University of West Bohemia
Faculty of Applied Sciences
Department of Cybernetics

MASTER'S THESIS

Systematic control of metabolite availability using models and methods of synthetic biology

Pilsen, 2014 **Pavel Zach**

Prohlášení

Předkládám tímto k posouzení a obhajobě diplomovou práci zpracovanou na závěr studia na Fakultě aplikovaných věd Západočeské univerzity v Plzni.

Prohlašuji, že jsem diplomovou práci vypracoval samostatně a výhradně s použitím odborné literatury a pramenů, jejichž úplný seznam je její součástí.

V Plzni dne 16. května 2014 Pavel Zach

Declaration

I hereby declare that this master's thesis is completely my own work and that I used only the cited sources.

Acknowledgements

Here I would like to express my gratitude to all the persons who helped me during the writing of this thesis.

I cannot find words to express my gratitude to Daniel Georgiev, PhD., my great advisor. Without his guidance, persistent help, patience and enthusiasm this thesis would not have been possible. I consider it an honor to work with Prof. Jesús Picó and other great people at Universitat Politècnica de València. Our discussions have given rise to many bright ideas. I would also like to thank Guillermo Rodrigo, PhD., for being at the right time in the right place and for giving me the opportunity to collaborate on his project.

Finally, I wish to thank my dearest ones. My family for their unlimited support during all the years of my study. Special thanks go to my beloved girlfriend, for her tolerance of my false promises that I will come from the lab soon and encouragement in all difficult situations I've gone through so far.

Anotace

Cílem této práce je návrh univerzální programovatelné platformy pro regulaci operonů s využitím měření hladin metabolitů. V první části je uveden přehled mechanismů, které využívají bakterie, společně s alternativami ze syntetické biologie. Jednotlivé role těchto mechanismů jsou vyšetřeny simulacemi fyzikálně relevantních matematických modelů. Ve druhé části jsou uvedeny výhody využití měření metabolitů. Třetí část je pak věnována systémovému návrhu syntetických částí využívajících simultánní transkripce-translace v prokaryotech a měření hladiny metabolitů. V poslední části této práce je pro tyto části vytvořen a validován obecný model.

Klíčová slova: syntetická biologie, regulace operonu, simultánní transkripce a translace, RNA regulační elementy, Escherichia coli

Abstract

The goal of this thesis is to design a versatile and programmable operon regulation platform using metabolic intermediate measurements. In the first part, the regulatory mechanisms implemented by bacteria are reviewed, together with the synthetic biology design alternatives. The roles of these mechanisms are investigated using simulations of physically relevant mathematical models. In the second part, the advantages of using the metabolic measurements are shown. The third part is dedicated to systematic design of transcriptional-translational coupling devices that regulate operon expression using metabolic measurements. In the last section, a corresponding general model is developed and validated.

Keywords: synthetic biology, operon regulation, transcriptional-translational coupling, RNA regulatory elements, Escherichia coli

Contents

1 Introduction

Utilization of cellular organisms for production of valuable compounds became reality after the invention of different techniques for DNA synthesis and DNA transfection into living cells. These methods, such as restriction digest for cutting DNA at specific places, DNA ligation for joining of different DNA fragments, and cell transfection for introduction of new DNA into cells have allowed genetic engineering to become reality.

In 1977, Genentech, Inc., reported the production of the first human protein made in bacteria, using a synthetic recombinant gene for the first time. Only a year later, in 1978, the same company announced a successful production of human insulin in a bacterium *Escherichia coli* [1]. In the treatment of diabetes, this synthesized insulin replaced animal insulin, which was more difficult to obtain, therefore more costly and which was causing undesired reactions in some patients. This key moment has started a new age of biotechnology.

1.1 Genetic engineering

After these first accomplishments in genetic engineering, the potential of using multiple genes to create new pathways for production of different compounds from various substrates was soon discovered. With an increase in complexity, the need for a better understanding of these networks emerged. This has been shown by several cases where alterations, improving the product formation (such as increasing the activity of product-forming enzymes) have only resulted in small improvements in the product yields. These results could be expected, since most of the changes that decrease cell fitness often cause the cell's regulatory network to divert resources back to processes contributing to cellular fitness [2]. Since almost every process in the cell is strictly controlled and regulated, every disturbance in this balance can have severe consequences.

Some methods to at least partially overcome these difficulties were developed, such as Classical Strain Improvement (CSI). This method relies on random mutagenesis causing genomic alterations, followed by an extensive screening in order to identify phenotypes with desired improvements. Although this approach has been widely used [3], mainly by industry, it has some major disadvantages [2]. First, the information why the strain is improved is not known (in some cases, this information can be provided by sequencing of the altered genome). Second,

the majority of introduced mutations are deleterious and therefore less efficient. Third, this method doesn't use the potential of using specialized genes from different organisms. These drawbacks became the subject of a new field called metabolic engineering.

1.2 Metabolic engineering

Metabolic engineering aims at intelligent design, based on an extensive knowledge of the underlying metabolic network (i.e. the set of chemical reactions leading to the formation of a product, as well as the genes coding for enzymes catalyzing these reactions). This is often achieved by development of relevant mathematical models and their analysis, pointing out parts of the network and its surroundings that constrain the production of a target metabolite. The predictions from these models are often used to find the optimal trade-offs between the cell's ability to produce desired substance and its ability to survive.

In order to achieve the optimal trade-off, not only the single genes, but also the cell regulatory networks need to be targeted and engineered.

1.3 Synthetic biology

As metabolic engineering added a systematic approach and analysis to genetic engineering, synthetic biology introduced another level to engineering of genetic circuits and metabolic pathways. It aims to apply engineering principles, such as standardization, modularity and reusability of all the parts used, ultimately creating libraries of standardized well-characterized parts involved in the cell regulation.

Its focus is on creating genetic circuits or metabolic pathways *de novo*, exploring new options and new designs which can't be found in nature. This means an extensive need for computational modeling, accompanied by understanding of all the biological processes involved and the ability to synthesize new DNA parts at a reasonable price.

1.4 Biotechnology industry

After more than 40 years of using methods of genetic and metabolic engineering in development of commercial products, synthetic biology has begun to spread

into commercial applications. One of the first industrial enterprises, which used synthetic biology methods in their production, is the life sciences and materials company DSM in optimizing its process of synthetic antibiotic cephalexin production [4]. Although the number of biotechnology companies which have used synthetic biology techniques to speed up and improve their research is growing, full potential has yet to be realized.

1.5 Thesis objective

Joint regulation of multiple genes is realised by grouping the genes into clusters called operons. At present, this is not easily programmable as it relies on existing regulatory pathways. The general objective of this thesis is to develop a general programmable method for operon regulation using metabolic intermediate measurements. The specific objectives are:

1. develop a model of the regulatory mechanism,
2. propose a set of validation experiments for the model and for the developed regulatory mechanism,
3. validate the developed model.

The thesis is organized as follows: In Section 2, an overview of a model system (*trp* operon), its regulatory mechanisms and design alternatives is presented. In Section 3, different loop feedback designs and different controller inputs are compared. In Section 4, a regulatory mechanism based on transcriptional-translational coupling and its model is developed. In Section 5, the proposed model is validated. Discussion and future work can be found in Section 6.

2 Metabolic pathway for the biosynthesis of L-tryptophan

Bacteria use operons for the joint expression of enzymes catalyzing production of essential metabolites. In this section the regulatory mechanisms implemented by bacteria are reviewed. An excellent motivational example is the biosynthesis of amino acid L-tryptophan in the bacterium *E. coli*, which was selected for the following reasons:

1. It has been extensively studied, both experimentally [5] and mathematically [6].

2. It features three independent negative feedback loops (see Section 2.1). This suggests that regulation is very important.

3. Since L-tryptophan is essential for the cell, but its synthesis is expensive in terms of energy and nutrient requirements, these control mechanisms and their structure have supposedly evolved to maintain proper intracellular tryptophan concentration under various conditions [7] - therefore we expect this pathway to be robustly regulated.

4. There is a need to jointly regulate expression of different genes coding for the enzymes, as evidenced by the grouping of pathway genes in a single operon[1].

Below we will discuss the details of this pathway related to regulation, the existing mathematical models of this pathway, and the design alternatives.

2.1 Biology and regulation

The tryptophan biosynthetic pathway consists of multiple steps, which are catalyzed by the enzymes encoded in the *trp* operon (Fig. 1).

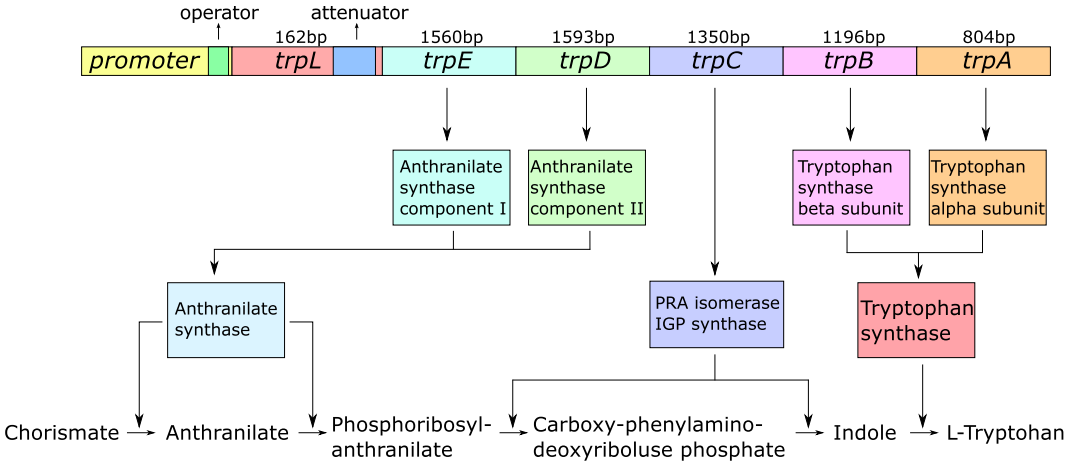

Figure 1: Scheme of the *trp* operon and tryptohan biosynthesis pathway.

[1]Operon can be defined as a group of genes, functionally related, being physically close to each other and under one common control mechanism [8].

Trp operon consists of five genes (*trpE*, *trpD*, *trpC*, *trpB* and *trpA*), which are transcribed together, from left (*trpE*) to right (*trpA*). This organization ensures that enzymes needed at the beginning of the metabolic pathway are expressed first.

Transcription of the operon is controlled by two different negative control mechanisms - transcriptional repression and attenuation. Also, third negative control mechanism is present in the form of allosteric enzymatic inhibition. A schematica of the corresponding nested control loops is shown in Fig. 2.

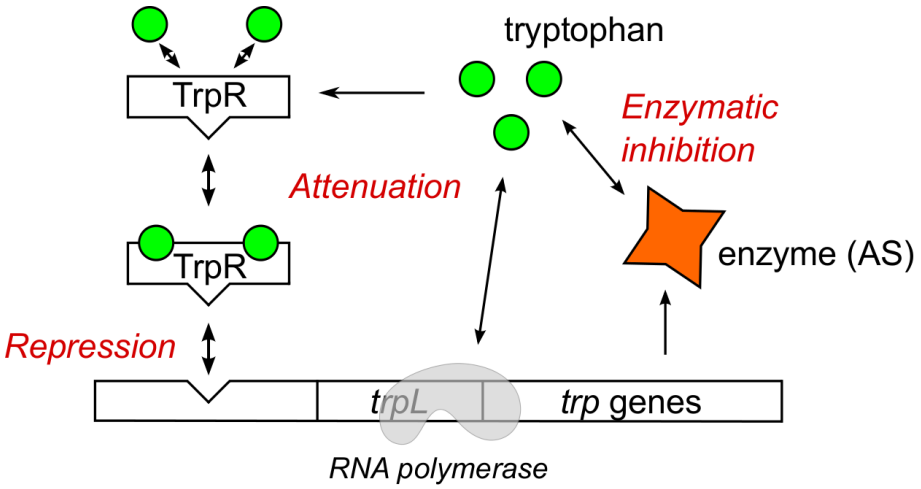

Figure 2: Control mechanisms in the *trp* operon.

