

Rapid prototyping of biomolecular circuits through module characterization in cell-free expression systems

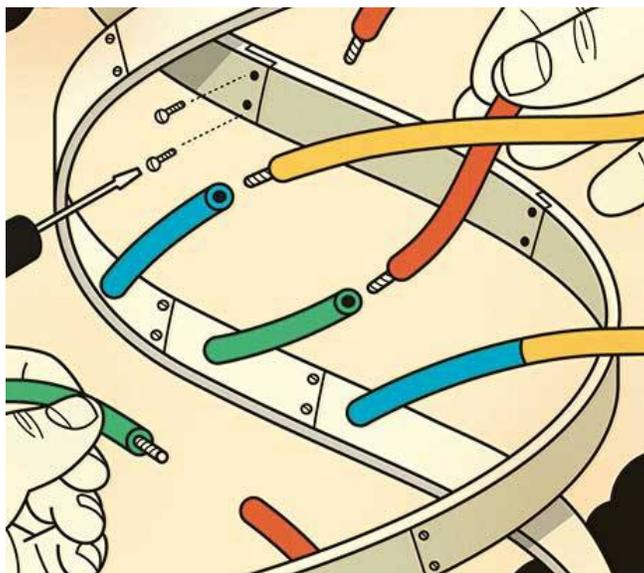
Anton Frisk

Mentors: Richard M. Murray and Johan Svensson-Bonde

Opponent: Leif Bülow

Co-Mentors: Clarmyra Hayes and Vipul Singhal

Collegues: Tiffany Zhou and Pulkit Malik



Abstract

Over the past years, the field of synthetic biology has gained a significant array of tools and parts, making way for increasingly complex bio-molecular circuits to be constructed. The development of biocircuits can be facilitated by assembling parts in a less complex, cell-free, environment which contains only the machinery for gene transcription (TX) and translation (TL), which have been extracted from bacteria. In this project, a part library was collected and used to assemble DNA constructs for a newly designed biocircuit. An *in vitro* TX-TL extract was used to test the circuit modules using linear DNA, and in parallel with predictive modeling of the biomolecular reactions, the overall circuit design was evaluated. The results have given valuable insight into the performance of the circuit modules in a much shorter time than conventional *in vivo* cloning and testing would have achieved.

Table of Contents

1. Introduction.....	3
1.1. Background to synthetic molecular biology.....	3
1.2. Genetic biocircuits.....	4
1.3. Cell-free transcription-translation (TX-TL) systems.....	8
1.4. Synthetic biology: research and applications.....	9
1.5. Cloning and DNA assembly.....	10
1.6. Computer modeling.....	12
1.7. Project aims.....	13
1.8. Circuit Designs.....	13
2. Methodology and materials.....	17
2.1. TX-TL modeling toolbox.....	17
2.2. TX-TL extract making.....	17
2.3. Golden Braid (GB) cloning and PCR.....	18
2.4. TX-TL experiments.....	19
2.5. Transformations and plasmid purification.....	19
3. Results and discussion.....	21
3.1. Characterization.....	21
3.2. Circuit design B: Attenuator-controlled NOR-gate.....	25
3.3. AND-gate 3: pSicA with invF and sicA.....	38
3.4. Continuation of the project.....	40
3.5. Future work.....	41
4. Conclusions.....	43
5. Acknowledgments.....	45
6. References.....	46
7. Appendicies.....	49
7.1. Glossary.....	49
7.2. Protocols and methods.....	51
7.3. Additional results.....	52

1. Introduction

The discovery of the DNA helix structure in 1953 marked the starting point of the endeavor to break down and understand how living things store genetic information and how it is used to control their biology on both a micro and macro scale [1], [2]. With increased understanding of the complex control network which governs how and when genetic information is used, the ability to genetically alter various life-forms, ranging from viruses to mammals, emerged. This has paved way for the research field of synthetic biology where novel biological systems, such as genetically modified organisms, are created by manipulating the biomolecular network that controls all life [3]. In this thesis, the latest tools of synthetic biology have been used to design and prototype a circuit of genetic regulator modules. This biocircuit, once carefully prototyped and tuned, could be implemented in bacteria to create a genetically modified bacterial strain which respond to inputs according to the circuit logic.

This chapter covers the fundamental ideas behind the research of this thesis. To aid the reader with field specific terminology, a glossary section (Appendix 7.1) is provided for terms marked in italics.

1.1. Background to synthetic molecular biology

Synthetic biology has previously been defined as the endeavor to design and construct new biological parts, devices and systems, and even re-designing the existing, natural biological systems, for useful purposes [4], [5]. This can be achieved on different scales, ranging from animal and plant engineering, to the individual cell and bacteria level, and all the way down to the biomolecular level involving protein and nucleic acid manipulations. The initial phase of biomolecular research was mainly concerned with understanding and describing the underlying mechanisms and made great progress in developing the tools necessary to study these systems. New biomolecular tools, such as the Nobel Prize winning development of restriction digest enzymes [6], paved the way for engineers and scientists to build new biological networks, resulting in both increased understanding of existing systems and a growing library of parts and modules which have been engineered towards perfection.

Although synthetic biology has largely focused on creating genetically modified organisms (GMOs) for various industrial, agricultural and scientific purposes, some work is also done on *in vitro* use of molecular biology systems with the aim of empowering technology with the flexibility and specificity of biology while by-passing the limitations of living cells, such as crosstalk with competing cell processes, tuning of growth

conditions and the spontaneous generation of mutations [3], [7].

1.2. Genetic biocircuits

Developing biological circuits is a key process when engineering advanced genetically-modified organisms (GMOs). With an increasing ability to design and assemble complex genetic circuits, fields such as bio-medicine, environmental monitoring and fuel production, among others, could potentially be made cleaner and more efficient with the tools of synthetic biology. The knowledge gained from implementing biochemical circuits in living organisms can also be used to provide further understanding about the complex network of gene expressions that controls cell functions [8].

In recent years, biocircuits with increasing complexity have been developed, and the building blocks that have been created can now be used for new, even more complex, circuits. The synthetic biology community has over the years collected and characterized a vast number up genetic parts, such as promoters, hybrid promoters, *transcription factors* (such as *repressors* and *activators*), fluorescent mRNAs and proteins, interfering mRNAs, enzymes and kinases to allow for circuit development. This library of parts can be used to, in theory, build any circuit logic function wanted. However, the performance of each part in terms of efficiency, cell toxicity and *orthogonality* to other parts, among other features, needs to be evaluated. Therefore there is a bias towards using well documented parts rather than introducing new ones that have not been thoroughly characterized [3].

By using this array of characterized parts, many circuits have been designed and realized, both *in vivo* and in cell-free breadboards *in vitro*. Logic gates, such as AND-, NOT- and NAND-gates, have previously been created within living cells [9]. These logic gates, when implemented in bacteria, can be used to steer bacterial behavior depending on external input such as hormones, inducer molecules, temperature and even light stimuli [10], [11]. This can be achieved by using transcription factors which change properties when exposed to certain molecules or conditions, an example of which is the repressor protein LacI which forms a complex with the lactose derivative Isopropyl-b-D-thiogalactopyranoside (*IPTG*) and, in doing so, loses its ability to repress specific promoters when the inducer is present in solution [12]. The growing library of well characterized parts and modules gives scientists new tools to further engineer bacteria for improved performance in various applications [3].

Biocircuit diagrams are used to provide an overview of the network of biochemical reactions that create the circuit logic. Figure 1 shows an example of an activator-repressor cascade diagram, with the relevant reactions included. Circuit diagrams, such as

this, will be used extensively throughout this thesis to illustrate the prototyped circuits and their expected output.

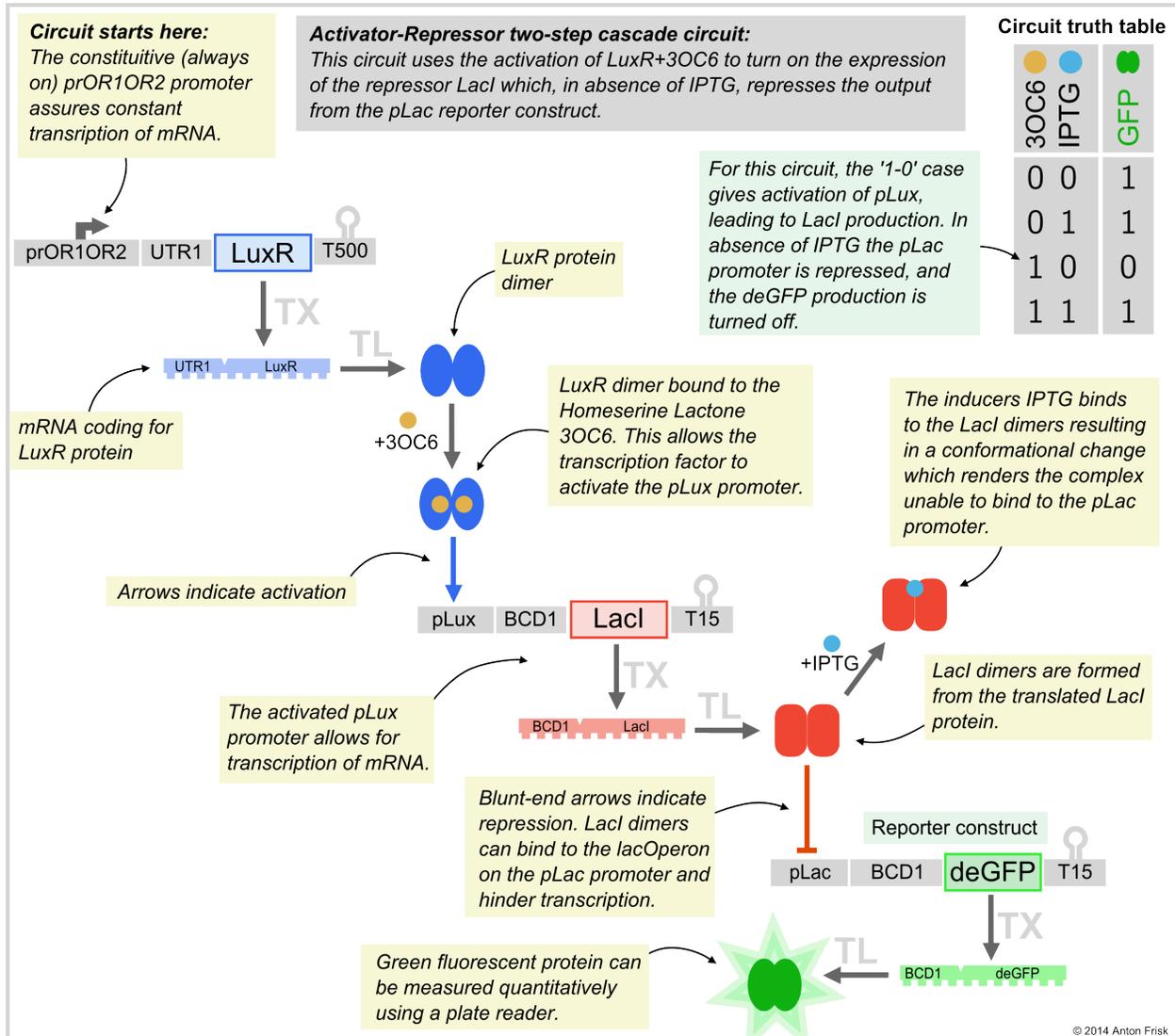


Figure 1: An example of a biocircuit diagram; an activator-repressor two-step cascade. The output from the reporter construct is controlled by the concentration of inducers IPTG and 3OC6 in accordance with the circuit truth table. For more details about DNA constructs, see section '1.5. Cloning and DNA assembly'.

To simplify these diagrams, the transcriptional and translational details are often excluded and only the activating and repressing arrows are shown. Circuit truth tables show how the output is expected to change depending on the input concentrations. For any given species, '0' denotes low or zero concentration and '1' denotes high or saturating concentration.

Combining circuit modules, like logic gates, has allowed for complex multi-input circuits such as the Four-input AND-gate (by Moon *et al*, shown in Figure 2) to be created. This circuit uses inducible transcription factors to control the AND-gates modules consisting of hetero-dimer transcription factors [13]. This circuit is of interest, not only because of its impressive number of integrated parts, but also because of the three new AND-gate modules which were used to build it. In section 1.8.2, these AND-gate modules are further explained and evaluated as candidate building blocks for circuit designs.

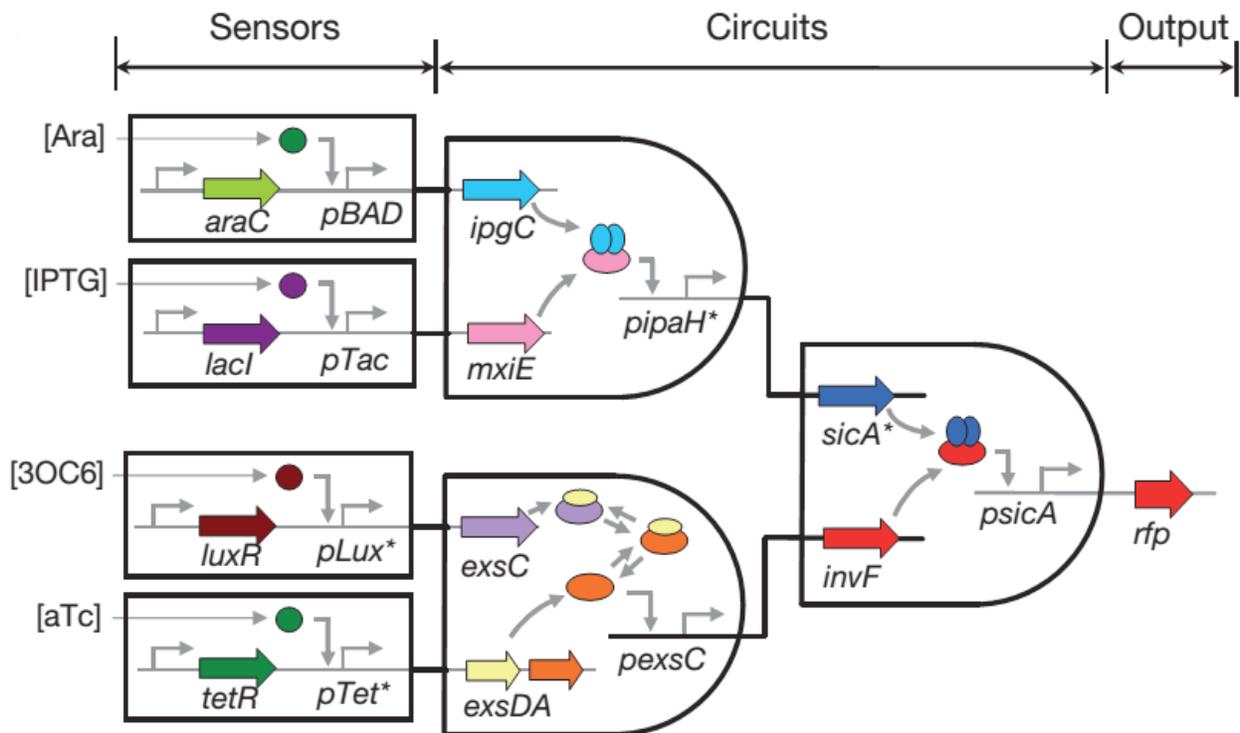


Figure 2: The Four-input AND-gate created by Moon *et al*. [13], reprinted with permission. Four signal molecules; Ara, IPTG, 3OC6 and aTc are used to activate (or release the repression from) four different promoters (*pBAD*, *pTac*, *pLux* and *pTet* respectively). This results in the production of transcription factors and chaperones which together activate the *ppaH** and *pexsC* promoters. When the first two AND-gate modules are activated, the two proteins needed to activate *pSicA* are produced which results in a detectable *RFP* production.

1.2.1. Logic gates

Logic gates are devices that perform logic operations with one or more inputs resulting

in a logic output. In modern day computers, logic gates are the basis on which memory and computation is built, and the electronic circuits in a microprocessor can contain more than 100 million gates. However, the concept of logic gates are not limited to electronic circuits; logic gates have been created using mechanical, optical, fluidic and even biomolecular systems. Figure 3 shows a set of the most fundamental logic gates which take one or two inputs and turns those into an output. In the circuit truth tables, provided to the right of each gate, the logic is displayed. When designing biocircuits, the concept of logic gates is often applied to describe the idealized circuit response to certain inputs. For instance, in this project biomolecular logic AND-gates have been tested and used to build a larger gate with a NOR-gate logic.

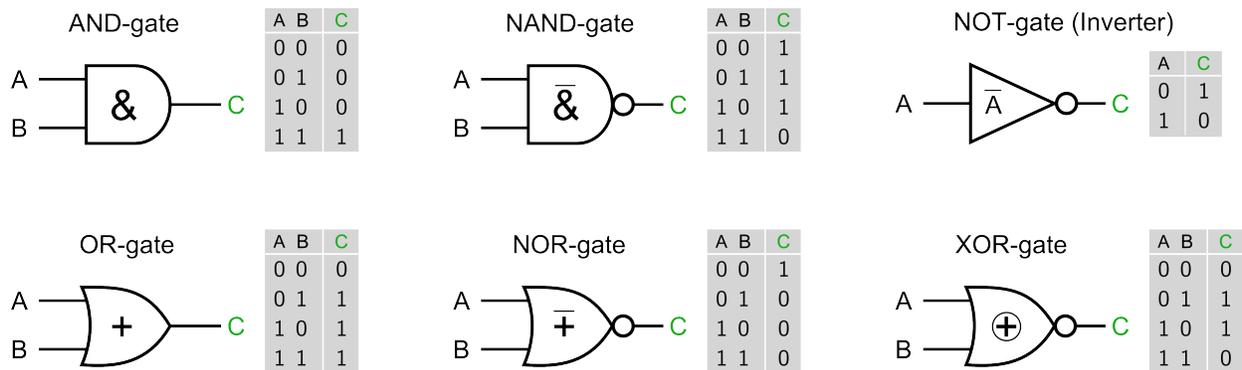


Figure 3: Logic gates with their symbols and corresponding truth tables. These logic gates have one or two inputs (A&B) and one output (C). In the truth tables, 1 denotes a high signal and 0 denotes a low/zero signal.

1.2.2. Measurable outputs

When designing biocircuits it is important to consider how the system can be observed. If the circuit output cannot be easily measured and quantified, both trouble-shooting and benchmarking will become difficult. Hence, the output of a circuit usually include a biomolecular reporter. Commonly used reporters are for instance fluorescent proteins, fluorescent mRNA aptamers and luminescent proteins.

