tRNA mimicking structures to control and monitor biological processes

Paul, Avishek

DOI:
10.33612/diss.166342562

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 18-01-2022
Chapter 1

General Introduction
1.1. INTRODUCTION

Synthetic biology has introduced new molecular functions into cells by harbouring a forward-engineering approach. At the early stage of scientific evolution of the realm of synthetic biology, the potential of biological parts to achieve network like architectures and to exhibit collective logical function was realized\(^1\). Since then, many ground-breaking researches in genomics and genetic engineering paved the way to construct and assemble biomolecular parts to engineer biological networks\(^2\)\(^-\)^\(^5\). Synthetic biology, with modularization, rationalization and modelling, has filled the inventory of modern biological research with ever-increasing suite of genetic devices and biological modules to programme a living cell.

Programming living cells often requires adoption of an electrical engineering framework to design an artificial genetic switch: each functional unit in the switch is called “genetic part” which can be interconnected to build a layer of complex and sophisticated genetic switches. All the triggers involved in stimulating a genetic switch are called inputs and the response generated from the switch in presence of inputs is called output. These inputs and outputs can be characterized as analog and digital signals depending on whether the signals are composed of a continuous set of values or binary values, respectively\(^6\)\(^,\)\(^7\).

Programmed cells can be implemented as therapeutic agents to rectify genetic disease or colonize niches in human microbiome to exhibit a therapeutic function\(^8\)\(^-\)^\(^12\). Synthetic genetic switches have emerged as an imperative tool to uncover new natural products. These products can be characterized by synthetic regulation since the induction conditions of many relevant gene clusters are unknown\(^13\)\(^-\)^\(^17\).

Despite its enormous potential, the advancement of genetic switch design often encounters significant hurdles because of several undefined, complex and dynamic interactions that can occur between the switches and the biological environment\(^18\)\(^,\)\(^19\). All genetic switches are composed of several regulatory components and a precise channelling of biological information among these components is necessary to develop a well characterized genetic switch\(^20\)\(^,\)\(^21\). In this regard, different computational tools and part libraries have been built to regulate the function of each genetic parts in the switch\(^22\)\(^-\)^\(^25\). Desirable attributes of part libraries include specificity, orthogonality (strong interaction among cognate genetic parts without encountering any crosstalk) and high dynamic response.
1.2. FUNCTIONAL UNITS OF A GENETIC SWITCH

The composition of a genetic switch is desired to be highly modular\textsuperscript{26} so that the genetic parts can be freely combined without altering the functionality of the switch. In principle, a genetic switch architecture is composed of two parts – (1) the recognition unit or aptamer, which binds to a ligand and (2) an expression platform which facilitates a structural change in the switch architecture upon binding of the inputs to the recognition unit to influence gene expression.

1.2.1. The aptamer/recognition unit

1.2.1.1. Aptamers generated by SELEX

For synthetic genetic switch applications, there is growing interest in identifying suitable aptamer units specific for a given target molecule that, in many cases, does not represent a natural cell component to ensure orthogonality. Such aptamers for non-natural ligands can be isolated by an \textit{in vitro} evolution and selection process (SELEX)\textsuperscript{27,28}. However, only a handful of selected aptamers have been successfully implemented into riboswitch constructs due to the incompatibility and non-functionality \textit{in vivo}\textsuperscript{29-32}. Hence, the identification of \textit{in vitro} selected aptamers that can be incorporated as recognition unit into a functional riboswitch is a very challenging task.

The most prominent and widely used example of an \textit{in vitro} selected aptamer that represents a functional ligand sensing platform in genetic switches \textit{in vivo} is a theophylline binding aptamer\textsuperscript{33} which has a very high affinity toward its ligand (320 nM) and shows a high selectivity against structurally related molecules like caffeine\textsuperscript{34}. In the unbound state, the theophylline aptamer shows multiple conformations but folds into a defined hairpin like structure upon ligand interaction\textsuperscript{35,36}. Furthermore, the good cell permeability renders theophylline a very attractive ligand for a variety of switch constructs in prokaryotes and eukaryotes. To construct the switch, the aptamer was inserted as a recognition domain into the switch architecture regulating the transcription process in cells\textsuperscript{37,38}. In absence of theophylline, the switch inhibits the transcription of a target gene. But upon binding of theophylline to its aptamer, the switch undergoes a significant change in structural conformation and restores the transcription process of the gene (Figure1).
Figure 1. Engineering a genetic switch to control transcription of a target gene (bgaB) with theophylline as input. The theophylline aptamer (red) was fused to a short spacer (cyan) followed by a sequence complementary to the 3’ part of the aptamer (blue) and a U stretch (black). The ribosome binding site (black box) and the open reading frame of the reporter gene bgaB are located at the downstream region. The terminator structure formation gets inhibited when theophylline binds to the aptamer, resulting in initiation of the transcription process.39

Even though the SELEX process is designed to find aptamers with high affinity against the target molecule, there are cases in which in vitro selected aptamers fail to work in vivo illustrating that binding affinity itself is not the only feature necessary for switch activity in cells. Rather, it is crucial that binding of the ligand to the aptamer domain also induces a change in secondary structure of the switch.