2.1.1 Transcriptional repression

Transcriptional repression is mediated by a dimeric repressor protein TrpR, encoded by the *trpR* gene. This gene is not adjacent to the operon, and its transcription is controlled by a moderate promoter, also negatively regulated by TrpR [9].

When the tryptophan is present at low concentration in the cell, TrpR protein is in an inactive form and cannot bind to the operator site of the *trp* operon (see Fig. 3). This allows RNA polymerase to bind to the *trp* promoter and transcribe the genes. This elevates the levels of enzymes in the metabolic pathway and in turn increases the level of tryptophan in the cell. When the tryptophan concentration is sufficient, the excess tryptophan binds noncooperatively to the TrpR repressor (two molecules per one protein), causing allosteric change in the protein

conformation and thus allowing the protein to bind to the operator [10]. When the repressor protein is bound, the RNA polymerase cannot transcribe the genes, decreasing corresponding in enzyme levels and causing tryptophan production to decrease.

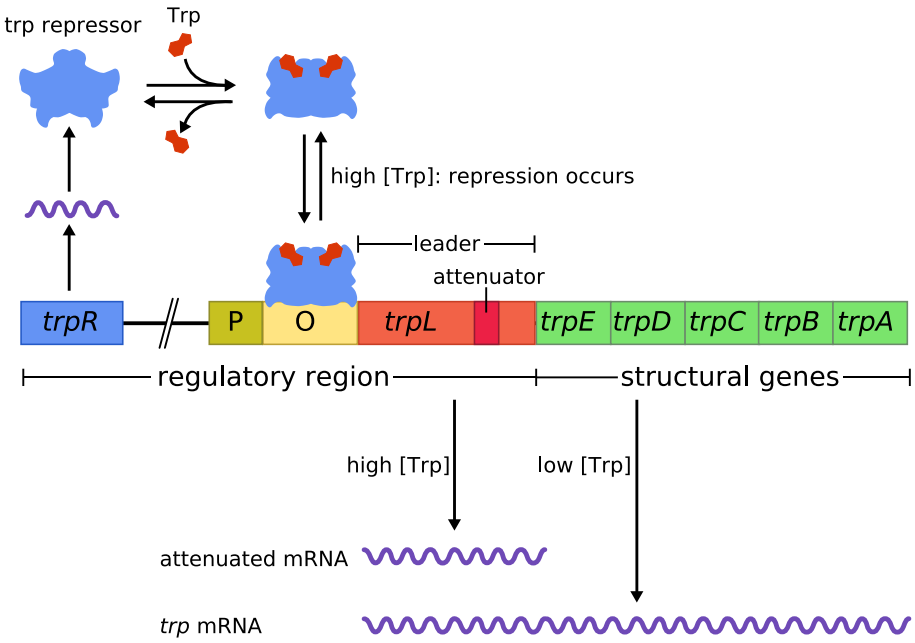

Figure 3: Repression mechanism in *trp* operon. *Source: Wikimedia Commons, file Trpoperon.svg*

This mechanism regulates *trp* operon expression over approximately a 100-fold range [11] and its strength depends on the intracellular concentration of tryptophan molecules [12].

Synthetic biology alternatives Negative feedback for the regulation of operon expression may be introduced by expressing a repressor protein from one of the genes in the operon. Specificity is determined by the fact that the produced repressor must be able to recognize and bind to a specific sequence (called the operator) in the promoter of the operon. This implicates that, when introducing negative feedback, a specific tandem of promoter and repressor protein must be used. Although there are some well characterized examples of these pairs commonly used in synthetic circuits (such as repressor protein LacI and promoter

pLac promoter [13]), their set is quite limited. These limitations can be overcame by the use of transcription activator-like effectors (TAL effectors or TALE).

TAL effectors were discovered in 2009 in *Xanthomonas* bacteria, which use them while infecting various plant species [14]. The purpose of these proteins is to modulate transcription of the host plant cells in order to promote bacterial infection. Main feature of these TALEs is their central repeat domain, consisting of typically 34 amino acid repeats.

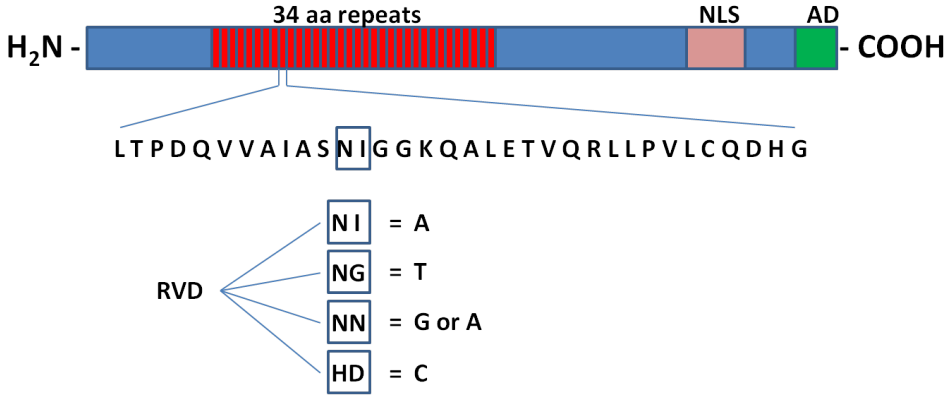

Figure 4: Scheme of a TAL effector. Amino acids NI in the central repeat recognize nucleotide adenin. *Source: Freiburg 2012 iGEM wiki*

This domain is able to recognize and bind specific DNA sequences, based on the identity of two critical amino acids in each of these repeats. Since there is a known relation between these two amino acids and recognized nucleotides (see Fig. 4), these proteins can be easily engineered in order to bind to a DNA sequence of our interest. As was shown by the Slovenian iGEM 2012 team [15] and many other research groups, these proteins can be also engineered to work as repressors and therefore can be used for introduction of a genetic negative feedback loop.

However, the use of general transcription factors has one disadvantage in comparison to regulation in the *trp* operon - the resulting feedback is not derived from the final metabolite, but from the level of enzymes produced by this operon.

2.1.2 Attenuation

Transcriptional attenuation is a mechanism for premature termination of transcription, based on the transcriptional-translational coupling. It has been found

in a number of operons in *E. coli* and other prokaryotes [10]. It consists of a relatively short sequence, located between promoter and the operon structural genes (in the 5' UTR).

In the *trp* operon, this sequence is a 162bp long leader region *trpL* (see Fig. 1) [5]. This leader region consists of four sequences, numbered 1 to 4 from left to right. Since each sequence is partially complementary to the next one, three different secondary structures (hairpins) can form on the transcribed RNA: 1-2, 2-3 and 3-4. However, the transcribed RNA can adopt only two conformations, called termination and anti-termination:

- **Termination conformation** - in this case, structure 3-4 (terminator hairpin preventing transcription) is formed. This is possible only when the sequence 3 is not paired with sequence 2, i.e. sequence 2 is paired to sequence 1 or sequence 2 is sequestered by a ribosome.
- **Anti-termination conformation** - in this case, termination structure 3-4 is not formed. This is possible only when complex 2-3 is formed, i.e. when sequence 2 is available and is not part of complex 1-2.

The final conformation depends on the intracellular concentration of tryptophan charged tRNA. It was discovered that the leader region contains an open reading frame (ORF)[2] [11], coding for a short polypeptide of 14 amino acids and termed the leader peptide [16].

The mechanism by which attenuation regulates transcription is shown in Fig 5. First, RNAP begins with a transcription of the leader region with sequences 1 and 2. Almost immediately after this sequence is transcribed, ribosome bounds and begins with the translation of the produced RNA. If the tryptophan concentration is low, ribosome stalls at sequence 1 with two consecutive Trp codons, blocking the formation of complex 1-2 and thus forming the anti-termination conformation (Fig. 5b).

Otherwise, if tryptophan concentration is high, the sequence 2 is sequestered instead of sequence 1, so the terminator conformation is formed (Fig. 5a), preventing the transcription of the remaining operon genes. This mechanism regulates *trp* operon expression over a 6 to 8-fold range [11].

[2]ORF denotes a region of nucleotides, bounded by the start and stop codon.

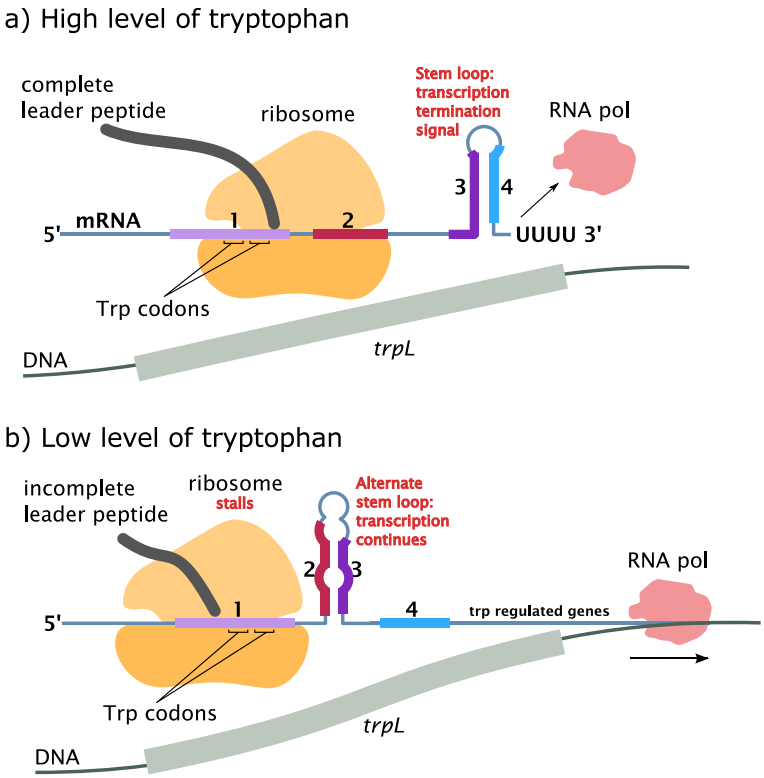

Figure 5: Attenuation mechanism in *trp* operon. *Source: Wikimedia Commons, file Trp operon attenuation.svg*

Synthetic biology alternatives The synthetic biology alternative to attenuators are transcriptional riboswitches. According to [17], transcriptional riboswitches can be considered as a special class of attenuators, although some authors consider riboswitches as a separate group. The main difference between attenuators and riboswitches is in the sensing RNA element. In attenuators, this part is a general element sensing a variety of inputs (e.g. proteins, temperature), while in riboswitches, this part is called an aptamer that is bound by different small molecules (called ligands). After the aptamer part, riboswitch consists of an expression platform that performs a specific action upon a ligand binding (see Fig. 6). This action can be the formation of an intrinsic terminator preventing transcription, or formation of an anti-terminator that enables transcription of the downstream genes.