Fluorescent proteins of different colors have been engineered from the wild-type genes, such as green fluorescent protein (GFP) which is found in a certain species of jellyfish (*Aequorea victoria*). The GFP protein, and its derivatives, forms a cylinder shape with a chromophore center which can absorb light in the violet spectra and emit in the green spectra. The process of forming the structure necessary to fluoresce is called maturation and normally take between 40 and 120 min (with great variability depending on conditions). Many variants of GFP have been developed to enhance its performance and

usability, such as the super-folder variant 'sfGFP' which matures much faster than the wild-type GFP in a range of conditions. Multiple fluorescence colors of GFP mutants have also been created, with emission peaks ranging from blue to yellow [14].

Similarly to the fluorescent proteins there are reporters consisting of RNA aptamers (short single-stranded RNA pieces) and specific dyes which come together to form a fluorescent complex. These reporters make it possible to measure transcriptional events rather than the output of the whole transcription-translation process as with fluorescent proteins. This results in a faster generation of output compared to GFP and its mutants [15].

Luminescent proteins, such as 'Luciferase', generate light through an oxidative enzymatic reaction with a substrate molecule, Luciferin. These type of proteins are the source of bioluminescence for many organisms, such as fireflies, and certain species of beetles, worms, mushrooms and photo-bacteria. The Luciferase enzyme catalyses the oxidation of Luciferin to form a product in an excited state which then emits light to relax down to its ground state. The reaction consumes both ATP and oxygen [16].

For detection of gene expression in bacteria, a great variety of reporters exist, all with distinct advantages and disadvantages. Fluorescent proteins do not require any dye or ATP for measurement, but the maturation time introduces an unwanted delay in output signal. RNA aptamer reporters give a fast response time, but the quick degradation of mRNA in cells can be a limiting factor. Luminescent proteins can give a strong output signal for low-light measuring. However, since the luminescence process consumes ATP, these reporters may be less suitable in situations where resources are scarce [16].

In this project, the fluorescent proteins *deGFP* and *sfGFP* have been used exclusively. These are lab standard reporters in the Murray Biocircuits lab where research was conducted, providing a highly useful basis for comparison with previous experiments as well as plate-reader calibration data which allowed for the quantification of output in nM (nanomolar units) rather than RFUs (*relative fluorescent units*).

1.3. Cell-free transcription-translation (TX-TL) systems

The development of biocircuits can be facilitated by assembling parts in a less complex, cell-free, environment which contains only the machinery for gene transcription (TX) and translation (TL) which have been extracted from bacteria. Such a 'breadboard' for synthetic biocircuits has been developed and successfully used for prototyping increasingly complex circuits [7], [8], [17]. This cell-free TX-TL system (here on referred to as just 'TX-TL') has the advantage that individual circuit parts can be tested without the adverse effects of '*cross-talk*' with other cell functions, and it allows for

rapid parts testing and *part mining*. In addition, cell parameters (such as ATP concentration, ion balance etc) can easily be controlled [7].

One challenge when prototyping biocircuits in a cell-free system is to find out how the test tube results correlate to circuit behavior within cells. For TX-TL to become a widely used circuit design tool this *in vitro* to *in vivo* mapping needs to be thoroughly explored and ultimately quantified. This work aims to contribute to an increased understanding of this mapping by re-characterizing modules in TX-TL which have previously been characterized *in vivo*.

1.4. Synthetic biology: research and applications

Synthetic biology is the focus subject for many prominent research groups world-wide. In bio-computation in mammalian cells, the 'Department of Biosystems Science and Engineering' at ETH Zurich is in the fore-front. Their work on creating modular systems for high performance bio-computer, with applications in, for example, creating circuits capable of detecting cancer mutations within cells (cell differentiation detection), allowing gene therapy to heal organisms from the disease by selectively inducing self-destruction mechanisms in dangerously mutated cells. The group has published several papers on bio-computation and new tool-sets for modular circuit development [18]–[20]. Other prominent biocircuit labs are the 'Voigt lab' at MIT and UCSF Biochemistry department, which jointly have developed the previously mentioned Four-input AND-gate [13].

At University of Minnesota, the Noireaux lab has worked with extending the use of cell-free extracts to explore synthetic biology systems from an *in vitro* perspective. With the use of TX-TL extracts, they have achieved self-assembly of artificial cytoskeletons in liposomes [21] and synthesized complete bacteriophages viruses from their genome [22]. Thus, they have shown the possibilities of reverse engineering life-forms from their genetic code.

TX-TL can also be used to build sensors and diagnostics devices, an example of which is a recently developed paper-based Ebola test created at Boston University. This system utilizes an RNA responsive toggle switch which produces a red marker in the presence of Ebola RNA and green in its absence. By freeze-drying the circuit DNA together with TX-TL extract and buffer, the whole circuit can be stored in a filter paper and used for cheap and versatile diagnostics [23], [24].

There are also commercially available cell free expression systems, which can be used for other purposes than circuit prototyping, often aimed at high yield protein synthesis for

purification.

In the 'Murray biocircuits lab' at Caltech, several complex biocircuits have been developed by prototyping parts and modules in the cell-free TX-TL environment [25], [26]. The ongoing projects strive to create circuits with more complex functions that involve a high number of parts and modules. One milestone is to be able to rapidly prototype circuits consisting of 8-16 promoters in a matter of weeks. The reduced complexity of the TX-TL environment also simplifies computer modeling of biocircuits, which has enabled the creation of a MATLAB-based modeling toolbox that allows testing of circuit logic *in silico* (see section 1.6).

1.5. Cloning and DNA assembly

Every DNA construct in a biocircuit needs to contain a certain set of DNA parts to functionally fill its role in the circuit. The following parts are typically included in constructs:

- Promoter (P): allows for transcription of the DNA by providing binding sites for RNA polymerases (RNAP).
- Untranslated region (U): allows for translation of RNA by providing a binding site for ribosomes (therefore also called RBS – ribosomal binding site).
- Coding sequence (C): provides the code to be read by ribosomes and thereby determines the amino acid sequence of the final proteins.
- Terminator (T): stops the transcription of RNA from DNA by forming a hair-pin secondary structure which stops RNAP.

To assemble the DNA constructs for a circuit, parts that are compatible with the chosen assembly method needs to be made (see Figure 4, arrow A). The templates used for part creation can be plasmids or linear pieces which have been derived from a variety of organisms using the tools of bioengineering and genetic modification. These DNA templates are then used to amplify the specific sequences needed using standard Polymerase Chain Reaction (PCR). For this project, more than 70 parts were created and for several of these functionality has been tested.

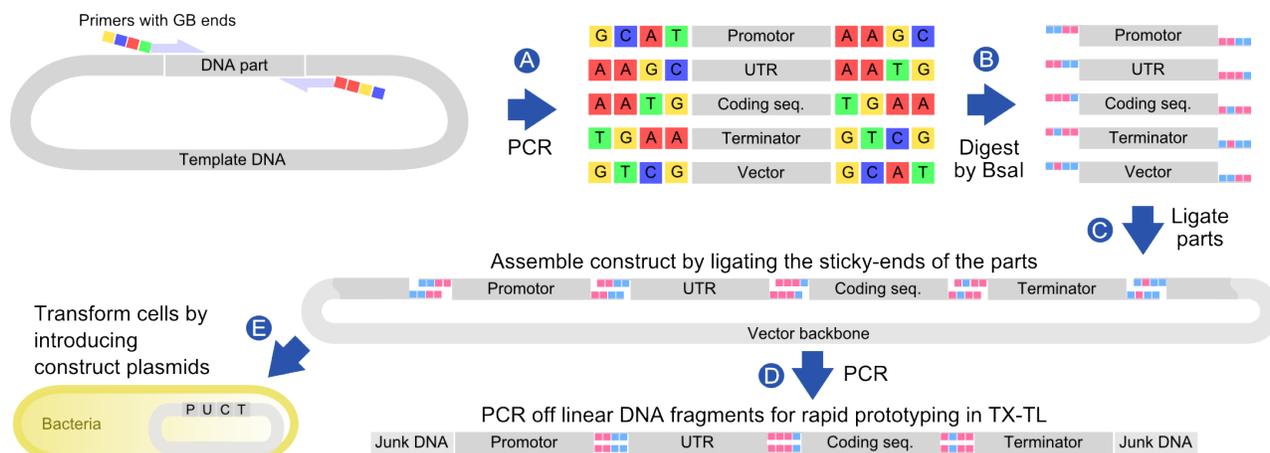


Figure 4: The process of Golden Braid cloning, from parts creation (A) and construct assembly (B&C), to testing *in vitro* (D) and *in vivo* (E).

To combine the individual parts into constructs (by some referred to as 'cassettes') the Golden Braid (GB) method used [27]. The advantage of this method is that it allows for several DNA pieces to be combined in one single reaction. This is achieved by having the ends of each part tailored so that it forms 'sticky ends' which then bind to the complementary DNA sticky ends of other pieces, after being digested by the enzymes BsaI or BbsI (Figure 4, arrow B). This forms a sequence of parts that attach to the vector backbone to form circular plasmids (Figure 4, arrow C). Construct assembly with the Golden Braid method is generally fast, compared to other assembly methods, but requires a library of compatible parts. While the assembled plasmids can be used to transform cells for immediate *in vivo* testing (Figure 4, arrow E), circuit debugging in a cell-free TX-TL system, prior to *in vivo* testing, is in many cases a far more time-efficient strategy [8]. The GB products can be used as PCR templates together with primers that include some 'junk DNA' to the ends; which helps to protect against DNA exonucleases that would otherwise destroy the region of interest during TX-TL construct testing. The PCR product can be readily used for testing in TX-TL (Figure 4, arrow D), and sequenced to determine that the correct DNA piece has been assembled.

The rate at which every DNA construct produces protein is dependent on a vast number of variables, of which only a few can be controlled with ease. In bacterial cells, the copy-number of a plasmid gives a rough estimate (low, medium or high) of how many plasmids are present within the cell, which depends on how strong the replication of origin on the plasmid is. In TX-TL experiments, the concentration of individual constructs can be tuned much more precisely to determine the optimal working conditions. Since there is no way to fine-tune DNA concentration ratios *in vivo*, other methods are used to change the expression strength of constructs, primarily by altering

promoter strength, ribosomal binding site (RBS) strength or inducer levels. The promoter part of a construct usually has some connection to the logic function of the circuit (e.g. it allows for activation or repression by transcription factors), and therefore tuning the strength of the promoter is sometimes not practical since it may alter its logic function. However, the RBS part of a construct often plays the sole role of initiating translation of mRNA, and can thus be changed to allow tuning of construct expression strength. By using an RBS of bicistronic design architecture (BCDs), the translational strength can be regulated in a predictable manner, with minimal influence from the upstream promoter and downstream coding sequence [28].

1.6. Computer modeling

In synthetic biology, like many other engineering disciplines, an interest for solving research questions through computer modeling and simulations has emerged with the aim of accelerating the progress and minimizing labor intensive lab work. Today there is only limited capabilities of simulating the vast number of interconnected biochemical reactions that occur within living cells. In comparison, the cell-free TX-TL environment is a much less complicated system to model, free from *cross-talk* with cell functions. As such, it is far easier to build an accurate but simplified model of the biomolecular reactions within a TX-TL experiment, compared to an equivalent model of an *in vivo* experiments [29]. To be able to achieve some predictive capabilities from a model, sufficient characterization experiments need to be done to determine some of the input parameters for the model. Ideally, a series of fairly simple characterization experiments would give the model the data needed to predict the behaviour of a larger circuit based on characterized parts and modules.

In an effort to achieve predictive modeling, researchers at 'Murray biocircuits lab' and their collaborators have developed a MATLAB toolbox for modeling of reactions in TX-TL. The 'TX-TL modeling toolbox' [30] is based on the MATLAB add-on SimBiology which allows the user to define chemical reactions and sets up ordinary differential equations for a system of reactions. This means that every reaction between species in the model will be assigned with a rate and the concentrations of species will vary over time. This way of modeling takes into account resource limitations and other features inherent to the TX-TL environment [8]. Naturally, all starting concentrations and reaction parameters in the system cannot be known with exactitude, which inevitably becomes a source of modeling error. However, the current toolbox parameters are largely based on characterization experiments and should therefore have values reasonably close to reality.

With increasing characterization data and growing knowledge of the mechanisms of biomolecular systems, predictive modeling holds the promise of accelerating synthetic biology research by shortening design iteration time. It will also allow researchers to theoretically explore questions that are otherwise too labor intensive to work with experimentally [31].

1.7. Project aims

The goal of this project has been to use the method of rapid prototyping in TX-TL to test and design a biocircuit of moderate complexity and characterize the modules that it consists of. The designed circuit should be possible to clone into E. Coli to enable *in vivo* testing, but the main focus has been on the prototyping and design phase of biocircuit development.

The proposed design process for rapid prototyping of biocircuits is as follows:

Stage 0: Modeling: Simulation of the main circuit and its modules.

Stage 1: Component prototyping: TX-TL implementation of the circuit constructs.

Stage 2: Component assembly: TX-TL implementation of the circuit using plasmid DNA.

Stage 3: In vivo implementation of the circuit: Transform cells with circuit plasmids.

Stage 2 and 3 were never reached during the project time and is therefore not described in detail in this report.

1.8. Circuit Designs

Finding a circuit design that allows for rapid prototyping and which satisfies the project aims requires great knowledge and planning. For this project, several different designs were proposed and evaluated from a theoretic stand-point. The main one is described here.

1.8.1. Circuit design B: Attenuator-controlled NOR-gate

The Attenuator-controlled NOR-gate was designed to explore how AND-gate modules could be controlled from stacked layers of repression. This circuit, shown in Figure 5, uses the repression from TetR and LacI to control the production of two antisense mRNAs (Anti 1 & 2) which subsequently attenuates the two AND-gate proteins NRI and NRII. The background of each of these circuit parts will be explained in this section.

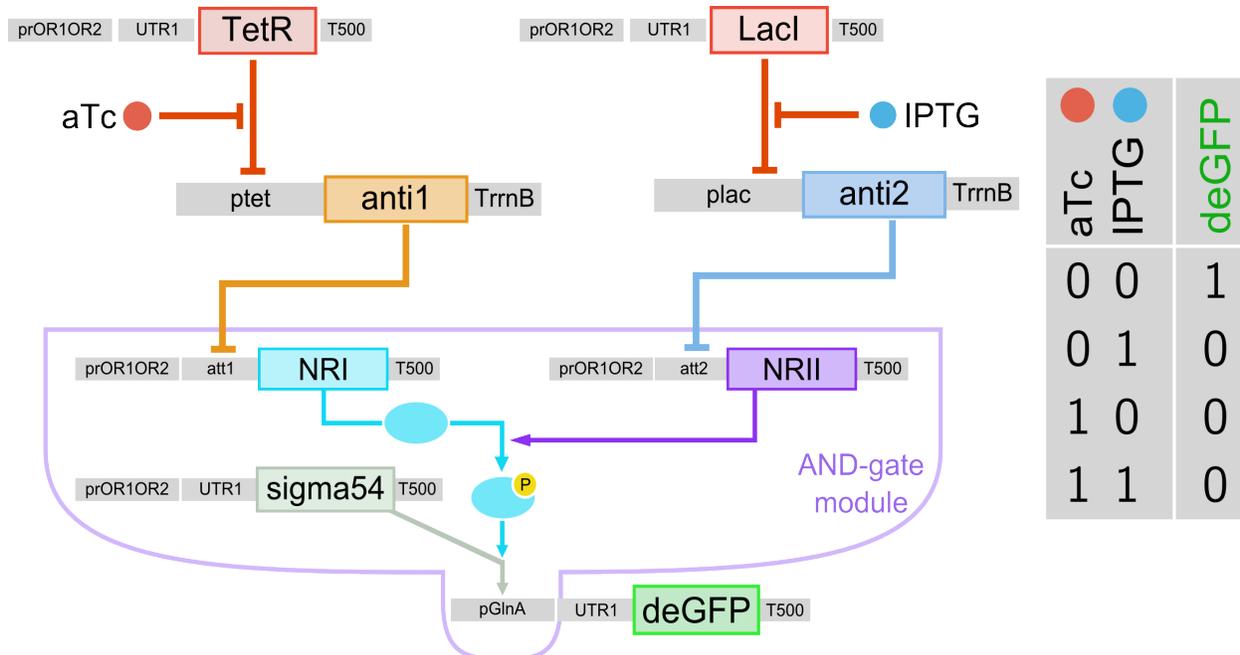


Figure 5: Circuit diagram for the 'Attenuator-controlled NOR-gate' in its original design proposal. The circuit consists of an AND-gate module based on the nitrogen regulator promoter pGlnA and a stacked layer of repression from antisense 1 and 2 followed by an inducible repression from TetR and LacI. The circuit output can be controlled with inducers aTc and IPTG in accordance with the truth table to the right.

The proposed AND-gate module comes from the nitrogen regulatory system of *E.Coli* and uses the pGlnA promoter which is activated by the presence of phosphorylated NRI (Nitrogen Regulator One), denoted NRI-P. The promoter is used within bacterial cells to regulate certain nitrogen-related metabolic pathways (such as glutamine synthetase) and is controlled by the nitrogen-limitation sigma factor; σ^{54} (sig54). When nitrogen-deficiency occurs within the cell, the kinase NRII will get phosphorylated by consuming ATP, and subsequently it can phosphorylate NRI which then activates the pGlnA promoter (Figure 6). NRI-P degrades over time (half-life measured *in vitro* to 3.5 min) to unphosphorylated NRI, which means that the NRII kinase needs to consistently consume ATP to keep the promoter activated [32], [33]. With sigma factor 54 naturally occurring in cells under the right conditions, this module forms a logic AND-gate, since both NRI and NRII are needed for activation.

To control this AND-gate module, the design uses attenuator-antisense pairs to repress the production of NRI and NRII. Premature transcriptional termination occurs when antisense mRNA binds to attenuator mRNA from the ongoing transcription of a construct. The antisense-attenuator interaction forces the mRNA strands into a downstream hair-pin loop which blocks RNAP from continuing transcription along the DNA strand and eventually lead to termination when RNAP breaks off [34]. Thus, when antisense mRNA is transcribed it can interact with the transcription process of constructs containing the corresponding attenuator region, and lead to premature termination and decreased expression from those constructs.

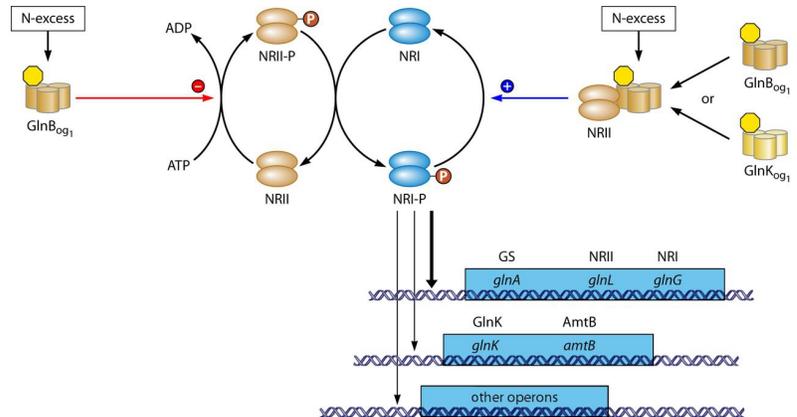


Figure 6: Transcription regulation of nitrogen-regulated genes by NRI-NRII. NRII consumes ATP, and the resulting NRII-P can phosphorylate NRI into NRI-P which activates certain promoters. Reprinted with permission from Heeswijk *et al* [46].