Moreover, aptamers are selected through SELEX under conditions where magnesium or sodium are present at high concentration. As a consequence, their functionality within cells is often hampered. SELEX is generally performed in 5–10 mM Mg\(^{2+}\), while the in vivo concentration of magnesium is much lower. Thus, a selection at such low Mg\(^{2+}\) levels might increase the probability to isolate aptamers that are functional in vivo. In addition, potassium is the most abundant monovalent cation in prokaryotic and eukaryotic cells and a corresponding adjustment of the SELEX conditions might be very useful as well.

1.2.1.2. Aptamers found in cellular environment

Due to the limitations of in vitro SELEX derived aptamers to be implemented in cellular context, an in vivo selection would be a desirable alternative. However, the sequence library size for ligand screening in vivo is very small as transformation efficiency is limited. Another elegant solution would be an in silico selection of aptamers or even complete genetic switches where a given sequence is screened for the formation of suitable ligand-binding pockets. Even though in silico design of a genetic switch can deliver great promise to control a
target gene, currently very few promising results have been reported\textsuperscript{50-52}. So, considering all the strategies of designing synthetic switches, the only promising approach is to use aptamers found \textit{in vivo}.

Synthetic genetic switches that harbour aptamers occurring in natural riboswitches can recognize any endogenous molecule\textsuperscript{53} which could result in unwanted background activity. But an ideal synthetic switch should act as an orthogonal regulator that is activated only by a specific non-natural ligand. To overcome this problem, natural riboswitch-derived aptamers were reengineered to recognize a synthetic ligand\textsuperscript{54-56}. This reengineering process was conducted by mutagenesis of the ligand-binding site of the aptamer. The original target ligand was also chemically modified to fabricate a ligand library. Thus, two different libraries were created – aptamer and target ligand. The mutated switch library was cloned upstream of a reporter gene and tested against the ligand library\textsuperscript{57}. The combination of a genetic switch and its cognate ligand can only activate the target gene.

Even though this process generates an orthogonal genetic switch, it is undoubtedly a cumbersome process since both the aptamer and the target molecule need to be modified to create libraries and they need to be tested against each other. To make the process simple, an external molecule can be used as a target ligand against a genetic switch, which contains an aptamer binding specifically to the target molecule. For example, an antibiotic derivative could be used as a ligand since it was reported that some antibiotic derivatives are non-toxic for the cells under certain concentrations\textsuperscript{58}. It was observed that an RNA switch accommodating an aptamer that binds specifically to the original antibiotic can control expression of a target gene with a derivative of the original antibiotic as target ligand\textsuperscript{58}.

\textbf{1.2.2. The expression platform}

Apart from the aptamer domain, a synthetic genetic switch is also composed of an expression platform. The expression platform is generally responsible to control the expression of a target gene. The expression platform is attached to the aptamer by a genetic unit containing certain number of nucleotides. This unit of the switch is often referred to as “transmitter domain”\textsuperscript{59}. When the ligand binds to its cognate aptamer, the information regarding structural change of the switch is relayed from the aptamer to the expression platform through the transmitter module. While the aptamer of a genetic switch exhibits a high degree of conservation at the sequence and structural level, the adjacent expression
platform usually varies in regard to its sequence and secondary structure. For example, translation is differently regulated in bacterial and eukaryotic systems. In bacteria, a ribosomal binding site (RBS) is located upstream of the AUG start codon in an open reading frame\(^6\). The switch can control the accessibility of the RBS for the ribosome, simply by masking or presenting this sequence in a ligand-dependent way\(^6\). In eukaryotic translation, the small ribosomal subunit attaches to the 5’-cap of the mRNA and scans for the start codon, where the complete ribosome assembly occurs and consequently translation is initiated\(^6\). Hence, a eukaryotic translational switch can be realized by controlling the scanning process of the ribosomal subunit.

