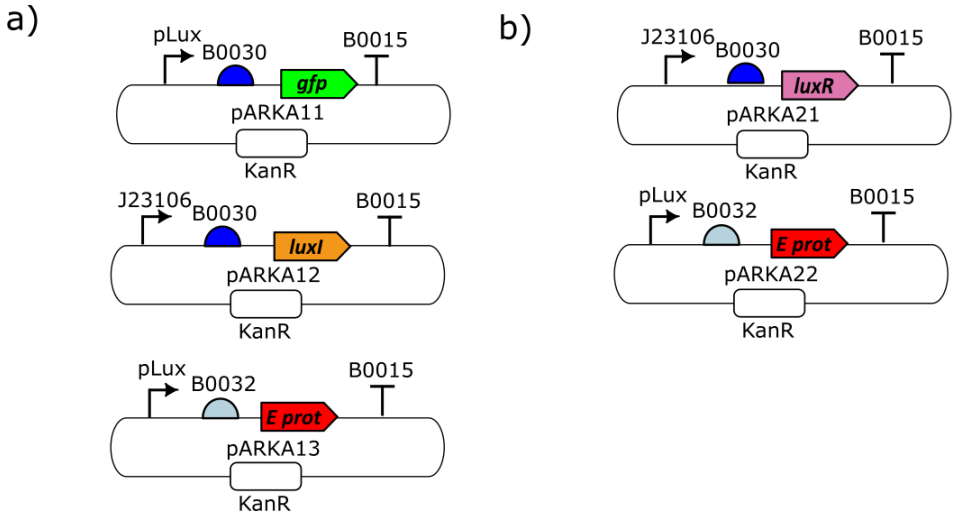

Figure 4.3. Representation of the different transcriptional units assembled in the new GoldenBraid Level 1 destination plasmids. a) Transcriptional units assembled in pARK α1 destination plasmids. b) Transcriptional units assembled in pARK α2 destination plasmids.

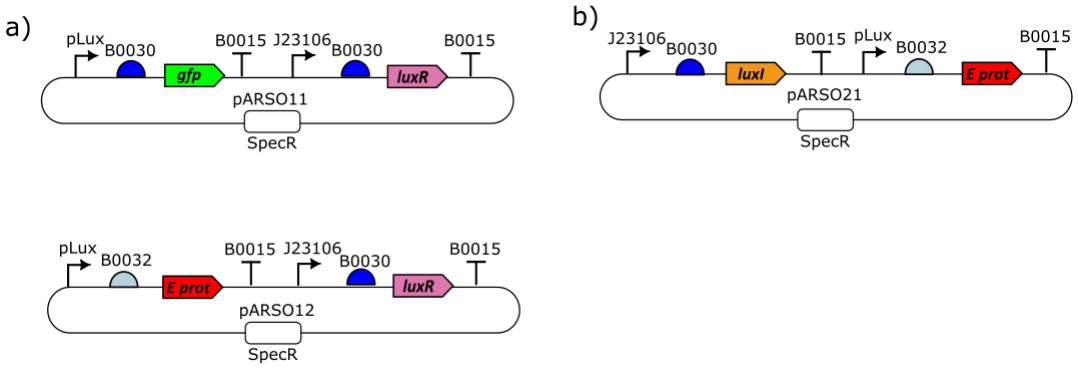

Figure 4.4. Representation of the different transcriptional units combined in GoldenBraid Level 2 destination plasmids. a) Transcriptional units combined into GoldenBraid Ω1 destination plasmids. b) Transcriptional units combined into GoldenBraid Ω2 destination plasmids.

4.2.4 Construction of the genetic circuit by combination of multiple transciptional units in a Level 1 destination plasmid

One construction with four transcriptional units was assembled into pARK α1 destination plasmid using the plasmids of the previous section as indicated in Materials an Methods. This plasmid is pARKA1_o1 and is the result of joining the construction of pARSO11 with the construction of pARSO21 into pARK α1 destination plasmid. This plasmid contains the entire genetic circuit designed in this work to give *E. coli* the oscillatory population density. It is represented in Figure 4.5 and its sequence is in Appendix I .

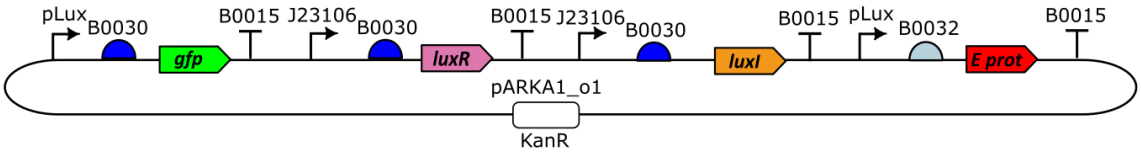

Figure 4.5. Representation of the genetic circuit designed in this work into a GoldenBraid Level 1 destination plasmid.

4.3 Test: Characterization of DNA basic parts and test of the designed genetic circuit

4.3.1 Characterization of the interaction of LuxR with the pLux promoter

The objective of this characterization was to obtain the unknown parameter values associated with the transcription in the transcriptional units that have the pLux promoter. These parameters are α_g, k_{G1} and k_{G2}. As the length of the transcripts and the context of the sequences are not being taken into account when evaluating the transcription ratios, the result obtained for these parameters will be the same that will be used for α_e, k_{E1} and k_{E2} which are the parameters that control the induced mRNA production of the E-lysis protein.

Experimental results

To obtain these parameter values, *E. coli* carrying the pARSO11 plasmid (Figure 4.4) were used in a experiment with the addition of different concentrations of external AHL. This plasmid contains a GFP transcriptional unit whose expression is activated by LuxR associated with AHL. In this experiment, the OD_{600} and the fluorescence of the cultures over time was measured as indicated in Materials and Methods. The results of the measurement were trimmed, avoiding the lag and stationary phases. Then the number of cells at each time was calculated using Equation 4.14 with the OD_{600} data (Boada et al., 2019). The next step was to calculate the molecules of equivalent fluorescein (MEFL) with the fluorescence data and with Equation 4.15 (Boada et al., 2019). Finally, the ratio between MEFL and the number of cells was obtained. These results are shown in Figure 4.6.

$$NCells = 10^{8.062}(OD_{600})^{1.185} \tag{4.14}$$

$$MEFL = 10^{12.28}(F)^{1.038} \tag{4.15}$$

Where NCells is the number of cells in a culture, OD_{600} is the optical density of the culture at λ= 600nm, MEFL is the number of molecules of equivalent fluorescein of the culture and F is the green fluorescence measured in the culture (arbitrary units).

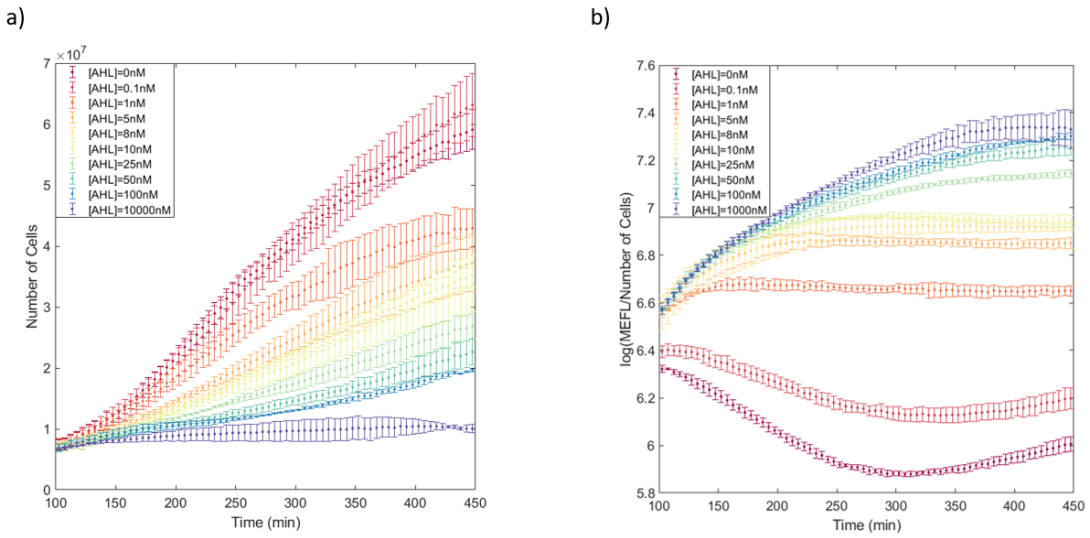

Figure 4.6. Temporal evolution of the population density (number of cells) and the expression of the GFP per cell of *E. coli* carrying the pARSO11 plasmid at different AHL concentrations. a) *E. coli* population density (number of cells) over time. b) Expression of GFP over time as decimal logarithm of the ratio between the molecules of equivalent fluorescein and the number of cells over time.

As can be seen in Figure 4.6, the greater the addition of AHL to the medium, the greater is the expression of GFP performed by each cell. An effect of AHL on growth can also be seen: the higher the concentration of AHL, the lower the growth. This may be due to a possible toxic effect of high concentrations of the inducer on the bacteria. This effect will be taken into account in the model parameter values estimation. The data of the bacteria with [AHL]=1000 nM will not be used in the parameter values estimation process because they did not grow.

Model parameter values estimation

To estimate α_g, k_{G1} and k_{G2}, a mathematical model that does not have the ODEs of the mRNA of LuxI, the LuxI protein, the mRNA of the E-lysis protein and the E-lysis protein species was developed. Nor does it have the effects of the lethal gene on growth or the production of intracellular AHL. This model is in Appendix III.

The first step was to obtain the parameter values of the growth equation. To do this, a value for the parameter K_{mx} which represents the maximum growth capacity and should be the same for all conditions, was obtained by performing the multi-objective optimization process explained in

Materials and Methods and choosing the value that minimizes the error between the simulated and the experimental number of cells for the entire data set. Then, using this value, the value of the specific growth rate (μ) for each experimental condition was obtained using the same process but selecting the value which minimizes the error for each particular case. The values selected are in Table 4.2, and the result of a simulation with these values and its comparison with the experimental results is in Figure 4.7.

Table 4.2. Estimated values of the parameters of the growth equation for the different AHL amounts used in the characterization of the interaction of LuxR with the pLux promoter.

Parameter	**Value**
K_{mx}	$7.29 \cdot 10^7$ cells
μ_0	0.0118 min^{-1}
$\mu_{0.1}$	0.0115 min^{-1}
μ_1	0.0093 min^{-1}
μ_5	0.0074 min^{-1}
μ_8	0.0068 min^{-1}
μ_{10}	0.0067 min^{-1}
μ_{25}	0.0055 min^{-1}
μ_{50}	0.0050 min^{-1}
μ_{100}	0.0050 min^{-1}

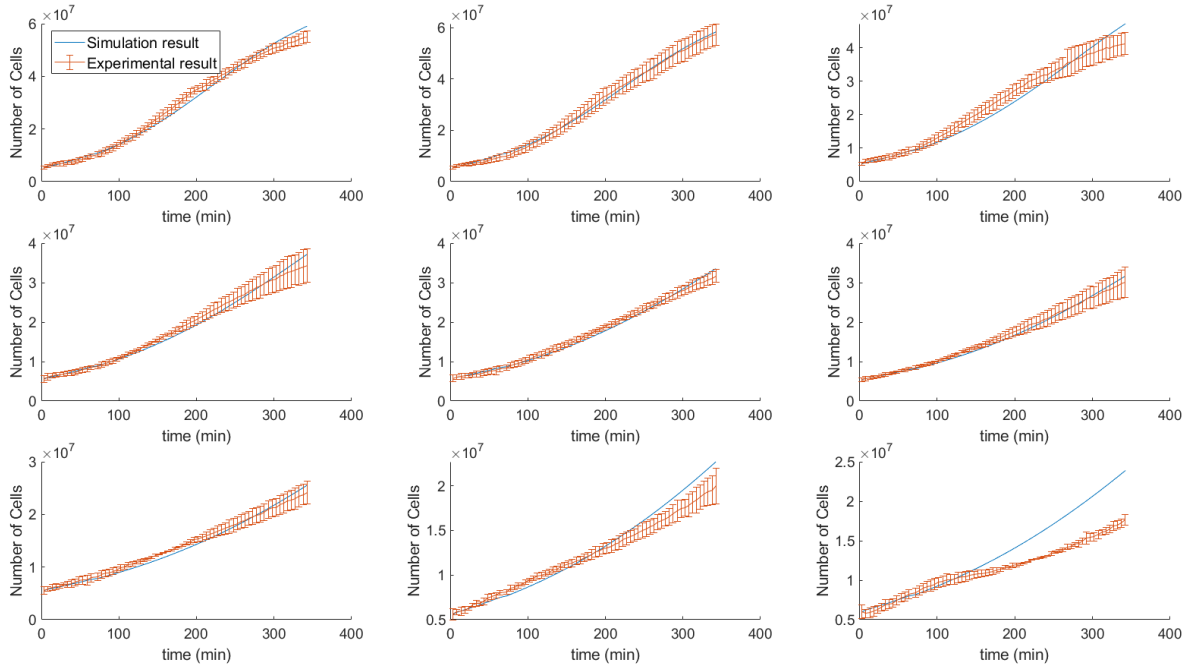

Figure 4.7. Comparison between the experimental and the simulated results for the population density (number of cells) over time of *E. coli* carrying the pARSO11 plasmid at different AHL concentrations. From left to right and from top to bottom the concentrations of AHL are: 0 nM, 0.1 nM, 1 nM, 5 nM, 8 nM, 10 nM, 25 nM, 50 nM and 100 nM.

As it can be seen in Figure 4.7, simulations with the model using the estimated parameters values faithfully represent reality, except in the case of [AHL]=100 nM, which presents an experimental growth curve very different from the expected one.

The next step was obtaining the parameters α_g, k_{G1} and k_{G2}. In this case, the objective was to minimize the error between the simulated and the experimental fluorescence data in steady state (a long time after induction). With this process, the values obtained were: α_g=0.0849 min^{-1}, k_{G1}=20.7767 min^{-1} and k_{G2}=0.0100 molecules. With these estimated values, a representation of the steady state with both the simulation and the experimental data was done. It is in Figure 4.8.