To control the antisense production, an additional layer of repression from TetR and LacI (Figure 5, top) is used. Since TetR and LacI respond to the inducers *anhydrotetracycline* (aTc) and *Isopropyl-b-D-thiogalactopyranoside* (IPTG) respectively, the circuit can be controlled with these inducers levels. These stacked layers of attenuation and repression is expected to give the circuit a NOR-gate logic where either aTc or IPTG will sequester the deGFP output (see circuit truth table in Figure 5).

1.8.2. Alternative AND-gate modules: *pipaH* and *pSicA*

AND-gate modules, such as the pGlnA system, can be found in other regulatory functions within cells. The Four-input AND-gate (Figure 2) uses three different promoters which are activated by a combination of two or more proteins interacting. These AND-gate modules were gleaned from *type-III secretion systems* (T3SS) of three different bacterial strains; *Shigella flexneri* (*pipaH*), *Pseudomonas aeruginosa* (*pExsC*) and *Salmonella typhimurium* (*pSicA*). The *pipaH* and *pSicA* AND-gate systems are based on transcription factors which form complexes with dimers of a smaller chaperone protein. The transcription factor-chaperone complexes can bind to, and thereby activate,

their respective promoters. AND-gate orthogonality to one another, and an increased dynamic range of the modules, was accomplished by editing the different DNA sequences, resulting in excellent *in vivo* performance for all three of the AND-gate modules [13]. These T3SS also involve other proteins which can bind the chaperones and inhibiting the activation [35], [36], but for the purpose of prototyping the *pipaH* and *pSicA* AND-gates, the full range of T3SS proteins were not needed.

2. Methodology and materials

The following sections aim to describe the methods used to go through *Stage 0: Modeling* and *Stage 1: Component Prototyping* of the rapid prototyping process. This involved building a parts library, *in vitro* cloning of constructs using the Golden Braid (GB) method, linearizing GB products through PCR reactions, and finally conducting TX-TL experiments using linear DNA. In addition, several TX-TL extracts were made and characterized.

2.1. TX-TL modeling toolbox

Computer modeling was done using the 'TX-TL modeling toolbox' [30] (trunk166), a MATLAB package based on the add-on SimBiology that allows for simulation of biocircuit designs (see section 1.6). All simulation runs were made with no native repressors (such as LacI) present, and with a virtual 'gamS' concentration of 20 μM . MATLAB 2014a was used.

2.2. TX-TL extract making

In order to run biocircuits *in vitro*, an extract containing the necessary transcription and translation machinery is needed. This TX-TL extract is made from bacteria cultures and processed through several steps to make an extract which is free of endogenous DNA but rich in enzymes that transcribe RNA from DNA, and translate proteins from RNA. Since this extract is derived from living bacteria, and due to the many processing steps, there may be great batch-to-batch variability. This could be due to cell growth conditions, the various cleaning processes, the cell lysis method and the particular bacteria strain used. This section aims to summarize the process of extract making, which followed a modified version of [37] that uses a homogenizer to lyse cells instead of bead beating.

To produce a well performing TX-TL extract, a healthy bacteria culture was needed. Mini-cultures were started by picking a single cell colony from a plate containing the desired bacteria strain and placing it in 4 ml fresh LB media containing nutrients and antibiotics which the particular strain was resistant to (typically Carbenicillin or Chloramphenicol). These cultures were grown for 5-8 hours at 37°C before 100 μl were transferred into conical flasks with 50 ml media and grown for another 8 hours in 37°C. In the final growth step, 6 ml of the cell culture was transferred to 4L conical flasks with 600 ml of LB media and 12.5 ml 50% glucose. Cells were grown for 4 to 7 hours at 37°C, until the liquid reached $\text{OD}_{600} = 3$. The incubation was then stopped to ensure that cells did not get effected by resource limitations in the media and leave the 'log-phase'

growth state. To extract the cells from the remaining media, the cultures were centrifuged at 5000 G for 12 min at 4°C, and the supernatant was decanted. The remaining cell pellets were cleaned by adding 200 ml of S30 buffer, dissolving the pellet, centrifuge for 12 min, and removing the supernatant. The cleaning steps were repeated twice before the remaining pellet was stored at -80°C.

On the day of extract making, the cell pellets were thawed on ice and re-suspended by adding 1 ml of S30 buffer per 1 g of pellet and vortexing. A homogenizer was used to lyse the cells and the resulting liquid was centrifuged at 12000G for 10 min at 4°C. The supernatant was transferred into a falcon tube and put in 37°C for run-off incubation. The run-off time was varied between 0 and 200 min to test the effect of this processing step. The extract was then centrifuged once more and the supernatant (the ready extract) was transferred into smaller tubes and flash-frozen in liquid nitrogen for long term storage.

2.3. Golden Braid (GB) cloning and PCR

PCR was done to amplify the parts needed for GB assembly, and to linearize the assembly product. All reagents were sourced from New England Biolabs[®] (NEB). The PCR reactions used 0.5 to 10 ng of template DNA and otherwise followed the protocol for 'NEB Phusion[®] Hot Start Flex 2X Master Mix' as described by the manufacturer [38]. A 'Bio-Rad C1000 Touch Thermal Cycler' was used to control temperatures in both PCR and GB reactions.

After PCR amplification of linear parts, the products were DpnI digested by adding 1 μ l of 'NEB DpnI Enzyme' and incubating in a water bath at 37°C for 2 hours to completely digest remaining methylated DNA. All PCR products were purified using 'QIAquick PCR Purification Kit' from Qiagen. The final DNA concentration was measured using a 'Nanodrop 2000' from Thermo Scientific. The accuracy proved to be dependent on the volume used for measuring, and for most concentrations measurements 2 μ l (which gave more stable reads) was used rather than 1 μ l.

Golden braid assembly was done using DpnI digested and purified DNA parts that were diluted so that 1 μ l of each part gave equimolar amounts of DNA to 50 ng of vector backbone. The process for GB assembly followed the protocol described in Appendix 7.2.1. When linearizing GB product, 1 μ l of the product was used as PCR template.

Some of the templates and primers needed for parts creation were already available in the lab, but many primers had to be designed and ordered as oligos from Integrated

DNA technologies (IDT). Primer design was facilitated by Geneious 7 (software package) which also was used to organize all DNA sequence files of parts and constructs. To systematically organize the parts, a letter and number was assigned to each part (e.g. the prOR1OR2 promoter used was designated 'P33') to properly distinguish them from other promoters with similar names. Following the naming convention for the parts library, every construct was assigned a number, preceded by an 'A' (for 'Assembled construct'), e.g. A50: pGlnA--UTR1--deGFP--T500 : P67—U32--C41--T11. During this project over 80 constructs were assembled, and the majority was used in different part of the project. All the constructs needed for the parts testing and circuit prototyping were made using the GB method, except constructs A66, A67, CH81 and CH92 which were made using Gibson assembly prior to the start of this project.

2.4. TX-TL experiments

Experiments with the cell-free TX-TL system largely followed a protocol developed at the 'Murray biocircuits lab'. Purified PCR products of linear DNA were used in all TX-TL experiments, with the sole exception of the standard 'positive control' (PC) plasmid which was added at 1 nM final concentration for experiment verification. To avoid linear DNA degradation by extract native RecBCD enzyme, 3 μ M of the RecBCD-inhibiting protein 'gamS' was included in the TX-TL master mix (consisting of extract and buffer).

DNA solutions and inducers were mixed in 500 μ l tubes at ratios determined using a TX-TL calculations spreadsheet [39] prior to adding the TX-TL master mix consisting of extract and buffer. Proper mixing was achieved by flicking the tubes or quickly vortexing before spin-down with a desktop centrifuge. 10 μ l of reaction mixture was transferred to each well in a 384 well plate which was then sealed with a protective plastic film. Typically the content of each tube was divided into two wells creating two repeat measurements from which a mean could later be calculated. After spin down of the 384 well plate at 2000 rpm for 10 seconds, the plate was placed in a 'BioTek Synergy H1 Hybrid Multi-Mode Microplate Reader' [40] to measure the green fluorescent protein production (gain: 61) from each well over at least 8 hours. Measurement temperature was held at 29°C. Data collected from TX-TL experiments were processed using 'LibreOffice Calc' and 'OriginPro 9' to produce graphs.

2.5. Transformations and plasmid purification

Construct that had been successfully used in circuit prototyping were selected for cell transformation and colony screening. Z-Competent (from Zymo Research) chemically competent cells were exposed to 5 μ l of GB product following the recommended transformation protocol. After overnight growth, single colonies were picked and used

for colony PCR to screen for transformed cells with the correct plasmid. Colonies which displayed the correct band were grown to make glycerol stocks and to purify the construct plasmids.

Purified linear DNA and plasmids were used for sequencing together with primers according to the guidelines from the sequencing company; Eurofins Genomics. The sequencing tubes were bar-coded and sent to Eurofins for DNA sequencing and the results were analyzed with Geneious 7.

3. Results and discussion

This section will describe the results from prototyping various circuit modules to gradually build up the designed circuit. Six different TX-TL extracts have been used during this project. The extracts mainly differ in the lysis methods and the bacteria strain used, which is summarized in Table 1.

Table 1: The TX-TL extracts used during this project and their respective bacteria strain and lysis method. *Run-off used indicates how long the incubation time was for the extract used in subsequent TX-TL experiments, see section 3.1 for more details.

Name	Cell strain	Lysis method	Maker	Run-off used* /min
Es1	BL21rosetta	Homogenizer	Clare Hayes	40-120 (mix)
Es2	BL21rosetta	Homogenizer	SURF14 students	120
Es3	BL21rosetta	Homogenizer	SURF14 students	120
Es4	DH5-alpha	Homogenizer	SURF14 students	40
Es5	ExpressIQ	Homogenizer	Zachary Sun	80 (20+60)
eZS3	BL21rosetta	Bead-Beating	Zachary Sun	80

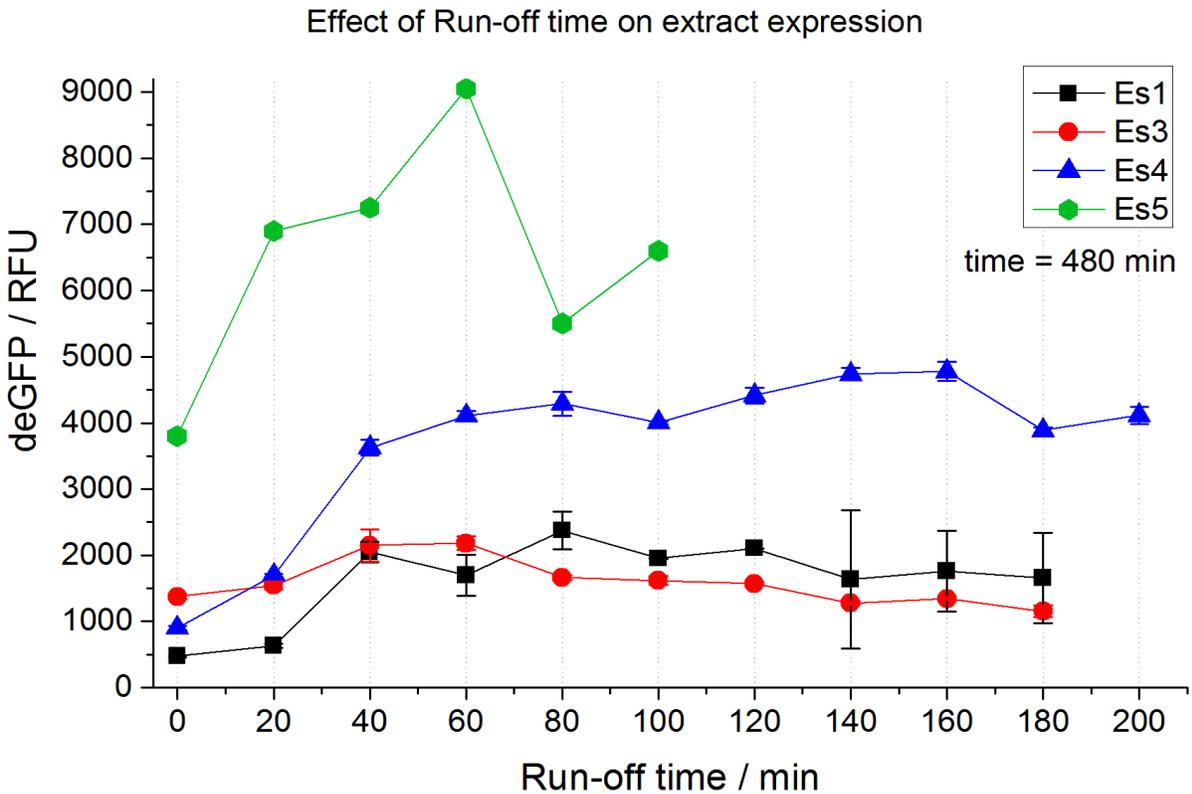
3.1. Characterization

For each extract, a buffer solution was made by calibrating Mg^{2+} and K^{+} concentrations for maximum expression with a high-expressing 'positive control' (PC) plasmid.

Additionally, experiments were designed to determine the optimal time for the 'run-off' process step. This processing step, where extract is placed in a shaker at 37°C (see section 2.2), is known to increase the signal from extracts, but the process by which this is achieved is largely unknown. The run-off time experiments for extracts Es1, Es3, Es4 and Es5, where each run-off time point was tested with the control plasmid and fluorescence at $t=480$ min was recorded, is shown in Figure 7. Extract Es2 was determined unsuccessful since it consistently produced fluorescence lower than 100 RFU at gain 61 for all run-off times. eZS3 was made prior to the start of this project and could therefore not be included.

The results show that there was an optimal run-off time for each extract which yielded an up to 4.5 times higher signal than before the run-off incubation. While all three extracts benefit from the first 40 minutes of run-off incubation, the effect of prolonged incubation seems to have little or adverse effect on extract performance. To determine

whether the optimal run-off time is a function of cell strain, growth condition or processing methods, this experiment would have to be replicated over a large number of extract batches.



Biotek 1&2 @ 485/515 nm, g61.

2014-07-25 & 2014-07-29

Figure 7: Data from run-off experiments with extract batches Es1:BL21ros:h, Es3:BL21ros, Es4:DH5a:h and Es5:ExpIQ:h. The optimum run-off incubation time varies between extracts. Es5 data is from single repeats, hence no error bars are shown.

3.1.1. Repressor-inducer characterization

Es5, based on the ExpressIQ strain, proved to give highest signal and was therefore used for further circuit prototyping. A 'C1-cassette' in the ExpressIQ genome makes it constitutively produce LacI repressor protein, so experiments were carried out to determine the native LacI repression in the extract. Figure 8 shows the deGFP production from increasing concentration of reporter constructs with the pLac and pTet promoters respectively (top graphs). The signal from the pTet reporter is approximately six-fold of that of the pLac reporter, which can be explained by the extract native LacI. The bottom graph of Figure 7 shows the effect of adding repressor producing constructs, A55 or A53, to 7 nM of the reporter constructs A71 and A46 respectively. As expected, the repressed deGFP signal (orange bars) could be regained by adding inducers aTc and

IPTG, respectively. In the pLac case (A46), additional IPTG also binds to the extract native LacI, hence the de GFP concentration goes up beyond the 'reporter only' level (blue bar). The signal from the pTet reporter (A71) was only regained to approximately 78% of the 'reporter only' case when the highest concentration (10 $\mu\text{g/ml}$) of aTc was used. This shows that all the expressed TetR dimers cannot be bound to aTc, even with relatively high inducer concentrations. Since aTc exhibits some fluorescence on its own, a negative control with equal concentration of the inducer was subtracted from the measured deGFP signal in the experiments where aTc was used.

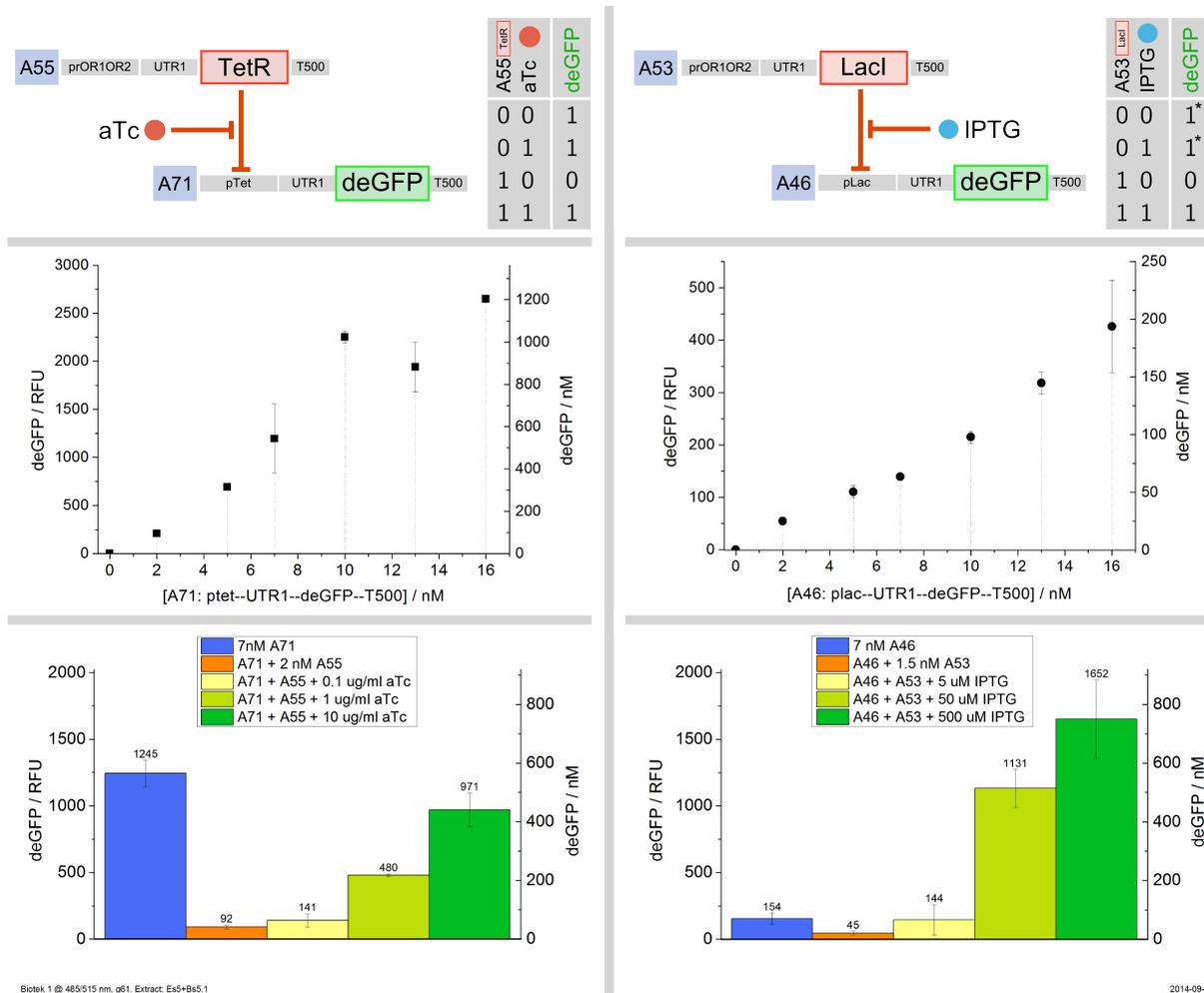


Figure 8: Characterizing TetR and LacI repression to determine construct functionality and the extract native LacI repression. The top graphs show how increasing reporter concentration leads to higher output deGFP in the absence of A55 and A53. Bottom graphs show how the presence of TetR and LacI lowers the deGFP levels, which can then be regained by increasing the inducer concentration with aTc and IPTG respectively. (* Without native LacI in the extract, the logic would be identical to the TetR repression).