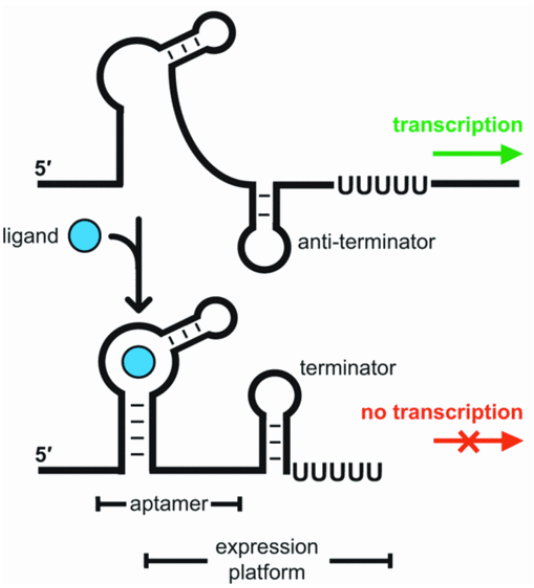

Figure 6: Mechanism of transcriptional riboswitches. In this case, upon a ligand binding to the aptamer, terminator is formed, thus transcription is stopped. Original image taken from [18].

The discovery and characterization of these RNA-based regulatory elements is dated to 2002, when the first comprehensive study was published [19]. Since then, many natural riboswitches were discovered and a lot of work has been done in the study of their mechanism.

The problem of engineering synthetic riboswitches can be generally[3] divided into two separate parts:

1. **Aptamer design** - practically all approaches to obtain a new aptamer are based on a method called SELEX (systematic evolution of ligands by exponential enrichment). This method involves generating of a large pool of randomized RNA sequences (up to 10^{14}), and their mixing with immobilized target ligands. After washing, the bound sequences are eluted, amplified by PCR[4] and subjected to further round of this selection process. After each iteration, aptamers with higher affinity (ability to bound the ligand) are found [20].

[3]In some cases, aptamer itself can work as a sensing and regulatory domain.
[4]Polymerase chain reaction.

2. **Expression platform design** - involves design of an RNA sequence, which alters its secondary structure upon a ligand binding to the aptamer. Therefore this sequence can be divided into two parts - module performing the desired change in its secondary structure, and communication module, which serves as an interlink between an aptamer and this module [21].

Unfortunately, there are some limitations in designing new riboswitches. First, aptamers obtained by the SELEX procedure are aptamers that work *in vitro*. However, they often show reduced or disabled functionality in live cells, mainly due to a different concentration of Mg ions. Also, a little success has been achieved in designing expression platforms for transcriptional riboswitches, due to more complicated interactions in the produced RNA molecule. However, this situation has began to change, since in 2013 some authors reported *de novo* design of a synthetic riboswitches that regulates transcription termination [22], [23].

2.1.3 Enzymatic inhibition

The third regulation mechanism in the biosynthesis of tryptophan does not influence the process of transcription of the enzymes involved, instead it targets their activity, specifically the activity of anthranilate synthase (AS), a heterotetramer formed by two TrpE and two TrpD polypeptides [12]. This enzyme catalyzes the first step of the pathway leading to the production of tryptophan. If the tryptophan concentration is high, it binds to the AS, causing allosteric inhibition of its activity and therefore decreasing its production.

Synthetic biology alternatives Although significant advances are being made in the field of protein engineering every year, to the author's knowledge, no methods enabling design of new enzymes, or rational redesign of existing enzymes (in order to introduce this kind of inhibition) exist. Current level of protein engineering in synthetic biology involves increasing of the activity, altering protein specificity and altering regulatory elements of proteins (mainly deleting allosteric inhibition by destroying the specific site by site directed mutagenesis [24] or by directed evolution) [25].

2.2 Mathematical modeling

Several deterministic models have been developed to understand the *trp* operon mechanisms of gene regulation. These models are based on different mathematical methods and different levels of simplification, but only a few of them have also considered attenuation (stating that the influence of attenuation is negligible; this argument is based on the results from [26]). These models are presented in Table 1.

Table 1: Overview of deterministic *trp* operon models

Year, Author	**Type**[a]	**Validated experimentally**[b]	**Reference**
2001 Santillán	DDE	Yes	[27]
2004 Santillán	DDE	Yes	[12]
2004 Venkatesh	ODE	Yes	[7]
2006 Bhartiya	ODE	Yes	[6]
2008 Nguyen	S-Systems	Yes	[10]

[a]DDE=delay differential equations; ODE=ordinary differential equations; for S-Systems, see [28].

[b]Interestingly, all listed models were validated against experimental results presented in [11], making a standard for *trp* operon modeling from this reference. However, as some authors pointed out, the time step in dynamic measurements is so large that the dynamics before reaching a steady-state is not well known and some models show considerable discrepancies between their initial dynamics and these data [29].

Although all models were experimentally validated (see Table 1 for details), some used parameters differed significantly across parameter sets. For example, the dissociation constant value of 810μM for enzymatic inhibition used in model by Venkatesh practically eliminated the influence of this negative feedback.

For this reason, authors of one of the most recent models have surveyed the existing models, using the most reasonable value of each parameter [30]. Their model, based on Bhartiya et al. [6], considers this set of reactions:

$$O_t \longrightarrow O_p \longrightarrow \emptyset, \tag{1}$$

$$O_p \longrightarrow mRNA \longrightarrow \emptyset, \tag{2}$$

$$mRNA \longrightarrow Enz \longrightarrow \emptyset, \tag{3}$$

$$Enz \longrightarrow Trp \longrightarrow \emptyset. \tag{4}$$

These equations stand for the production and loss (i.e. degradation and dilution due to cell growth) of free operator (1), production and loss of mRNA from the free operator (2), production and loss of anthranilate synthase from mRNA (3) and production and consumption of tryptophan (4). To expand the model capabilities, we added another reaction, representing tryptophan uptake from the environment:

$$\emptyset \longrightarrow Trp. \tag{5}$$

This set of reactions can be modeled using ODEs as follows:

$$\frac{dO_p}{dt} = k_1 \frac{K_1^{n_1}}{K_1^{n_1} + T^{n_1}} O_t - (k_{d1} + \mu) O_p, \tag{6}$$

$$\frac{dM}{dt} = k_2 \frac{K_2^{n_2}}{K_2^{n_2} + T^{n_2}} O_p - (k_{d2} + \mu) M, \tag{7}$$

$$\frac{dE}{dt} = k_3 M - (k_{d3} + \mu) E, \tag{8}$$

$$\frac{dT}{dt} = k_4 \frac{K_3^{n_3}}{K_3^{n_3} + T^{n_3}} E + d \frac{T_{ext}}{e + T_{ext}\left(1 + \frac{T}{f}\right)} - \left(\frac{g}{T + K_g} + \mu\right) T. \tag{9}$$

Values of all used parameters are listed in Appendix A. It is also important to note that because the level of $\mathrm{tRNA}^{\mathrm{Trp}}$ is dependent on the level of tryptophan in the cell, it is possible to use tryptophan concentration as controller inputs for all three types of regulation (i.e., transcriptional repression, attenuation, allosteric inhibition).

This model was validated by its authors, using experimental data from [11], both for wild-type cells and different mutant strains. Authors reported and demonstrated that this model confidently reflects the *trp* operon behavior.

3 Engineering operon regulation based on the *trp* operon

Some important questions arise from the structure of tryptophan system. Many engineering control strategies rely on sophisticated and intense computational algorithms to achieve robustness. But cellular organisms have to rely only on simple

chemical interactions. One of the questions, which many authors [6], [7] asked during the study of the *trp* operon, is why nature uses three different feedback loops instead of one. Is this design somehow advantageous? If so, could this motif be transferred to other biological or engineering applications?

3.1 Single versus nested loop feedback design

Bhartiya et al. answered this question by comparing two different designs of the *trp* operon system, conceptualizing the system as three processes in series (see Fig. 7). First design involved all three negative feedback loops (repression, attenuation and inhibition) while the second one used a single feedback loop strategy (repression). We will refer to these designs as nested feedback loop (NFL) and single feedback loop (SFL). Note different names are used in [7].

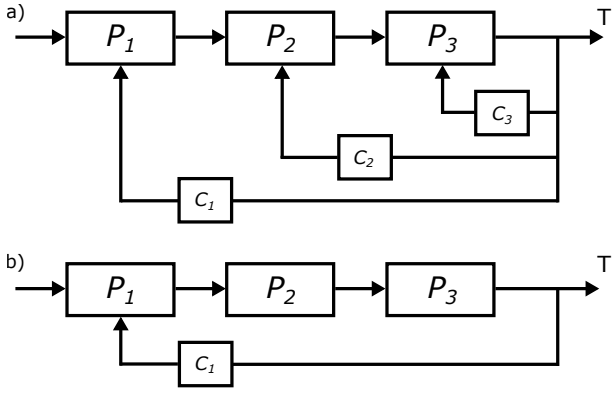

Figure 7: Schematic of designs used for the comparison of NFL and SFL. Processes P_1, P_2 and P_3 represent transcription, translation and tryptophan synthesis, respectively. a) NFL with genetic regulation (C_1), transcriptional attenuation (C_2) and enzyme inhibition (C_3). b) SFL with genetic regulation only.

Dynamic response of the tryptophan system must satisfy number of conditions. Tryptophan concentration must be maintained in a very narrow range. In addition, short rise time is needed, accompanied with short settling time. Generally, short rise time is associated with longer settling times and excessive overshoot. Assuming that tryptophan system has evolved for efficient cellular functioning, the authors in [6] introduced a simple metric to capture the deviances from the wild-type response:

$$I(p,s) = \frac{1}{\sqrt{t_f}} \sqrt{\int_0^{t_f} [T(t,p,s) - T^*(t)]^2 \, \mathrm{d}t}, \tag{10}$$

where $T^*(t)$ is the wild-type response at time t. Using this metric, the authors searched for optimal SFL parameter values realted to genetic regulation (i.e., cooperativity constant n_1 and dissociation constant K_1). Figure 8 shows the resulting plot showing the outputs (tryptophan concentration) of wild-type NFL and the optimized SFL design.

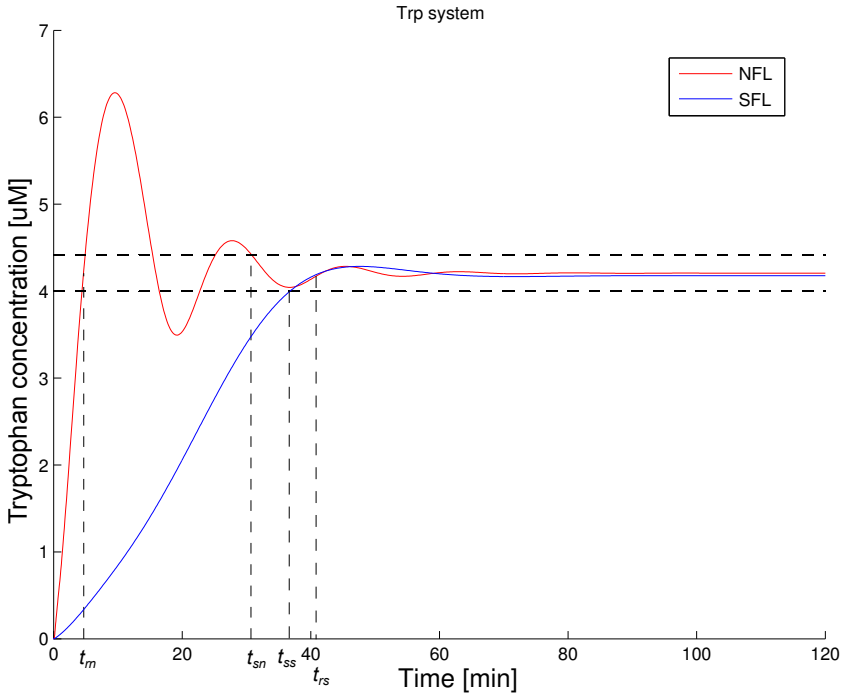

Figure 8: Comparison of NFL and SFL of tryptophan system, as shown in [6]. Even with the altered parameters of genetic regulation, the SFL shows significantly longer rise (t_{rs}) and settling times (t_{ss}) in comparison to NFL.