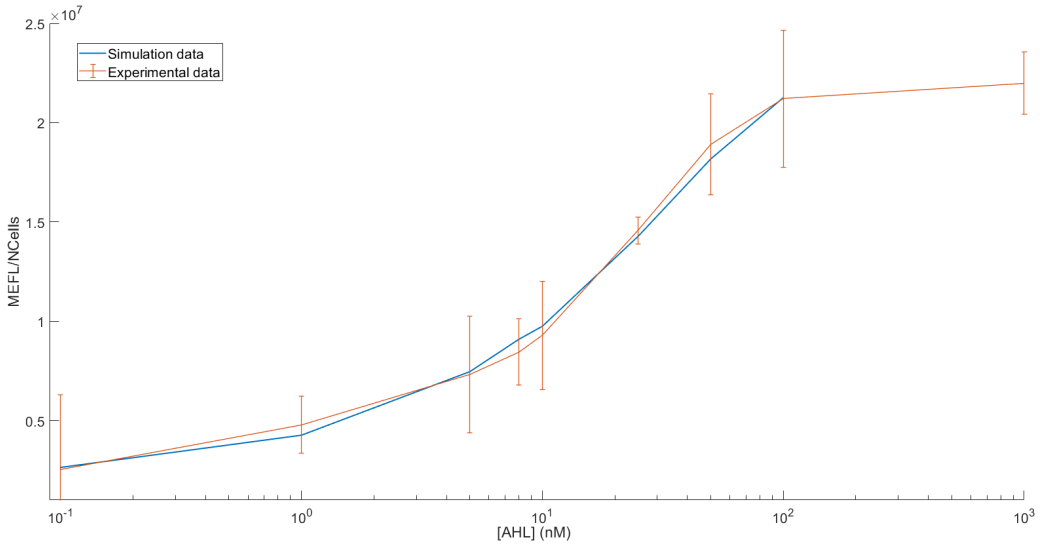

Figure 4.8. Comparison between the experimental and the simulated results of the GFP expression per cell of *E. coli* carrying the pARSO11 plasmid in steady state at different AHL concentrations. The fluorescence per cell is represented, where the fluorescence is indicated in molecules of equivalent fluorescein.

4.3.2 Characterization of the lethality of the E-lysis protein

The objective of this characterization was to obtain those unknown parameter values that define the lethality of the E-lysis protein: K_{me}, n and Θ. K_{E5} which is the degradation rate of the E-lysis protein was also obtained with this characterization.

Experimental results

To perform this characterization, *E. coli* carrying the plasmid pARSO12 (Figure 4.4) were used in a experiment with the addition of different concentrations of external AHL. This plasmid contains an E-lysis protein transcriptional unit whose expression is activated by LuxR associated with AHL. The OD_{600} of the cultures over time was measured as indicated in Materials and Methods. The results were used to calculate the number of cells using Equation 4.14. This reflects the *E. coli* population density and consequently the lethality of the E-lysis protein. The results are represented in Figure 4.9.

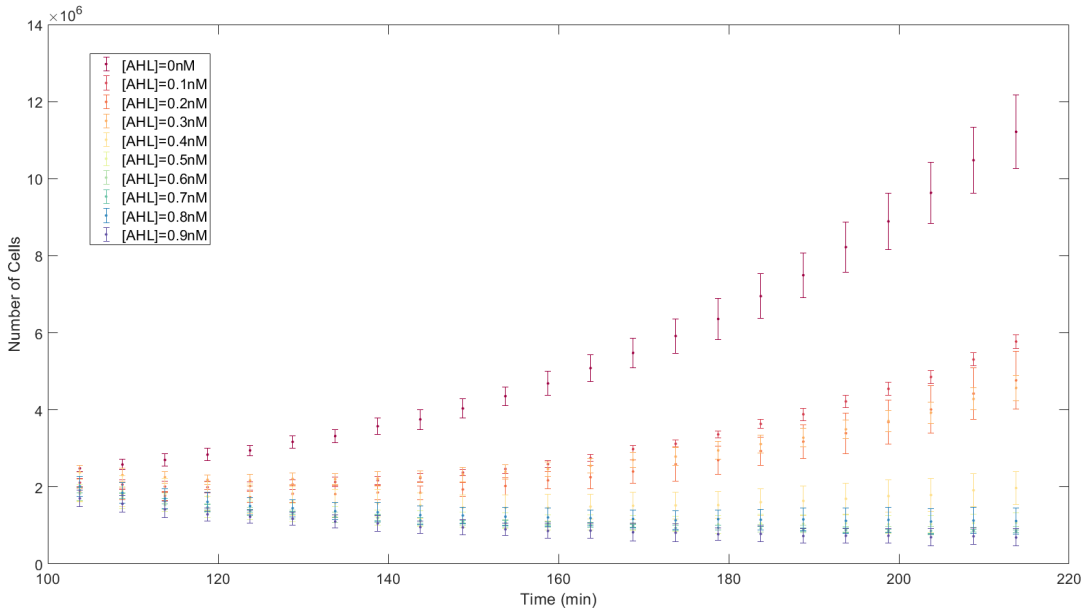

Figure 4.9. Temporal evolution of the *E. coli* population density with the pARSO11 plasmid incorporated over time at different AHL concentrations.

In Figure 4.9, it can be observed that the higher the concentration of AHL, and therefore the higher the expression of E-lysis protein, the higher the cell death. In addition, by using low inductor concentrations, the effect of AHL toxicity on bacteria can be considered to be negligible. Furthermore, the fact that lysis of only part of the population occurs at low concentrations leads to not considering the term of AHL toxicity in the growth equation in the model of the complete genetic circuit.

As a future perspective to improve this characterization a quantification of the E-lysis protein could be carried out by means of the fusion of this protein with a fluorescent protein, so that it would be possible to know in a more exact way the number of proteins that there are. In this way, it would not be necessary to depend on the parameters obtained in the characterization of the interaction of LuxR with the pLux promoter to obtain an approximation of the number of proteins there are.

Model parameter estimation

To estimate K_{me}, n, Θ and K_{E5} a mathematical model that does not have the ODEs of the LuxI mRNA, the LuxI protein, the GFP mRNA and the GFP species was developed. Nor does it have the production of intracellular AHL. This model is in Appendix III. To obtain the parameter values, an optimization process was performed as indicated in Material and Methods in order to reduce the error between the experimental an the simulated number of cells over time. In the optimization process, 3 groups with different behavior were identified corresponding to three AHL concentration ranges: 0 nM, 0.1-0.4 nM; >0.4 nM. In this way, experimental data representative of each group were chosen to simplify the optimization process carried out. These representatives were the data from the experiments at [AHL]=0 nM, [AHL]=0.4 nM and [AHL]=0.7 nM. With these data, other step of optimization was performed and the values estimated for the parameters were K_{me}=0.366 min^{-1}, n=2.125, Θ=7359.91 molecules and k_{E5}=0.478 min^{-1}. μ and K_{mx} were

also estimated and the values obtained were μ=0.0170 min^{-1} and K_{mx}=3.313·10^7 cells. With these value and the mathematical model designed for this case, a simulation was performed, and the comparison between the simulated and the experimental number of cells over time is in Figure 4.10.

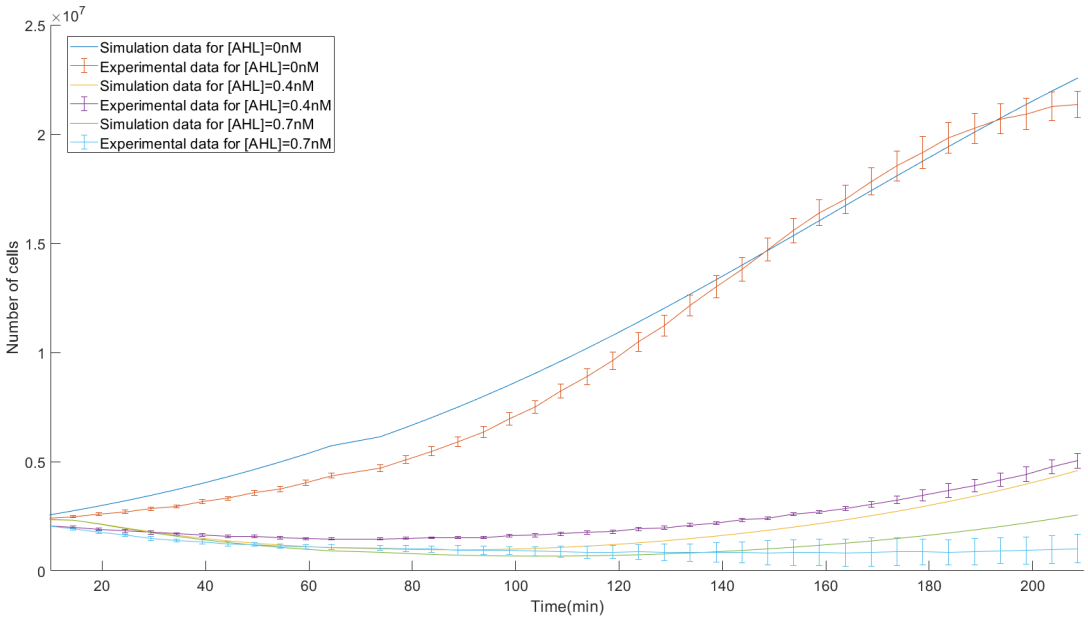

Figure 4.10. Comparison between the experimental and the simulated population density (number of cells) over time of *E. coli* carrying the pARSO13 plasmid at different AHL concentrations.

4.3.3 Test of the genetic oscillator circuit

To test if the designed genetic circuit gives *E. coli* population density an oscillatory behaviour, *E. coli* carrying this genetic circuit (pARK1_o1 plasmid) were used in a experiment. This plasmid contains four transcriptional units: an E-lysis protein transcriptional unit whose expression is activated by LuxR associated with AHL, a GFP transcriptional unit whose expression is activated by LuxR associated with AHL, a LuxI constitutive transcriptional unit and a LuxR constitutive transcriptional unit. The OD_{600} of the *E. coli* cultures over time was measured as indicated in Materials and Methods. The results were used to was used to calculate the number of cells using Equation 4.14. The results obtained are represented in Figure 4.11. It is important to note that having used OD_{600} as a measure to obtain the number of cells has the limitation that dead cells also contribute to OD_{600}. Therefore, the observed oscillation in Figure 4.11 is likely to be less pronounced than it is supposed to be.

These results can be compared with the results of studies with similar conditions of other genetic circuits that give the bacteria an oscillatory population density. In the work of Din et al. (2016), whose genetic circuit was the starting point of this work, they carried out a study with *Salmonella enterica* in a batch culture when they were trying to tune the circuit by changing the LuxI degradation rate. The result of the OD_{600} over time obtained in this work is very similar to the one obtained by Din et al. (2016), confirming that it is possible to obtain oscillations with the proposed genetic circuit in *E. coli*.

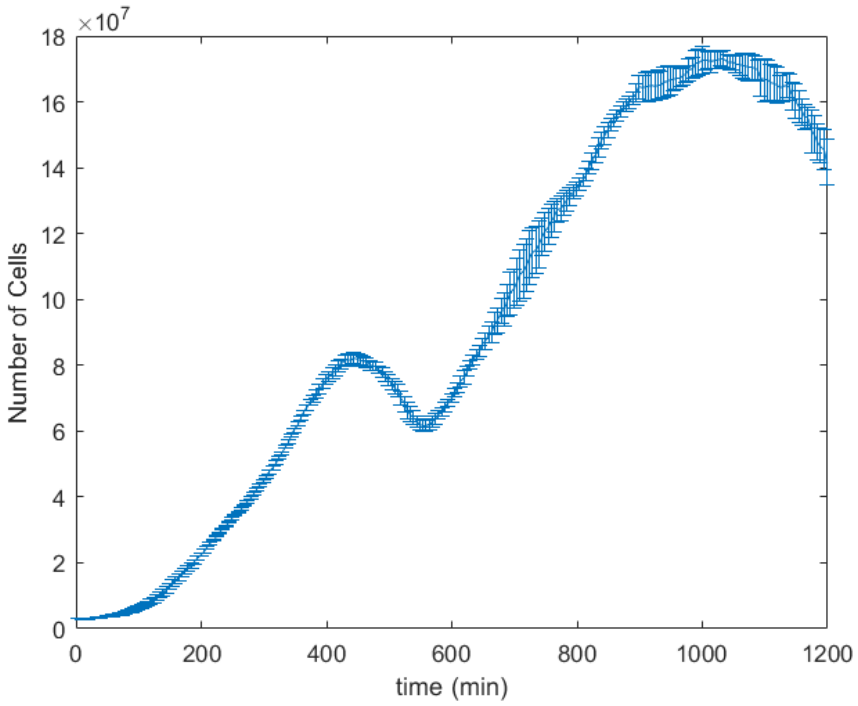

Figure 4.11. Evolution of the population density (number of cells) over time of *E. coli* carrying the pARSO13 plasmid at different AHL concentrations.

In spite of this, the results and interpretations that can be obtained from a genetic circuit of this type in a batch culture are limited since the number of hours that the cultures can be maintained without exhausting the resources of the medium are not high enough to reach at least two cycles of oscillation. In addition, from the first time that the threshold at which cells die is reached, the contents of the interior of the cells will be released into the medium and, by that time, part of the resources of the medium will have been consumed. For this reason, the conditions of the medium from the second cycle of oscillation are different from the initial ones. Finally, extracellular AHL has a very slow degradation rate, causing that the time that the concentration of AHL remains high after a cell death event is long, and that the time it takes for population density to increase again is even longer.

For all these reasons, the use of a microfluidic device as used by Din et al. (2016) to evaluate the genetic circuit, or the use of a continious culture in a turbidostat are proposed as a future consideration to test the behavior of this genetic circuit.