These characterization experiments of extract Es5 shows that simple repressor and inducer functions behave as expected and that the extract expression was sufficiently high to be used for further circuit prototyping.

3.1.2. Tuning expression strength with ribosomal binding sites

To evaluate how gene expression can be tuned by the strength of its ribosomal binding site, four of the bicistronic designs (BCDs) from Mutalik *et al.* [28] were used to assemble reporter constructs. These reporters were benchmarked against a reporter with the lab standard 'UTR1' in a TX-TL experiment. The tested constructs were:

A36: prOR1OR2--UTR1--deGFP--T500 : P33--U32--C41--T11

A37: prOR1OR2--BCD2--deGFP--T500 : P33--U17(zs)--C41--T11

A5: prOR1OR2--BCD1--deGFP--T500 : P33--U35--C41--T11

A38: prOR1OR2--BCD18--deGFP--T500 : P33--U43--C41--T11

A39: prOR1OR2--BCD8--deGFP--T500 : P33--U46--C41--T11

When compared in TX-TL, the constructs show the same general pattern of RBS strength as describe in the original paper, as shown in Figure 9. The lab standard 'UTR1', which is not of BCD type, is known to give strong translation and is closely followed by BCD2 in RBS strength. Notably, the data from Mutalik is from *in vivo*

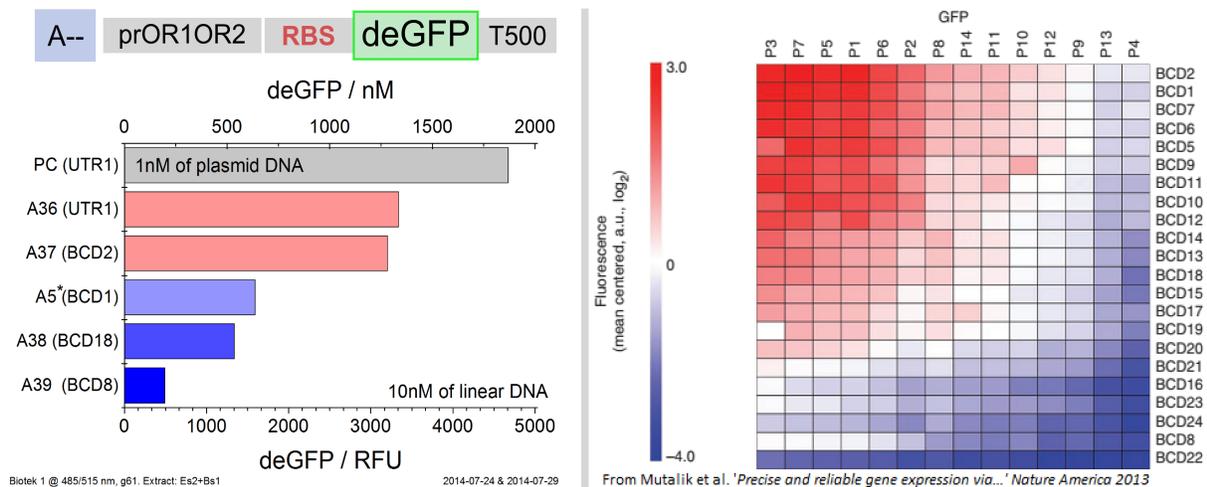


Figure 9: The deGFP expression can be modulated by changing the ribosomal binding site. To the left, four different BCDs were compared to the lab standard 'UTR1' and a positive control (PC) plasmid (shown in gray). Measured at $t=420$ min. All linear DNAs were at 10 nM concentration. *A5 (BCD1) data was from a separate run and has been normalized to the level of PC expression in that experiment.

To the right; adapted from Mutalik *et al* [9] and reprinted with permission. Heat-map of GFP expression for different combinations of promoters (columns) and BCDs (rows). The BCDs show predictable translational strength over a wide range of promoters.

experiments, but the *in vitro* TX-TL experiments gave a fairly similar pattern of RBS strength.

3.2. Circuit design B: Attenuator-controlled NOR-gate

The attenuator-controlled NOR-gate was chosen for prototyping with TX-TL experiments, for which the results will be presented in this section.

3.2.1. AND-gate 1: pGlnA with NRI, NRII and sig54

The first module to be tested was the pGlnA AND-gate. As previously described, the module should only activate in the presence of sigma 54 and phosphorylated NRI (which requires the kinase NRII to phosphorylate it). The pGlnA construct (A50), was tested with NRI (A61), NRII (A62) and sig54 (A63). All constructs were assembled with the GB method.

A50: pGlnA--UTR1--deGFP--T500 : P67--U32--C41--T11

A61: prOR1OR2--UTR1--NRI--T500 : P33--U32--C56--T11

A62: prOR1OR2--UTR1--NRII--T500 : P33--U32--C57--T11

A63: prOR1OR2--UTR1--sig54--T500 : P33--U32--C58--T11

In contrast to what previous work has found, the first test shows that the promoter was active even without adding DNA coding for sigma 54 (Figure 10). Furthermore, the signal dropped when 2nM of A63 was added. Sequencing of the linear DNA construct A50 showed that the pGlnA promoter sequence was intact, but A63 sequencing did not give a comprehensive read, which could be an indication of a failed DNA assembly. The permanent activation of A50 could be due to residual sigma54 and NRI-P (and possibly NRII) in the extract. Extracts produced with the bead-beating method from cells under slightly different growth conditions has shown to not contain significant amounts of

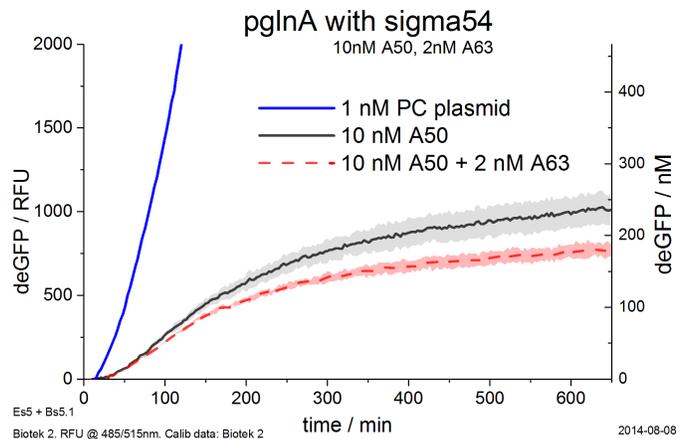


Figure 10: The pGlnA promoter on construct A50 is active even without the presence of DNA coding for NRI, NRII and sigma54. Surprisingly, the introduction of A63 results in lowered deGFP signal. Shaded areas show the signal from the two repeats in the experiment (std), and the mean is shown as lines.

sigma54 [41]. However, with a different growth protocol and lysis method, the extract used for these tests (Es5) could possibly contain some or all proteins needed to activate pGlnA. Further testing, possibly in different extracts, could be done to resolve this question.

When A61 and A62 were introduced to the system, the initial rate of deGFP production increased, but the added DNA load made the GFP production stop earlier (Figure 11). The experiments were done for both 10 nM and 20 nM of the reporter construct; A50. From the data, a time derivative was taken using the 'differentiate function' in 'OriginPro 9' with an applied Savitzky-Golay smooth filter (poly-order: 2, Points of window: 20). In Figure 11, the right hand graphs show the first derivative of the deGFP signal which reveals that the rate of production from the pGlnA promoter is indeed higher when A61 and A62 are added, compared to the case with A50 only (with or without A63).

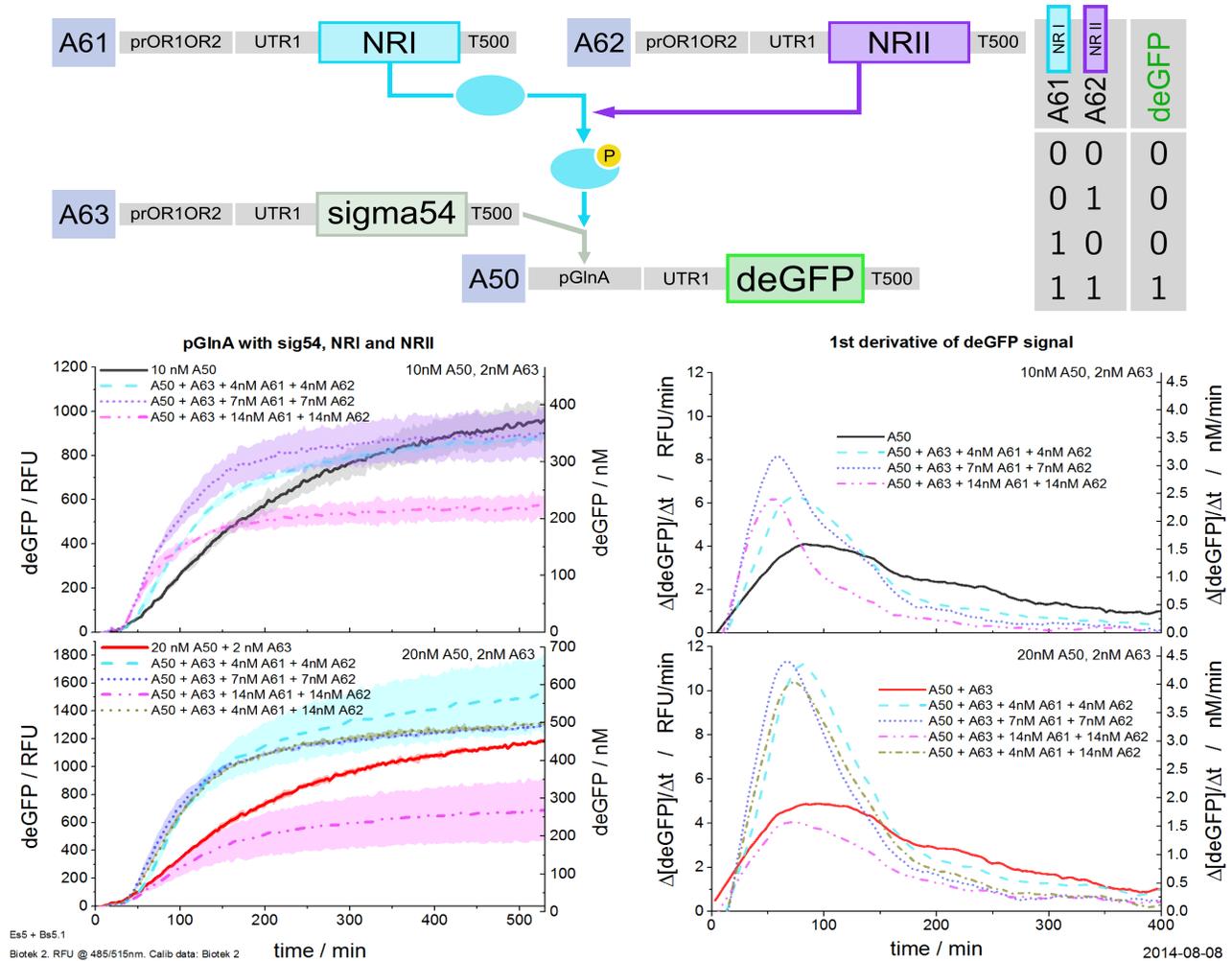


Figure 11: Prototyping experiment of the pGlnA AND-gate module. The deGFP production rate increases when A61 and A62 were added. However, the production stops at an earlier time which indicates that resource limitations greatly effected the experiment. Shaded areas on line plots show the standard deviation of two wells within the same experiment.

With an increasing amount of DNA in the system, it has been shown that resources such as NTP and amino acids gets consumed more rapidly [8], which explains the early stop in deGFP production when adding more of A61 and A62. Therefore, a more interesting metric when analyzing this AND-gate module might be the peak rate of deGFP production since resource limitations are less dominant during the initial phase of the TX-TL run. To correlate the experiment to *in vivo* function, the peak rate might also be a more suitable metric because resources are generally replenished in living cells, giving less pronounced loading effects.

In Figure 12, the peak rate of production from 10nM of A50 is shown over the concentration space of A61 and A62. This clearly displays that the promoter activation and resource loading both come into play resulting in an optimum working concentration point at 4 nM A61 and 14 nM A62. Overall, the functionality test of this AND-gate module shows that it can be activated to some degree, but it lacks a proper OFF-state. It is highly possible that the module would function differently *in vivo*, or in extracts prepared with different protocols, since previous lab experiences suggest this. More experiments would be needed to determine this.

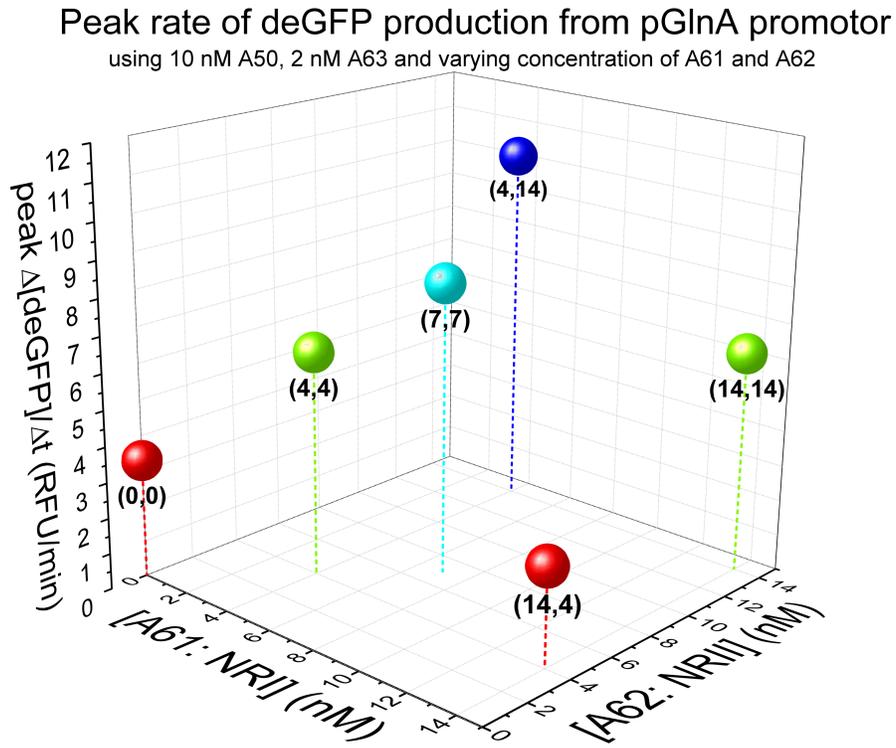
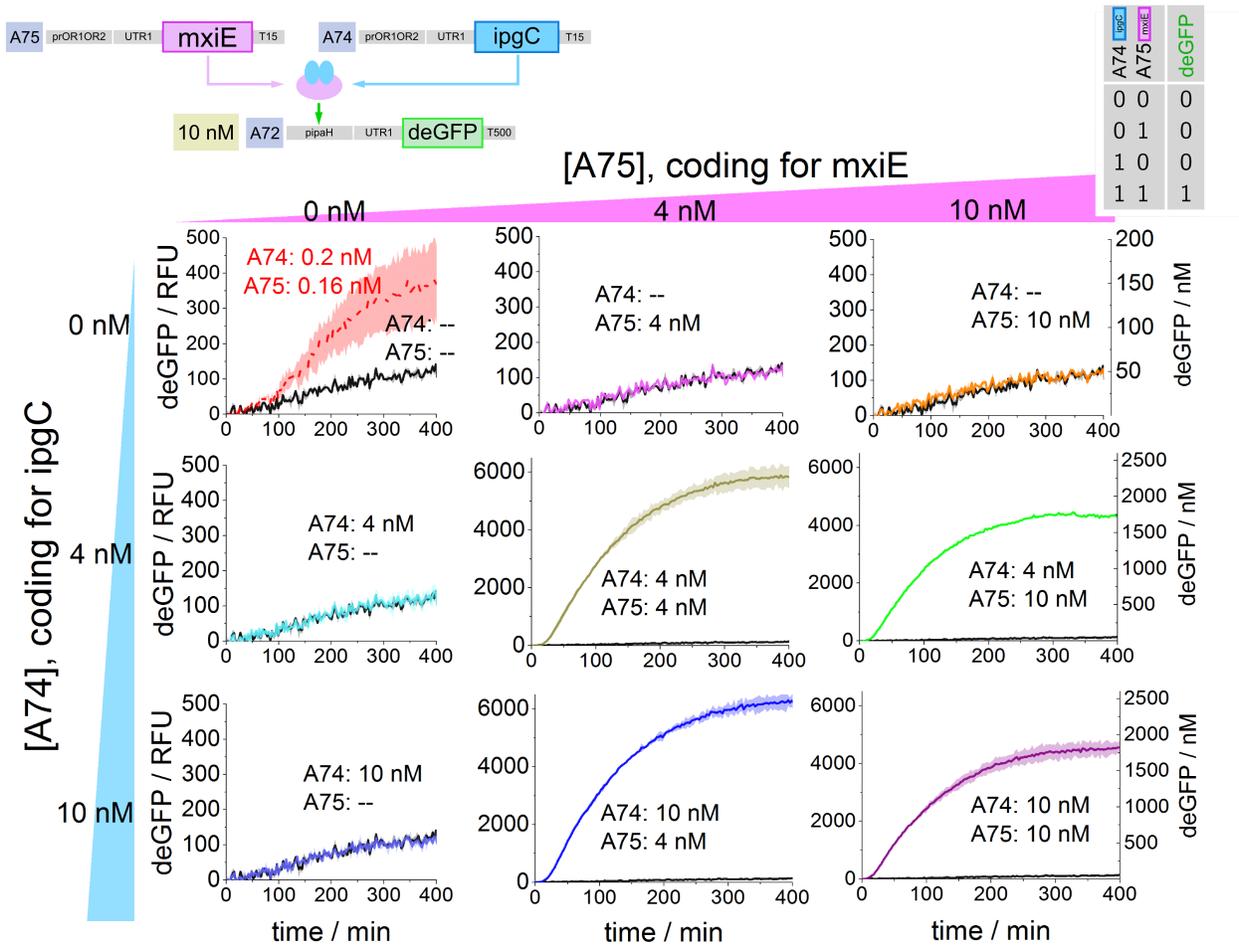


Figure 12: Mapping out the concentration space to find a optimum activation of the pGlnA promotor without putting heavy resource loads on the system. The highest peak rate of deGFP production is reached with 4 nM of A61 and 14 nM of A62.