The expression platform of a switch is difficult to engineer. The prediction of precise information channelling between an aptamer domain and expression platform in terms of ligand-induced structural rearrangements is challenging\(^6,64\). For example, the expression platform can function as co-transcriptional regulator (Figure 2a). Ligand-dependent premature termination of transcription can turn off gene expression by preventing RNA polymerase from transcribing the coding region. Ligand binding typically destabilizes the anti-terminator and permits formation of a terminator hairpin that disrupts the active transcription complex and leads to RNA polymerase dissociation from the mRNA. Alternatively, ligand binding may affect rho-dependent transcription termination. The expression platform can also regulate translation (Figure 2b). Protein translation in bacteria starts with binding of the 30S ribosome to the ribosome binding site (RBS), hence blocking the RBS prevents translation. Ligand binding to a riboswitch aptamer can influence stem formation either to block the RBS or to reveal the RBS. Finally, the expression platform can involve ligand-induced ribozyme activation (Figure 2c), which has been observed for the glucosamine-6-phosphate riboswitch class. It was found that ligand binding activates self-cleavage of the glmS RNA, leading to mRNA degradation.

\(\text{Figure 2.}\) Schematic to present different expression platforms of a switch. The ligand specific for the aptamer domain is denoted as a red circle. (a) Expression
platform for transcriptional attenuation: In the ligand-bound form, a terminator hairpin is formed causing transcription termination. (b) Expression platform for translation inhibition: Ligand binding causes an alternative structure, which occludes the ribosome-binding site (RBS) and therewith prevents initiation of translation. (c) Messenger RNA (mRNA) degradation caused by ribozyme activity: Ligand-induced self-cleavage by a ribozyme destabilizes the mRNA, which leads to rapid mRNA degradation. As mentioned above, the structure of the aptamer was maintained the same in all of the three cases but the structures of the expression platforms vary.

There are mainly two ways to engineer the expression platform,

1.2.2.1. Trial and error method

On some occasions, it becomes difficult to predict the secondary structure of the switch before and after ligand binding to the recognition domain. In those instances, trial and error method is usually adopted to engineer an expression platform by altering the distance between the aptamer and the expression platform or by changing the length of the transmitter module. The trial and error method was first implemented on a genetic switch controlling the expression of a LacZ gene. The switch contains a theophylline aptamer located five nucleotides upstream of the ribosomal binding site (RBS) of the LacZ gene. The aptamer insertion led to a theophylline-dependent expression of β-galactosidase, monitored by blue-white screen of bacterial colonies. The functionality of the RNA switch was further optimized by increasing the distance between the aptamer and RBS to eight nucleotides, resulting in a 10-fold activation ratio. Later, the region between the aptamer and the RBS was also randomized and varied in size from four to eight residues. It was observed that some of these variable sequences between the aptamer and RBS were able to mask the ribosomal binding site in hairpin like secondary structures where the aptamer sequence was involved. When the ligand bound to the aptamer, the aptamer folded into a structure that disrupted the original conformation of the construct. As a result, the RBS was freely accessible, allowing ribosome assembly and translation.

An RNA switch was also designed to control translation process from OFF to ON regulation. The E. coli TPP switch binds thiamine pyrophosphate (TPP) and masks the ribosomal binding site. This terminates translation of the downstream encoded ThiM gene that is involved in TPP production. In the unbound state, the RBS is freely accessible and ThiM is expressed. Later, 30
random nucleotides were inserted between the TPP switch and the RBS located downstream. To select for upregulation (gene expression induced by TPP binding), tetracycline resistance gene tetA was used as a reporter. In a dual-selection strategy, *E. coli* cells expressing constructs with a randomized insert were first grown in Tet medium in the presence of the TPP ligand, so that only tetA expression ensures survival. The resulting clones were then subjected to a second selection on NiCl₂ medium without ligand. Here, tetA expression is lethal, and only those clones that express tetA in a TPP-dependent way could survive. As a result, riboswitches that show a TPP-dependent upregulation were isolated.

### 1.2.2.2. Rational design method

Apart from strategies relying on randomized sequence elements with subsequent screening and selection, approaches based on educated guesses have proven to be equally promising. The pivotal factor for this approach is detailed knowledge of the switch mechanism. Genetic switches generally control the target gene expression by influencing either the transcription or the translation process. Bacterial transcriptional switches first encode aptamer domain. After encoding the aptamer, the switches restrict movement of the RNA polymerase for certain time period so that the RNA polymerase cannot encode the coding region of the target gene. This pausing gives the aptamer the necessary time to bind to the corresponding ligand and the regulatory domain to adopt the desired structural conformation. Through the structural changes, the transcriptional switches control the gene expression by forming either terminator structure or anti-terminator structure. In this competition between ligand binding and speed of RNA polymerization, the switch has to bind to its ligand very fast to decide between transcription and termination during the time when the movement of the RNA polymerase halts. Hence, the functionality of the transcriptional switch depends on the kinetics of RNA folding, ligand binding, and speed of transcription. In contrast, bacterial translational switches bind their target molecule and reveal or mask the ribosomal binding site after RNA synthesis is completed. For a rational design approach, these different regulatory principles of genetic switches have to be considered, especially when new synthetic expression platforms are designed.