Even an optimally tuned SFL with the lowest value of $I(p,s)$ cannot achieve required response characteristics - the SFL rise time value of over 40 min is not sufficient for the cell in order to survive.

This comparison was based on the model published by Venkatesh et al. [7], using the original set of parameters. In addition, one of the two parameters authors varied (n_1) is practically impossible to tune in real systems. Therefore,

we performed a new comparison using the model and the parameter set from Section 2.2 and varying instead the rate of transcription k_1 and the dissociation constant K_1. These two parameters are experimentally tunable by changing the DNA sequence of the promoter. The results are shown in Fig. 9, where NFL is a model of wild-type tryptophan system based from Section 2.2 and SFL represents the tryptophan system with only transcriptional repression feedback, with altered parameters k_1 and K_1 in order to achieve minimal value of the metric $I(p, s)$. Over 500 000 different values of k_1, each with 500 different values of K_1 were evaluated in this search.

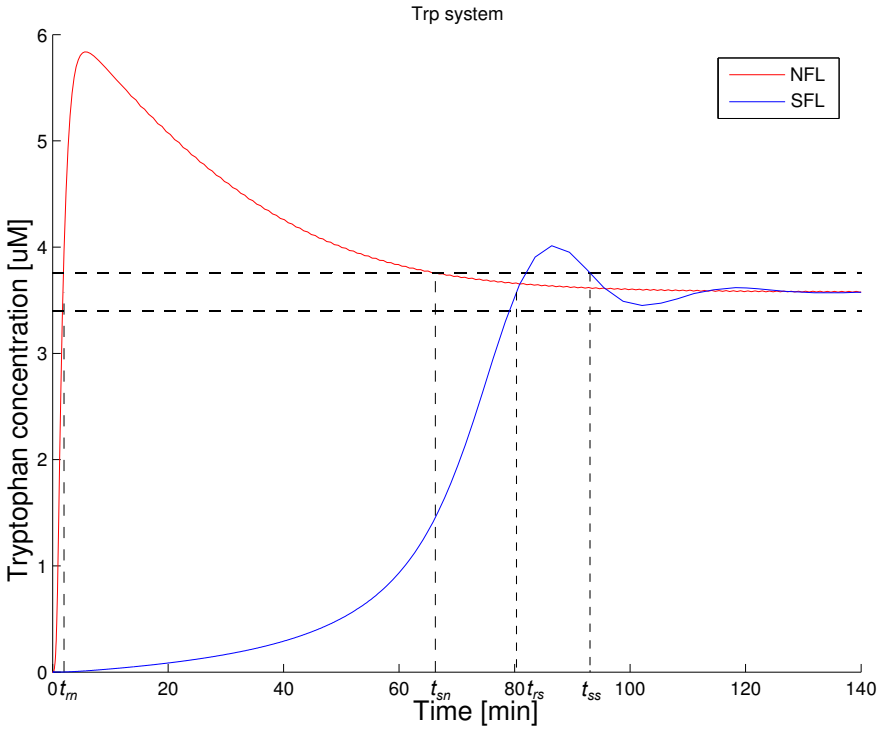

Figure 9: Comparison of NFL and SFL of tryptophan system, using the model from this chapter. Parameters of genetic regulation in SFL were altered in order to minimize the value of metric $I(p, s)$. Again, significantly longer rise (t_{rs}) and settling times (t_{ss}) were observed.

As we can see in Fig. 9, the performance of SFL based on the new model is even worse, rise time of 80 min wouldn't be enough for a cell to survive.

Motivated by the fact that transcriptional regulation is one of the basic and most common regulatory mechanisms, we wanted to find its main advantage. From the other two regulatory mechanisms in the *trp* operon, only enzymatic

inhibition doesn't work on transcriptional level. Therefore, we made another SFL design using only the enzymatic inhibition, and used again the metric (10) in order to optimize its dynamics.

As can be seen in Fig. 10, the performance of this design is the worst in terms of the rise and settling times. But one important thing can be seen in Fig. 10, when enzyme levels of the wild-type *trp* operon and two different SFL designs are compared. For the same steady-state level of the target metabolite, a significantly lower concentration of the enzymes is needed. This is due to the fact that in SFL with transcriptional repression only, all the produced enzymes are active, while in the wild-type tryptophan system and SFL with enzymatic inhibition a fraction of the enzymes are always inhibited.

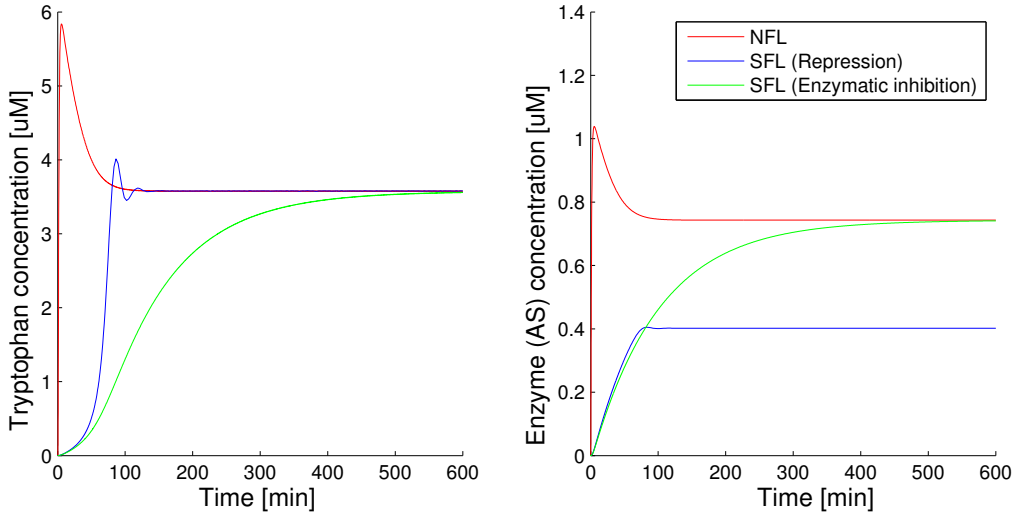

Figure 10: Comparison of the dynamic of the wild-type *trp* operon system and two different SFL designs that use transcriptional repression or enzymatic inhibition.

Overall, our simulations confirm the results from other authors [29], [10], [6], that an optimal structural design has evolved in the tryptophan system in order to meet necessary characteristics and that transcriptional or enzymatic regulation alone cannot outperform the wild-type NFL design.

3.2 Translational control as an alternative to transcriptional regulation

One regulatory mechanism used in synthetic biology is not involved in the regulation of *trp* operon. This mechanism works on a translational level and is implemented in form of translational riboswitches.

Translational riboswitches are based on the same principle as the riboswitches controlling transcription. They also have an aptamer and expression platform (see Sec. 2.1.2), which changes its conformation with ligand binding. The difference is that translational riboswitches only sequestre or free an RBS site on the mRNA, using usually simple pairing (see Fig. 11).

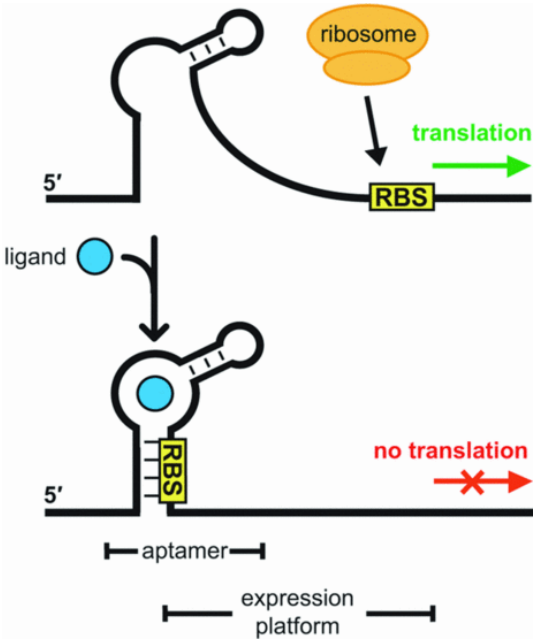

Figure 11: Mechanism of translational riboswitches. In this case, upon a ligand binding to the aptamer, RBS is sequestered, thus the ribosome cannot bind and begin with translation. Original image taken from [18].

Since no complex secondary structures are formed, the design of these elements is simpler than that of transcriptional riboswitches [22], [31], and many artificial riboswitches have been designed [21]. Disadvantage of their use in the control of

gene expression in operons is that a full mRNA molecule is always produced. This production generates unnecessary burden for the cell.

One disadvantage is that a single translational riboswitch can only regulate translation of genes belonging to a single RBS (see Fig. 12). This regulation is plausible for operons like the *trp* operon, where a single RBS controls all the genes using codon overlapping (last nucleotide of the stop codon is also the first nucleotide of the start codon (see Fig. 12) and [32]). The majority of existing operons, however, are controlled by multiple RBSs, one for each gene [33].

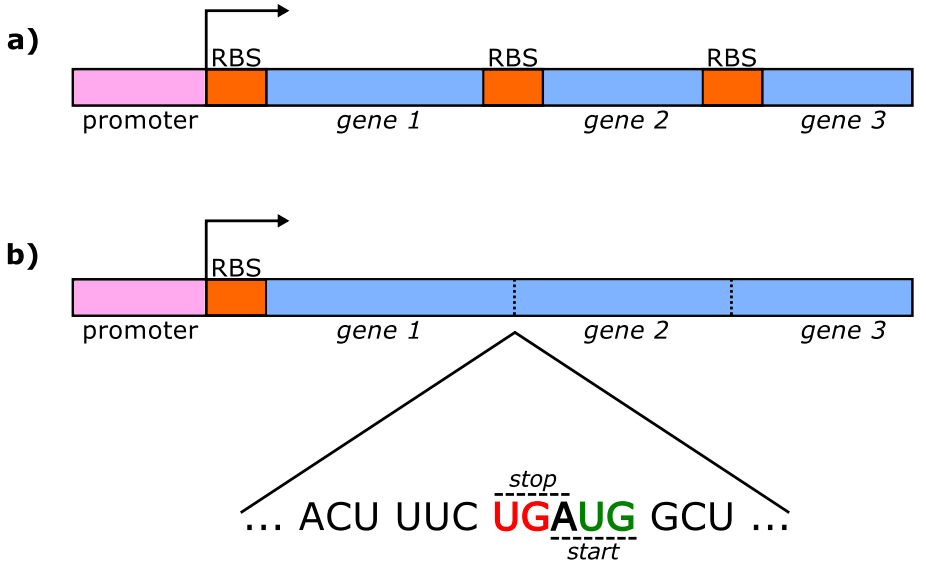

Figure 12: a) Common structure of bacterial operon - every gene has its own RBS. b) Operon with only one RBS for all genes (example is the *trp* operon). In this case, one ribosome can translate all genes at once without leaving the mRNA. When ribosome encounters stop codon, it also detects start codon, shifts frame and begins to translate the next gene.