3.2.2. AND-gate 2: *pipaH* with *mxiE* and *ipgC*

Since the results of the TX-TL testing of the pGlnA AND-gate module did not show a very convincing AND-gate logic, alternative AND-gate modules were considered. The *Shigella* system from Moon *et al* [13], with the transcription factor *mxiE* binding to the chaperone protein *ipgC* to activate the promoter *pipaH*, was chosen for initial TX-TL testing.



Biotek 1 @ 485/515 nm. Extract: Es5+Bs5.1

2014-08-18

Figure 13: TX-TL testing of the *pipaH* promoter with varying concentrations of DNA coding for transcription factor *mxiE* and chaperone *ipgC*. Promoter activity is low in presence of only one of the proteins, but up to 50-fold higher when both *mxiE* and *ipgC* are present.

Coding sequences for *mxiE* and *ipgC*, and the promoter sequence for *pipaH* (the modified sequence denoted *pipaH** in the original paper), were extracted from the original Moon *et al* plasmids, and the necessary GB ends were added. This allowed for three test constructs to be assembled;

A72: *pipaH*--UTR1--*deGFP*--T500 : P73--U32--C41--T2
A74: *prOR1-OR2*--UTR1--*ipgC*--T500 : P33--U32--C71--T2
A75: *prOR1-OR2*--UTR1--*mxiE*--T500 : P33--U32--C72--T2

After PCR amplification, the linear DNA constructs were tested in TX-TL using a constant reporter concentration (10 nM of A72) but varying the concentration of the *mxiE* and *ipgC* producing DNA pieces (0 to 10 nM of A74 and A75). The result, shown in Figure 13, shows a very low leakiness from the *pipaH* promoter and strong activation when both *mxiE* and *ipgC* are present (50-fold difference after 400 min).

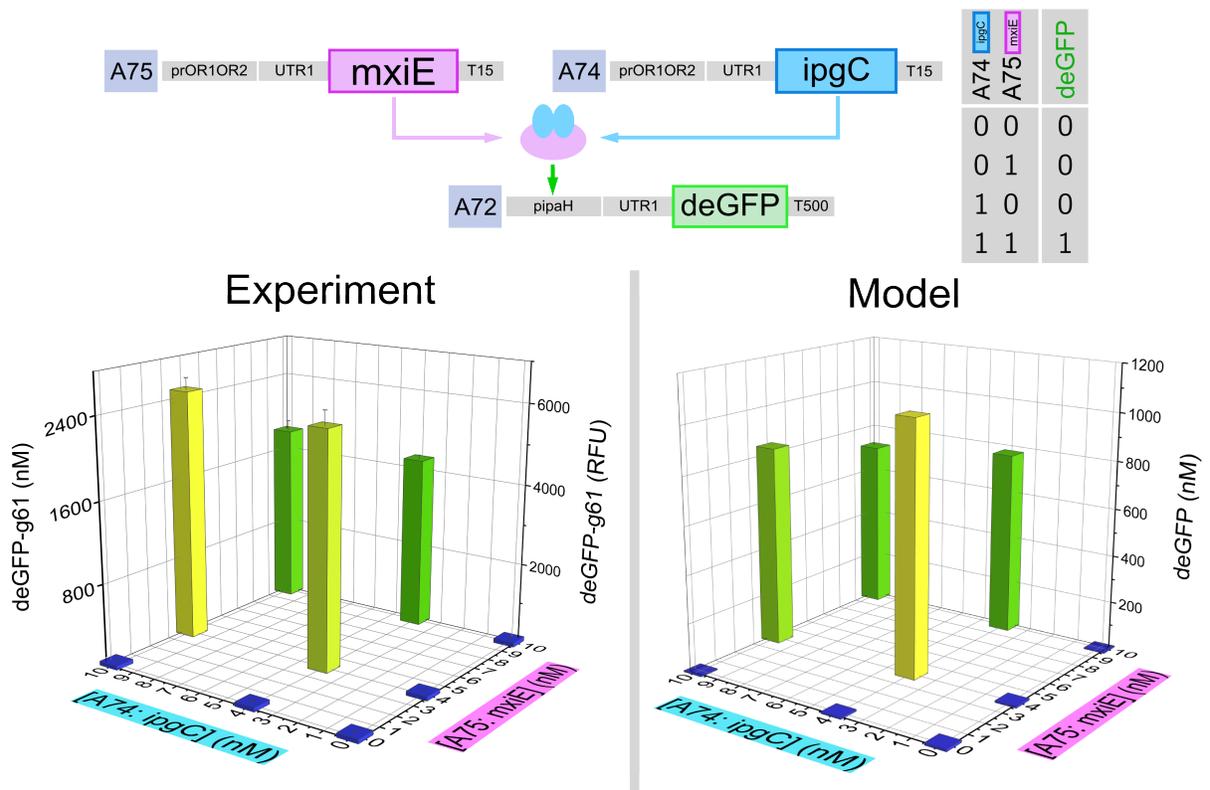


Figure 14: GFP signal from linear *pipaH* reporter construct (A72), $t=400$ min, shown over the A74-A75 concentration space. To the left: Experimental data (from Figure 12). To the right: Modeling results of the same setup using parameters tuned to fit the experiment. In Appendix 7.3.1, the full modeling result of the concentration space is shown as a 3D surface.

Loading of the TX-TL machinery is believed to be the cause for the drop in signal when adding 10 nM of A75, compared to the higher signal for both cases with 4 nM of A75. To reach the highest possible activation of 10 nM *pipaH* reporter, a A75 concentration lower than 10 nM should therefore be used. The optimal A74 concentration appears to be 10 nM but an eventual signal plateau was never quite reached, meaning that an even higher concentration of A74 could result in even stronger promoter activation.

The characterization of the *pipaH* module gave sufficient data to allow for a model to be created with the TX-TL MATLAB toolbox. The system parameters, such as the *mxiE*-*ipgC* binding rate, were tuned to fit the experimental data. As seen in Figure 14, the parameters could not be tuned to fit the experiment exactly and the magnitude of *deGFP* output was roughly 2.1-fold lower in the model, compared to the experiment. Naturally, there should be a parameter set for the model that more accurately fits the experimental data, but with a large multi-dimensional parameter space of reaction variables, the ideal parameter set is hard to obtain without using sophisticated search algorithms. This model was later used to step by step build the complete circuit and predict the responses, see section 3.2.5.

Figure 15 shows the rate of fluorescence increase, which can be used to compare *pipaH* activation with *pGlnA* activation. It is apparent that the fold change in rate for *pipaH* is much greater than that from *pGlnA*. The concentration space plot for *pipaH* also looks very similar to the end-point measurements which shows that end-point measurement is proportional to the rate in this case.

All in all the TX-TL prototyping of the *pipaH* AND-gate shows excellent ON/OFF-ratio and a strict AND-gate logic which makes it a suitable replacement for the *pGlnA* system in the main circuit.

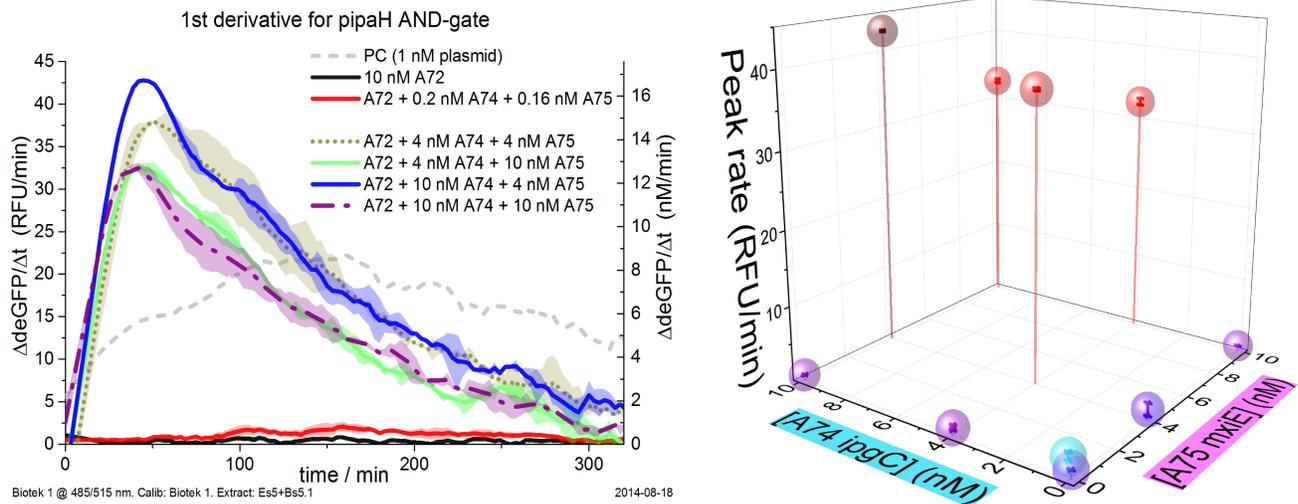


Figure 15: The time derivative of the fluorescence signal (RFU/min) from the *pipaH* TX-TL testing. The promoter is greatly activated in the presence of *mxiE* and *ipgC*, reaching a peak production after approximately 50 min. For comparison, the positive control plasmid at 1 nM is included in dashed gray. To the right, the peak rate is shown over A74-A75 concentration space.

3.2.3. Attenuator-antisense pairs

The next layer in circuit design B uses attenuator-antisense pairs to control the input to the *pipaH* AND-gate. To prototype this in TX-TL, several new constructs were assembled and their linear DNA was amplified to sufficiently high concentrations. The constructs used for functionality testing were:

A66: pJ23119--anti1--TrrnB : JBL004 (Directly linearized from pre-made plasmid)

A67: pJ23119--anti2--TrrnB : JBL008 (Directly linearized from pre-made plasmid)

A59: prOR1OR2--att1--deGFP--T500 : P33--U4s--C41--T11

A60: prOR1OR2--att2--deGFP--T500 : P33--U5s--C41--T11

First, the functionality of the attenuator parts (*att1* and *att2*) by the antisense parts (*anti1* and *anti2*) was tested. This was done by combining antisense constructs that had previously been tested (JBL004 and JBL008), with the untested attenuator constructs (A59 and A60). The results, shown in Figure 16 suggest that some attenuation effect is present, since the signal decreases when the correct antisense-attenuator pairs were combined.

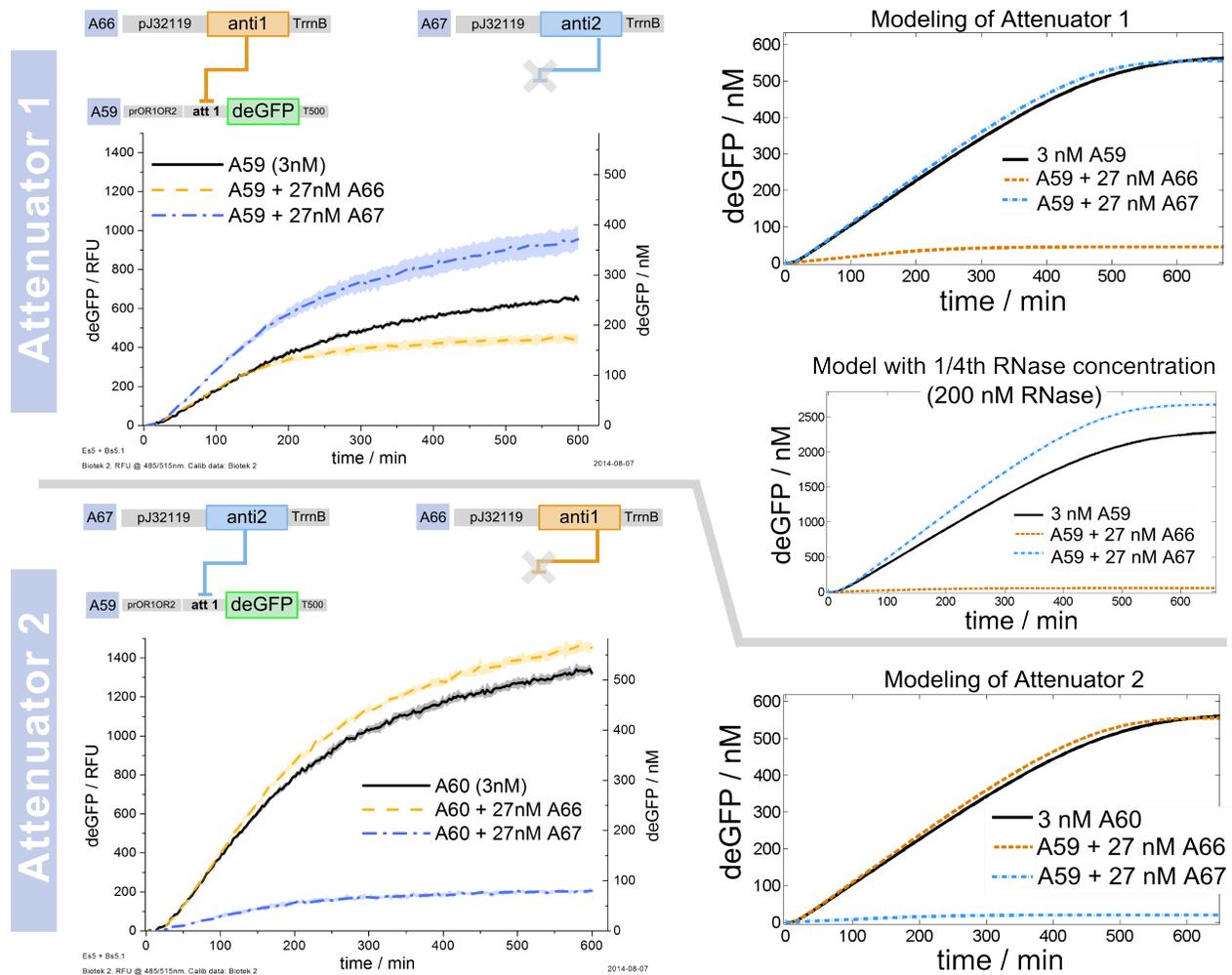


Figure 16: Functionality test of attenuator-antisense pairs. The top graphs (1) shows the attenuator 1 response to the two antisense constructs, and the bottom graphs (2) shows the same for attenuator 2. On the right; output from models of the attenuator-antisense experiments. For Attenuator 1, the results from a model with a lower RNase starting concentration is included to display how the effect of RNase loading can influence the result.

From these tests it can be concluded that the attenuator constructs A59 and A60 are functional, indicating that GB assembly with the attenuators as untranslated regions (UTRs) gives functional constructs capable of interacting with the antisense mRNA. For both attenuator constructs, an increased signal is gained when adding the 'wrong' antisense, which was supported by the predictive modeling results. This effect may seem counter intuitive but can be explained by the loading of RNases from the additional mRNA from these construct. By occupying RNases, the deGFP mRNAs in the solution are degraded much slower, resulting in a higher output signal. Figure 16 also includes a modeling result of attenuator 1 where a lower concentration (200 nM) of RNase was used. With 1/4th the RNase in the model, the effects of RNase loading from 27 nM of

A67 is greatly enhanced, as seen in the large signal increase. The actual RNase concentration in the extract is difficult to measure and was not determined in this work.

The RNase loading effect is in competition with the effects of increased DNA loading on the TX-TL system. In this first functionality test, RNase loading seems to be the dominant effect, but in subsequent tests it was not (see Figure 17). One possible explanation is that the concentrations used in this experiment were not measured with the same accuracy as in the rest of the experiments and the conditions may therefore be slightly different than reported.

3.2.4. Attenuator controlled *pipaH* AND-gate

To test the effect of attenuators on the *pipaH* AND-gate the following constructs were used:

A36: prOR1-OR2--UTR1--deGFP--T500 : P33--U32--C41--T2

A72: *pipaH*--UTR1--deGFP--T500 : P73--U32--C41--T2

A74: prOR1-OR2--UTR1--ipgC--T500 : P33--U32--C71--T2

A75: prOR1-OR2--UTR1--mxiE--T500 : P33--U32--C72--T2

A91: J23119--att1--ipgC--T15 : P68--U4s--C71--T15

A92: J23119--att2--mxiE--T15 : P68--U5s--C72--T15

A66: pJ23119--anti1--TrrnB : JBL004

A67: pJ23119--anti2--TrrnB : JBL008

To be able to separate the effects of DNA loading from the actual attenuator-antisense interaction, the two antisense constructs (A66 and A67) were tested on A36, A72+A74+A75 and A72+A91+A92, as shown in Figure 17 (circuit diagrams A to C). In the first two cases (A and B), there should be no attenuation from the antisense mRNA since none of the constructs contain either of the two attenuator regions. The top row of graphs in Figure 17 shows how the addition of antisense DNA lowers the deGFP signal for all three cases with case 'C' having the largest decrease. This was an expected result since 'C' will be affected both by DNA loading and by the antisense attenuation of the AND-gate constructs A91 and A92, while A and B will be affected by DNA loading only.