4.4 Redesign of the genetic circuit

Once the information from the test of the system is obtained, it is possible to determine which model parameters, and, consequently, which experimental components should be adjusted to optimize the behavior of the system. In our particular case, some of these parameters are the following:

- k_{E3}: it represents the translation rate of the E-lysis protein. Reducing this parameter would increase the amplitude of the oscillations as it would reduce the expression of E-lysis protein. This parameter could be changed modifying the corresponding RBS.
- k_{E5}: it represents the degradation rate of the E-lysis protein. Increasing this parameter would increase the amplitude of the oscillations as it would reduce the amount of E-lysis protein. Regarding the valley of the oscillations, its length would decrease because the amount of protein after a death event would decrease faster. This parameter could be changed adding a degradation tag to the E-lysis protein coding sequence.
- The parameters of the pLux promoter (k_{E1}, k_{E2} and α_e) can be modify to change the affinity of the TF to the DNA. In this way, increasing the affinity of the TF to the DNA would increase the sensitivity of the cells to AHL and would decrease the amplitude of the oscillations, as it would need less cells producing AHL to produce a death event. These parameters could be changed doing site-directed mutagenesis in the pLux promoter sequence.
- The growth equation parameters (μ and K_{mx}) can be changed to imodify the amplitude or the period of the oscillations. They could be changed, modifying the culture media or the bacterial strain.

5 Conclusions

A synthetic biology methodology, based on the design, build, and test cycle, was successfully used for the rational design of a genetic oscillator to establish an internal control of population density in *E. coli* cultures. This genetic oscillator is based on synchronized cycles of cell lysis, using the cellular communication of *Vibrio fischeri* and the E-lysis protein of phage φ174. The successful implementation of the synthetic biology methodology included the following achievements:

1. The *in silico* reproduction of the desired behavior of the cell population containing the designed genetic circuit regarding population density dynamics was achieved. This was accomplished with the creation of a deterministic mathematical model of the designed genetic circuit which is a system of 13 ODEs with 38 parameters.

2. The characterization of DNA basic parts carried out in this work allowed the estimation of the unknown parameter values of the mathematical model. In particular, for the pLux promoter, the parameters associated with the binding affinity of the TF to the pLux promoter were estimated. For example, the value of the maximum rate of transcription with this promoter (k_{G1}) was estimated to be 20.78 min^{-1}. In the other hand, the parameters estimated for the E-lysis protein defined the lethality of this protein.

3. Upon assembly of the designed genetic circuit using the GoldenBraid methodology, with a total of 8 new intermediate plasmids, *E. coli* showing an oscillatory population density behavior were obtained. Moreover, some of these new assemblies allowed to characterize the DNA basic parts used in this work, as mentioned before. Furthermore, the obtained oscillatory behavior showed to be compatible to the data from similar genetic circuits from the bibliography (Din et al., 2016).

4. Taking into account the information obtained in this work, including both simulations with the model and experimental data, future considerations to optimize the design of this genetic circuit could be taken. In particular, parameters that affect the expression of the E-lysis protein, the degradation of the E-lysis protein, the sensitivity of cells to AHL or the growth of bacteria can be changed to modify the characteristics of the oscillation.

With all this, this work proposes a methodology that can be used to design and develop new genetic circuits based on synchronized cycles of cell lysis that achieve the control of the population density and the control of the expression of a protein of interest such as a therapeutic protein for *in vivo* drug delivery in a way that it is expressed periodically.

Bibliography

ANDERSON, J.C.; VOIGT, C.A. and ARKIN, A.P (2007). *Environmental signal integration by a modular AND gate.* Molecular systems biology 3, pp. 133–133.

ANDRIANATOANDRO, E.; BASU, S.; KRAIG, D.K. and WEISS, R. (2006). *Synthetic biology: new engineering rules for an emerging discipline.* Molecular Systems Biology 2,2006.0028.

ARKIN, A.P. (2001). *Synthetic cell biology.* Current Opinion in Biotechnology 12.6, pp. 638–644.

BLASCO, X.; HERRERO, J.M. and REYNOSO-MEZA, G. (2017). *Interactive Tool for Analyzing Multiobjective Optimization Results with Level Diagrams.* Proceedings of the Genetic and Evolutionary Computation Conference Companion. GECCO '17. Berlin, Germany: ACM, pp. 1689–1696.

BOADA ACOSTA, Y. F. (2018). *A systems engineering approach to model, tune and test synthetic gene circuits.* PhD thesis.

BOADA, Y.; VIGNONI, A.; ALARCON-RUIZ, I.; ANDREU-VILARROIG, C.; MONFORT-LLORENS, R.; REQUENA, A. and PICÓ, J.(2019). *Characterization of gene circuits parts based on multi-objective optimization with standard calibrated measurement.* ChemBioChem. Accepted author manuscript.

BOADA, Y.; VIGNONI, A.; REYNOSO-MEZA, G.; PICÓ, J.*Parameter identification in synthetic biological circuits using multiobjective optimization.* 49.26. Foundations of Systems Biology in Engineering - FOSBE 2016, pp. 77–82.

BORNHOLDT, S. (2005). *Less Is More in Modeling Large Genetic Networks.* Science 310.5747, pp. 449–451.

CAMERON, D.; BASHER, C.and COLLINS, J.J. (2014). *A brief history of synthetic biology.* Nature Reviews Microbiology 12, pp. 381-.

CANTON, B.; LABNO A.; and ENDY D. (2008). *Refinement and standardization of synthetic biological parts and devices.* Nature Biotechnology 26, pp. 787–.

CERONI, F.; FURINI, S.; GIORDIANO, E. and CAVALCANTI, S. (2010). *Rational design of modular circuits for gene transcription: A test of the bottom-up approach.* Journal of Biological Engineering 4.1, pp. 14–.

CHANDRAN, D.; COPELAND, W.B.; SLEIGHT, S.C. and SAURO, H.M. (2008). *Mathematical modeling and synthetic biology.* Drug Discovery Today: Disease Models 5.4. Animal models of addiction / Kinetic models, pp. 299–309.

CHAO, R.; YUAN, Y. and ZHAO, H. (2015). *Recent advances in DNA assembly technologies.* FEMS Yeast Research 15.1, pp. 1–9.

CHELLABOINA, V.; BHAT, S.J.; HADDAD, W.M. and BERNSTEIN, D.S. (2009). *Modeling and analysis of mass-action kinetics.* IEEE Control Systems Magazine 29.4, pp. 60–78.

CHEN, W.W.; NIEPEL, M. and SORGER, P.K. (2010). *Classic and contemporary approaches to modeling biochemical reactions.* Genes development 24.17. 24/17/1861[PII], pp. 1861–1875.

DE LORENZO, V. and DANCHIN, A. (2008). *Synthetic biology: discovering new worlds and new words.* EMBO reports 9.9, pp. 822–827.

DECOENE, T.; DE PAEPE, B.; MAERTES, J.; COUSSEMENT, P.; PETERS, G.; DE MAESENEIRE, S.L. and DE NEY, M. (2018). *Standardization in synthetic biology: an engineering discipline coming of age.* Critical Reviews in Biotechnology 38.5. pp. 647–656.

OMAR DIN, M.; DANINO, T.; PRINDLE, A.; SKALAK, M.; SELIMKHANOV, J.; ALLEN, K.; JUKIO, E.; ATOLIA, E.; TSIMRING, L.S.; BHATIA, S.M. and HASTY, J. (2016). *Synchronized cycles of bacterial lysis for in vivo delivery.* Nature 536, PP. 81-85.

ELLIS, T.; ADIE, T. and BALDWIN G.S. (2011). *DNA assembly for synthetic biology: from parts to pathways and beyond.* Integrative Biology 3.2, pp. 109–118.

ELOWITZ, M.B. and LEIBLER S. (2000). *A synthetic oscillatory network of transcriptional regulators.* Nature 403.6767, pp. 335–338.

ENDLER, L.; RODRIGUEZ, N.; JUTY, N.; CHELLIAH, V.; LAIBE, C.; LI, C. and LE NOVÈRE, N. (2009). *Designing and encoding models for synthetic biology.* Journal of The Royal Society Interface 6.suppl4, S405–S417.

ENDY, D. (2005). *Foundations for engineering biology.* Nature 438.7067, pp. 449–453.

ENGLER, C.; KANDZIA, R. MARILLONNET, S. (2008). *A One Pot, One Step, Precision Cloning Method with High Throughput Capability.* PLOS ONE 3.11, pp. 1–7.

GARDNER, T.S.; CANTOR, C.R. and COLLINS, J.D. (2000). *Construction of a genetic toggle switch in Escherichia coli.* Nature 403.6767, pp. 339–342.

GOLDBETER, A.; GÉRARD, C.; GONZE, D.;LELOUP, J.C and DUPONT, G. (2012). *Systems biology of cellular rhythms.* In: FEBS Letters 586.18, pp. 2955–2965.

KNIGHT, T. (2008). *Draft Standard for Biobrick BB-2 Biological Parts.* url: https://dspace.mit.edu/handle/1721.1/45139 (visited on 30/05/2019).

LILACCI, G. and KHAMMASH, M. (2010). *Parameter Estimation and Model Selection in Computational Biology.* In: PLOS Computational Biology 6.3, pp. 1–17.

LIU, Z.; ZHANG, J.;, JIN, J.; GENG Z.; QI, Q. and LIANG, Q. (2018). *Programming Bacteria With Light-Sensors and Applications in Synthetic Biology.* Frontiers in microbiology 9, pp. 2692–2692.

LOENEN, W.A.; DRYDEN, D.T.; RALEIGH, E.A.; WILSON, G.G.; and MURRAY, N.E. (2014). *Highlights of the DNA cutters: a short history of the restriction enzymes.* Nucleic acids research, 42(1), pp. 3–19.

LUCIGEN (2019). *E. cloni® 10G and 10GF' Electrocompetent Cells.* url: https://www.lucigen.com/E.-cloni-10G-and-10GF-Electrocompetent-Cells/ (visited on 04/06/2019).

MARGUET, P.; TANOUCHI, Y.; SPITZ, E.; SMITH, C. and You, L. (2010). *Oscillations by minimal bacterial suicide circuits reveal hidden facets of host-circuit physiology.* PloS one, 5(7), e11909.

MORGAN, K. (2014). *Plasmids 101: Origin of Replication.* url: https://blog.addgene.org/plasmid-101-origin-of-replication (visited on 03/06/2019).

MUKHERJI, S. and VAN OUDENAARDEN, A. (2009). *Synthetic biology: understanding biological design from synthetic circuits.* Nature Reviews Genetics 10, pp. 859–.

OTERO-MURAS, I. and BANGA, J.R. (2016). *Design Principles of Biological Oscillators through Optimization: Forward and Reverse Analysis.* PLOS ONE 11.12, pp. 1–26.

PICÓ, J.; VIGNONI, A.; PICÓ-MARCO, E. y BOADA, Y. (2015). *Modelado de sistemas bioquímicos: De la Ley de Acción de Masas a la Aproximación Lineal del Ruido.* Revista Iberoamericana de Automática e Informática Industrial RIAI 12.3, pp. 241–252.

PIERCE, B. A. (2016). Genética : un enfoque conceptual. spa. Madrid. ISBN: 9788498353921.

POMERENING, J.R.; KIM, S.Y. and FERRERL J.E. (2005). *Systems-Level Dissection of the Cell-Cycle Oscillator: Bypassing Positive Feedback Produces Damped Oscillations.* Cell 122.4, pp. 565–578.

PURCELL, O.M; SAVERY, N.J.; GRIERSON, C.S. and DI BERNARDO, M. (2010). *A comparative analysis of synthetic genetic oscillators.* Journal of The Royal Society Interface 7.52, pp. 1503–1524.

PURNICK, P.E.M. and WEISS, R. (2009). *The second wave of synthetic biology: from modules to systems.* Nature Reviews Molecular Cell Biology 10, pp. 410–.

QIN, N.; CALLAHAN, S.M.; DUNLAP, P.V. and STEVENS, A.M. (2007). A*nalysis of LuxR Regulon Gene Expression during Quorum Sensing in Vibrio fischeri.* Journal of Bacteriology 189.11, pp. 4127–4134.

REYNOSO-MEZA, G.; GARCÍA-NIETO, S; SANCHÍS, J. and BLASCO, F.X. (2013). *Controller Tuning by Means of Multi-Objective Optimization Algorithms: A Global Tuning Framework.* IEEE Transactions on Control Systems Technology 21.2, pp. 445–458.

SARRION-PERDIGONES, A.; FALCONI, E.E.; ZANDALINAS, S.I.; JUAREZ, P.; FERNÁNDEZ-DEL-CARMEN, A.; GRANDELL, A. and ORZAEZ, D.(2011). *GoldenBraid: An Iterative Cloning System for Standardized Assembly of Reusable Genetic Modules.* PLOS ONE 6.7, pp. 1–11.

SISO-NADAL, F.; OLLIVER, J.F. and SWAIN, P.S. (2007). *Facile: a command-line network compiler for systems biology.* BMC Systems Biology 1.1, pp. 36–.

SLOMOVIC, S.; KEITH, P. and COLLINS, J.J. (2015). *Synthetic biology devices for in vitro and in vivo diagnostics.* Proceedings of the National Academy of Sciences 112.47, pp. 14429–14435.

STRICKER, J.; COOKSON, S.; BENNET, M.R.; MATHER, W.H.; TSIMRING, L.S. and HASTY, J.(2008). *A fast, robust and tunable synthetic gene oscillator.* Nature 456, 516-519.

THE ROYAL ACADEMY OF ENGINEERING (2009). Synthetic Biology: scope, applications and implications. Technical Report. The Royal Academy of Engineering. ISBN: 1-903496-44-6

TOMAZOU , M.; BARAHOMA, M.; POLIZZI, K.M. and STAN, G.B. (2018). *Computational Re-design of Synthetic Genetic Oscillators for Independent Amplitude and Frequency Modulation.* Cell Systems 6.4, 508–520.e5.

TYSON, J.J.; ALBERT, R.; GOLDBETER, A.; RUOFF, P. and SIBLE, J. (2008) *Biological switches and clocks.* In: Journal of The Royal Society Interface 5.suppl_1, S1–S8.

URIU, K.(2016). *Genetic oscillators in development.* Development, Growth Differentiation 58.1, pp. 16–30.

VALENCIA UPV IGEM (2018a). Valencia UPV 2018 Model. url: http://2018.igem.org/Team:Valencia_UPV/Model (visited on 13/06/2019).

VALENCIA UPV IGEM (2018b). Valencia UPV 2018 part collection. url: http://2018.igem.org/Team:Valencia_UPV/Part_Collection (visited on 24/05/2019).