### 1.3. BACTERIAL SENSOR WITH EMBEDDED READ OUT SYSTEM: A DIFFERENT FORM OF RNA SWITCH

Apart from developing bacterial RNA based switches that can tune a target gene expression, there are other RNA based devices that were engineered to detect
ligands in cells without influencing any gene expression. Such RNA structures or RNA sensors represent a special type of riboswitch, because their ligand dependent alternative structures allow the sensing of intracellular molecules like metabolites or proteins\textsuperscript{78-83}. Like RNA switches, bacterial RNA sensors are composed of a recognition unit or aptamer that binds to the target molecule and a transmitter module. But RNA sensors do not contain an expression platform to control a gene expression, instead they contain a signalling domain that binds a dye to generate a fluorescent output. There are many bacterial RNA sensors present, but the most prominent one is the “spinach aptamer containing sensor”. The sensor contains an \textit{in vitro} SELEX generated aptamer (Figure 3b) that can bind to 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI)\textsuperscript{84} (Figure 3a), a molecule mimicking the GFP chromophore. DFHBI is a nonfluorescent molecule but demonstrates 1000-fold increase in fluorescence upon binding to the aptamer (Figure 3c). The fluorescence signal looks like a GFP signal and hence the aptamer was assigned the name “spinach”. The spinach aptamer-DFHBI complex was implemented in bacteria to visualize the DFHBI signal inside cells (Figure 3d). Spinach aptamer exhibits signal in cells only in presence of DFHBI, but in absence of DFHBI the aptamer does not show any signal. To fabricate a sensor, the spinach aptamer was fused to an RNA aptamer (Figure 3e)\textsuperscript{85}. Here the RNA aptamer acts as a recognition unit. In absence of the target molecule, the spinach aptamer cannot fold into the right conformation and it cannot bind to the DFHBI molecule. When the RNA aptamer binds the target molecule, the whole switch architecture gets altered and the spinach aptamer gains its correct folding to bind to the DFHBI molecule and exhibits GFP like fluorescence (Figure 3f). The sensor with the spinach aptamer has been implemented to recognize adenosine, guanosine, ADP, GTP, or SAM in cells\textsuperscript{78}. 
Figure 3. Detection of small molecules with spinach sensor. (a) Chemical structure of DFHBI compound. (b) Secondary structure of the spinach aptamer. (c) DFHBI generates a GFP like signal upon binding with the spinach aptamer. (d) Activation of the DFHBI molecule by the spinach aptamer in cells. Control RNA sample does not contain any spinach aptamer, so it is unable to produce any GFP like signal. (e) Fabrication of spinach-based sensor by replacing the stem loop 3 of the spinach aptamer with the recognition module specific for the target compound. (f) Mechanism of the spinach sensor to detect a target molecule. Like spinach aptamer, malachite green (MG) aptamer was also used to detect target ligands. It was observed that the MG aptamer generates around 2000-fold increase in fluorescence when bound to its target fluorophore. Several sensors were fabricated by fusing the MG aptamer to ligand binding aptamer unit to detect molecules like theophylline, ATP and FMN.
1.4. OPTIMIZING SWITCH RESPONSE

When the first synthetic circuits were built, there were few options available for tuning circuits\textsuperscript{20,21}. However, to engineer a sophisticated genetic switch, the switch needs to be tuned to meet the specifications required for a particular application. For example, a large dynamic range may be required to control an important biological pathway. Similarly, low OFF states are desirable when expressing toxic proteins\textsuperscript{88}. New libraries of well-characterized parts and computational tools have made it easier to design and tune genetic circuits. There are mainly two strategies that can be adopted to hone the functionality of a bacterial switch – implanting genetic parts with desired characteristics in a switch and to understand the interactions between the switch components and the host cells.