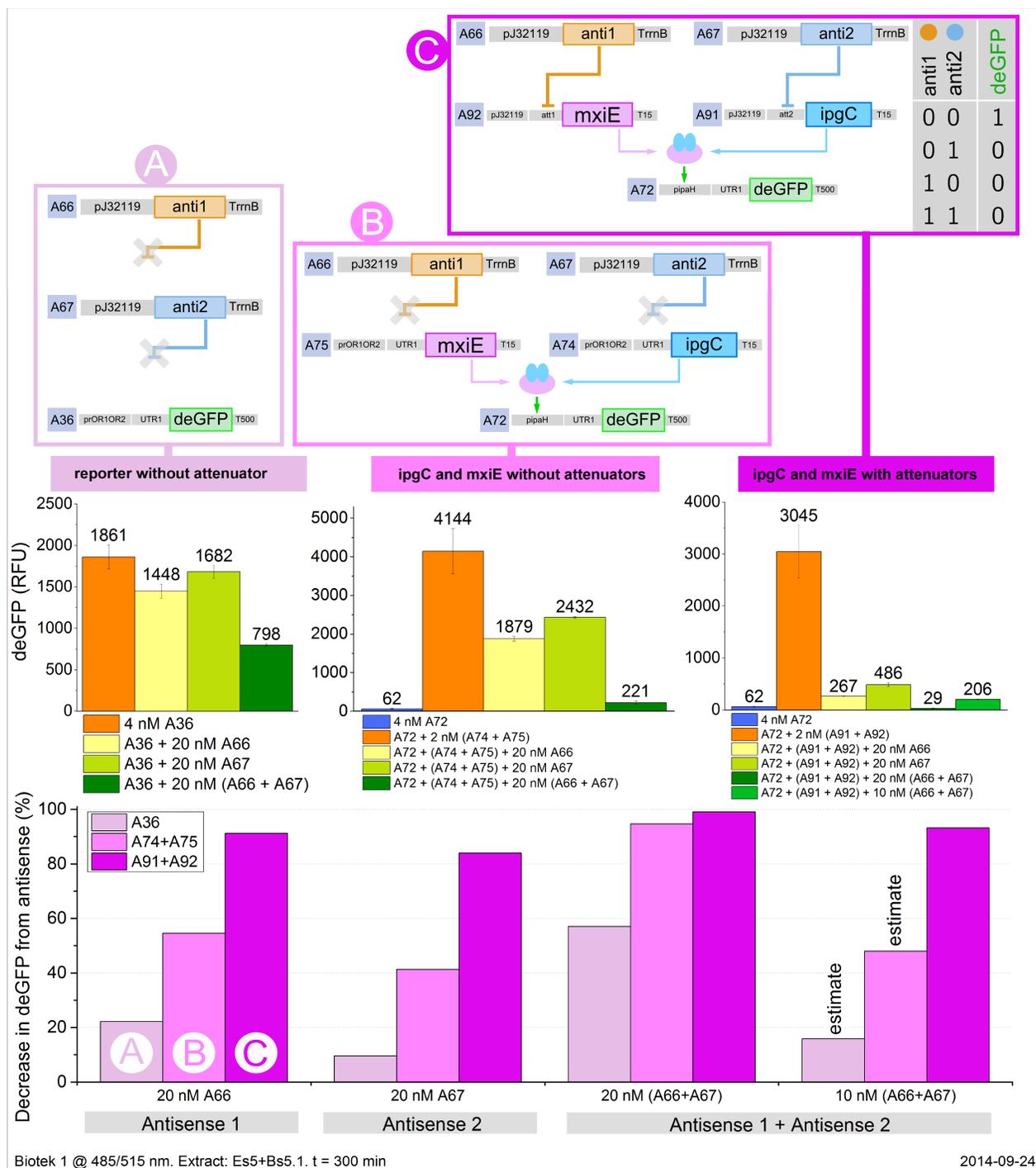


Figure 17: Testing the effect of antisense coding DNA on the expression of a downstream AND-gate module. To control for TX-TL loading, a simple reporter construct (A) and a attenuator-free AND-gate (B) was used. The attenuator controlled sub-circuit (C) is supposedly effected by both loading and attenuation from antisense mRNA. t=300 min.

In the bottom graphs, the decrease as compared to the highest signal (orange bars in upper graphs) is illustrated for all three data sets. The decrease in signal with A36 is consistently smaller than the decrease for the two AND-gate cases since 'A' contained 4 nM less DNA and the reporter does not need any transcription factors to be activated. In the two AND-gate cases, the two-step process (production of ipgC and mxiE first, followed by deGFP production) is greatly effected by the increased loading, thus there was no significant difference in the decrease between 'B' and 'C' when adding 20nM of both A66 and A67. However, when adding 20 nM of either A66 or A67 the attenuator controlled AND-gate (with A91+A92) decreases significantly more than case 'B'.

Assuming that DNA loading is solely responsible for the decrease in 'A' and 'B', the decrease from adding 10nM each of A66 and A67 can be estimated as the mean between the decrease from 20 nM of either A66 or A67. However, for C it is apparent that the addition of 10 nM A66+A67 results in a decrease slightly greater than the mean of the two cases with 20 nM of either A66 or A67. This suggests that attenuation of both the ipgC and the mxiE production is more effective, in decreasing the activation of the pipahH promoter, than stopping the production of only one of these AND-gate proteins.

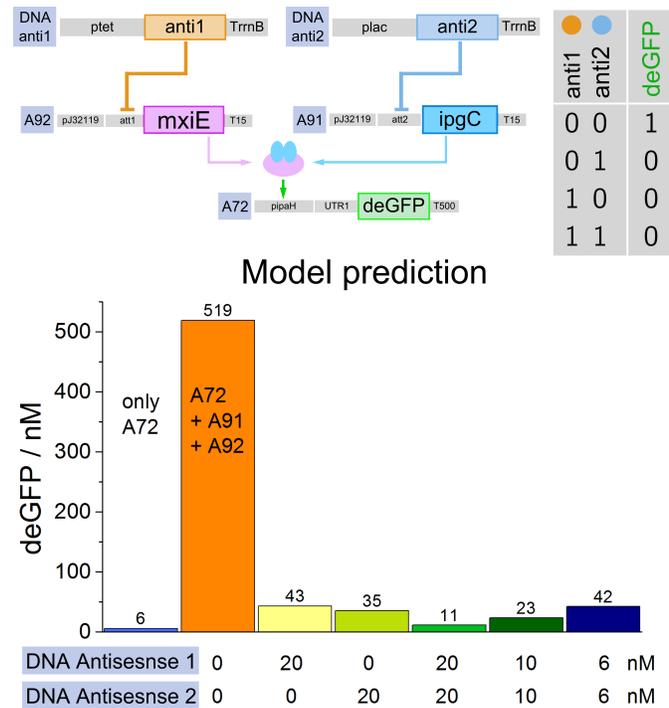


Figure 18: Modeling result of antisense controlling the pipahH AND-gate. The result should be compared to Figure 17, circuit C. The model uses 4 nM A72 and 2 nM each of A91 and A92.

A model was built to fit the experiment with case 'C' using the parameters tuned to fit the previous experiments with pipahH and attenuator-antisense pairss. The result in Figure 18 shows a circuit response close to the experiment with signal decreases of at least 92% upon addition of antisense coding DNA. As previously mentioned, the model does not fully encompass the effect of DNA loading, hence the signal decrease in the model is manly due to the antisense attenuation effect.

To be able to test the next circuit layer in TX-TL, attempts were made to make antisense construct which could be controlled by repressors TetR and LacI. However, these attempts were largely unsuccessful, partly due to the high DNA concentrations

needed to fit all the relevant circuit constructs in the reaction solution. Results from this can be found in Appendix 7.3.2.

3.2.5. Modeling of complete circuit

The complete 'Attenuator-controlled NOR-gate'-circuit was never successfully assembled in a TX-TL experiment. Instead, the modeling toolbox was used to predict the circuit logic at different stages.

A concentration of 6 nM of each of the antisense constructs was deemed to give a sufficient signal attenuation (as seen in Figure 17) and was used for further modeling. To determine the optimal working concentrations of TetR and LacI expressing constructs ('DNA TetR' and 'DNA LacI') a grid of simulations was made with varying concentrations of the two DNA constructs. Figure 18.A shows that much of the signal can be regained when adding the two repressor coding DNAs (compare (0,0): 42 nM deGFP to (2,2): 412 nM deGFP). However, when too much DNA is added the signal drops due to DNA overloading the system (e.g. (7,7): 252 nM deGFP). At this stage of the circuit, a clear AND-gate logic is displayed, where the presence of only one of the repressors (TetR or LacI) is not enough to significantly increase the signal since only one of the two AND-gate proteins will be expressed in abundance in these cases.

With fixed concentrations of 2 nM each of 'DNA TetR' and 'DNA LacI', the full circuit response to inducer levels could be modeled, as shown in Figure 18.B. When either aTc (TetR-inducer) or IPTG (LacI-inducer) are introduced, the repression on either of the antisense constructs is lifted and the resulting antisense mRNA attenuates the expression of the AND-gate proteins. The signal decrease is largest when both of the inducers are at high concentrations, which for TX-TL experiments mean up to 10 μ M aTc (= 4.6 μ g/ml) and 500 μ M IPTG. These inducer concentration levels result in a 8-fold decrease in signal (compare (1,1): 411 nM deGFP to (10000,10000): 51 nM deGFP). Such a fold-change in deGFP expression would be readily detectable in a TX-TL experiment, provided that the signal-to-noise ratio is sufficiently high. The full circuit exhibits a clear NOR-gate logic, as anticipated when designing the circuit.

It remains uncertain how well this circuit model relates to a real circuit assembled *in vitro* or *in vivo*. In TX-TL, the effect of DNA loading would most likely bring down the overall signal, resulting in a smaller fold-change response. In cells, where resources are continuously replenished, the loading effects are expected to be less dominant. However, the inherent possibility of cross-talk with cell functions might affect the circuit performance.

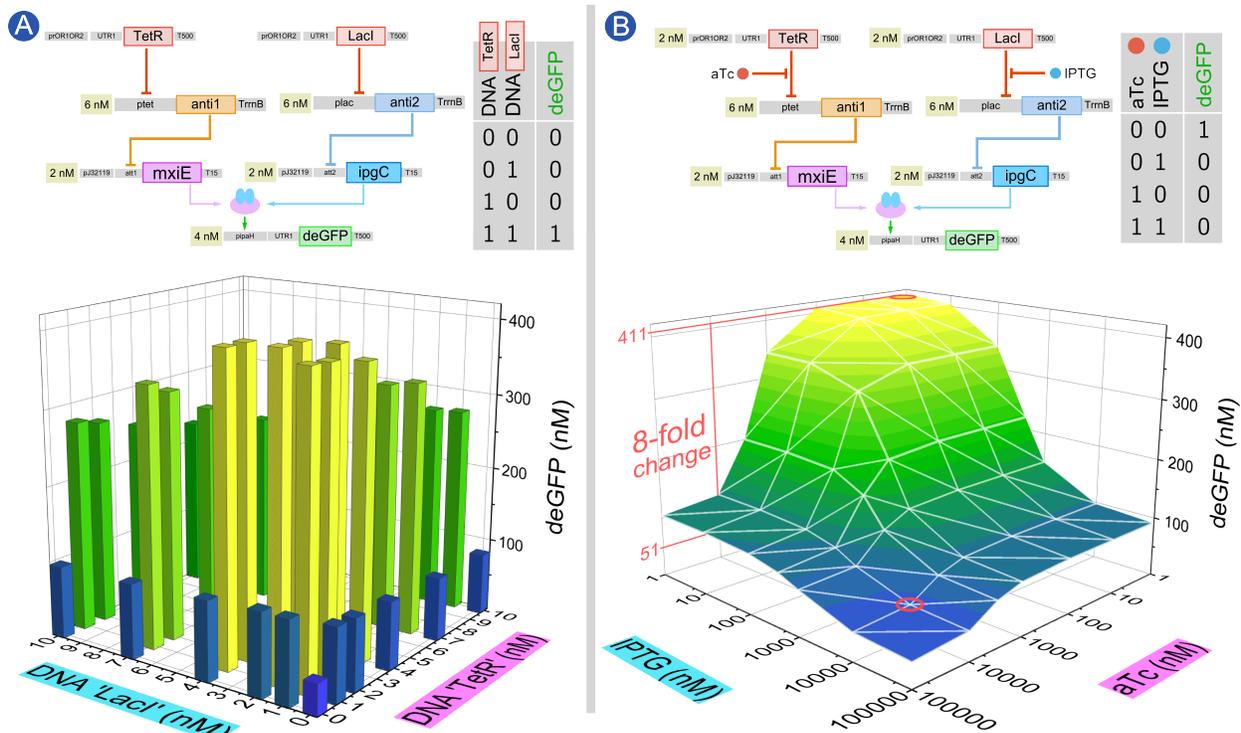


Figure 19: Modeling of the complete 'Attenuator-controlled NOR-gate'. A: Modeling of the circuit with varying repressor concentrations to find an optimal repression levels. B: The full circuit response to inducers aTc and IPTG. The model predicts an 8-fold change in deGFP expression from the circuit. In Appendix 7.3.3, the result of 'A' is provided from a different angle and as a heat-map.

3.3. AND-gate 3: pSicA with invF and sicA

To test the pSicA AND-gate module, which had the largest dynamic range according to 'Four-input AND-gate'-paper by Moon *et al*, the DNA parts had to be extracted from the original plasmids. The promoter, pSicA, and the chaperone CDS, sicA, were PCR'd up using newly designed primers, while the transcription factor CDS, InvF, had to undergo extra cloning to make it compatible with GB assembly (see inlet).

***In vitro* cloning of a GB compatible InvF coding sequence**

The coding sequence for the transcription factor 'InvF' contained cut-sites sequences for both of the restriction digest enzymes BsaI and BbsI, making GB assembly with this CDS impossible. To remove these cut-sites, a three step process was used. First, primers were designed to PCR out three sub-segments of the CDS so that the two cut-sites were removed. Secondly, the three segments were merged together in a GB reaction to form a CDS free of cut-sites. This was done by designing the first set of primers so that the three segments all got compatible GB sticky-ends which merged them together with a vector backbone in the reaction.

The edits in the CDS naturally meant that some of the codons in the sequence were altered, but by using the E.Coli codon library [45], these alterations could be made so that the codons still coded for the same amino acids as before. The end result was an InvF CDS without cut-sites that in theory should be translated to the exact same protein by the E.Coli translation machinery. In the third step, the new coding sequence was PCR from the intermediate GB product with primers that added the standard GB compatible ends. This part was used to assemble new constructs, such as A118: prOR1-OR2--UTR1--InvF--T16.

After sequencing of linearized A118 containing the new InvF CDS, it was apparent that the cloning procedure was successful. TX-TL prototyping of the whole pSicA AND-gate would reveal if the new InvF CDS actually resulted in a protein with the same functionality as before the cloning.

The newly developed parts were used to make three test constructs that were linearized for rapid TX-TL prototyping:

A104: pSic--UTR1--sfGFP--T16 : P75--U32--C83--T16

A105: prOR1-OR2--UTR1--sicA--T16 : P33--U32--C85--T16

A118: prOR1-OR2--UTR1--InvF--T16 : P33--U32--C86--T16

The concentrations of the transcription factor construct, A118, and the chaperone construct, A105, were varied from 0 nM to 6 nM, while the concentration of the reporter construct, A104, was kept constant at 6 nM. The result, using extract Es5, is shown to the left in Figure 20. An ON/OFF ratio of 740 can be seen from the (0,0) case to the highest expressing cases, and the off-states signals were low, ranging from 39 to 435 RFU at gain 61. To investigate the effect of using different extract preparation methods, the bead-beating extract eZS3 with corresponding buffer bZS3 was used for a replicate experiment. The results from this second run, shown to the right in Figure 20, were similar to those obtained with homogenizer extract Es5, though the overall magnitude of signal was about 1/3 lower. Thus it could be concluded that the extract preparation method does not greatly effect the functionality of the pSicA module. Compared to the

pipaH AND-gate, the pSicA module shows an even greater ON/OFF ratio. The activation of the pSicA promoter seems to very drastic as only small additions (0.5 nM) of A105 and A118 gives a significant activation.

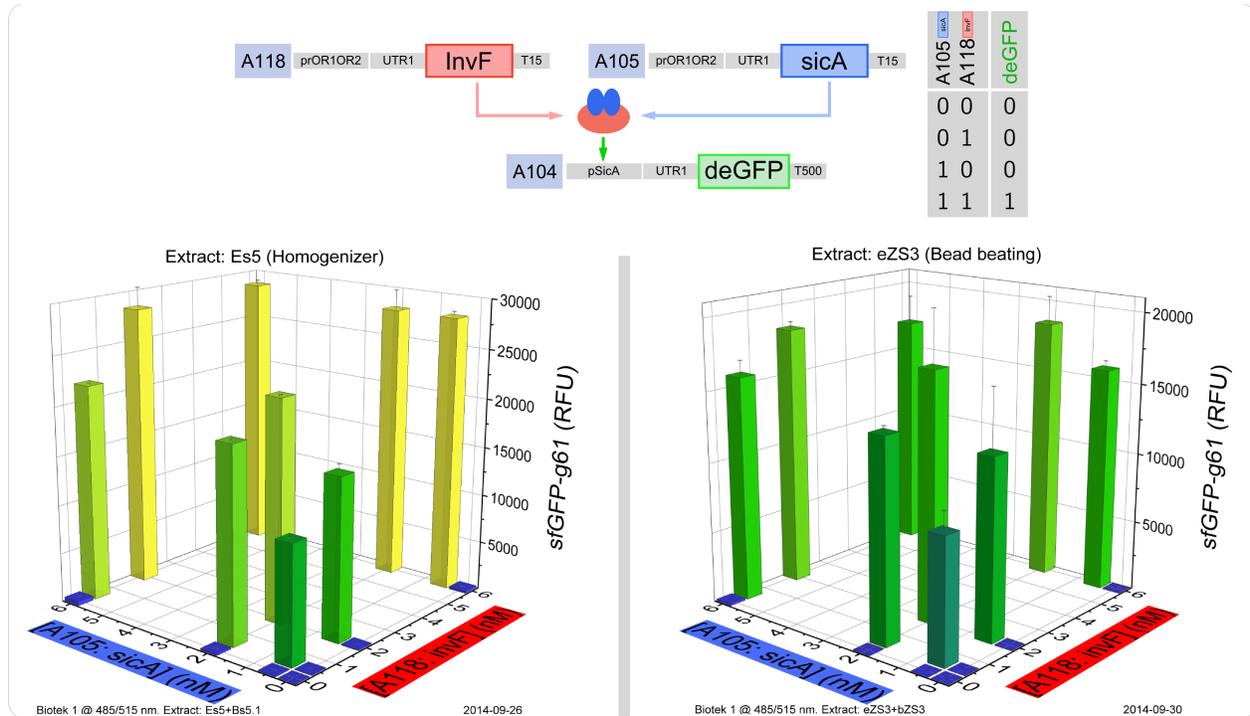


Figure 20: Prototyping of the pSicA AND-gate module using two different extracts: homogenizer extract Es5 (left) and bead-beating extract eZS3 (right). Reporter construct A104 was at 6 nM concentration. In Appendix 7.3.4, the same data is shown with rescaled Z-axes to show the level of the OFF-states.