3.3 Choice of controller input

In order to investigate the difference between different controller inputs, a simple model was made:

$$\frac{dM}{dt} = k_1 C_1(E) - (k_{d2} + \mu)M, \tag{11}$$

$$\frac{dE}{dt} = k_3 C_2(P)M - \mu E, \tag{12}$$

$$\frac{dP}{dt} = k_4 E - \mu P. \tag{13}$$

These equations refer to the transcription of mRNA (M) from an operon (11) and production of an enzyme (E) (12) catalyzing the formation of a product/metabolite(P) (13). Default values for all the parameters were taken from the *trp* operon model and can be found in Appendix A. Functions $C_1(E)$ and $C_2(P)$ were defined as

$$C_1(E) = \frac{K_1^{n_1}}{K_1^{n_1} + E^{n_1}}, \tag{14}$$

$$C_2(P) = \frac{K_p}{K_p + P}. \tag{15}$$

Function C_1 incorporates the effect of genetic repression by the produced enzyme, function C_2 incorporates the effect of translational repression from a riboswitch (therefore controller input to C_2 is the metabolite concentration, whereas controller input to C_1 is the enzyme concentration). Parameters K_1 and n_1 were again taken from Appendix A, parameter K_p was set to 10μM. Three different designs were compared, using single controllers C_1 or C_2, or their combination.

In this comparison (Fig. 13), parameter k_1 in the model with the controller C_1 was lowered to $k_1 = 3$ min^{-1}. For other two models, parameters k_1 and k_3 were tuned in order to get the same steady-state level of the metabolite. These parameters were chosen because they are easily tunable in real systems (promoter and RBS strength, codon usage). Simulation time was 3 hours. At time $t = 70$ min, after all models reach their steady-state, parameter k_4 was raised two times, for a duration of one hour. This perturbation can be caused in real systems by a sudden rise in a substrate concentration, an addition of a metabolite to the medium or by a change in enzymatic activity.

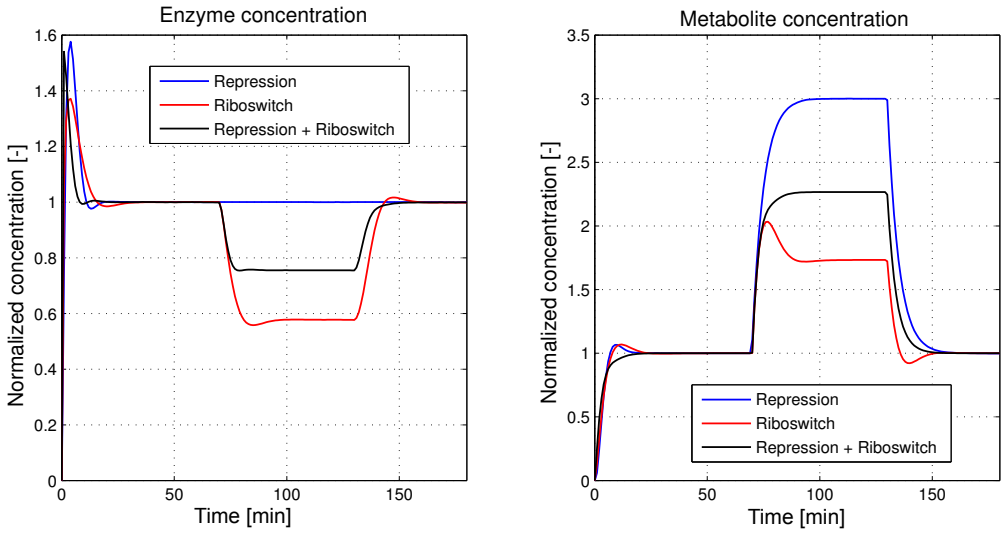

Figure 13: Comparison of three different synthetic designs for control of operons. *Repression* is a design which uses only controller C_1, *riboswitch* design includes only the controller C_2 and *repression + riboswitch* uses controllers C_1 and C_2.

All three designs show overshoot in their enzyme levels at the beginning. Upon introduction of a change in parameter k_4, we can observe that enzyme levels of repression-only design haven't changed. This was predicted, since the feedback signal is not sensitive to metabolite concentration change as caused by the introduced perturbation. The level of metabolite for this design is in accordance with this behavior - new temporal steady state was three times higher.

Other two designs have a feedback derived directly from the metabolite levels. As can been seen, the single riboswitch design outperforms the combined one in terms of sensitivity to the introduced perturbation - riboswitch design shows 1.7-fold change in its metabolite steady-state, meanwhile combined design shows 2.3-fold change. This can be explained by the respective enzyme levels - although these levels were lowered upon the induction, the combined design maintain higher levels. This is due to the fact that in this design, when the perturbation is introduced, the controller C_2 acts to lower the level of the enzymes. This change in the enzymatic concentration is noticed by the controller C_1, which tries to compensate for this loss.

4 Transcription-translation coupling as a way of controlling gene expression

In Section 3 it was shown that the designs with feedback on final metabolites measurements outperforms feedback on enzymatic levels. From Section 2.1, we know that the control elements using metabolite concentration as an controller input are riboswitches. Since the transcriptional riboswitches are very difficult to engineer, the only way now is to use translational riboswitches that are rationally designable using the current knowledge of RNA interactions. However, as was presented in Sec. 3.2, they are not able to control the transcription of the operon mRNA (which can be over 10kbp[5] long and its production from constitutive promoter can be heavy burden for the cell). Also, introducing this riboswitch to the RBS of every gene could significantly increase the experimental workload and make their design harder (because of possible RNA interactions in the long transcript). Last, but very important thing is that these elements cannot be simply linked together in order to form more complex regulatory networks.

The question we asked was, is there a way to overcome these limitations? Is it possible to use the advantage of the simple design of translational riboswitches and use them for a regulation on the transcriptional level?

From the study of the *trp* operon, we know about one mechanism that basically links together translational control with transcription - the attenuation (Sec. 2.1.2). In order the transcription to stop or continue, translation of the leader peptide must be completed or interrupted. This phenomenon is based on the transcription-translation coupling [34], [35], and is specific to prokaryotes, where translation can begin practically immediately after the start of transcription.

4.1 Adaptors design based on the *tna* operon

The transcriptional-translational coupling for the control of the gene expression has been already used by Liu et al. [31]. In this work, the authors developed an adaptor for converting translational regulators into regulators of transcription, based on the *tna* (tryptophanase) operon. This operon codes for the enzymes involved in the tryptophan conversion. The transcription of this operon is con-

[5]kilo base pairs = 1000 nucleotides

trolled by a rho-dependent terminator[6] in the leader sequence *tnaC*. The RBS of the *tnaC* is always active, which means that the leader region is being constantly translated. If the level of tryptophan is sufficient, the ribosome is able to translate the whole leader sequence. At the end of this sequence, the ribosome stalls at the terminator, physically sequestering it, therefore allowing polymerase to transcribe the coding genes. If the tryptophan concentration is low, ribosome is not able to translate the leader sequence and reach the terminator in order to sequester it. In this case, the transcription is terminated.

Authors hypothesized that replacing the native *tnaC* RBS with an RBS which can be regulated in order to switch between 'on' and 'off' states could be the key to the adaptor design. In the 'off' state, the ribosome wouldn't be able to even start with the translation of the leader region, much less obstruct the terminator. In the 'on' state in the presence of a sufficient level of tryptophan, the leader region would be translated and the terminator site sequestered. This hypothesis was tested and proven, using different RNA elements and translational riboswitches as the regulated RBS.

4.2 Adaptors design based on the *trp* operon

We have identified several disadvantages in the adaptors design based on the *tna* operon. First, since the egineered adaptors are based on the original *tnaC* sequence, sufficient levels of tryptophan must be present in order to block the trasncription termination even if the RBS is in the 'on' state. Second, while designing more complex circuits such as NOR gates controlling the operon transcription, one *tnaC* element must be used for every gate (Fig. 14a). This makes the whole design also more complex. And third, the change of the native RBS must be done in the *tnaC* element, which requires precise predictions of the secondary structures induced by these changes (Fig. 14b).

[6]This type of terminators needs a special protein, called Rho factor, to be bound in order to terminate transcription.

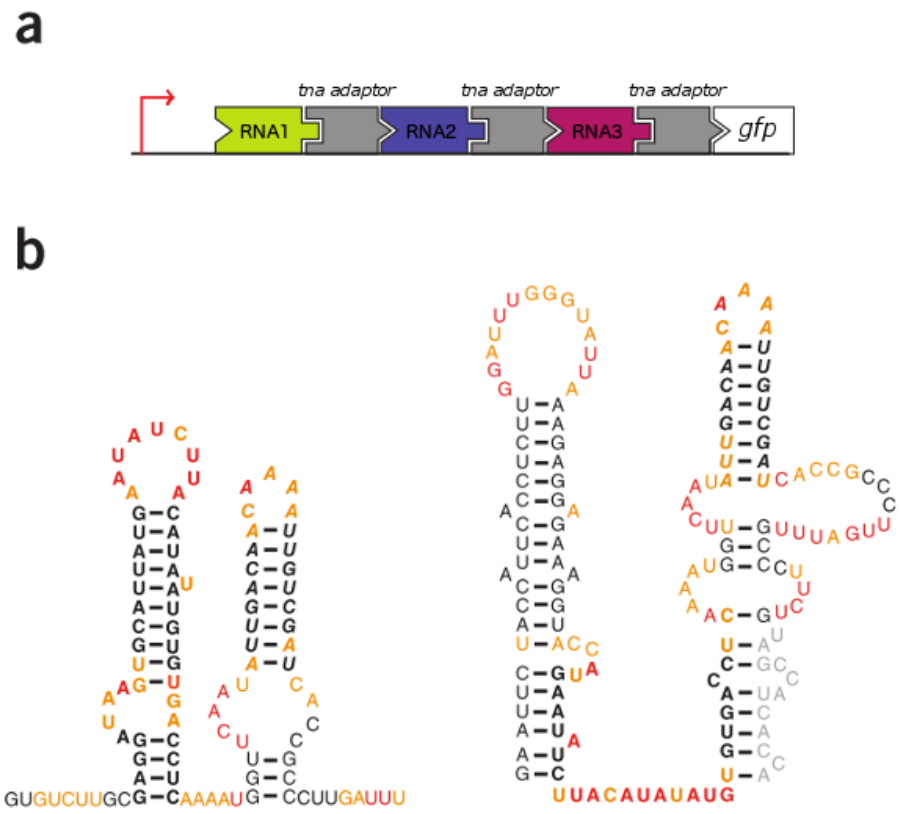

Figure 14: a) Scheme of a design with NOR gates controlling transcription of the gfp gene. For every gate, one tna adaptor (grey) must be used together with corresponding RNA element. b) Wild-type *tnaC* sequence on the left. On the right is the *tnaC* sequence with an RNA element controlling RBS state. Notice that for an introduction of an RBS control element, the changes must be done inside the *tnaC* sequence. Images adopted from [31].