VAZQUEZ-VILAR, M.; QUIJANO-RUBIO, A.; FERNANDEZ-DEL-CARMEN, A.; SARRION-PERDIGONES, A.; OCHOA-FERNANDEZ, R.; ZIARSOLO, P.; BLANCA, J.; GRANELL, A. and ORZAEZ, D. (2017). *GB3.0: a platform for plant bio-design that connects functional DNA elements with associated biological data.* Nucleic Acids Research 45.4, pp. 2196–2209,

VILLAVERDE, A.F. and BANGA, J.R. (2014). *Reverse engineering and identification in systems biology: strategies, perspectives and challenges.* Journal of The Royal Society Interface 11.91, p. 20130505

WEBER, E.; ENGLER, C; GRUETZNER, R.; WERNER, S and MARILLONNET, S. (2011). *A Modular Cloning System for Standardized Assembly of Multigene Constructs.* PLOS ONE 6.2, pp. 1–11.

WENINGER, A.; KILLINGER, M. and VOGL, T. (2016). *Key Methods for Synthetic Biology: Genome Engineering and DNA Assembly.* In: Synthetic Biology. Ed. by Anton Glieder et al. Cham: Springer International Publishing, pp. 101–141.

WIRLING, C.; HERWIG, R., and LEHRACH, H. (2007). *Resources, standards and tools for systems biology.* Briefings in Functional Genomics 6.3, pp. 240–251.

YOU, L.; COX III, R.S; WEISS, R. and ARNOLD, F.H. (2004). *Programmed population control by cell-cell communication and regulated killing.* Nature 428.6985, pp. 868–871.

Appendix I: DNA sequences

Oligonucleotides

bbfw:

TATCTGCAGTGACAGGATATATTGGCGGG

bbrev:

TATGAATTCGGGTTACACCACAATATATGGTGCC

mrfpalpha1fw:

TATGAATTCCGTCTCAGGAGAGAGACCTTTACGGCTAGCTCAGTCCTAG

mrfpalpha1rev:

ATACTGCAGCGTCTCATGACAGCGAGAGACCTATAAACGCAGAAAGGCCCACC

mrfpalpha2fw:

TATGAATTCCGTCTCAGTCAGGAGAGAGACCTTTACGGCTAGCTCAGTCCTAG

mrfpalpha1rev:

ATACTGCAGCGTCTCAAGCGAGAGACCTATAAACGCAGAAAGGCCCACC

VerTUfw:

GCAACCTCTCGGGCTTCTGGAT

VerTUrv:

ACAGCGACTTAGTTTACCCGCCA

Plasmids

pARK α1:

CCTTGGCTTGTGGACAATGCGCTACGCGCACCGGCTCCGCCCGTGGACAACCGCAA
GCGGTTGCCCACCGTCGAGCGCCTTTGCCCACAACCCGGCGGCCGGCCGCAACAGA
TCGTTTTATAAATTTTTTTTTTTGAAAAAGAAAAAGCCCGAAAGGCGGCAACCTCT
CGGGCTTCTGGATTTCCGATCCCCGGAATTAGAGATCTTGGCAGGATATATTGTGG
TGTAACCCGAATTCCGTCTCAGGAGAGAGACCTTTACGGCTAGCTCAGTCCTAGGT
ATAGTGCTAGCTACTAGAGATTAAAGAGGAGAAATACTAGATGGCTTCCTCCGAA
GACGTTATCAAAGAGTTCATGCGTTTCAAAGTTCGTATGGAAGGTTCCGTTAACGG
TCACGAGTTCGAAATCGAAGGTGAAGGTGAAGGTCGTCCGTACGAAGGTACCCAG
ACCGCTAAACTGAAAGTTACCAAAGGTGGTCCGCTGCCGTTCGCTTGGGACATCCT
GTCCCCGCAGTTCCAGTACGGTTCCAAAGCTTACGTTAAACACCCGGCTGACATCC
CGGACTACCTGAAACTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGTTATGAAC
TTCGAAGACGGTGGTGTTGTTACCGTTACCCAGGACTCCTCCCTGCAAGACGGTGA
GTTCATCTACAAAGTTAAACTGCGTGGTACCAACTTCCCGTCCGACGGTCCGGTTA
TGCAGAAAAAAACCATGGGTTGGGAAGCTTCCACCGAACGTATGTACCCGGAAGA
CGGTGCTCTGAAAGGTGAAATCAAAATGCGTCTGAAACTGAAAGACGGTGGTCAC
TACGACGCTGAAGTTAAAACCACCTACATGGCTAAAAAACCGGTTCAGCTGCCGGG
TGCTTACAAAACCGACATCAAACTGGACATCACCTCCCACAACGAAGACTACACCA
TCGTTGAACAGTACGAACGTGCTGAAGGTCGTCACTCCACCGGTGCTTAATAACGC
TGATAGTGCTAGTGTAGATCGCTACTAGAGCCAGGCATCAAATAAAACGAAAGGC
TCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCT
ACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAGGTCTCTC
GCTGTCATGAGACGCTGCAGTGACAGGATATATTGGCGGGTAAACTAAGTCGCTG
TATGTGTTTGTTTGAGATCTCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACC
GTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCAT
CACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGAT
ACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCG
CTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAG
CTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTG
TGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTT
GAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAG
GATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCT
AACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGT
TACCTTCGGAAGAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA
GCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCA
AGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCAC
GTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTA
AATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGTGTAACATTGGTCTAG
TGATTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGAT
TATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCG
AGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCC
AACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGA
AATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCT

TTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATC
AACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGTCGAAATACGCGATCGC
TGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGC
CAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATG
CTGTTTTCCCTGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATA
AAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCAT
CTCATCTGTAACAACATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTG
GCGCATCGGGCTTCCCATACAATCGGTAGATTGTCGCACCTGATTGCCCGACATTA
TCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGG
CCTTGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGT
TTATGTAAGCAGACAGTTTTATTGTTCATGATGATATATTTTTATCTTGTGCAATG
TAACATCAGAGATTTTGAGACACAACGTGGCTTTGTTGAATAAATCGAACTTTTGC
TGAGTTGAAGGATCAGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAA
AGCAAAAGTTCAAAATCACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGT
GGCTCCCTCACTTTCTGGCTGGATGATGGGGCGATTCAGGCGATCCCCATCCAACA
GCCCGCCGTCGAGCGGGCTTTTTTATCCCCGGAAGCCTGTGGATAGAGGGTAGTT
ATCCACGTGAAACCGCTAATGCCCCGCAAAGCCTTGATTCACGGGGCTTTCCGGCC
CGCTCCAAAAACTATCCACGTGAAATCGCTAATCAGGGTACGTGAAATCGCTAATC
GGAGTACGTGAAATCGCTAATAAGGTCACGTGAAATCGCTAATCAAAAAGGCACG
TGAGAACGCTAATAGCCCTTTCAGATCAACAGCTTGCAAACACCCCTCGCTCCGGC
AAGTAGTTACAGCAAGTAGTATGTTCAATTAGCTTTTCAATTATGAATATATATAT
CAATTATTGGTCGC

pARK α2:

CTCAGTCAGGAGAGAGACCTTTACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCTA
CTAGAGATTAAAGAGGAGAAATACTAGATGGCTTCCTCCGAAGACGTTATCAAAG
AGTTCATGCGTTTCAAAGTTCGTATGGAAGGTTCCGTTAACGGTCACGAGTTCGAA
ATCGAAGGTGAAGGTGAAGGTCGTCCGTACGAAGGTACCCAGACCGCTAAACTGA
AAGTTACCAAAGGTGGTCCGCTGCCGTTCGCTTGGGACATCCTGTCCCCGCAGTTC
CAGTACGGTTCCAAAGCTTACGTTAAACACCCGGCTGACATCCCGGACTACCTGAA
ACTGTCCTTCCCGGAAGGTTTCAAATGGGAACGTGTTATGAACTTCGAAGACGGT
GGTGTTGTTACCGTTACCCAGGACTCCTCCCTGCAAGACGGTGAGTTCATCTACAA
AGTTAAACTGCGTGGTACCAACTTCCCGTCCGACGGTCCGGTTATGCAGAAAAAAA
CCATGGGTTGGGAAGCTTCCACCGAACGTATGTACCCGGAAGACGGTGCTCTGAA
AGGTGAAATCAAAATGCGTCTGAAACTGAAAGACGGTGGTCACTACGACGCTGAA
GTTAAAACCACCTACATGGCTAAAAAACCGGTTCAGCTGCCGGGTGCTTACAAAAC
CGACATCAAACTGGACATCACCTCCCACAACGAAGACTACACCATCGTTGAACAGT
ACGAACGTGCTGAAGGTCGTCACTCCACCGGTGCTTAATAACGCTGATAGTGCTAG
TGTAGATCGCTACTAGAGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGA
CTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACA
CTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATAGGTCTCTCGCTTGAGACGCT
GCAGTGACAGGATATATTGGCGGGTAAACTAAGTCGCTGTATGTGTTTGTTTGAG
ATCTCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT
GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCT
CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCC

TGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGT
CCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTAT
CTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGT
TCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAA
GACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAG
GTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACT
AGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAG
AGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTG
TTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC
TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGT
CATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTT
TTAAATCAATCTAAAGTATATATGTGTAACATTGGTCTAGTGATTAGAAAAACTCA
TCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTT
TTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGG
ATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACC
TATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGA
CGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCA
ACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATT
CATTCGTGATTGCGCCTGAGCGAGTCGAAATACGCGATCGCTGTTAAAAGGACAA
TTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAAT
ATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCTGGGA
TCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGT
CGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACAA
CATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTC
CCATACAATCGGTAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTT
ATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTTGAGCAAGACG
TTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGAC
AGTTTTATTGTTCATGATGATATATTTTTATCTTGTGCAATGTAACATCAGAGATT
TTGAGACACAACGTGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGATC
AGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAA
ATCACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTT
CTGGCTGGATGATGGGGCGATTCAGGCGATCCCCATCCAACAGCCCGCCGTCGAG
CGGGCTTTTTTATCCCCGGAAGCCTGTGGATAGAGGGTAGTTATCCACGTGAAACC
GCTAATGCCCCGCAAAGCCTTGATTCACGGGGCTTTCCGGCCCGCTCCAAAAACTA
TCCACGTGAAATCGCTAATCAGGGTACGTGAAATCGCTAATCGGAGTACGTGAAA
TCGCTAATAAGGTCACGTGAAATCGCTAATCAAAAAGGCACGTGAGAACGCTAATA
GCCCTTTCAGATCAACAGCTTGCAAACACCCCTCGCTCCGGCAAGTAGTTACAGCA
AGTAGTATGTTCAATTAGCTTTTCAATTATGAATATATATATCAATTATTGGTCGC
CCTTGGCTTGTGGACAATGCGCTACGCGCACCGGCTCCGCCCGTGGACAACCGCAA
GCGGTTGCCCACCGTCGAGCGCCTTTGCCCACAACCCGGCGGCCGGCCGCAACAGA
TCGTTTTATAAATTTTTTTTTTTGAAAAAGAAAAAGCCCGAAAGGCGGCAACCTCT
CGGGCTTCTGGATTTCCGATCCCCGGAATTAGAGATCTTGGCAGGATATATTGTGG
TGTAACCCGAATTCCGT

pARKA11:

CGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAG
GAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATA
CCTGGAATGCTGTTTTCCCTGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGA
GTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTA
GTCTGACCATCTCATCTGTAACAACATTGGCAACGCTACCTTTGCCATGTTTCAGA
AACAACTCTGGCGCATCGGGCTTCCCATACAATCGGTAGATTGTCGCACCTGATTG
CCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAAT
TTAATCGCGGCCTTGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTT
GTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATATTTTTATC
TTGTGCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTGTTGAATAAATC
GAACTTTTGCTGAGTTGAAGGATCAGATCACGCATCTTCCCGACAACGCAGACCGT
TCCGTGGCAAAGCAAAAGTTCAAAATCACCAACTGGTCCACCTACAACAAAGCTCT
CATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGCGATTCAGGCGATCC
CCATCCAACAGCCCGCCGTCGAGCGGGCTTTTTTATCCCCGGAAGCCTGTGGATAG
AGGGTAGTTATCCACGTGAAACCGCTAATGCCCCGCAAAGCCTTGATTCACGGGGC
TTTCCGGCCCGCTCCAAAAACTATCCACGTGAAATCGCTAATCAGGGTACGTGAAA
TCGCTAATCGGAGTACGTGAAATCGCTAATAAGGTCACGTGAAATCGCTAATCAAA
AAGGCACGTGAGAACGCTAATAGCCCTTTCAGATCAACAGCTTGCAAACACCCCTC
GCTCCGGCAAGTAGTTACAGCAAGTAGTATGTTCAATTAGCTTTTCAATTATGAAT
ATATATATCAATTATTGGTCGCCCTTGGCTTGTGGACAATGCGCTACGCGCACCGG
CTCCGCCCGTGGACAACCGCAAGCGGTTGCCCACCGTCGAGCGCCTTTGCCCACAA
CCCGGCGGCCGGCCGCAACAGATCGTTTTATAAATTTTTTTTTTTGAAAAAGAAAA
AGCCCGAAAGGCGGCAACCTCTCGGGCTTCTGGATTTCCGATCCCCGGAATTAGAG
ATCTTGGCAGGATATATTGTGGTGTAACCCGAATTCCGTCTCAGGAGTACTAGACC
TGTAGGATCGTACAGGTTTACGCAAGAAAATGGTTTGTTATAGTCGAATAAATACT
ATTAAAGAGGAGAAATACTAGAATGCGTAAAGGAGAAGAACTTTTCACTGGAGTT
GTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAG
TGGAGAGGGTGAAGGTGATGCAACATACGGAAAACTTACCCTTAAATTTATTTGC
ACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGGTTATGG
TGTTCAATGCTTTGCGAGATACCCAGATCATATGAAACAGCATGACTTTTTCAAGA
GTGCCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGG
GAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGA
ATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATT
GGAATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATG
GAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAGCGTTCAACTA
GCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGA
CAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACGAAAAGAGAG
ATCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGAT
GAACTATACAAATAATAAGCTTCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAA
AGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTC
ACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATACGCTGTCATGAGACGCT
GCAGTGACAGGATATATTGGCGGGTAAACTAAGTCGCTGTATGTGTTTGTTTGAG
ATCTCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT
GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCT

CAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCT
GGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTC
CGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATC
TCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT
CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAG
ACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGG
TATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTA
GAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAGA
GTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT
TTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCT
TTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTC
ATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTT
TAAATCAATCTAAAGTATATATGTGTAACATTGGTCTAGTGATTAGAAAAACTCAT
CGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTT
TGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGA
TGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCT
ATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGAC
GACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAA
CAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTC
ATTCGTGATTGCGCCTGAGCGAGTCGAAATA

pARKA12:

GACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAA
CAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTC
ATTCGTGATTGCGCCTGAGCGAGTCGAAATACGCGATCGCTGTTAAAAGGACAAT
TACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAAT
ATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCTGGGA
TCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGT
CGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACAA
CATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTC
CCATACAATCGGTAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTT
ATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTTGAGCAAGACG
TTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGAC
AGTTTTATTGTTCATGATGATATATTTTTATCTTGTGCAATGTAACATCAGAGATT
TTGAGACACAACGTGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGATC
AGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAA
ATCACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTT
CTGGCTGGATGATGGGGCGATTCAGGCGATCCCCATCCAACAGCCCGCCGTCGAG
CGGGCTTTTTTATCCCCGGAAGCCTGTGGATAGAGGGTAGTTATCCACGTGAAACC
GCTAATGCCCCGCAAAGCCTTGATTCACGGGGCTTTCCGGCCCGCTCCAAAAACTA
TCCACGTGAAATCGCTAATCAGGGTACGTGAAATCGCTAATCGGAGTACGTGAAA
TCGCTAATAAGGTCACGTGAAATCGCTAATCAAAAAGGCACGTGAGAACGCTAATA
GCCCTTTCAGATCAACAGCTTGCAAACACCCCTCGCTCCGGCAAGTAGTTACAGCA
AGTAGTATGTTCAATTAGCTTTTCAATTATGAATATATATATCAATTATTGGTCGC
CCTTGGCTTGTGGACAATGCGCTACGCGCACCGGCTCCGCCCGTGGACAACCGCAA

GCGGTTGCCCACCGTCGAGCGCCTTTGCCCACAACCCGGCGGCCGGCCGCAACAGA
TCGTTTTATAAATTTTTTTTTTTGAAAAAGAAAAAGCCCGAAAGGCGGCAACCTCT
CGGGCTTCTGGATTTCCGATCCCCGGAATTAGAGATCTTGGCAGGATATATTGTGG
TGTAACCCGAATTCCGTCTCAGGAGTTTACGGCTAGCTCAGTCCTAGGTATAGTGC
TAGCTACTATTAAAGAGGAGAAATACTAGAATGACAATCATGATTAAGAAAAGCG
ATTTCTTAGCTATCCCGTCAGAAGAGTATAAAGGTATCCTGAGCCTGCGTTATCAA
GTCTTCAAACAGCGTCTGGAATGGGACTTGGTTGTGGAGAACAATCTTGAAAGCG
ACGAGTATGACAATTCAAATGCGGAGTACATTTACGCATGTGACGATACTGAGAAT
GTTTCAGGATGTTGGCGTCTTTTACCGACGACCGGGGATTACATGCTGAAGTCGG
TATTTCCCGAATTGTTAGGTCAACAATCCGCTCCGAAGGACCCGAACATCGTCGAA
CTTTCTCGTTTCGCGGTAGGGAAGAACAGCTCTAAAATTAACAATTCTGCCTCCGA
GATCACGATGAAGCTGTTCGAGGCTATCTATAAGCATGCGGTTTCGCAAGGTATTA
CGGAGTACGTAACTGTTACATCGACTGCCATCGAGCGTTTCCTGAAACGTATCAAG
GTGCCCTGTCACCGTATCGGTGATAAGGAAATCCATGTCTTAGGAGATACAAAATC
TGTAGTACTTTCCATGCCTATCAACGAACAGTTCAAAAAGGCAGTATTAAACTAAT
AAGCTTCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGT
TTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTC
GGGTGGGCCTTTCTGCGTTTATACGCTGTCATGAGACGCTGCAGTGACAGGATAT
ATTGGCGGGTAAACTAAGTCGCTGTATGTGTTTGTTTGAGATCTCATGTGAGCAA
AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCA
TAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC
GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTG
CGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTC
GGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAG
GTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG
CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGC
CACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGC
TACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTG
GTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAGAGTTGGTAGCTCTTGA
TCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGAT
TACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG
ACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAA
AGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAG
TATATATGTGTAACATTGGTCTAGTGATTAGAAAAACTCATCGAGCATCAAATGAA
ACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTC
TGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGT
ATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCG
TCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGAC

pARKA13:

CCTTGGCTTGTGGACAATGCGCTACGCGCACCGGCTCCGCCCGTGGACAACCGCAA
GCGGTTGCCCACCGTCGAGCGCCTTTGCCCACAACCCGGCGGCCGGCCGCAACAGA
TCGTTTTATAAATTTTTTTTTTTGAAAAAGAAAAAGCCCGAAAGGCGGCAACCTCT
CGGGCTTCTGGATTTCCGATCCCCGGAATTAGAGATCTTGGCAGGATATATTGTGG
TGTAACCCGAATTCCGTCTCAGGAGTACTAGACCTGTAGGATCGTACAGGTTTACG

CAAGAAAATGGTTTGTTATAGTCGAATAAATACTTCACACAGGAAAGTACTAGAAT
GGTACGCTGGACTTTGTGGGATACCCTCGCTTTCCTGCTCCTGTTGAGTTTATTGC
TGCCGTCATTGCTTATTATGTTCATCCCGTCAACATTCAAACGGCCTGTCTCATCA
TGGAAGGCGCTGAATTTACGGAAAACATTATTAATGGCGTCGAGCGTCCGGTTAA
AGCCGCTGAATTGTTCGCGCTTACCTTGCGTGTACGCGCAGGAAACACTGACGTTC
TTGCTGACGCAGAAGAAAACGTGCGTCAAAAATTACGTGCAGAAGGAGTGAGCTT
CCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATC
TGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGG
GCCTTTCTGCGTTTATACGCTGTCATGAGACGCTGCAGTGACAGGATATATTGGCG
GGTAAACTAAGTCGCTGTATGTGTTTGTTTGAGATCTCATGTGAGCAAAAGGCCA
GCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTC
CGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCC
GACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTC
CTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGC
GTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCG
CTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTAT
CCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCA
GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT
TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGC
GCTCTGCTGAAGCCAGTTACCTTCGGAAGAAGAGTTGGTAGCTCTTGATCCGGCAA
ACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCA
GAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAG
TGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTT
CACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATG
TGTAACATTGGTCTAGTGATTAGAAAAACTCATCGAGCATCAAATGAAACTGCAAT
TTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGA
AGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCT
GCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAAT
AAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGC
AAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTC
ATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGA
GTCGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAA
CCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATCAGGATATT
CTTCTAATACCTGGAATGCTGTTTTCCCTGGGATCGCAGTGGTGAGTAACCATGCA
TCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCA
GCCAGTTTAGTCTGACCATCTCATCTGTAACAACATTGGCAACGCTACCTTTGCCA
TGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAATCGGTAGATTGTCGC
ACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATCAGCATCCA
TGTTGGAATTTAATCGCGGCCTTGAGCAAGACGTTTCCCGTTGAATATGGCTCATA
ACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATAT
ATTTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTGTT
GAATAAATCGAACTTTTGCTGAGTTGAAGGATCAGATCACGCATCTTCCCGACAAC
GCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAATCACCAACTGGTCCACCTACAA
CAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGCGATTC
AGGCGATCCCCATCCAACAGCCCGCCGTCGAGCGGGCTTTTTTATCCCCGGAAGCC

TGTGGATAGAGGGTAGTTATCCACGTGAAACCGCTAATGCCCCGCAAAGCCTTGAT
TCACGGGGCTTTCCGGCCCGCTCCAAAAACTATCCACGTGAAATCGCTAATCAGGG
TACGTGAAATCGCTAATCGGAGTACGTGAAATCGCTAATAAGGTCACGTGAAATC
GCTAATCAAAAAGGCACGTGAGAACGCTAATAGCCCTTTCAGATCAACAGCTTGCA
AACACCCCTCGCTCCGGCAAGTAGTTACAGCAAGTAGTATGTTCAATTAGCTTTTC
AATTATGAATATATATATCAATTATTGGTCGC

pARKA21:

CTGAGCGAGTCGAAATACGCGATCGCTGTTAAAAGGACAATTACAAACAGGAATC
GAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACAATATTTTCACCTGAATC
AGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCTGGGATCGCAGTGGTGAGTA
ACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGGTCGGAAGAGGCATAAA
TTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACAACATTGGCAACGCTAC
CTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAATCGGTAG
ATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAATC
AGCATCCATGTTGGAATTTAATCGCGGCCTTGAGCAAGACGTTTCCCGTTGAATAT
GGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCAT
GATGATATATTTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACAACGTG
GCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGATCAGATCACGCATCTTC
CCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAATCACCAACTGGTCC
ACCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGG
GGCGATTCAGGCGATCCCCATCCAACAGCCCGCCGTCGAGCGGGCTTTTTTATCCC
CGGAAGCCTGTGGATAGAGGGTAGTTATCCACGTGAAACCGCTAATGCCCCGCAA
AGCCTTGATTCACGGGGCTTTCCGGCCCGCTCCAAAAACTATCCACGTGAAATCGC
TAATCAGGGTACGTGAAATCGCTAATCGGAGTACGTGAAATCGCTAATAAGGTCA
CGTGAAATCGCTAATCAAAAAGGCACGTGAGAACGCTAATAGCCCTTTCAGATCAA
CAGCTTGCAAACACCCCTCGCTCCGGCAAGTAGTTACAGCAAGTAGTATGTTCAAT
TAGCTTTTCAATTATGAATATATATATCAATTATTGGTCGCCCTTGGCTTGTGGAC
AATGCGCTACGCGCACCGGCTCCGCCCGTGGACAACCGCAAGCGGTTGCCCACCGT
CGAGCGCCTTTGCCCACAACCCGGCGGCCGGCCGCAACAGATCGTTTTATAAATTT
TTTTTTTTGAAAAAGAAAAAGCCCGAAAGGCGGCAACCTCTCGGGCTTCTGGATTT
CCGATCCCCGGAATTAGAGATCTTGGCAGGATATATTGTGGTGTAACAAGCTTCGT
CTCAGTCAGGAGTTTACGGCTAGCTCAGTCCTAGGTATAGTGCTAGCTACTATTAA
AGAGGAGAAATACTAGAATGAAAAACATAAATGCCGACGACACATACAGAATAAT
TAATAAAATTAAAGCTTGTAGAAGCAATAATGATATTAATCAATGCTTATCTGATA
TGACTAAAATGGTACATTGTGAATATTATTTACTCGCGATCATTTATCCTCATTCT
ATGGTTAAATCTGATATTTCAATCCTAGATAATTACCCTAAAAAATGGAGGCAATA
TTATGATGACGCTAATTTAATAAAATATGATCCTATAGTAGATTATTCTAACTCCA
ATCATTCACCAATTAATTGGAATATATTTGAAAACAATGCTGTAAATAAAAAATCT
CCAAATGTAATTAAAGAAGCGAAAACATCAGGTCTTATCACTGGGTTTAGTTTCCC
TATTCATACGGCTAACAATGGCTTCGGAATGCTTAGTTTTGCACATTCAGAAAAAG
ACAACTATATAGATAGTTTATTTTTACATGCGTGTATGAACATACCATTAATTGTT
CCTTCTCTAGTTGATAATTATCGAAAAATAAATATAGCAAATAATAAATCAAACAA
CGATTTAACCAAAAGAGAAAAAGAATGTTTAGCGTGGGCATGCGAAGGAAAAAGC
TCTTGGGATATTTCAAAAATATTAGGTTGCAGTGAGCGTACTGTCACTTTCCATTT

AACCAATGCGCAAATGAAACTCAATACAACAAACCGCTGCCAAAGTATTTCTAAAG
CAATTTTAACAGGAGCAATTGATTGCCCATACTTTAAAAATTAATAAGCTTCCAGG
CATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTG
TTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTT
TCTGCGTTTATACGCTTGAGACGAAGCTTGACAGGATATATTGGCGGGTAAACTAA
GTCGCTGTATGTGTTTGTTTGAGATCTCATGTGAGCAAAAGGCCAGCAAAAGGCC
AGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGA
CGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTA
TAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGAC
CCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTT
CTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTG
GGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTA
TCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTG
GTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTG
GTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGA
AGCCAGTTACCTTCGGAAGAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACC
GCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAG
GATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAA
AACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGAT
CCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGTGTAACATT
GGTCTAGTGATTAGAAAAACTCATCGAGCATCAAATGAAACTGCAATTTATTCATA
TCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAA
CTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGA
CTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCA
AGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTAT
GCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCA
CTCGCATCAACCAAACCGTTATTCATTCGTGATTGCGC

pARKA22:

TTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCA
GCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCT
GTTTTCCCTGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAA
AATGCTTGATGGTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATC
TCATCTGTAACAACATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGG
CGCATCGGGCTTCCCATACAATCGGTAGATTGTCGCACCTGATTGCCCGACATTAT
CGCGAGCCCATTTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGC
CTTGAGCAAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTT
TATGTAAGCAGACAGTTTTATTGTTCATGATGATATATTTTTATCTTGTGCAATGT
AACATCAGAGATTTTGAGACACAACGTGGCTTTGTTGAATAAATCGAACTTTTGCT
GAGTTGAAGGATCAGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAA
GCAAAAGTTCAAAATCACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTG
GCTCCCTCACTTTCTGGCTGGATGATGGGGCGATTCAGGCGATCCCCATCCAACAG
CCCGCCGTCGAGCGGGCTTTTTTATCCCCGGAAGCCTGTGGATAGAGGGTAGTTAT
CCACGTGAAACCGCTAATGCCCCGCAAAGCCTTGATTCACGGGGCTTTCCGGCCCG
CTCCAAAAACTATCCACGTGAAATCGCTAATCAGGGTACGTGAAATCGCTAATCGG