3.4. Continuation of the project

Additional cloning and TX-TL testing has been done by researchers C. Hayes and T. Zhou to further prototype the 'Attenuator-controlled NOR-gate' before eventual *in vivo* implementation in *E.Coli*. As a first step, the antisense control of the AND-gate was tested again using a 'blank' antisense to compensate for DNA loading. This 'blank' was a sequence with the same length as the real antisense but without the ability to effect attenuator regions in any way. By keeping the total antisense DNA concentration constant at 40 nM, the true attenuation effect could be quantified. The results, shown in Figure 19, confirms that the attenuation works as suggested by the previously presented experiments (section 3.2.4).

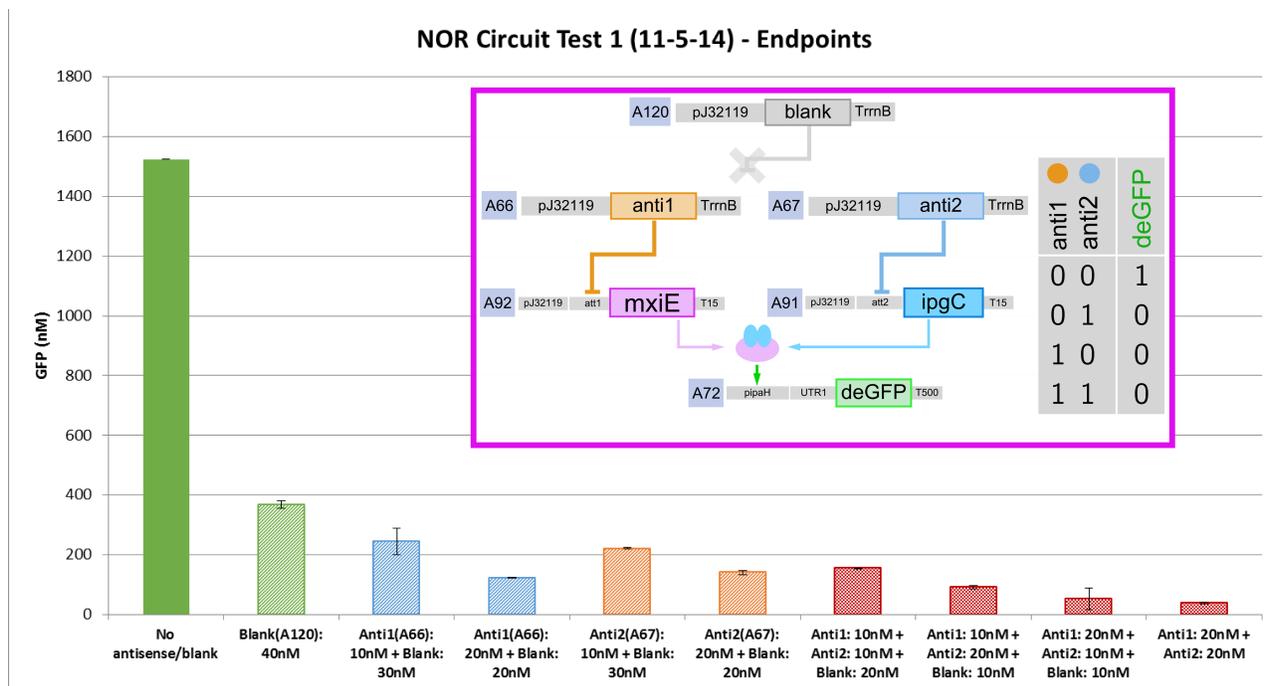


Figure 21: Attenuator control of the *pipaH* AND-gate tested using a blank antisense to distinguish loading effects from mRNA attenuation. Concentrations: A72 at 4nM, A91 and A92 at 2 nM each. Courtesy of T. Zhou and C. Hayes, 2014 (unpublished work)

3.5. Future work

To finalize the TX-TL prototyping of circuit design B, the complete circuit, further attempts to assemble the complete circuit in a TX-TL experiment could be made. However, it remains uncertain if such an experiment will give any further insight for design improvement since DNA loading effect will be very dominant due to the large total DNA concentration. On the other hand, there are extracts that are known to have increased expression and overall performance, suggesting that a sufficient signal-to-noise ratio could be obtained with lower DNA concentrations and thereby opening up the possibility to test circuits of this size.

Another approach to evaluate the circuit design would be to clone the circuit construct onto one or two plasmids and transform cells for *in vivo* testing of the complete circuit. This is generally a time consuming approach due to the several cloning steps and colony screenings required. The TX-TL prototyping has led up to some design improvements (e.g. swapping out the AND-gate module) and helped in tuning the DNA ratios for some of the constructs. An *in vivo* implementation of the circuit is therefore more likely to succeed when this new insight is used.

By confirming the function of the *pipaH* and *pSicA* modules *in vitro*, this work opens up the possibility to design new circuits for TX-TL testing. As proven by the implementation of the 'Four-Input AND-gate' by Moon *et al*, these modules can be used to sum the outputs of two sub-circuits. This type of sum function is not limited to the outputs of static sub-circuits, such as logic gates, but could also be applied to circuits with dynamic responses such as oscillators and pulse generators. Figure 22 illustrates how two oscillators (3-node repressilators) could be combined to form an interference pattern logic. When the two repressilators are out of phase the two AND-gate proteins are not abundantly present at the same time, resulting in no significant reporter activation. Eventually, circuits like these could allow for complex patterns of time-triggered events to be implemented in living cells and further push the boundaries of synthetic biology progress.

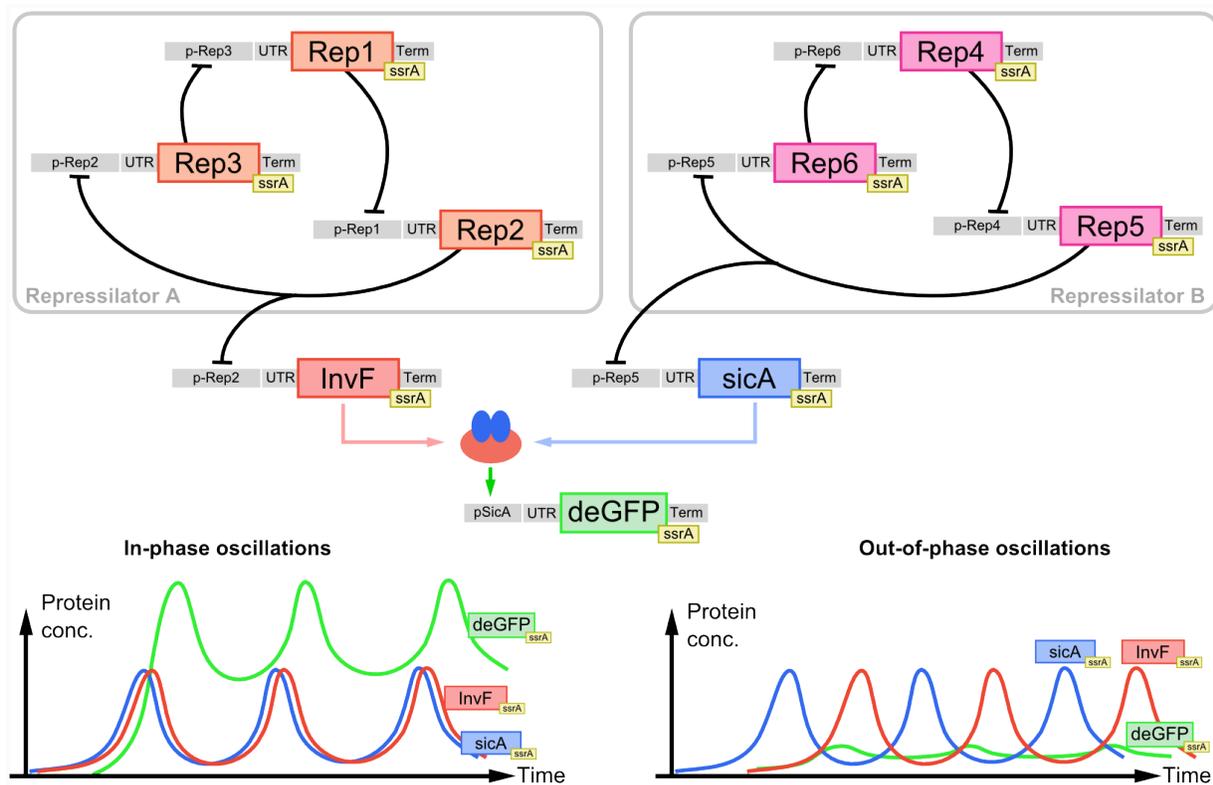


Figure 22: Schematic illustration of a potential dual-oscillator circuit. The output of two repressilators is combined in a pSicA AND-gate to form a circuit output which depends on the phase shift of the oscillators. *ssrA*-tags on each of the proteins assures fast degradation to achieve a dynamic response [17]. The purpose of this design is to show the potential use of AND-gates in future generation dynamic biocircuits.

4. Conclusions

During this project, three AND-gate modules, two attenuator-antisense pairs and several circuit motifs combining these elements have been prototyped in a cell-free TX-TL system. This has given valuable insight for design improvements and for biocircuit designs implementation in general.

The pGlnA promoter AND-gate testing showed an unexpected leaky signal even in the absence of sigma 54, NRI and NRII. Furthermore, the heavy resource loading of the TX-TL system made the module difficult to evaluate. For assembly of the complete circuit in TX-TL, this modules performance was deemed inadequate, but previous unpublished work show that the response could be better when working *in vivo*. Further testing would be needed to determine if pGlnA would be functional in this circuit design. As an alternative, the pipaH AND-gate module was tested, which showed excellent ON/OFF ratio and AND-gate logic, which made it a stronger candidate for further circuit development.

Attenuation by antisense mRNA was achieved to some extent, but a large amount (up to 10-fold concentration) of antisense DNA was needed to properly sequester the signal. When testing these parts, one of the limitations of using a cell-free system became apparent as the system became overloaded when more than approximately 30 nM of DNA was used in a single reaction. To quantify the effect of attenuator-antisense pairs, these loading effect needs to be compensated for. This is best done by introducing a 'blank antisense' which does not effect either of the attenuators but assures a constant DNA load in all experiments, making the attenuation easily measurable. Though this approach was not done in the first experiments, later work by T. Zhou *et al* (see section 3.4) confirmed the results of the initial tests.

In the published literature, no previous attempt of *in vitro* characterization of Moon *et al*'s engineered AND-gates modules, pipaH and pSicA, has been found. Some characterization has previously been done with the naturally occurring T3SS system components which has revealed their mechanism for gene regulation [36]. However, this work introduces the novelty of *in vitro* characterization of the engineered AND-gate variants in TX-TL, and thereby lays the foundation for rapid prototyping of new circuit design incorporating these modules.

In a broader sense, this thesis work can serve as a reference for continued development of biocircuits using the methodology of rapid prototyping using TX-TL in synergy with computer modeling. Regardless of biocircuit design, a rational approach to testing and trouble-shooting will be needed to minimize the time and cost of realization. Thus, the

rapid prototyping process used in this thesis might be a suitable strategy for future work in the field.

Designing moderately complex circuits, like the ones proposed in this project, will help build towards better control of engineered bacteria and other life forms, as well as give a better understanding of the processes governing gene expression and biocircuits within living cells. This is the spring board needed to accelerate biotechnology to new achievements within many different fields, such as medicine, biochemical production, drug delivery and gene therapy.

The field of synthetic biology has been advancing steadily over the past years, and with some truly impressive technical achievements, the foundation for advancement in the field has been laid, piece by piece. With automation replacing manual labor, and a growing library of enhanced genetic parts, it is reasonable to think that synthetic biology will continue to accelerate over the coming years. Naturally, the new capabilities raises important questions about the responsibility of manufacturers and labs who engineer GMOs. Perhaps one of the mostly widely debated issues in synthetic biology is the use of GMOs in agriculture and food production, where concerns have been raised about its impact on human health, food safety and ecosystems [42], [43]. As with any powerful technology, the risks of implementation has to be weighed against the benefits of progress. The power of biological systems can undoubtedly be used to solve pressing problems, as illustrated by the paper-based Ebola diagnostics tool (mentioned in section 1.3). Eventually, the extended use of biological systems in technological applications may bridge the gap between the *in vitro* and *in vivo* world which, if used wisely, can help us create a more sustainable future.

5. Acknowledgments

I would like to thank my mentor Professor Richard M. Murray at California Institute of Technology (Pasadena, CA, USA). Professor Murray has provided me with a unique research opportunity and excellent conditions for me to provide my contribution to the synthetic biology research at the 'Murray biocircuits lab', for which I am very grateful. My co-mentors Clare Hayes and Vipul Singhal, together with Professor Murray, have given me extended guidance in my research and invaluable support through-out this project. I would also like to thank my fellow SURF students, Tiffany Zhou and Pulkit Malik, for a good collaboration and the work we have done together in the lab. Last but not least, all the members of Murray biocircuits lab at Caltech have given me great support and advice in all aspects of this project; Thank you all.

At Lund University, I would like to thank Johan Svensson-Bonde at the Applied Biochemistry department, LTH, for mentoring me through my project planning and thesis writing.

6. References

- [1] Fredholm, Lotta, “The Discovery of the Molecular Structure of DNA - The Double Helix,” *Nobelprize.org. Nobel Media AB 2014.*, 09-Jan-2015. [Online]. Available: http://www.nobelprize.org/educational/medicine/dna_double_helix/readmore.html. [Accessed: 09-Jan-2015].
- [2] J. D. Watson and F. H. C. Crick, “A Structure for Deoxyribose Nucleic Acid.,” *Nature*, vol. 171, pp. 737–738, 1953.
- [3] E. Andrianantoandro, S. Basu, D. K. Karig, and R. Weiss, “Synthetic biology: new engineering rules for an emerging discipline,” *Mol. Syst. Biol.*, vol. 2, May 2006.
- [4] “What is synthetic biology?,” *Synthetic biology project*. [Online]. Available: <http://www.synbioproject.org/topics/synbio101/definition/>. [Accessed: 13-Jan-2015].
- [5] “About and FAQ,” *Syntheticbiology.org*. [Online]. Available: <http://syntheticbiology.org/>. [Accessed: 13-Jan-2015].
- [6] “Physiology or Medicine 1978 - Press Release,” *Nobelprize.org. Nobel Media AB*, 09-Jan-2015. [Online]. Available: http://www.nobelprize.org/nobel_prizes/medicine/laureates/1978/press.html. [Accessed: 09-Jan-2015].
- [7] V. Noireaux, R. Bar-Ziv, and A. Libchaber, “Principles of cell-free genetic circuit assembly,” *Proc. Natl. Acad. Sci.*, vol. 100, no. 22, pp. 12672–12677, Oct. 2003.
- [8] D. Siegal-Gaskins, V. Noireaux, and R. M. Murray, “Biomolecular resource utilization in elementary cell-free gene circuits,” in *American Control Conference (ACC), 2013*, 2013, pp. 1531–1536.
- [9] B. Wang, R. I. Kitney, N. Joly, and M. Buck, “Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology,” *Nat. Commun.*, vol. 2, p. 508, Oct. 2011.
- [10] A. Levskaya, “Engineering Escherichia colito see light,” *NATURE*, vol. 438, p. 24, 2005.
- [11] S. S. Golden, “Light-responsive gene expression in cyanobacteria.,” *J. Bacteriol.*, vol. 177, no. 7, p. 1651, 1995.
- [12] L. H. Hansen, S. Knudsen, and S. J. Sørensen, “The Effect of the lacY Gene on the Induction of IPTG Inducible Promoters, Studied in Escherichia coli and Pseudomonas fluorescens,” *Curr. Microbiol.*, vol. 36, no. 6, pp. 341–347, Jun. 1998.
- [13] T. S. Moon, C. Lou, A. Tamsir, B. C. Stanton, and C. A. Voigt, “Genetic programs constructed from layered logic gates in single cells,” *Nature*, vol. 491, no. 7423, pp. 249–253, Oct. 2012.
- [14] D. M. Chudakov, M. V. Matz, S. Lukyanov, and K. A. Lukyanov, “Fluorescent

- Proteins and Their Applications in Imaging Living Cells and Tissues,” *Physiol. Rev.*, vol. 90, no. 3, pp. 1103–1163, Jul. 2010.
- [15] J. S. Paige, K. Y. Wu, and S. R. Jaffrey, “RNA Mimics of Green Fluorescent Protein,” *Science*, vol. 333, no. 6042, pp. 642–646, Jul. 2011.
- [16] L. H. Naylor, “Reporter gene technology: the future looks bright,” *Biochem. Pharmacol.*, vol. 58, no. 5, pp. 749–757, 1999.
- [17] H. Niederholtmeyer, V. Stepanova, and S. J. Maerkl, “Implementation of cell-free biological networks at steady state,” *Proc. Natl. Acad. Sci.*, vol. 110, no. 40, pp. 15985–15990, Oct. 2013.
- [18] Y. Benenson, T. Paz-Elizur, R. Adar, E. Keinan, Z. Livneh, and E. Shapiro, “Programmable and autonomous computing machine made of biomolecules,” *Nature*, vol. 414, no. 6862, pp. 430–434, Nov. 2001.
- [19] N. Lapique and Y. Benenson, “Digital switching in a biosensor circuit via programmable timing of gene availability,” *Nat. Chem. Biol.*, vol. 10, no. 12, pp. 1020–1027, Oct. 2014.
- [20] L. Prochazka, B. Angelici, B. Haefliger, and Y. Benenson, “Highly modular bow-tie gene circuits with programmable dynamic behaviour,” *Nat. Commun.*, vol. 5, p. 4729, Oct. 2014.
- [21] Y. T. Maeda, T. Nakadai, J. Shin, K. Uryu, V. Noireaux, and A. Libchaber, “Assembly of MreB Filaments on Liposome Membranes: A Synthetic Biology Approach,” *ACS Synth. Biol.*, vol. 1, no. 2, pp. 53–59, Feb. 2012.
- [22] J. Shin, P. Jardine, and V. Noireaux, “Genome Replication, Synthesis, and Assembly of the Bacteriophage T7 in a Single Cell-Free Reaction,” *ACS Synth. Biol.*, vol. 1, no. 9, pp. 408–413, Sep. 2012.
- [23] K. Weintraub, “Detecting Ebola with a Piece of Paper,” *MIT Technology Review*, 24-Oct-2014. [Online]. Available: <http://www.technologyreview.com/news/532016/synthetic-biologists-create-paper-based-diagnostic-for-ebola/>. [Accessed: 15-Dec-2014].
- [24] K. Pardee, A. A. Green, T. Ferrante, D. E. Cameron, A. DaleyKeyser, P. Yin, and J. J. Collins, “Paper-Based Synthetic Gene Networks,” *Cell*, vol. 159, no. 4, pp. 940–954, Nov. 2014.
- [25] M. K. Takahashi, J. Chappell, C. A. Hayes, Z. Z. Sun, J. Kim, V. Singhal, K. J. Spring, S. Al-Khabouri, C. P. Fall, V. Noireaux, R. M. Murray, and J. B. Lucks, “Rapidly Characterizing the Fast Dynamics of RNA Genetic Circuitry with Cell-Free Transcription–Translation (TX-TL) Systems,” *ACS Synth. Biol.*, 2014.
- [26] Z. Z. Sun, E. Yeung, C. A. Hayes, V. Noireaux, and R. M. Murray, “Linear DNA for Rapid Prototyping of Synthetic Biological Circuits in an *Escherichia coli* Based TX-TL Cell-Free System,” *ACS Synth. Biol.*, vol. 3, no. 6, pp. 387–397, Jun. 2014.
- [27] A. Sarrion-Perdigones, E. E. Falconi, S. I. Zandalinas, P. Juárez, A. Fernández-del-