Therefore, we have designed our own adaptors, using the *trp* operon as an inspiration. As was presented in Sec. 2.1.2, the attenuation in the *trp* operon is based on the formation of an anti-termination or termination loop. This termination loop works as an intrinsic (rho-independent) terminator, therefore it doesn't need a Rho-factor to be present in order to terminate transcription.

Our goal was to design a new modular leader region, able to control the expression of the structural genes located downstream of this sequence. Assuming that this leader region is right after the promoter, we first put an intrinsic terminator at the beginning of this region. In a normal situation, this terminator would stop the majority of bound RNA polymerases, therefore greatly reducing

the rate of transcription. The ratio of the stopped RNAPs to all bound RNAPs is called termination efficiency and is known for the characterized terminators (like for the ones from the Registry of Standard Biological Parts [36]).

Based on the [37], where an extensive study of intrinsic termination was published, we proposed a novel mechanism of regulation based on the transcriptional-translational coupling. As in the adaptors by Liu et al., we introduced a translational control element with an RBS and a start codon at the beginning of the leader region (before the intrinsic terminator). Also, we put a stop codon right after the terminator. In this setup, ribosome can begin with the translation of the leader region only when the RBS is in the 'on' state. When this is the case, the transcriptional-translational coupling occurs, i.e. the transcribing RNAP is tightly followed by the translating ribosome. When the RNAP encounters the intrinsic terminator, it is not affected, because the translating ribosome pushes the RNAP through the terminator sequence, not allowing RNAP to unbound. After the RNAP is safely after the terminator, the ribosome leaves at the stop codon.

Therefore, the translational control element in the leader region is able to control the transcription of the genes located downstream of the terminator. The mechanism is illustrated in Fig. 15. The are two advantages in comparison to the adaptors from Liu et al.:

1. RNA elements in the leader region can be easily switched, without any need for modifications to the existing sequences.

2. Since the whole leader region was engineered *de novo*, the transcription of the downstream genes is not dependent on the levels of tryptophan available in the cell, although these dependencies can be engineered by the selection of the used codons.

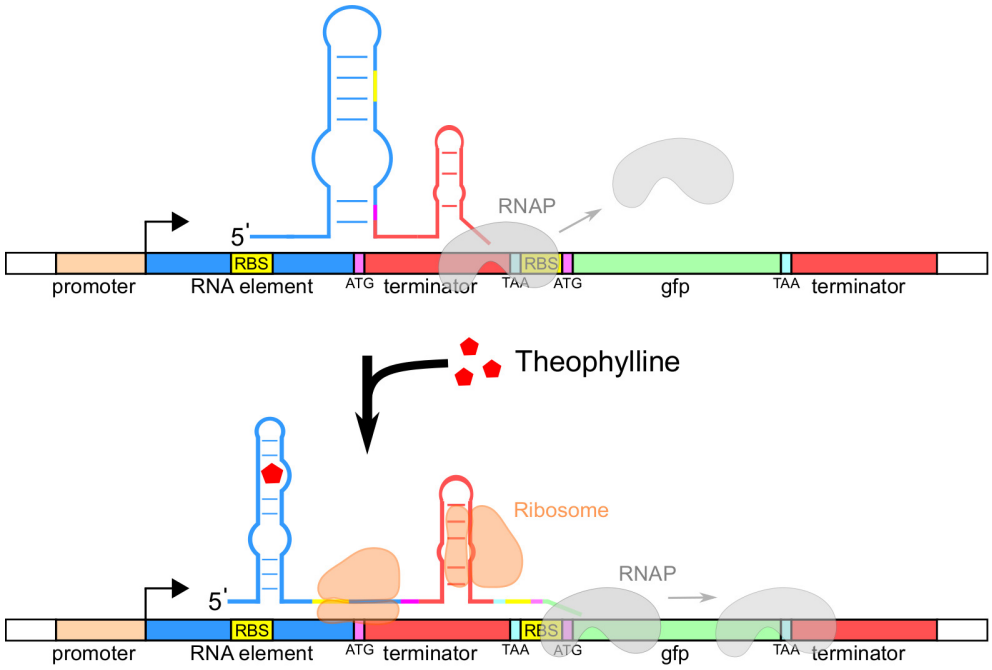

Figure 15: Mechanism of action of our adaptors design. In this example, when in the 'off' state (without theophylline), the translational riboswitch sequesters the RBS, therefore ribosome cannot bind and help the RNAP to go through the intrinsic terminator. In the 'on' state, the RBS is free and ribosomes interact with the transcribing RNAPs, allowing them to continue.

4.3 Overview of the designed adaptors

As a proof of concept, two different adaptors were designed and synthesized:

1. Adaptor using a theophylline riboswitch 12.1 [38] as an RNA element sequestering the RBS. Upon binding of a theophylline molecule, RBS is freed and ribosome can bound (see Fig. 15).

2. Adaptor using an RNA hairpin loop to sequester the RBS [39]. In the opposite direction, on the same plasmid, a trans-acting RNA (taRNA) is being produced under the control of a LacI promoter, which is repressed by a LacI proteins in default. Upon the IPTG induction, taRNA is produced. This taRNA interacts with the RNA hairpin loop in the adaptor, forces it to change its conformation and expose the RBS (see Fig. 16).

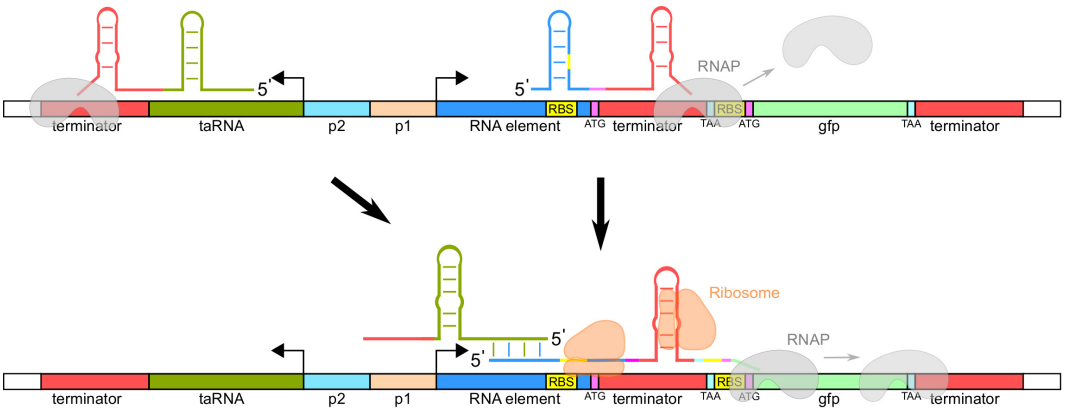

Figure 16: Mechanism of action of our adaptor with the RNA hairpin loop. In the default state, the hairpin loop is formed (blue) and RBS is sequestered. Upon the interaction with taRNA, produced by the promoter p2, the secondary structure changes and RBS is exposed.

Note that both these adaptors are positively regulated by the ligand (theophylline) or by the taRNA. This would correspond to a positive feedback loop. However, this could be easily changed by the use or design of different elements for translational control. We have chosen these two examples because both used RNA elements have been already experimentally tested in their native roles. This minimized the chance of a failure due to the conformational changes of the RNA elements.

4.4 Mathematical model

Before experiments, a set of ODEs was developed in order to capture the main properties of the adaptors designed:

$$\frac{dS}{dt} = CP_m(1-T_p) - (k_{dr}+\mu)S, \tag{16}$$

$$\frac{dM}{dt} = CP_mT_p - (k_{dm}+\mu)M, \tag{17}$$

$$\frac{dP}{dt} = rM\frac{m}{m+\mu+k_{dp}} - (k_{dp}+\mu)M. \tag{18}$$

Equation (16) stands for the production of a short RNA (i.e. RNA produced when the transcribing RNAP ends at the intrinsic terminator), equation (17) for

the production of mRNA (i.e. RNA which transcribed the coding gene) and finally, equation (18) for the production of a protein from the produced mRNA. Note the term T_p, which is a function defined as

$$T_p = f(K_{RBS}, Lig, T_{max}) = T_{max} K_{RBS} \frac{Lig}{K + Lig} + (1 - T_{max}), \tag{19}$$

where T_{max} is the intrinsic terminator efficiency, K_{RBS} is the relative strength of the RBS used in the RNA element, and Lig is a ligand concentration.

The above equations can be used directly for the adaptor with the theophylline riboswitch. For the adaptor with the hairpin loop, we must take into account the production of the taRNA:

$$\frac{dT}{dt} = CP_{ta} - (k_{dt} + \mu)T, \tag{20}$$

where P_{ta} is the transcription rate that depends on the concentration of the IPTG and LacI repressor proteins:

$$P_{ta} = P_{ta}^0 \frac{1 + \frac{1}{f}\left(\frac{L}{K_L\left(1+\frac{IPTG}{K_I}\right)}\right)^{nl}}{1 + \left(\frac{L}{K_L\left(1+\frac{IPTG}{K_I}\right)}\right)^{nl}}. \tag{21}$$

Then, the value of T can be used in the (19) as the ligand concentration Lig. Values and descriptions of all the used parameters can be found in Appendix B.

5 Simulations and experimental validation

In this section the above formulated model is characterized through simulation. Experimental data supporting the validity of the model is then presented.

5.1 Model simulations

The model with a theophylline riboswitch (Equations 16-18) was simulated to verify that it generates the expected behavior. The second model is an extension bof the first and hence its simulation is here omitted. Also note, zero induction of the RNA device cannot be achieved by the second model due to a basal transcription of taRNA.

First, the effects of the ligand (theophylline) concentration on the protein expression with for a fixed (maximal) relative RBS strength were characterised:

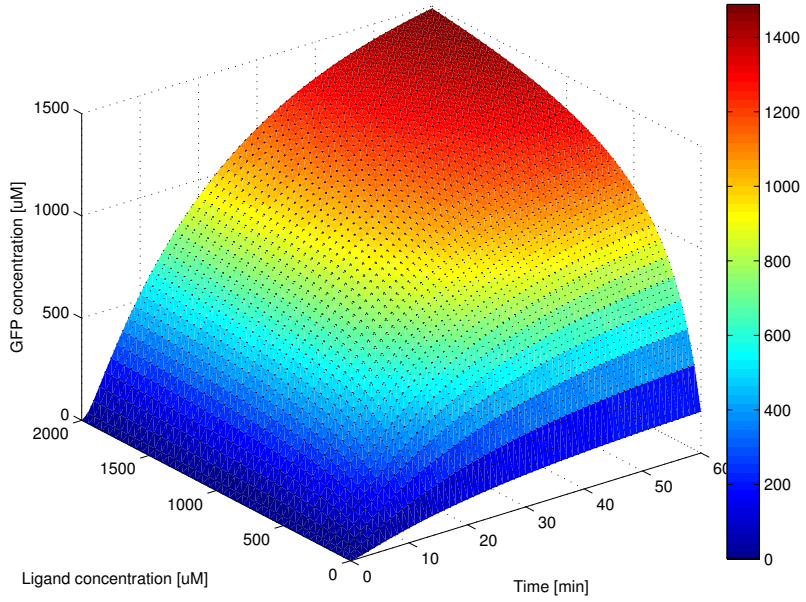

Figure 17: Protein concentration as a function of time and ligand concentration. Results from our model for the adaptor with the theophylline riboswitch.