AGTACGTGAAATCGCTAATAAGGTCACGTGAAATCGCTAATCAAAAAGGCACGTG
AGAACGCTAATAGCCCTTTCAGATCAACAGCTTGCAAACACCCCTCGCTCCGGCAA
GTAGTTACAGCAAGTAGTATGTTCAATTAGCTTTTCAATTATGAATATATATATCA
ATTATTGGTCGCCCTTGGCTTGTGGACAATGCGCTACGCGCACCGGCTCCGCCCGT
GGACAACCGCAAGCGGTTGCCCACCGTCGAGCGCCTTTGCCCACAACCCGGCGGCC
GGCCGCAACAGATCGTTTTATAAATTTTTTTTTTTGAAAAAGAAAAAGCCCGAAAG
GCGGCAACCTCTCGGGCTTCTGGATTTCCGATCCCCGGAATTAGAGATCTTGGCAG
GATATATTGTGGTGTAACCCGAATTCCGTCTCAGTCAGGAGTACTAGACCTGTAGG
ATCGTACAGGTTTACGCAAGAAAATGGTTTGTTATAGTCGAATAAATACTTCACAC
AGGAAAGTACTAGAATGGTACGCTGGACTTTGTGGGATACCCTCGCTTTCCTGCTC
CTGTTGAGTTTATTGCTGCCGTCATTGCTTATTATGTTCATCCCGTCAACATTCAA
ACGGCCTGTCTCATCATGGAAGGCGCTGAATTTACGGAAAACATTATTAATGGCGT
CGAGCGTCCGGTTAAAGCCGCTGAATTGTTCGCGCTTACCTTGCGTGTACGCGCAG
GAAACACTGACGTTCTTGCTGACGCAGAAGAAAACGTGCGTCAAAAATTACGTGCA
GAAGGAGTGAGCTTCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGG
GCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGG
CTCACCTTCGGGTGGGCCTTTCTGCGTTTATACGCTTGAGACGCTGCAGTGACAGG
ATATATTGGCGGGTAAACTAAGTCGCTGTATGTGTTTGTTTGAGATCTCATGTGAG
CAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTT
CCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGT
GGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTC
GTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCC
TTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGT
AGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGC
TGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC
GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT
GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATT
TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAGAGTTGGTAGCTCTT
GATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAG
ATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTC
TGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAA
AAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAA
AGTATATATGTGTAACATTGGTCTAGTGATTAGAAAAACTCATCGAGCATCAAATG
AAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTT
TCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTG
GTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCT
CGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGG
TGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCAT
TACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTTATTCATTCGTGATTGC
GCCTGAGCGAGTCGAAATACGCGATCGCTG

pARSO11:

TAACGGATCCGGTCTCAGGAGTACTAGACCTGTAGGATCGTACAGGTTTACGCAA
GAAAATGGTTTGTTATAGTCGAATAAATACTATTAAAGAGGAGAAATACTAGAAT
GCGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATG

GTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAAC
ATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCAT
GGCCAACACTTGTCACTACTTTCGGTTATGGTGTTCAATGCTTTGCGAGATACCCA
GATCATATGAAACAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGTACA
GGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTC
AAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAA
AGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATG
TATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGA
CACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCC
AATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTG
CCCTTTCGAAAGATCCCAACGAAAAGAGAGATCACATGGTCCTTCTTGAGTTTGTA
ACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAATAATAAGCTTCCAG
GCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTT
GTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCT
TTCTGCGTTTATACGCTGTCAGGAGTTTACGGCTAGCTCAGTCCTAGGTATAGTGC
TAGCTACTATTAAAGAGGAGAAATACTAGAATGAAAAACATAAATGCCGACGACA
CATACAGAATAATTAATAAAATTAAAGCTTGTAGAAGCAATAATGATATTAATCAA
TGCTTATCTGATATGACTAAAATGGTACATTGTGAATATTATTTACTCGCGATCAT
TTATCCTCATTCTATGGTTAAATCTGATATTTCAATCCTAGATAATTACCCTAAAA
AATGGAGGCAATATTATGATGACGCTAATTTAATAAAATATGATCCTATAGTAGAT
TATTCTAACTCCAATCATTCACCAATTAATTGGAATATATTTGAAAACAATGCTGT
AAATAAAAAATCTCCAAATGTAATTAAAGAAGCGAAAACATCAGGTCTTATCACTG
GGTTTAGTTTCCCTATTCATACGGCTAACAATGGCTTCGGAATGCTTAGTTTTGCA
CATTCAGAAAAAGACAACTATATAGATAGTTTATTTTTACATGCGTGTATGAACAT
ACCATTAATTGTTCCTTCTCTAGTTGATAATTATCGAAAAATAAATATAGCAAATA
ATAAATCAAACAACGATTTAACCAAAAGAGAAAAAGAATGTTTAGCGTGGGCATG
CGAAGGAAAAAGCTCTTGGGATATTTCAAAAATATTAGGTTGCAGTGAGCGTACT
GTCACTTTCCATTTAACCAATGCGCAAATGAAACTCAATACAACAAACCGCTGCCA
AAGTATTTCTAAAGCAATTTTAACAGGAGCAATTGATTGCCCATACTTTAAAAATT
AATAAGCTTCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTT
CGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACC
TTCGGGTGGGCCTTTCTGCGTTTATACGCTGTCATGAGACCGGATCCTGACAGGAT
ATATTGGCGGGTAAACTAAGTCGCTGTATGTGTTTGTTTGAGATCTCATGTGAGCA
AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC
ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGG
CGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGT
GCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTT
CGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAG
GTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG
CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGC
CACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGC
TACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTG
GTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAGAGTTGGTAGCTCTTGA
TCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGAT
TACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG

ACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAA
AGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAG
TATATATGTGTAACATTGGTCTAGTGATTATTTGCCGACTACCTTGGTGATCTCGC
CTTTCACGTAGTGGACAAATTCTTCCAACTGATCTGCGCGCGAGGCCAAGCGATCT
TCTTCTTGTCCAAGATAAGCCTGTCTAGCTTCAAGTATGACGGGCTGATACTGGGC
CGGCAGGCGCTCCATTGCCCAGTCGGCAGCGACATCCTTCGGCGCGATTTTGCCGG
TTACTGCGCTGTACCAAATGCGGGACAACGTAAGCACTACATTTCGCTCATCACCA
GCCCAGTCGGGCGGCGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAG
ATCCTGTTCAGGAACCGGATCAAAGAGTTCCTCCGCCGCTGGACCTACCAAGGCAA
CGCTATGTTCTCTTGCTTTTGTCAGCAAGATAGCCAGATCAATGTCGATCGTGGCT
GGCTCGAAGATACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGCAGTTC
GCGCTTAGCTGGATAACGCCACGGAATGATGTCGTCGTGCACAACAATGGTGACT
TCTACAGCGCGGAGAATCTCGCTCTCTCCAGGGGAAGCCGAAGTTTCCAAAAGGTC
GTTGATCAAAGCTCGCCGCGTTGTTTCATCAAGCCTTACGGTCACCGTAACCAGCA
AATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACTGCGGAGCCGTACAAATGT
ACGGCCAGCAACGTCGGTTCGAGATGGCGCTCGATGACGCCAACTACCTCTGATAG
TTGAGTCGATACTTCGGCGATCACCGCTTCCCTCATGATGTTTAACTTTGTTTTAG
GGCGACTGCCCTGCTGCGTAACATCGTTGCTGCTCCATAACATCAAACATCGACCC
ACGGCGTAACGCGCTTGCTGCTTGGATGCCCGAGGCATAGACTGTACCCCAAAAAA
ACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGTTACCACCGCTGCGTTCGG
TCAAGGTTCTGGACCAGTTGCGTGAGCGCATACGCTACTTGCATTACAGCTTACGA
ACCGAACAGGCTTATGTCCACTGGGTTCGTGCCTTCATCCGTTTCCACGGTGTGCG
TCACCCGGCAACCTTGGGCAGCAGCGAAGTCGAGGCATTTCTGTCCTGGCTGGAAC
ACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATAT
TTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTGTTGA
ATAAATCGAACTTTTGCTGAGTTGAAGGATCAGATCACGCATCTTCCCGACAACGC
AGACCGTTCCGTGGCAAAGCAAAAGTTCAAAATCACCAACTGGTCCACCTACAACA
AAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGCGATTCAG
GCGATCCCCATCCAACAGCCCGCCGTCGAGCGGGCTTTTTTATCCCCGGAAGCCTG
TGGATAGAGGGTAGTTATCCACGTGAAACCGCTAATGCCCCGCAAAGCCTTGATTC
ACGGGGCTTTCCGGCCCGCTCCAAAAACTATCCACGTGAAATCGCTAATCAGGGTA
CGTGAAATCGCTAATCGGAGTACGTGAAATCGCTAATAAGGTCACGTGAAATCGCT
AATCAAAAAGGCACGTGAGAACGCTAATAGCCCTTTCAGATCAACAGCTTGCAAAC
ACCCCTCGCTCCGGCAAGTAGTTACAGCAAGTAGTATGTTCAATTAGCTTTTCAAT
TATGAATATATATATCAATTATTGGTCGCCCTTGGCTTGTGGACAATGCGCTACGC
GCACCGGCTCCGCCCGTGGACAACCGCAAGCGGTTGCCCACCGTCGAGCGCCTTTG
CCCACAACCCGGCGGCCGGCCGCAACAGATCGTTTTATAAATTTTTTTTTTTGAAA
AAGAAAAAGCCCGAAAGGCGGCAACCTCTCGGGCTTCTGGATTTCCGATCCCCGGA
ATTAGAGATCTTGGCAGGATATATTGTGGTG

pARSO12:

TAACGGATCCGGTCTCAGGAGTACTAGACCTGTAGGATCGTACAGGTTTACGCAA
GAAAATGGTTTGTTATAGTCGAATAAATACTTCACACAGGAAAGTACTAGAATGGT
ACGCTGGACTTTGTGGGATACCCTCGCTTTCCTGCTCCTGTTGAGTTTATTGCTGC
CGTCATTGCTTATTATGTTCATCCCGTCAACATTCAAACGGCCTGTCTCATCATGG

AAGGCGCTGAATTTACGGAAAACATTATTAATGGCGTCGAGCGTCCGGTTAAAGC
CGCTGAATTGTTCGCGCTTACCTTGCGTGTACGCGCAGGAAACACTGACGTTCTTG
CTGACGCAGAAGAAAACGTGCGTCAAAAATTACGTGCAGAAGGAGTGAGCTTCCA
GGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGT
TGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCC
TTTCTGCGTTTATACGCTGTCAGGAGTTTACGGCTAGCTCAGTCCTAGGTATAGTG
CTAGCTACTATTAAAGAGGAGAAATACTAGAATGAAAAACATAAATGCCGACGACA
CATACAGAATAATTAATAAAATTAAAGCTTGTAGAAGCAATAATGATATTAATCAA
TGCTTATCTGATATGACTAAAATGGTACATTGTGAATATTATTTACTCGCGATCAT
TTATCCTCATTCTATGGTTAAATCTGATATTTCAATCCTAGATAATTACCCTAAAA
AATGGAGGCAATATTATGATGACGCTAATTTAATAAAATATGATCCTATAGTAGAT
TATTCTAACTCCAATCATTCACCAATTAATTGGAATATATTTGAAAACAATGCTGT
AAATAAAAAATCTCCAAATGTAATTAAAGAAGCGAAAACATCAGGTCTTATCACTG
GGTTTAGTTTCCCTATTCATACGGCTAACAATGGCTTCGGAATGCTTAGTTTTGCA
CATTCAGAAAAAGACAACTATATAGATAGTTTATTTTTACATGCGTGTATGAACAT
ACCATTAATTGTTCCTTCTCTAGTTGATAATTATCGAAAAATAAATATAGCAAATA
ATAAATCAAACAACGATTTAACCAAAAGAGAAAAAGAATGTTTAGCGTGGGCATG
CGAAGGAAAAAGCTCTTGGGATATTTCAAAAATATTAGGTTGCAGTGAGCGTACT
GTCACTTTCCATTTAACCAATGCGCAAATGAAACTCAATACAACAAACCGCTGCCA
AAGTATTTCTAAAGCAATTTTAACAGGAGCAATTGATTGCCCATACTTTAAAAATT
AATAAGCTTCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTT
CGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACC
TTCGGGTGGGCCTTTCTGCGTTTATACGCTGTCATGAGACCGGATCCTGACAGGAT
ATATTGGCGGGTAAACTAAGTCGCTGTATGTGTTTGTTTGAGATCTCATGTGAGCA
AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC
ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGG
CGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGT
GCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTT
CGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAG
GTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTG
CGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGC
CACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGC
TACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTG
GTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAGAAGAGTTGGTAGCTCTTGA
TCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGAT
TACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTG
ACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAA
AGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAG
TATATATGTGTAACATTGGTCTAGTGATTATTTGCCGACTACCTTGGTGATCTCGC
CTTTCACGTAGTGGACAAATTCTTCCAACTGATCTGCGCGCGAGGCCAAGCGATCT
TCTTCTTGTCCAAGATAAGCCTGTCTAGCTTCAAGTATGACGGGCTGATACTGGGC
CGGCAGGCGCTCCATTGCCCAGTCGGCAGCGACATCCTTCGGCGCGATTTTGCCGG
TTACTGCGCTGTACCAAATGCGGGACAACGTAAGCACTACATTTCGCTCATCACCA
GCCCAGTCGGGCGGCGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAG
ATCCTGTTCAGGAACCGGATCAAAGAGTTCCTCCGCCGCTGGACCTACCAAGGCAA