- Carmen, A. Granell, and D. Orzaez, “GoldenBraid: An Iterative Cloning System for Standardized Assembly of Reusable Genetic Modules,” *PLoS ONE*, vol. 6, no. 7, p. e21622, Jul. 2011.
- [28] V. K. Mutalik, J. C. Guimaraes, G. Cambray, C. Lam, M. J. Christoffersen, Q.-A. Mai, A. B. Tran, M. Paull, J. D. Keasling, A. P. Arkin, and D. Endy, “Precise and reliable gene expression via standard transcription and translation initiation elements,” *Nat. Methods*, vol. 10, no. 4, pp. 354–360, Mar. 2013.
- [29] Z. A. Tusa, V. Singhal, J. Kim, and R. M. Murray, “An in silico modeling toolbox for rapid prototyping of circuits in a biomolecular ‘Breadboard’ system,” in *2013 Conference on Decision and Control*, 2013.
- [30] “Biomolecular Breadboards: Models,” *OpenWetWare: Biomolecular Breadboards: Models*, 02-Dec-2014. [Online]. Available: http://www.openwetware.org/wiki/Biomolecular_Breadboards:Models. [Accessed: 02-Dec-2014].
- [31] E. Karzbrun, J. Shin, R. H. Bar-Ziv, and V. Noireaux, “Coarse-Grained Dynamics of Protein Synthesis in a Cell-Free System,” *Phys. Rev. Lett.*, vol. 106, no. 4, Jan. 2011.
- [32] B. Magasanik, “The regulation of nitrogen utilization in enteric bacteria,” *J. Cell. Biochem.*, vol. 51, no. 1, pp. 34–40, 1993.
- [33] V. Weiss and B. Magasanik, “Phosphorylation of nitrogen regulator I (NRI) of *Escherichia coli*,” *Proc. Natl. Acad. Sci.*, vol. 85, no. 23, pp. 8919–8923, 1988.
- [34] J. B. Lucks, L. Qi, V. K. Mutalik, D. Wang, and A. P. Arkin, “Versatile RNA-sensing transcriptional regulators for engineering genetic networks,” *Proc. Natl. Acad. Sci.*, vol. 108, no. 21, pp. 8617–8622, 2011.
- [35] M. Lunelli, R. K. Lokareddy, A. Zychlinsky, and M. Kolbe, “IpaB–IpgC interaction defines binding motif for type III secretion translocator,” *Proc. Natl. Acad. Sci.*, vol. 106, no. 24, pp. 9661–9666, Jun. 2009.
- [36] K. H. Darwin and V. L. Miller, “Type III secretion chaperone-dependent regulation: activation of virulence genes by SicA and InvF in *Salmonella typhimurium*,” *EMBO J.*, vol. 20, no. 8, pp. 1850–1862, Apr. 2001.
- [37] Z. Z. Sun, C. A. Hayes, J. Shin, F. Caschera, R. M. Murray, and V. Noireaux, “Protocols for Implementing an *Escherichia coli* Based TX-TL Cell-Free Expression System for Synthetic Biology,” *J. Vis. Exp.*, no. 79, Sep. 2013.
- [38] New England Biolabs Inc., “Protocol for Phusion® Hot Start Flex 2X Master Mix.” [Online]. Available: <https://www.neb.com/protocols/2012/08/27/protocol-for-phusion-hot-start-flex-2x-master-mix-m0536>. [Accessed: 03-Feb-2015].
- [39] “Biomolecular Breadboards:Protocols.” [Online]. Available: http://www.openwetware.org/wiki/Biomolecular_Breadboards:Protocols. [Accessed: 13-Jan-2015].
- [40] “BioTek Synergy H1 Multi-Mode Reader.” [Online]. Available:

http://www.biotek.com/products/microplate_detection/synergyh1_hybrid_multimode_microplate_reader.html. [Accessed: 02-Dec-2014].

- [41] J. Shin and V. Noireaux, “An E. coli Cell-Free Expression Toolbox: Application to Synthetic Gene Circuits and Artificial Cells,” *ACS Synth. Biol.*, vol. 1, no. 1, pp. 29–41, 2011.
- [42] “The Truth about Genetically Modified Food.” [Online]. Available: <http://www.scientificamerican.com/article/the-truth-about-genetically-modified-food/>. [Accessed: 03-Feb-2015].
- [43] E. Waltz, “GM crops: Battlefield,” *Nature*, vol. 461, no. 7260, pp. 27–32, 2009.
- [44] S. Ehrt, “Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor,” *Nucleic Acids Res.*, vol. 33, no. 2, pp. e21–e21, Jan. 2005.
- [45] “Escherichia coli/Codon usage,” *OpenWetware.org*. [Online]. Available: http://openwetware.org/wiki/Escherichia_coli/Codon_usage. [Accessed: 12-Jan-2015].
- [46] W. C. van Heeswijk, H. V. Westerhoff, and F. C. Boogerd, “Nitrogen Assimilation in Escherichia coli: Putting Molecular Data into a Systems Perspective,” *Microbiol. Mol. Biol. Rev.*, vol. 77, no. 4, pp. 628–695, Dec. 2013.

7. Appendicies

7.1. Glossary

Glossary entries are provided for words marked in *italics* on their first appearance in the text.

activator – a protein or protein complex which activates the expression from a construct by allowing RNAP to bind to the promoter. Some activators require an inducer molecule to function.

anhydrotetracycline (aTc) – an inducer molecule which bind the TetR repressor and through a conformational change drastically lowers its binding affinity to the Tet operon on promoters such as pTet. Due to its conjugated system the molecule exhibits some fluorescence [44]. MW = 462.88 g/mol. 1 µg/ml = 2.16 µM.

constitutive promoter – a promoter which is always activated, resulting in continuous transcription of the downstream DNA.

cross-talk – unwanted transfer of signals between communication channels. In this context; unwanted interference, due to lack of orthogonality, between a biocircuit and other systems such as the regulatory functions of a transformed host cell.

GFP – Green fluorescent protein, a protein which fluoresces in the green light spectra. Different versions are available, such as **deGFP** (destabilized enhanced GFP) and

sfGFP (superfolder GFP) used in this project. In TX-TL experiments, sfGFP result in a much higher fluorescence than deGFP.

inducer – a molecule that activated gene expression, by either disabling a repressor or inducing an activator.

in silico – term used to describe an experiment of process conducted or produced by means of computer modeling or computer simulation.

in vitro – term used to describe a process performed or taking place in a test tube, culture dish, or elsewhere outside a living organism.

in vivo – term used to describe a process taking place in a living organism.

Isopropyl-b-D-thiogalactopyranoside (IPTG) – a chemical mimic of lactose which acts as an inducer for LacI repressor. By binding to the LacI dimer the complex undergoes a conformation change which makes it unable to bind to the Lac operon on promoters such as pLac and pTac [12].

orthogonality – in this context, orthogonality means that any given set of modules do not overlap in their regulatory functions or interfere with one another. E.g. the repressors TetR-pTet and LacI-pLac can be said to be orthogonal since TetR protein cannot repress the pLac promoter, and vice versa, and because the two modules can function along side one another without any significant interactions disturbing the logic.

positive control – in this context, positive control refers to a construct known to constitutively express GFP and can be used as a reference to determine if an experiment was prepared correctly.

reporter construct (reporter) – a construct which expresses a marker, such as GFP, which makes the output measurable by available methods. Reporters are used as a way of quantifying gene expression, and provides the researcher with a means to “see” what genes are expressing under certain conditions.

repressor – a protein or protein complex which stops the expression from a construct by hindering RNAP from binding to the promoter. Repressors used in this project are TetR and LacI.

restriction digest enzyme – Enzymes which have the property of cleaving DNA molecules at or near specific sequences of bases. Enzymes used in this project are BsaI, BbsI and DpnI.

RFP – Red fluorescent protein. (see GFP)

RFU – Relative Fluorescence Units, the measuring unit given by a plate-reader when measuring fluorescence (e.g. from fluorescent reporters). Can be converted to nM units with sufficient calibration data.

transcription factor – a protein or protein complex which effects the expression of a gene. Either a repressor or an activator.

type-III secretion systems (T3SS) – an invasion system used by pathogenic bacteria to

secrete effector proteins into host the cells of host organisms [36]. These T3SS contain a large number of proteins which can be used in other contexts to build biocircuits [13].

7.2. Protocols and methods

7.2.1. Golden Braid reaction protocol

Golden Braid (GB) reactions were carried out in accordance with the following protocol. For each reaction the following reactants were added to a 500 μ l tube:

- 1 μ l NEB BsaI-HF (restriction digest enzyme)
- 1 μ l NEB T4 DNA Ligase
- 1.5 μ l 10X T4 DNA Ligase Buffer
- 1.5 μ l 10X BSA protein

50 ng of vector backbone, along with equimolar amount of the inserts (P,U,C and T) were added to the tube. Nuclease free water was added to a total volume of 30 μ l. The mixture was spun down in a table-top spinner, flicked to mix and spun down again. A 'Bio-Rad C1000 Touch Thermal Cycler' was used to control the reaction temperature during digest and ligation.

7.3. Additional results

7.3.1. Model of PipaH AND-gate

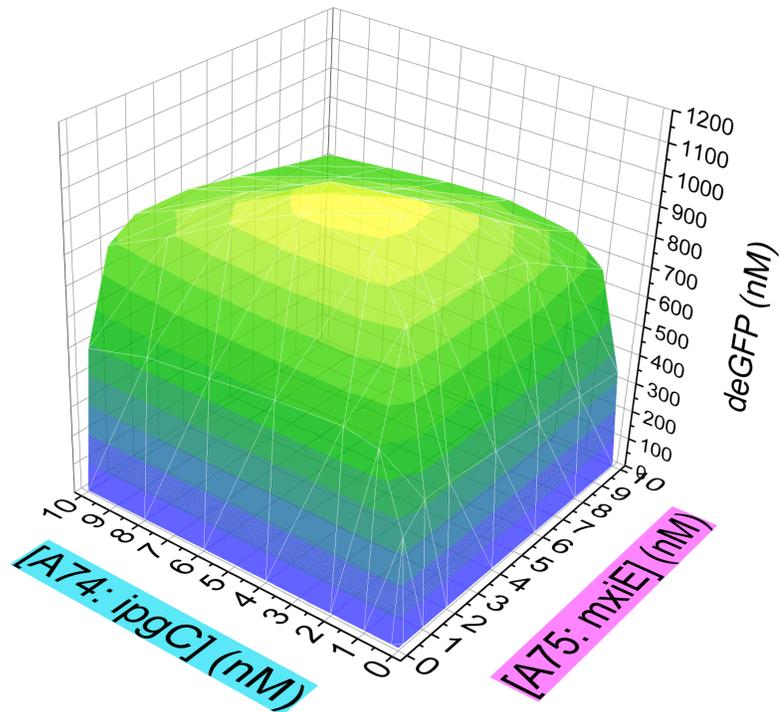


Figure 23: Surface plot of the modeled pipaH AND-gate system clearly displays a peak expression around 2-4 nM of each of the two DNA constructs.

7.3.2. Full circuit TX-TL assembly

An attempt was made at assembling the full circuit in TX-TL, as shown in Figure 22.

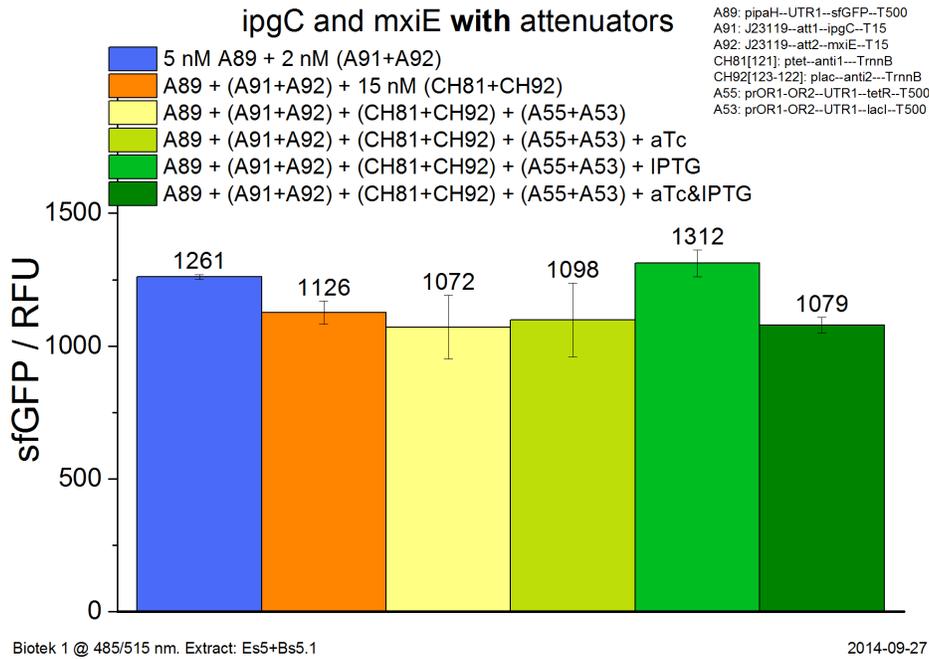


Figure 24: Full circuit assembled in a TX-TL experiment. The result does not follow the predicted circuit logic.

7.3.3. Complete circuit model results

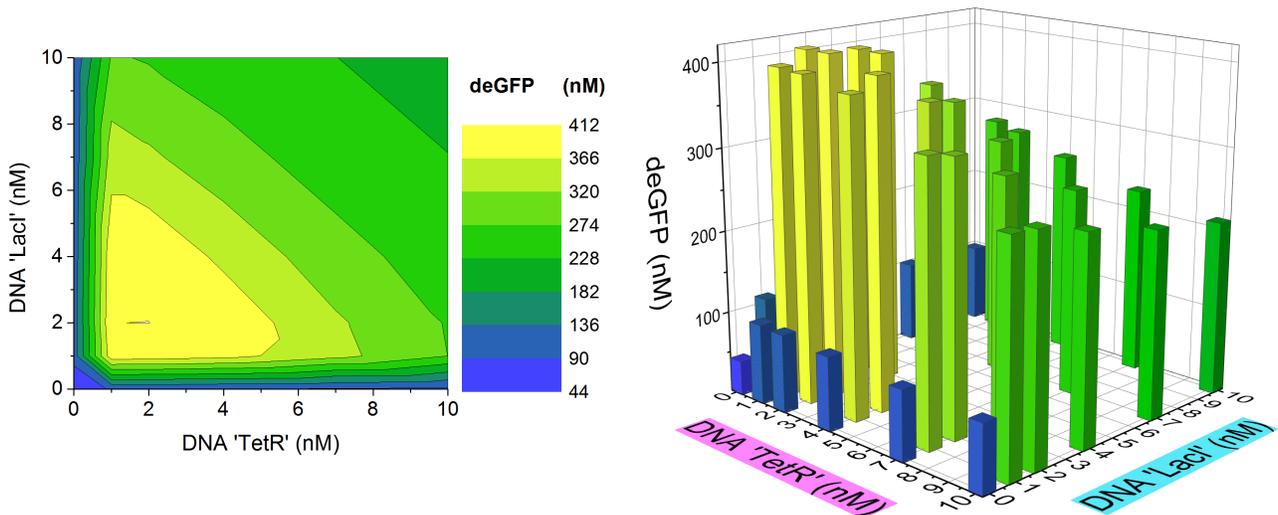


Figure 25: Modeling result of varying repressor concentrations, shown as heat map (left) and 3D bars from a different angle (right).

7.3.4. pSicA AND-gate

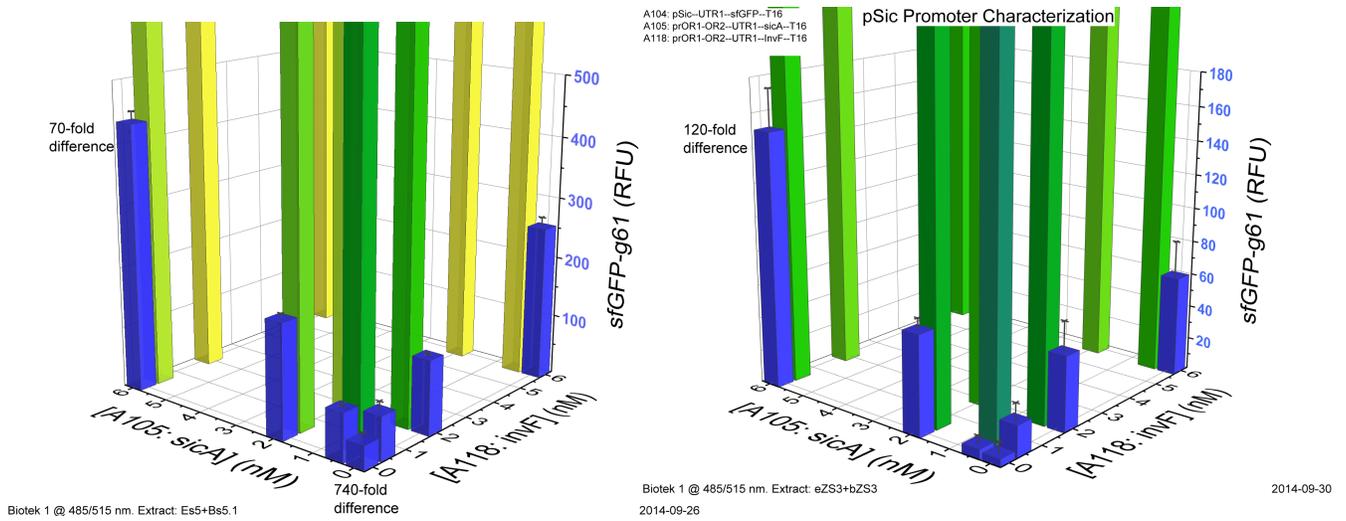


Figure 26: Data from pSicA AND-gate characterization rescaled to show the expression from the OFF-states of the logic. Left: Es5 homogenizer extract. Right: eZS3 bead-deating extract