1.4.1. Selecting genetic parts

Bacterial switches are often characterized by their response function, which captures how the output of the switch changes as a function of input. The shape of this function is defined by: (i) the ON and OFF states, which define the dynamic range of the corresponding switch, (ii) the amount of input required to reach the half-maximum output (also referred to as the threshold) and (iii) the cooperativity of the switch\textsuperscript{89,90}. The response of a switch can be regulated up or down by changing promoter strengths\textsuperscript{91}, RBS strengths or the degradation rates of proteins\textsuperscript{92}. The expression of different switch components is often regulated by incorporating those components in multiple plasmids at different copy numbers in order to synthesize each component at the necessary level.
Figure 4. Mechanism to regulate transcription process. (Top) Transcriptional activation where the promoter-bound TFs promote the recruitment of RNAPs, increasing the probability per unit time that a RNAP will bind. (Bottom) Transcriptional repression via steric inhibition, wherein one or more TFs physically block RNAP binding to the promoter or impede its progress along the template DNA strand. RNAP = RNA polymerase, O = operator site (TF binding sequence), TATA = RNAP binding sequence (TATA box in eukaryotes and archaea, −10 and −35 consensus sequences in bacteria), RBS = ribosome binding site.

Transcriptional activation can be triggered by promoter-bound transcription factors (TF) that promote the recruitment of RNA polymerase (RNAP), increasing the probability per unit time that a RNAP will bind. On the other hand, transcriptional repression can be induced by steric inhibition, wherein one or more TFs physically block RNAP binding to the promoter or impede its progress along the template DNA strand (Figure 4). In general, mechanisms for both activation and repression vary and can involve more complex actions including altering DNA secondary structure and recruiting additional coregulator proteins. A sequestering molecule can be expressed that binds a circuit component and prevents it from functioning. Sequestration has been achieved using sRNAs that bind to mRNA and proteins that bind to transcription factors.
1.4.2. Interactions between switch components and the host organism

Genetic switches often promote biochemical interactions within living cells. Most switches use host resources to function, including transcription and translation machinery (e.g., ribosomes and RNAP), DNA-replication equipment and metabolites. The availability of these resources depends on different strain backgrounds, environmental conditions, media, cell density and growth rate. It was observed that expression of some switch regulators slowed *E. coli* growth. Bacterial switches based on protein-protein interactions can also exhibit toxicity when the proteins bind to off-target partners. Small RNA with RBS-like sequences can also cause toxicity by titrating ribosomes (Figure 5). Complex genetic switches, which involve many genetic parts, are particularly sensitive to cell toxicity that can arise from individual regulators because their effects are compounded when they are expressed together in the host. Slow growth can put pressure on the host organism to evolve away the switch, via either homologous recombination, point mutations, deletions or copy-number reduction.
Figure 5. Cellular mRNAs (blue) associate with ribosomes at a rate dependent upon the concentration of the mRNA and ribosomes in that cell. (Top) Wild-type gene expression scenario. (Bottom) An exogenous rcRNA (Ribosome competing RNA) competes with bulk cellular mRNAs for translational machinery. A reduction in the number of available ribosomes results in a decreased probability of a given cellular mRNA associating with a ribosome. Switches can exhibit an unexpected output when they overuse a limited resource that is shared with other cellular processes. Overburdening resources causes queuing, which results in a delay or reduction in switch activity.

1.5. IMPORTANCE OF DIFFERENT REGULATORS, OTHER THAN RIBOSWITCH, WITH DIFFERENT REGULATORY MECHANISMS

The channelling of biological information through different components of a switch relies on regulators used in designing the switch. For example, DNA binding proteins can be incorporated in an engineered switch to recruit RNA polymerase or to impede the introduction of RNA polymerase. Modulation in RNA polymerase recruitment can influence the transcription of the genetic switch. The fate of the transcription of the switch can also be altered by CRISPRi system which includes Cas9 protein into a specific site of DNA and hamper the transcription process. Besides transcriptional modulators, there are regulators present that can also influence the bacterial translation process.

1.5.1. DNA binding proteins

There are many proteins present in bacteria that can bind to specific DNA sequences and subsequently alter the transcription process. These proteins can be implemented as regulators in a genetic switch by simply designing a promoter containing a specific sequence where the proteins can bind and influence the transcription (Figure 6a). These proteins can be designed by engineering zinc-finger proteins, transcription activator-like effectors (TALE), TetR and LacI homologs etc. Currently the library of DNA binding proteins has been significantly expanded. But the expansion of the protein library faces several challenges since the proteins need to be highly orthogonal so that they interact with only the required sequence and not with other switch components or endogenous factors. Several genetically encoded logic gates have been engineered using the DNA binding proteins.
1.5.2. CRISPRi