Figure 17 shows the expected behavior. Zero ligand concentration resulted in minimal protein (GFP) concentration. This can be attributed to the terminator efficiency, which is around 88%. When the ligand concentration increases, the protein concentration also goes up, reaching a maximum at the maximal ligand concentration. Further increasing of the ligand concentration doesn't have any significant effect on the protein concentration, suggesting that the transcription rate is almost maximal at the ligand concentration of 2000μM. This corresponds to a situation when all riboswitches (in case of multiple plasmids) are already in the 'on' state.

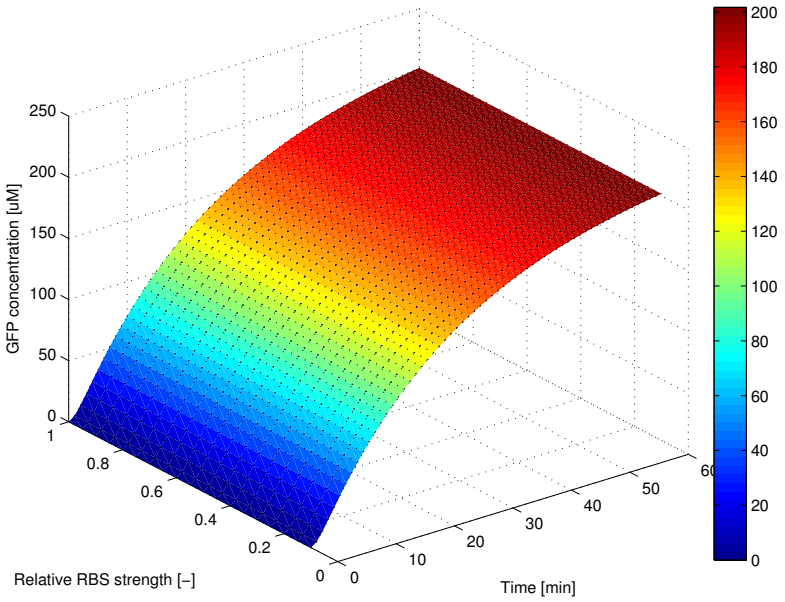

Figure 18: Protein concentration as a function of time and relative RBS strength. Ligand concentration was set to zero.

Second, we explored the model behavior with no ligand present. When no ligands are present, varying the RBS strength in the RNA element shouldn't have any impact on the protein expression (since the RBS is sequestered). The same protein expression, corresponding to the basal level of transcription should be seen for all values of K_{RBS}. This was also confirmed (see Fig. 18).

Finally, the relationship between the ligand concentration and relative RBS strength was investigated. For the same non-zero values of the ligand concentration, the protein steady-state concentration should be increasing with increasing RBS strength. The model behavior corresponds to this assumption, as shown in Figure 19.

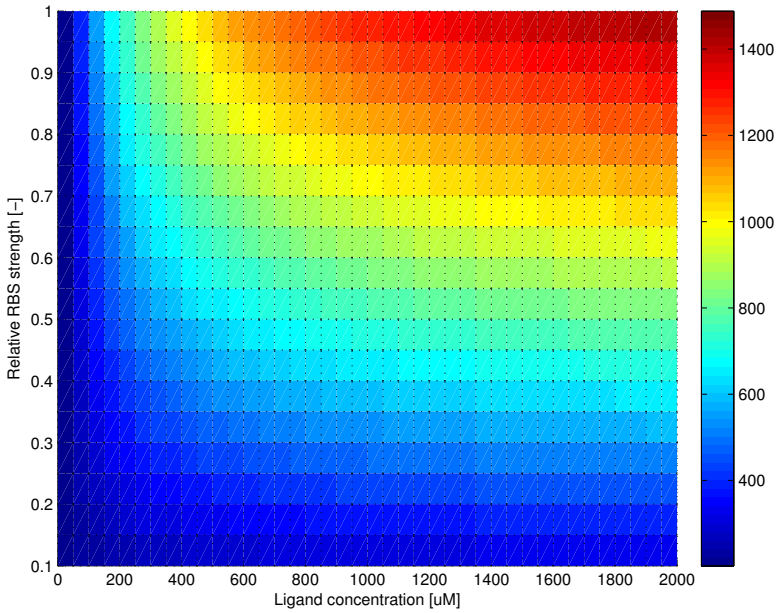

Figure 19: Steady-state protein concentration as a function of ligand concentration and relative RBS strength.

5.2 Experimental validation

A set of experiments was proposed to validate the results obtained from the simulations. A plasmid contaning the adaptor with a harpin loop (see Sec. 4.3) was transformed into *E. coli* cells. Transformed colonies were incubated overnight in the LB media with the appropriate antibiotic. After the overnight incubation, cultures were diluted into M9 media and transferred to a plate for fluorescence and absorbance measurements.

Three different concentrations of the inducer of the promoter p2 (IPTG) were tested, each in three replicates. These concentrations were chosen in order to achieve: zero induction (no IPTG), full induction (1mM IPTG) and partial (50%) induction (20µM IPTG). Measurements were taken every 10 minutes, for 200 minutes in total.

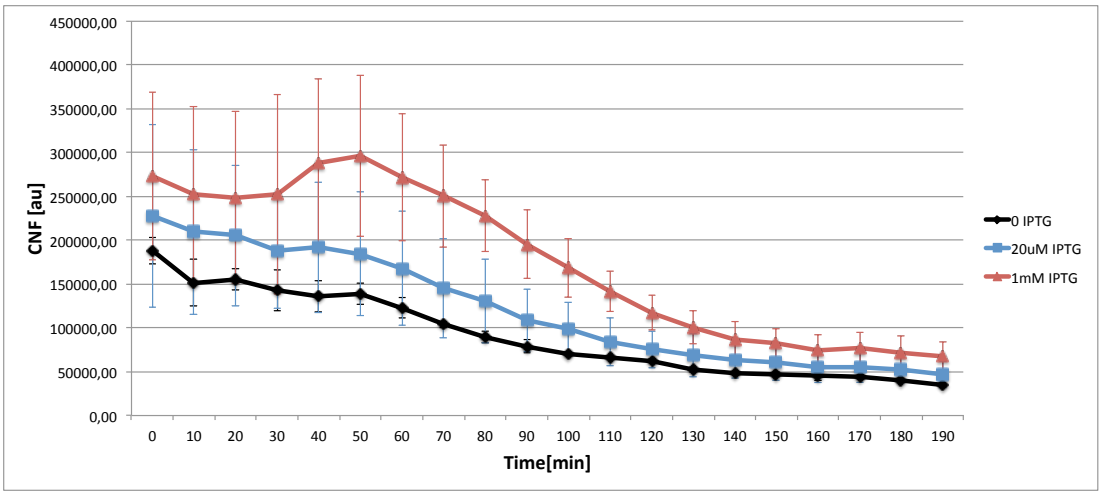

Figure 20: Results from experimental characterization. Mean value of CNF for all replicates is shown together with the standard deviation.

Resulting plot is shown in Fig. 20, where CNF is the Corrected Normalized Fluorescence, which was calculated for every sample as follows:

$$CNF(t) = \frac{F(t) - F_m}{A(t) - A_m} - F_c, \tag{22}$$

where $F(t)$ is the measured fluorescence at time t, F_m is the fluorescence of the M9 media, $A(t)$ is the measured absorbance at 600nm, A_m is the absorbance of M9 media and F_c is the average normalized natural fluorescence of the cells.

The results validate the hypothesis. Fully induced cells had the highest value of CNF for the duration of the experiment and cells with no induction maintained the lowest value of CNF. The CNF for all the three cases exhibited a decreasing trend with time. This is probably caused by the transfer of cells into M9 media from LB media. M9 media is minimal (i.e., contains only essential nutrients). Hence cells transferred from the optimal LB media to M9 media have a higher metabolic burden in producing other metabolites and thereby have a decreased capacity for expression of the transfected plasmid.

The transient phase hindered the validation of the dynamic characteristics of our model. New experiments are planned using prolonged incubation after the transfer from LB media to M9.

6 Discussion and future work

6.1 Analysis of the *trp* operon

In this work, the tryptophan biosynthetic pathway and the underlying operon were analysed using a developed physically relevant model. The roles of the implemented regulatory mechanisms were elucidated in terms in terms of structure and controller input. In addition to these mechanisms, alternatives from the field of synthetic biology were proposed, with a special focus on their programmability.

6.2 Engineering operon regulation, adaptors design

We showed that controller inputs based on metabolite measurements can outperform the controller inputs based on enzyme measurements and that transcriptional regulation is advantageous in the regulation of large operons. To address challenges of rational design of transcriptional regulation based on metabolite measurements, we proposed a novel mechanism motivated by easily designable elements from translational regulation. These adaptors use transcriptional-translational coupling in prokaryotes and are potentially scalable to more larger logic circuits.

6.3 Simulations and experiments

A general deterministic model was proposed in order to capture the behavior of the designed devices. In order to validate the proposed model, experimental characterization of one of the constructs was performed. Steady state predictions matched live cell measurements. Dynamic predictions remain to be tested.

6.4 Future work

In future work, new adaptors, using different RNA regulatory elements will be designed and synthesized in order to investigate their characteristics. Proposed elements include aptazymes, ribozymes and different types of rationally engineered RNA hairpin loops. The role of other structures and their properties in the designed leader region (RBS, intrinsic terminator) will be also evaluated.

Through a combined model based and evidence based design approach, the future ambition is to build on the regulatory concepts developed in this thesis to alleviate outstanding issues in future biotechnologies.

List of Figures

List of Tables

References

[1] First Successful Laboratory Production of Human Insulin Announced. In: *Genentech: Press Releases* [online]. 1978 [cit. 2014-04-16]. Available at: http://www.gene.com/media/press-releases/4160/1978-09-06/first-successful-laboratory-production-o

[2] RAAB, R. Michael, Keith TYO and Gregory STEPHANOPOULOS. Metabolic Engineering. *Biotechnology for the Future*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, n. 100, p. 1. DOI: 10.1007/b136411. Available at: http://www.springerlink.com/index/10.1007/b136411

[3] CONNORS, Neal. Classical Strain Improvement – Back to the Future!. *Society for Industrial Microbiology and Biotechnology* [online]. 2011 [cit. 2014-04-16]. Available at: http://www.simbhq.org/2011/09/08/classical-strain-improvement-%E2%80%93-back-to-the-future/

[4] ERICKSON, B., R. SINGH and P. WINTERS. Synthetic Biology: Regulating Industry Uses of New Biotechnologies. *Science*. 2011-09-01, vol. 333, issue 6047, s. 1254-1256. DOI: 10.1126/science.1211066. Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.1211066

[5] YANOFSKY, Charles, et al. The complete nucleotide sequence of the tryptophan operon of Escherichia coli. *Nucleic Acid Research*. 1981, vol 9, n. 24, 6647–6668. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC327632/

[6] BHARTIYA, Sharad, Nikhil CHAUDHARY, K.V. VENKATESH and Francis J. DOYLE. Multiple feedback loop design in the tryptophan regulatory network of Escherichia coli suggests a paradigm for robust regulation of processes in series. *Journal of The Royal Society Interface*. 2006-6-22, vol. 3, issue 8, p. 383-391. DOI: 10.1098/rsif.2005.0103. Available at: http://rsif.royalsocietypublishing.org/cgi/doi/10.1098/rsif.2005.0103

[7] VENKATESH, K.V., Sharad BHARTIYA and Anurag RUHELA. Multiple feedback loops are key to a robust dynamic performance of tryptophan regulation in Escherichia coli. *FEBS Letters*. 2004, vol. 563, 1-3, p. 234-240. DOI: 10.1016/S0014-5793(04)00310-2. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0014579304003102