CGCTATGTTCTCTTGCTTTTGTCAGCAAGATAGCCAGATCAATGTCGATCGTGGCT
GGCTCGAAGATACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGCAGTTC
GCGCTTAGCTGGATAACGCCACGGAATGATGTCGTCGTGCACAACAATGGTGACT
TCTACAGCGCGGAGAATCTCGCTCTCTCCAGGGGAAGCCGAAGTTTCCAAAAGGTC
GTTGATCAAAGCTCGCCGCGTTGTTTCATCAAGCCTTACGGTCACCGTAACCAGCA
AATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACTGCGGAGCCGTACAAATGT
ACGGCCAGCAACGTCGGTTCGAGATGGCGCTCGATGACGCCAACTACCTCTGATAG
TTGAGTCGATACTTCGGCGATCACCGCTTCCCTCATGATGTTTAACTTTGTTTTAG
GGCGACTGCCCTGCTGCGTAACATCGTTGCTGCTCCATAACATCAAACATCGACCC
ACGGCGTAACGCGCTTGCTGCTTGGATGCCCGAGGCATAGACTGTACCCCAAAAAA
ACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGTTACCACCGCTGCGTTCGG
TCAAGGTTCTGGACCAGTTGCGTGAGCGCATACGCTACTTGCATTACAGCTTACGA
ACCGAACAGGCTTATGTCCACTGGGTTCGTGCCTTCATCCGTTTCCACGGTGTGCG
TCACCCGGCAACCTTGGGCAGCAGCGAAGTCGAGGCATTTCTGTCCTGGCTGGAAC
ACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATAT
TTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTGTTGA
ATAAATCGAACTTTTGCTGAGTTGAAGGATCAGATCACGCATCTTCCCGACAACGC
AGACCGTTCCGTGGCAAAGCAAAAGTTCAAAATCACCAACTGGTCCACCTACAACA
AAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGCGATTCAG
GCGATCCCCATCCAACAGCCCGCCGTCGAGCGGGCTTTTTTATCCCCGGAAGCCTG
TGGATAGAGGGTAGTTATCCACGTGAAACCGCTAATGCCCCGCAAAGCCTTGATTC
ACGGGGCTTTCCGGCCCGCTCCAAAAACTATCCACGTGAAATCGCTAATCAGGGTA
CGTGAAATCGCTAATCGGAGTACGTGAAATCGCTAATAAGGTCACGTGAAATCGCT
AATCAAAAAGGCACGTGAGAACGCTAATAGCCCTTTCAGATCAACAGCTTGCAAAC
ACCCCTCGCTCCGGCAAGTAGTTACAGCAAGTAGTATGTTCAATTAGCTTTTCAAT
TATGAATATATATATCAATTATTGGTCGCCCTTGGCTTGTGGACAATGCGCTACGC
GCACCGGCTCCGCCCGTGGACAACCGCAAGCGGTTGCCCACCGTCGAGCGCCTTTG
CCCACAACCCGGCGGCCGGCCGCAACAGATCGTTTTATAAATTTTTTTTTTTGAAA
AAGAAAAAGCCCGAAAGGCGGCAACCTCTCGGGCTTCTGGATTTCCGATCCCCGGA
ATTAGAGATCTTGGCAGGATATATTGTGGTG

pARSO21:

TAGTTATCCACGTGAAACCGCTAATGCCCCGCAAAGCCTTGATTCACGGGGCTTTC
CGGCCCGCTCCAAAAACTATCCACGTGAAATCGCTAATCAGGGTACGTGAAATCGC
TAATCGGAGTACGTGAAATCGCTAATAAGGTCACGTGAAATCGCTAATCAAAAAG
GCACGTGAGAACGCTAATAGCCCTTTCAGATCAACAGCTTGCAAACACCCCTCGCT
CCGGCAAGTAGTTACAGCAAGTAGTATGTTCAATTAGCTTTTCAATTATGAATATA
TATATCAATTATTGGTCGCCCTTGGCTTGTGGACAATGCGCTACGCGCACCGGCTC
CGCCCGTGGACAACCGCAAGCGGTTGCCCACCGTCGAGCGCCTTTGCCCACAACCC
GGCGGCCGGCCGCAACAGATCGTTTTATAAATTTTTTTTTTTGAAAAAGAAAAAGC
CCGAAAGGCGGCAACCTCTCGGGCTTCTGGATTTCCGATCCCCGGAATTAGAGATC
TTGGCAGGATATATTGTGGTGTAACGATATCGGTCTCAGTCAGGAGTTTACGGCT
AGCTCAGTCCTAGGTATAGTGCTAGCTACTATTAAAGAGGAGAAATACTAGAATG
ACAATCATGATTAAGAAAAGCGATTTCTTAGCTATCCCGTCAGAAGAGTATAAAGG
TATCCTGAGCCTGCGTTATCAAGTCTTCAAACAGCGTCTGGAATGGGACTTGGTTG

TGGAGAACAATCTTGAAAGCGACGAGTATGACAATTCAAATGCGGAGTACATTTA
CGCATGTGACGATACTGAGAATGTTTCAGGATGTTGGCGTCTTTTACCGACGACCG
GGGATTACATGCTGAAGTCGGTATTTCCCGAATTGTTAGGTCAACAATCCGCTCCG
AAGGACCCGAACATCGTCGAACTTTCTCGTTTCGCGGTAGGGAAGAACAGCTCTAA
AATTAACAATTCTGCCTCCGAGATCACGATGAAGCTGTTCGAGGCTATCTATAAGC
ATGCGGTTTCGCAAGGTATTACGGAGTACGTAACTGTTACATCGACTGCCATCGAG
CGTTTCCTGAAACGTATCAAGGTGCCCTGTCACCGTATCGGTGATAAGGAAATCCA
TGTCTTAGGAGATACAAAATCTGTAGTACTTTCCATGCCTATCAACGAACAGTTCA
AAAAGGCAGTATTAAACTAATAAGCTTCCAGGCATCAAATAAAACGAAAGGCTCAG
TCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTA
GAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATACGCTGTCAGGAG
TACTAGACCTGTAGGATCGTACAGGTTTACGCAAGAAAATGGTTTGTTATAGTCGA
ATAAATACTTCACACAGGAAAGTACTAGAATGGTACGCTGGACTTTGTGGGATACC
CTCGCTTTCCTGCTCCTGTTGAGTTTATTGCTGCCGTCATTGCTTATTATGTTCAT
CCCGTCAACATTCAAACGGCCTGTCTCATCATGGAAGGCGCTGAATTTACGGAAAA
CATTATTAATGGCGTCGAGCGTCCGGTTAAAGCCGCTGAATTGTTCGCGCTTACCT
TGCGTGTACGCGCAGGAAACACTGACGTTCTTGCTGACGCAGAAGAAAACGTGCG
TCAAAAATTACGTGCAGAAGGAGTGAGCTTCCAGGCATCAAATAAAACGAAAGGC
TCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCT
ACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATACGCTTGAG
ACCGATATCTGACAGGATATATTGGCGGGTAAACTAAGTCGCTGTATGTGTTTGTT
TGAGATCTCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCG
CGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGA
CGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTC
CCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATAC
CTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAG
GTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCC
CCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCG
GTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAG
CGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTA
CACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA
GAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTT
TTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTT
TGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATT
TTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATG
AAGTTTTAAATCAATCTAAAGTATATATGTGTAACATTGGTCTAGTGATTATTTGC
CGACTACCTTGGTGATCTCGCCTTTCACGTAGTGGACAAATTCTTCCAACTGATCT
GCGCGCGAGGCCAAGCGATCTTCTTCTTGTCCAAGATAAGCCTGTCTAGCTTCAAG
TATGACGGGCTGATACTGGGCCGGCAGGCGCTCCATTGCCCAGTCGGCAGCGACA
TCCTTCGGCGCGATTTTGCCGGTTACTGCGCTGTACCAAATGCGGGACAACGTAAG
CACTACATTTCGCTCATCACCAGCCCAGTCGGGCGGCGAGTTCCATAGCGTTAAGG
TTTCATTTAGCGCCTCAAATAGATCCTGTTCAGGAACCGGATCAAAGAGTTCCTCC
GCCGCTGGACCTACCAAGGCAACGCTATGTTCTCTTGCTTTTGTCAGCAAGATAGC
CAGATCAATGTCGATCGTGGCTGGCTCGAAGATACCTGCAAGAATGTCATTGCGCT
GCCATTCTCCAAATTGCAGTTCGCGCTTAGCTGGATAACGCCACGGAATGATGTCG

TCGTGCACAACAATGGTGACTTCTACAGCGCGGAGAATCTCGCTCTCTCCAGGGGA
AGCCGAAGTTTCCAAAAGGTCGTTGATCAAAGCTCGCCGCGTTGTTTCATCAAGCC
TTACGGTCACCGTAACCAGCAAATCAATATCACTGTGTGGCTTCAGGCCGCCATCC
ACTGCGGAGCCGTACAAATGTACGGCCAGCAACGTCGGTTCGAGATGGCGCTCGA
TGACGCCAACTACCTCTGATAGTTGAGTCGATACTTCGGCGATCACCGCTTCCCTC
ATGATGTTTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCGTTGCTGCT
CCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGCTTGGATGCCCGAG
GCATAGACTGTACCCCAAAAAAACAGTCATAACAAGCCATGAAAACCGCCACTGCG
CCGTTACCACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGAGCGCATACG
CTACTTGCATTACAGCTTACGAACCGAACAGGCTTATGTCCACTGGGTTCGTGCCT
TCATCCGTTTCCACGGTGTGCGTCACCCGGCAACCTTGGGCAGCAGCGAAGTCGAG
GCATTTCTGTCCTGGCTGGAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGT
TTTATTGTTCATGATGATATATTTTTATCTTGTGCAATGTAACATCAGAGATTTTG
AGACACAACGTGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGATCAGA
TCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAATC
ACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCTG
GCTGGATGATGGGGCGATTCAGGCGATCCCCATCCAACAGCCCGCCGTCGAGCGG
GCTTTTTTATCCCCGGAAGCCTGTGGATAGAGGG

pARKA1_o1:

CCTTGGCTTGTGGACAATGCGCTACGCGCACCGGCTCCGCCCGTGGACAACCGCAA
GCGGTTGCCCACCGTCGAGCGCCTTTGCCCACAACCCGGCGGCCGGCCGCAACAGA
TCGTTTTATAAATTTTTTTTTTTGAAAAAGAAAAAGCCCGAAAGGCGGCAACCTCT
CGGGCTTCTGGATTTCCGATCCCCGGAATTAGAGATCTTGGCAGGATATATTGTGG
TGTAACCCGAATTCCGTCTCAGGAGTACTAGACCTGTAGGATCGTACAGGTTTACG
CAAGAAAATGGTTTGTTATAGTCGAATAAATACTATTAAAGAGGAGAAATACTAG
AATGCGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAG
ATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGC
AACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTC
CATGGCCAACACTTGTCACTACTTTCGGTTATGGTGTTCAATGCTTTGCGAGATAC
CCAGATCATATGAAACAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTATGT
ACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAA
GTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTT
TAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACA
ATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATT
AGACACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATAC
TCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAAT
CTGCCCTTTCGAAAGATCCCAACGAAAAGAGAGATCACATGGTCCTTCTTGAGTTT
GTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAATAATAAGCTTC
CAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCT
GTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGG
CCTTTCTGCGTTTATACGCTGTCAGGAGTTTACGGCTAGCTCAGTCCTAGGTATAG
TGCTAGCTACTATTAAAGAGGAGAAATACTAGAATGAAAAACATAAATGCCGACG
ACACATACAGAATAATTAATAAAATTAAAGCTTGTAGAAGCAATAATGATATTAAT
CAATGCTTATCTGATATGACTAAAATGGTACATTGTGAATATTATTTACTCGCGAT

CATTTATCCTCATTCTATGGTTAAATCTGATATTTCAATCCTAGATAATTACCCTA
AAAAATGGAGGCAATATTATGATGACGCTAATTTAATAAAATATGATCCTATAGTA
GATTATTCTAACTCCAATCATTCACCAATTAATTGGAATATATTTGAAAACAATGC
TGTAAATAAAAAATCTCCAAATGTAATTAAAGAAGCGAAAACATCAGGTCTTATCA
CTGGGTTTAGTTTCCCTATTCATACGGCTAACAATGGCTTCGGAATGCTTAGTTTT
GCACATTCAGAAAAAGACAACTATATAGATAGTTTATTTTTACATGCGTGTATGAA
CATACCATTAATTGTTCCTTCTCTAGTTGATAATTATCGAAAAATAAATATAGCAA
ATAATAAATCAAACAACGATTTAACCAAAAGAGAAAAAGAATGTTTAGCGTGGGC
ATGCGAAGGAAAAAGCTCTTGGGATATTTCAAAAATATTAGGTTGCAGTGAGCGT
ACTGTCACTTTCCATTTAACCAATGCGCAAATGAAACTCAATACAACAAACCGCTG
CCAAAGTATTTCTAAAGCAATTTTAACAGGAGCAATTGATTGCCCATACTTTAAAA
ATTAATAAGCTTCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCC
TTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTC
ACCTTCGGGTGGGCCTTTCTGCGTTTATACGCTGTCAGGAGTTTACGGCTAGCTCA
GTCCTAGGTATAGTGCTAGCTACTATTAAAGAGGAGAAATACTAGAATGACAATCA
TGATTAAGAAAAGCGATTTCTTAGCTATCCCGTCAGAAGAGTATAAAGGTATCCTG
AGCCTGCGTTATCAAGTCTTCAAACAGCGTCTGGAATGGGACTTGGTTGTGGAGA
ACAATCTTGAAAGCGACGAGTATGACAATTCAAATGCGGAGTACATTTACGCATGT
GACGATACTGAGAATGTTTCAGGATGTTGGCGTCTTTTACCGACGACCGGGGATT
ACATGCTGAAGTCGGTATTTCCCGAATTGTTAGGTCAACAATCCGCTCCGAAGGAC
CCGAACATCGTCGAACTTTCTCGTTTCGCGGTAGGGAAGAACAGCTCTAAAATTAA
CAATTCTGCCTCCGAGATCACGATGAAGCTGTTCGAGGCTATCTATAAGCATGCGG
TTTCGCAAGGTATTACGGAGTACGTAACTGTTACATCGACTGCCATCGAGCGTTTC
CTGAAACGTATCAAGGTGCCCTGTCACCGTATCGGTGATAAGGAAATCCATGTCTT
AGGAGATACAAAATCTGTAGTACTTTCCATGCCTATCAACGAACAGTTCAAAAAGG
CAGTATTAAACTAATAAGCTTCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAA
GACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCA
CACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATACGCTGTCAGGAGTACTAG
ACCTGTAGGATCGTACAGGTTTACGCAAGAAAATGGTTTGTTATAGTCGAATAAAT
ACTTCACACAGGAAAGTACTAGAATGGTACGCTGGACTTTGTGGGATACCCTCGCT
TTCCTGCTCCTGTTGAGTTTATTGCTGCCGTCATTGCTTATTATGTTCATCCCGTC
AACATTCAAACGGCCTGTCTCATCATGGAAGGCGCTGAATTTACGGAAAACATTAT
TAATGGCGTCGAGCGTCCGGTTAAAGCCGCTGAATTGTTCGCGCTTACCTTGCGTG
TACGCGCAGGAAACACTGACGTTCTTGCTGACGCAGAAGAAAACGTGCGTCAAAA
ATTACGTGCAGAAGGAGTGAGCTTCCAGGCATCAAATAAAACGAAAGGCTCAGTC
GAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGA
GTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCGTTTATACGCTGTCATGAGAC
GCTGCAGTGACAGGATATATTGGCGGGTAAACTAAGTCGCTGTATGTGTTTGTTT
GAGATCTCATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGC
GTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGAC
GCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCC
CCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACC
TGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGG
TATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCC
CGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGG

TAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGC
GAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAC
ACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAG
AAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT
TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTT
GATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTT
TGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGA
AGTTTTAAATCAATCTAAAGTATATATGTGTAACATTGGTCTAGTGATTAGAAAAA
CTCATCGAGCATCAAATGAAACTGCAATTTATTCATATCAGGATTATCAATACCAT
ATTTTTGAAAAAGCCGTTTCTGTAATGAAGGAGAAAACTCACCGAGGCAGTTCCAT
AGGATGGCAAGATCCTGGTATCGGTCTGCGATTCCGACTCGTCCAACATCAATACA
ACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAG
TGACGACTGAATCCGGTGAGAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGT
TCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCATCAACCAAACCGTT
ATTCATTCGTGATTGCGCCTGAGCGAGTCGAAATACGCGATCGCTGTTAAAAGGAC
AATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCGCATCAACA
ATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCCCTGG
GATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATG
GTCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAAC
AACATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCT
TCCCATACAATCGGTAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCAT
TTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTTGAGCAAGA
CGTTTCCCGTTGAATATGGCTCATAACACCCCTTGTATTACTGTTTATGTAAGCAG
ACAGTTTTATTGTTCATGATGATATATTTTTATCTTGTGCAATGTAACATCAGAGA
TTTTGAGACACAACGTGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGA
TCAGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCA
AAATCACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACT
TTCTGGCTGGATGATGGGGCGATTCAGGCGATCCCCATCCAACAGCCCGCCGTCGA
GCGGGCTTTTTTATCCCCGGAAGCCTGTGGATAGAGGGTAGTTATCCACGTGAAA
CCGCTAATGCCCCGCAAAGCCTTGATTCACGGGGCTTTCCGGCCCGCTCCAAAAAC
TATCCACGTGAAATCGCTAATCAGGGTACGTGAAATCGCTAATCGGAGTACGTGAA
ATCGCTAATAAGGTCACGTGAAATCGCTAATCAAAAAGGCACGTGAGAACGCTAAT
AGCCCTTTCAGATCAACAGCTTGCAAACACCCCTCGCTCCGGCAAGTAGTTACAGC
AAGTAGTATGTTCAATTAGCTTTTCAATTATGAATATATATATCAATTATTGGTCG
C

Appendix II: Reactions of the designed genetic circuit model

LuxR

$$\emptyset \xrightarrow{k_{R1}} \text{mLuxR} \tag{1}$$

$$\text{mLuxR} \xrightarrow{k_{R2}} \text{mLuxR} + \text{LuxR} \tag{2}$$

$$\text{mLuxR} \xrightarrow{k_{R3}} \emptyset \tag{3}$$

$$\text{LuxR} \xrightarrow{k_{R4}} \emptyset \tag{4}$$

Where (1) is the transcription of the LuxR mRNA, (2) is the translation of LuxR, (3) is the degradation of the LuxR mRNA and (4) is the degradation of the LuxR protein.

LuxI

$$\emptyset \xrightarrow{k_{I1}} \text{mLuxI} \tag{5}$$

$$\text{mLuxI} \xrightarrow{k_{I2}} \text{mLuxI} + \text{LuxI} \tag{6}$$

$$\text{mLuxI} \xrightarrow{k_{I3}} \emptyset \tag{7}$$

$$\text{LuxI} \xrightarrow{k_{I4}} \emptyset \tag{8}$$

Where (5) is the transcription of the LuxI mRNA, (6) is the translation of LuxI, (7) is the degradation of the LuxI mRNA and (8) is the degradation of the LuxI protein.

AHL

$$\mathrm{LuxI} \xrightarrow{k_{A1}} \mathrm{LuxI} + \mathrm{AHL_i} \tag{9}$$

$$\mathrm{AHL_i} \underset{k_{A2}*Vc}{\overset{k_{A2}}{\rightleftharpoons}} \mathrm{AHL_e} \tag{10}$$

$$\mathrm{AHL_i} \xrightarrow{k_{A3}} \emptyset \tag{11}$$

$$\mathrm{AHL_e} \xrightarrow{k_{A4}} \emptyset \tag{12}$$

Where (9) is the synthesis of AHL by the LuxI protein, (10) is the diffusion of AHL between the cell and the medium, (11) is the degradation of the intracellular AHL and (12) is the degradation of the extracellular AHL. It is important to note that $\mathrm{AHL_e}$ is a species that is in the culture medium and not inside the cell, so this will have to be taken into account in the differential equation that defines its temporal evolution.

Transcription factor

$$\mathrm{LuxR} + \mathrm{AHL_i} \underset{k_{T-1}}{\overset{k_{T1}}{\rightleftharpoons}} \mathrm{LuxRAHL} \tag{13}$$

$$2\,\mathrm{LuxRAHL} \underset{k_{T-2}}{\overset{k_{T2}}{\rightleftharpoons}} \mathrm{TF} \tag{14}$$

$$\mathrm{LuxRAHL} \xrightarrow{k_{T3}} \emptyset \tag{15}$$

$$\mathrm{TF} \xrightarrow{k_{T4}} \emptyset \tag{16}$$

Where (13) is the association of AHL with LuxR, (14) is the dimerization of LuxRAHL, forming the transcription factor, (15) is the degradation of LuxRAHL and (16) is the degradation of the transcription factor.

GFP

$$\emptyset \xrightarrow{\text{f(TF)}} \text{mGFP} \tag{17}$$

$$\text{TF} \xrightarrow{\text{f(TF)}} \text{TF} + \text{mGFP} \tag{18}$$

$$\text{mGFP} \xrightarrow{\text{k}_{\text{G3}}} \text{mGFP} + \text{GFP} \tag{19}$$

$$\text{mGFP} \xrightarrow{\text{k}_{\text{G4}}} \emptyset \tag{20}$$

$$\text{GFP} \xrightarrow{\text{k}_{\text{G5}}} \emptyset \tag{21}$$

Where (17) is the constitutive transcription of the GFP mRNA, (18) is the induced transcription of the GFP mRNA due to the association of the transcription factor to the DNA, (19) is the translation of the GFP, (20) is the degradation of the GFP mRNA and (21) is the degradation of the GFP.

E-lysis protein

$$\emptyset \xrightarrow{\text{f(TF)}} \text{mE} \tag{22}$$

$$\text{TF} \xrightarrow{\text{f(TF)}} \text{TF} + \text{mE} \tag{23}$$

$$\text{mE} \xrightarrow{\text{k}_{\text{E3}}} \text{E} \tag{24}$$

$$\text{mE} \xrightarrow{\text{k}_{\text{E4}}} \emptyset \tag{25}$$

$$\text{E} \xrightarrow{\text{k}_{\text{E5}}} \emptyset \tag{26}$$

Where (22) is the constitutive transcription of the E-lysis protein mRNA, (23) is the induced transcription of the E-lysis protein mRNA due to the association of the transcription factor to the DNA, (24) is the translation of the E-lysis protein, (25) is the degradation of the E-lysis protein mRNA and (26) is the degradation of the E-lysis protein.

Appendix III: Models used in DNA basic parts characterization

Characterization of the interaction of LuxR with the pLux promoter

$$\frac{dX}{dt} = \mu X - \frac{\mu X^2}{K_{\mathrm{mx}}} \tag{27}$$

$$\frac{d[mLuxR]}{dt} = k_{\mathrm{R1}} - (k_{\mathrm{R3}} + \mu)[mLuxR] \tag{28}$$

$$\frac{d[LuxR]}{dt} = k_{\mathrm{R2}}[mLuxR] - k_{\mathrm{T1}}[LuxR][AHL_i] + k_{\mathrm{T\text{-}1}}[LuxRAHL] - (k_{\mathrm{R4}} + \mu)[LuxR] \tag{29}$$

$$\frac{d[AHL_i]}{dt} = k_{\mathrm{A2}} Vc[AHL_e] - k_{\mathrm{T1}}[LuxR][AHLi] + k_{\mathrm{T\text{-}1}}[LuxRAHL] - (k_{\mathrm{A2}} + k_{\mathrm{A3}} + \mu)[AHL_i] \tag{30}$$

$$\frac{d[AHL_e]}{dt} = k_{\mathrm{A2}}[AHL_i] - X(k_{\mathrm{A2}} Vc[AHL_e]) - k_{\mathrm{A4}}[AHL_e] \tag{31}$$

$$\frac{d[LuxRAHL]}{dt} = k_{\mathrm{T1}}[LuxR][AHL_i] - k_{\mathrm{T\text{-}1}}[LuxRAHL] - 2k_{\mathrm{T2}}[LuxRAHL]^2 + 2k_{\mathrm{T\text{-}2}}[TF] - (k_{\mathrm{T3}} + \mu)[LuxRAHL] \tag{32}$$

$$\frac{d[TF]}{dt} = k_{\mathrm{T2}}[LuxRAHL]^2 - k_{\mathrm{T\text{-}2}}[TF] - (k_{\mathrm{T4}} + \mu)[TF] \tag{33}$$

$$\frac{d[mGFP]}{dt} = k_{\mathrm{G1}} CN_g \frac{[TF]}{[TF] + k_{\mathrm{G2}}} + \alpha_g \frac{k_{\mathrm{G2}}}{[TF] + k_{\mathrm{G2}}} - (k_{\mathrm{G4}} + \mu)[mGFP] \tag{34}$$

$$\frac{d[GFP]}{dt} = k_{\mathrm{G3}}[mGFP] - (k_{\mathrm{G5}} + \mu)[GFP] \tag{35}$$

Where (27) is the temporal evolution of the biomass (number of cells), (28) is the temporal evolution of the amount of the LuxR mRNA, (29) is the temporal evolution of the amount of the LuxR protein, (30) is the temporal evolution of the intracellular AHL, (31) is the temporal evolution of extracellular AHL, (32) is the temporal evolution of the heterodimer LuxR-AHL, (33) is the temporal evolution of the transcription factor, (34) is the temporal evolution of the GFP mRNA, (35) is the temporal evolution of the GFP. The concentration of all species is expressed in nM, which inside the cell is equivalent to molecules/cell (Boada Acosta, 2018).

Characterization of the lethality of E-lysis protein

$$\frac{dX}{dt} = \mu X - \frac{\mu X^2}{K_{\mathrm{mx}}} - \frac{K_{me}[E]^n}{\Theta^n + [E]^n} X \tag{36}$$

$$\frac{d[mLuxR]}{dt} = k_{\mathrm{R1}} - (k_{\mathrm{R3}} + \mu)[mLuxR] \tag{37}$$

$$\frac{d[LuxR]}{dt} = k_{\mathrm{R2}}[mLuxR] - k_{\mathrm{T1}}[LuxR][AHL_i] + k_{\mathrm{T\text{-}1}}[LuxRAHL] - (k_{\mathrm{R4}} + \mu)[LuxR] \tag{38}$$

$$\frac{d[AHL_i]}{dt} = k_{\mathrm{A2}} Vc[AHL_e] - k_{\mathrm{T1}}[LuxR][AHLi] + k_{\mathrm{T\text{-}1}}[LuxRAHL] - (k_{\mathrm{A2}} + k_{\mathrm{A3}} + \mu)[AHL_i] \tag{39}$$

$$\frac{d[AHL_e]}{dt} = k_{\mathrm{A2}}[AHL_i] - X(k_{\mathrm{A2}} Vc[AHL_e]) - k_{\mathrm{A4}}[AHL_e] \tag{40}$$

$$\frac{d[LuxRAHL]}{dt} = k_{\mathrm{T1}}[LuxR][AHL_i] - k_{\mathrm{T\text{-}1}}[LuxRAHL] - 2k_{\mathrm{T2}}[LuxRAHL]^2 + 2k_{\mathrm{T\text{-}2}}[TF] - (k_{\mathrm{T3}} + \mu)[LuxRAHL] \tag{41}$$

$$\frac{d[TF]}{dt} = k_{\mathrm{T2}}[LuxRAHL]^2 - k_{\mathrm{T\text{-}2}}[TF] - (k_{\mathrm{T4}} + \mu)[TF] \tag{42}$$

$$\frac{d[E]}{dt} = k_{\mathrm{G1}} CN_e \frac{[TF]}{[TF] + k_{\mathrm{G2}}} + \alpha_g \frac{k_{\mathrm{G2}}}{[TF] + k_{\mathrm{G2}}} - (k_{\mathrm{G4}} + \mu)[mE] \tag{43}$$

$$\frac{d[E]}{dt} = k_{\mathrm{G3}}[mE] - (k_{\mathrm{G5}} + \mu)[E] \tag{44}$$

Where (36) is the temporal evolution of the biomass (number of cells), (37) is the temporal evolution of the amount of the LuxR mRNA, (29) is the temporal evolution of the amount of the LuxR protein, (39) is the temporal evolution of the intracellular AHL, (40) is the temporal evolution of extracellular AHL, (41) is the temporal evolution of the heterodimer LuxR-AHL, (42) is the temporal evolution of the transcription factor, (43) is the temporal evolution of the E-lysis protein mRNA, (44) is the temporal evolution of the E-lysis protein. The concentration of all species is expressed in nM, which inside the cell is equivalent to molecules/cell (Boada Acosta, 2018).