Clustered regularly interspaced short palindromic repeat (CRISPR) is a component of bacterial immunity system that binds to a specific sequence of a foreign DNA molecule and cleaves the DNA\textsuperscript{117}. This DNA degradation mechanism of CRISPR involves a Cas (CRISPR-associated) nuclease and a short guide RNA (sgRNA) that directs the Cas protein to its binding site\textsuperscript{118}. Apart from DNA cleavage, CRISPR can also bind a specific DNA sequence to knock down target gene expression. But unlike the DNA cleavage process, knocking down a target gene involves a mutant Cas protein which is called deactivated Cas9 or dCas9\textsuperscript{119,120}. The dCas9 protein does not cleave the target sequence, instead it just binds to the target sequence and inhibits the gene expression. When the dCas9-sgRNA complex binds to the non-template (NT) DNA strand of the 5’UTR or the protein coding region, it can silence the gene expression by blocking the elongating RNAPs. On the other hand, when the dCas9-sgRNA complex binds to the promoter sequence (e.g., the –35 or –10 boxes of the bacterial promoter) or thecis-acting transcription factor binding site (TFBS), it can block transcription initiation by sterically inhibiting the binding of RNAP or transcription factors to specific DNA sequence\textsuperscript{121} (Figure 6b). In general, the properties of CRISPRi switch resembles the switches composed of the above-mentioned DNA binding proteins. The stability of the dCas9-sgRNA-DNA regulatory complex attributes the CRISPRi switches to function on timescales similar to the protein-based switches.

1.5.3. Thermodynamically stable 5’UTR structures

Protein expression in bacteria can be controlled through modulation of mRNA stability. There are several factors that contribute to overall bacterial mRNA stability, for example accessibility by cellular ribonucleases, ribosome binding to the RNA and presence of different RNA secondary structures. Bacterial mRNAs contain short 3’ untranslated region (3’UTR) that usually harbours a stem-loop structure such as Rho-independent transcription terminator to stall the transcription process\textsuperscript{122-124}. The 5’UTR of bacterial mRNA contains ribosome binding site (RBS) and translation enhancer signals. There are several complex secondary structures, like G-quadruplexes, that can be present at the 5’UTR site and that can influence the translation initiation rate due to thermodynamic energy changes that affect ribosome binding. On the other hand, these stable 5’UTR structures can guard the mRNA from ribonuclease attack, rendering the mRNA
more stable\textsuperscript{125} (Figure 6c). Modulation of mRNA stability appears to be a useful approach to control gene expression.

1.5.4. Small RNAs (sRNAs)

sRNAs can regulate gene expression at the transcriptional and post-transcriptional levels\textsuperscript{126-128}. The sRNAs are classified as \textit{cis-} or \textit{trans-}encoded according to their genomic location regarding their target. \textit{Cis-}encoded sRNAs are encoded from the opposite DNA strand of their target coding sequence and have perfect complementarity with it. The \textit{trans-}encoded sRNAs are encoded in a different genomic location. They regulate gene expression by creating short base-pairing interaction with the target mRNA.\textit{Trans-}encoded sRNAs often require an RNA chaperone, such as Hfq or ProQ, for their stability in RNase environment\textsuperscript{129-133} (Figure 6d, top). But recently different RNA scaffolds have been implemented where the small RNAs are embedded into those scaffolds to shun RNase degradation\textsuperscript{134}.

Binding of sRNAs to the RBS site or to the flanking sites of the RBS typically regulates ribosomal access\textsuperscript{135} (Figure 6d, bottom) but may also influence mRNA stability\textsuperscript{136-138}. Pairing farther upstream of the RBS or downstream of the start codon was observed to affect translation. sRNAs can upregulate the target gene expression by translational activation or enhancing mRNA stabilization\textsuperscript{139}. Upon binding, sRNAs may promote translation by relieving secondary structures blocking the RBS.
**Figure 6.** Mechanisms of different regulators. (a) The switch (toggle switch) entails DNA binding proteins. Here, repressor 1 protein inhibits transcription from promoter 1 and is induced by inducer 1. Repressor 2 protein inhibits transcription from promoter 2 and is induced by inducer 2. (b) Mechanism of CRISPR interference system. (c) Stable 5’ UTR structure prevents exonuclease binding to mRNA. (d) (Top) RNA structure bound with Hfq protein. (Bottom) different positions at which sRNAs can block ribosome binding^{21,120,126}.

**1.6. GENETIC SWITCHES V 2.0**

At the initial phase of development of synthetic biology, very simple biodevices (like toggle switch)^{21} were engineered. These engineered devices can be categorized as version 1.0 in synthetic biology. As the progress in synthetic biology has been made, synthetic biologists have successfully engineered a wide range of functionality into artificial gene switches, creating genetic devices version 2.0. Few examples of such devices are oscillators^{140-142}, digital logic evaluators^{143,144}, sensors^{145-147} etc. Some of these advanced engineered gene networks have been implemented to perform useful tasks such as population control^{148}, genetic timing for fermentation processes^{149} and image processing^{150-
Synthetic biologists have even begun to address important medical and industrial problems with engineered organisms, such as bacteria that invade cancer cells.