[8] RALSTON, Amy. Operons and Prokaryotic Gene Regulation. In: *Nature Education* [online]. 2008 [cit. 2014-04-18]. Available at: http://www.nature.com/scitable/topicpage/operons-and-prokaryotic-gene-regulation-992

[9] KELLEY, R. L. and C. YANOFSKY. Trp aporepressor production is controlled by autogenous regulation and inefficient translation. *Proceedings of the National Academy of Sciences*. 1982-05-15, vol. 79, issue 10, p. 3120-3124. DOI: 10.1073/pnas.79.10.3120. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.79.10.3120

[10] NGUYEN, Lan K. and Don KULASIRI. On Multiple Regulatory Mechanisms in the Tryptophan Operon System in Escherichia coli: In silico Study of Perturbation Dynamics. *In Silico Biology.* 2008, vol. 8, iss. 5, p. 485-510. Available at: http://iospress.metapress.com/content/F45185R18525RP16

[11] YANOFSKY, Charles and Virginia HORN. Role of regulatory features of the trp operon of Escherichia coli in mediating a response to a nutritional shift. *Journal of Bacteriology*. Baltimore: [American Society for Microbiology], 1994, vol. 154, n. 20, p 6245-54. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7928995

[12] SANTILLÁN, Moisés and Eduardo S. ZERON. Dynamic influence of feedback enzyme inhibition and transcription attenuation on the tryptophan operon response to nutritional shifts. *Journal of Theoretical Biology*. 2004, vol. 231, issue 2, p. 287-298. DOI: 10.1016/j.jtbi.2004.06.023. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0022519304003145

[13] LacI regulated promoter. In: Registry of Standard Biological Parts [online]. 2003 [cit. 2014-05-04]. Available at: http://parts.igem.org/Part:BBa_R0010

[14] MOSCOU, M. J. and A. J. BOGDANOVE. A Simple Cipher Governs DNA Recognition by TAL Effectors. Science. 2009-12-10, vol. 326, issue 5959, p. 1501-1501. DOI: 10.1126/science.1178817. Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.1178817

[15] iGEM 2012 - Team Slovenia [online]. 2012 [cit. 2014-05-04]. Available at: http://2012.igem.org/Team:Slovenia

[16] Trp operon leader peptide. In: *Uniprot* [online]. 2014 [cit. 2014-04-21]. Available at: http://www.uniprot.org/uniprot/P0AD92

[17] NAVILLE, M. and D. GAUTHERET. Transcription attenuation in bacteria: theme and variations. *Briefings in Functional Genomics and Proteomics*. 2009-10-15, vol. 8, issue 6, p. 482-492. DOI: 10.1093/bfgp/elp025. Available at: http://bfgp.oxfordjournals.org/cgi/doi/10.1093/bfgp/elp025

[18] KIM, Jane N. and Ronald R. BREAKER. Purine sensing by riboswitches. Biology of the Cell. 2008, vol. 100, issue 1, p. 1-11. DOI: 10.1042/BC20070088. Available at: http://doi.wiley.com/10.1042/BC20070088

[19] WINKLER, Wade C. and Ronald R. BREAKER. Genetic Control by Metabolite-Binding Riboswitches. ChemBioChem. 2003-10-06, vol. 4, issue 10, p. 1024-1032. DOI: 10.1002/cbic.200300685. Available at: http://doi.wiley.com/10.1002/cbic.200300685

[20] WITTMANN, Alexander and Beatrix SUESS. Engineered riboswitches: Expanding researchers' toolbox with synthetic RNA regulators. *FEBS Letters*. 2012, vol. 586, issue 15, p. 2076-2083. DOI: 10.1016/j.febslet.2012.02.038. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0014579312001615

[21] SUESS, B. A theophylline responsive riboswitch based on helix slipping controls gene expression in vivo. *Nucleic Acids Research*. 2004-02-23, vol. 32, issue 4, p. 1610-1614. DOI: 10.1093/nar/gkh321. Available at: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkh321

[22] WACHSMUTH, M., S. FINDEISS, N. WEISSHEIMER, P. F. STADLER and M. MORL. De novo design of a synthetic riboswitch that regulates transcription termination. *Nucleic Acids Research*. 2013-02-18, vol. 41, issue 4, p. 2541-2551. DOI: 10.1093/nar/gks1330. Available at: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gks1330

[23] CERES, P., J. J. TRAUSCH and R. T. BATEY. Engineering modular 'ON' RNA switches using biological components. Nucleic Acids Research. 2013-12-05, vol. 41, issue 22, p. 10449-10461. DOI: 10.1093/nar/gkt787. Available at: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkt787

[24] CARTER, Paul. Site-directed mutagenesis. *Biochemical Journal*. 1986, vol. 237, p. 1-7. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1146940/

[25] FOO, Jee Loon, Chi Bun CHING, Matthew Wook CHANG and Susanna Su Jan LEONG. The imminent role of protein engineering in synthetic biology. *Biotechnology Advances*. 2012, vol. 30, issue 3, p. 541-549. DOI: 10.1016/j.biotechadv.2011.09.008. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0734975011001613

[26] YANOFSKY, Charles, Richard L. KELLEY and Virginia HORN. Repression is relieved before attenuation in the trp operon of Escherichia coli as tryptophan starvation becomes increasingly severe. *Journal of Bacteriology*. Baltimore: [American Society for Microbiology], 1984, vol. 158, n. 3. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC215544/

[27] SANTILLAN, M. and M. C. MACKEY. Dynamic regulation of the tryptophan operon: A modeling study and comparison with experimental data. *Proceedings of the National Academy of Sciences*. 2001-02-13, vol. 98, issue 4, p. 1364-1369. DOI: 10.1073/pnas.98.4.1364. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.98.4.1364

[28] SAVAGEAU, Michael A. Introduction to S-systems and the underlying power-law formalism. Mathematical and Computer Modelling. 1988, vol. 11, p. 546-551. DOI: 10.1016/0895-7177(88)90553-5. Available at: http://linkinghub.elsevier.com/retrieve/pii/0895717788905535

[29] BHARTIYA, Sharad, Subodh RAWOOL and K. V. VENKATESH. Dynamic model of Escherichia coli tryptophan operon shows an optimal structural design. *European Journal of Biochemistry*. 2003, vol. 270, issue 12, p. 2644-2651. DOI: 10.1046/j.1432-1033.2003.03641.x. Available at: http://doi.wiley.com/10.1046/j.1432-1033.2003.03641.x

[30] NGUYEN, L.K. and D. KULASIRI. Distinct noise-controlling roles of multiple negative feedback mechanisms in a prokaryotic operon system. *IET Systems Biology*. 2011-03-01, vol. 5, issue 2, p. 145-156. DOI: 10.1049/iet-syb.2010.0020. Available at: http://digital-library.theiet.org/content/journals/10.1049/iet-syb.2010.0020

[31] LIU, Chang C, Lei QI, Julius B LUCKS, Thomas H SEGALL-SHAPIRO, Denise WANG, Vivek K MUTALIK and Adam P ARKIN. An adaptor from translational to transcriptional control enables predictable assembly of complex regulation. *Nature Methods*. 2012-9-30, vol. 9, issue 11, p. 1088-1094. DOI: 10.1038/NMETH.2184. Available at: http://www.nature.com/doifinder/10.1038/nmeth.2184

[32] OPPENHEIM, Daniel S. and Charles YANOFSKY. Translational coupling during expression of the tryptophan operon of Escherichia coli. *Genetics*. 1980, vol. 95, n. 5, p. 785-795. Available at: http://www.genetics.org/content/95/4/785.short

[33] WATSON, James Dewey. *Molecular biology of the gene*. 5th ed. San Francisco: Benjamin Cummings, 2004, 732 p. ISBN 03-212-2368-3.

[34] YANOFSKY, C. RNA-based regulation of genes of tryptophan synthesis and degradation, in bacteria. *RNA*. 2007-06-29, vol. 13, issue 8, p. 1141-1154. DOI: 10.1261/rna.620507. Available at: http://www.rnajournal.org/cgi/doi/10.1261/rna.620507

[35] RALSTON, Amy. Simultaneous Gene Transcription and Translation in Bacteria. In: *Nature Education* [online]. 2008 [cit. 2014-05-07]. Available at: http://www.nature.com/scitable/topicpage/simultaneous-gene-transcription-and-translation-in-bacteria-1025

[36] Registry of Standard Biological Parts [online]. 2006 [cit. 2014-05-12]. Available at: http://partsregistry.org

[37] YARNELL, W. S. Mechanism of Intrinsic Transcription Termination and Antitermination. Science. 1999, vol. 284, issue 5414, p. 611-615. DOI: 10.1126/science.284.5414.611. Available at: http://www.sciencemag.org/cgi/doi/10.1126/science.284.5414.611

[38] LYNCH, S. A. and J. P. GALLIVAN. A flow cytometry-based screen for synthetic riboswitches. Nucleic Acids Research. 2009-01-08, vol. 37, issue 1, p. 184-192. DOI: 10.1093/nar/gkn924. Available at: http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkn924

[39] RODRIGO, G., T. E. LANDRAIN and A. JARAMILLO. De novo automated design of small RNA circuits for engineering synthetic riboregulation in living cells. Proceedings of the National Academy of Sciences. 2012-09-18, vol. 109, issue 38, p. 15271-15276. DOI: 10.1073/pnas.1203831109. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1203831109

A *trp* operon model parameters

Parameter	Description	Value
k_1	Free operator synthesis rate	50 min^{-1}
k_2	Transcription rate	5.1 min^{-1}
k_3	Translation rate	20 min^{-1}
k_4	Tryptophan synthesis rate	59 min^{-1}
k_{d1}	Free operator degradation rate	0.5 min^{-1}
k_{d2}	mRNA degradation rate	0.96 min^{-1}
k_{d3}	Enzyme degradation rate	0 min^{-1}
K_1	Repression's dissociation constant	3.53 μM
K_2	Attenuation's dissociation constant	0.04 μM
K_3	Enzyme inhibition's dissociation constant	4.1 μM
n_1	Repression's Hill coefficient	1.92
n_2	Attenuation's Hill coefficient	1.72
n_3	Enzyme inhibition's Hill coefficient	1.2
g	Tryptophan uptake kinetic rate	25 μMmin^{-1}
K_g	Tryptophan uptake's dissociation constant	0.2 μM
O_t	Total operator concentration	0.00332μM
μ	Cell growth rate	0.01 min^{-1}
d	Tryptophan adsorption constant	23.5 μMmin^{-1}
e	Parameter in the Trp intake expression	0.9 μM
f	Parameter in the Trp intake expression	380 μM

B Adaptor model parameters

Parameter	Description	Value
C	Plasmid copy number	1
P_m	Transcription rate	3 min^{-1}
k_{dr}	Short RNA degradation rate	0.3 min^{-1}
μ	Cell growth rate	0.01 min^{-1}
k_{dm}	mRNA degradation rate	0.96 min^{-1}
r	Translation rate	50 min^{-1}
m	Protein maturation rate	0.05 min^{-1}
k_{dp}	Protein degradation rate	0.035 min^{-1}
K	Dissociation constant	300 μM
k_{dt}	taRNA degradation rate	0.3 min^{-1}
P_{ta}^0	Maximal transcription rate	3 min^{-1}
f	LacI repression fold	1000
L	LacI concentration	50 nM
K_L	LacI effective binding coefficient	50 μM
K_I	IPTG effective binding coefficient	400 μM
nl	Hill coefficient	2