But fabrication of the advanced gene switches requires thorough knowledge about the biological context in which the switch will be implemented. Usually, advanced gene networks should satisfy at least one of the following criteria: first, the switch should highlight design principles and provide modules that can be applied to the construction of other useful synthetic circuits; second, advance the tools available for novel scientific experiments; and third, enable real-world applications in different fields like medicine, industry, agriculture etc.

There are few notable advanced genetic switches presented below:

**1.6.1. Analog-digital converters**

Although the functionality of synthetic biological switches cannot be compared with the computing power of digital electronics, biological switches inspired by digital and analog electronics may significantly increase the programmability of biological behaviours. For example, biological analog-to-digital converters could translate external analog inputs, such as inducer concentrations, into internal digital representations for biological processing. Cells possessing analog-to-digital converters would be useful as biosensors in medical and environmental settings. Digital-to-analog converters, on the other hand, would translate digital representations back into analog outputs. Such systems might be useful in biotechnology applications, where reliable expression of different pathways is needed for programming different modes of operation in engineered cells. In addition, digital-to-analog converters may be useful in providing an approach for probing synthetic circuits. For example, because each analog level is associated with a distinct digital state, a single analog output can allow one to infer the internal digital state of a synthetic gene network.

**1.6.2. Biological containment switch**

Biological containment is necessary for ensuring that genetically modified organisms do not spread throughout the natural environment. The containment can be achieved by either passive or active techniques. In passive containment, cells are engineered to be dependent on exogenous supplementation to compensate for gene defects; whereas in active containment, cells are engineered to directly express toxic compounds when located outside their target environments.
1.6.3. Whole-cell biosensors

Programmable cells that act as whole-cell biosensors have been created by interfacing engineered gene networks with the cell’s natural regulatory circuitry\textsuperscript{147} or with other biological components, such as light-responsive elements\textsuperscript{150,151}. The development of novel or reengineered sensory modalities and components would expand the range of applications that programmable cells could address. In addition, magneto-responsive bacteria could play useful roles in environmental and medical applications\textsuperscript{154}. Synthetic bacteria, designed to form magnetosomes and seek out cancer cells, could be used to enhance imaging, and magnetic bacteria could be engineered to interact with nanoparticles to enhance the targeting of cancer cells. Ultimately, programmable cells possessing novel sensory modules could be integrated with mechanical, electrical and chemical systems to detect, process and respond to external stimuli, and exploited for a variety of environmental and biomedical applications. For example, bacteria could be engineered to seek out hazardous chemicals or heavy metals in the environment, perform cleanup and return to their origin to report on the number of hazardous sites encountered via analysis by microfluidic devices.

1.6.4. Synthetic genetic oscillator

Experimental and computational strategies have been adopted to develop robust genetic oscillators with tunable periods\textsuperscript{141,142,155}. These synthetic oscillators have significant utility in biotechnology applications, such as in the synthesis and delivery of biologic drugs. Bacteria that reside in the human gut can be engineered to synthesize an active drug at fixed time intervals. To realize such an application, one would need to develop and implement a signalling circuit for synchronizing synthetic genetic oscillators\textsuperscript{156,157}. Such circuits could be fabricated with modular components from bacterial quorum sensing systems.

In synthetic biology, genetic switches with pulse-based processing may open up exciting new methods for encoding information in engineered cells. For example, instead of transmitting information between cells by means of absolute levels of quorum-sensing molecules, the frequency of a robust genetic oscillator could be modulated. This might be useful in delivering information over longer distances, as frequency information may be less susceptible to decay over distance than absolute molecule levels.
1.7. APPLICATIONS OF SYNTHETIC GENETIC SWITCHES

Human health is afflicted through different ways including emergent drug-resistant microbes, cancer, obesity etc. Synthetic biology has made promising strides in overcoming all those health burdens. The rational and model-guided construction of biological parts is enabling new therapeutic platforms, from the identification of disease mechanisms and drug targets to the production and delivery of small molecules.

1.7.1. Unravelling disease mechanisms

Synthetic switches are being implemented to reconstruct natural biological systems to decipher how pathological behaviours may emerge. For example, researchers developed a synthetic testbed by systematically reconstructing the various components of the human B cell antigen receptor (BCr) signalling pathway\textsuperscript{158}. This allowed them to identify network topology features that trigger BCr signalling and assembly. A rare mutation in the immunoglobulin-\(\gamma\)-encoding gene was identified in one patient and introduced into the synthetic system in which it was shown to interfere the assembly of the BCr on the cell surface, thereby linking this faulty pathway component with disease onset. Pathogenic viral genomes can similarly be reconstructed for studying the molecular underpinnings of infectious disease pandemics.

1.7.2. Discovering new drug targets

Synthetic biology devices can be implemented to systematically probe the function of individual components of a biological pathway. For instance, modular RNA based regulators can be engineered to tune the expression of a toxic protein or any gene in a biological network\textsuperscript{159}. Engineered switches have been used to tightly regulate the expression of CcdB, a toxic bacterial protein that inhibits DNA gyrase, to comprehend the sequence of events leading to induced bacterial cell death\textsuperscript{160}. These synthetic biology studies, in conjunction with systems biology studies of quinolones (antibiotics that inhibit gyrase), led to the discovery that all major classes of bactericidal antibiotics induce a common cellular death pathway by stimulating oxidative damage\textsuperscript{161,162}. This approach offered new insights into how bacteria respond to lethal stimuli and paved the way for the development of more effective antibacterial therapies.
1.7.3. Delivery of therapeutic drugs

Synthetic genetic switches can be used for the controlled delivery of drugs as well as for gene therapy. In some cases, control over kinetics of drug release in the body may yield therapeutic advantages and reduce undesired side effects. Most hormones in the body are released in time-dependent pulses. Periodic synthesis and release of drugs can be autonomously achieved with synthetic oscillator circuits\textsuperscript{115,163,140-142}. Apart from this, synthetic switches offer a more controlled approach to gene therapy, such as the ability to dynamically silence, activate and tune the expression of desired genes. Additionally, the construction of non-native pathways offers a unique approach to gene therapy, such as for the treatment of metabolic disorders.

1.7.4. Agriculture and the environment

To support the production of isobutanol, a next-generation biofuel, synthetic gene circuits conferring ionic liquid resistance can increase the robustness and survival of bacterial production strains\textsuperscript{164}. Improved productivity for feedstock and raw materials has also been achieved by using a synthetic cell–cell communication strategy consisting of a fungal specialist population converting lignocellulosic biomass into soluble saccharides and a bacterial fermentation specialist that metabolizes soluble saccharides into desired products, such as isobutanol\textsuperscript{165}. Such synthetic consortia, in which a complex task is divided among multiple subpopulations that perform simple tasks with high robustness, are extremely effective for achieving high workforce productivity. To achieve control over reproduction of agricultural animals, a synthetic communication system was applied to achieve ‘synthetic artificial’ insemination\textsuperscript{166}.

1.8. THESIS MOTIVATION AND OVERVIEW

Synthetic biology seeks a variety of experimental approaches that either tend to modify or mimic biological systems. After early successes in engineering bacterial genetic switches, several genetic parts have been assembled together in different combinations to explore vast amount of new orthogonal switches, especially riboswitches. Chapter 1 provides an overview about the fundamental ingredients to build a synthetic bacterial genetic switch and the general procedure to fabricate those ingredients. Those ingredients or genetic parts play a pivotal role in channelling biological information through a genetic switch. But despite of mimicking biological functions, the synthetic switches often fail to provide the desired output due to unwanted interactions with different cellular components of
the host cell or by harbouring some regulatory components which are not characterized properly. Hence, proper knowledge and characterization of the genetic components are necessary to engineer an efficient genetic switch. Even though riboswitches are well-studied and exhibit high orthogonality, there are other regulators present which confer different regulatory mechanisms. By combining different regulatory mechanisms, it has been possible to develop many advanced genetic switches to perform important tasks, which would have been otherwise impossible to perform with earlier and simpler genetic switches. However, most of the RNA based genetic switches are cis encoded, meaning they are present in the same target mRNA to control target gene expression. Consequently, it is necessary to modify the 5’UTR region of the target mRNA to accommodate the switch making it difficult to target genomic DNA. To overcome this burden, in chapter 2, we have engineered a trans encoded RNA switch which mimics a tRNA architecture (tRNA mimicking switch or TMS switch). The TMS switch exhibits modularity and high orthogonality. The switch demonstrates its functionality with different types of inputs (RNA, small molecule and protein). Due to its high modular structure, the switch functioned as an OR logic gate with small molecule and RNA as inputs. Since the switch is trans encoded, it can easily modulate a gene present in the bacterial genome.

Apart from using small molecules and proteins as inputs, we also controlled gene expression by light. In complement to exploiting optogenetic tools, photochemical strategies mostly rely on the incorporation of photo-responsive small molecules into the corresponding biomacromolecular scaffolds. Therefore, generally large synthetic effort is required and the reversible switching of gene expression within a single system remains a challenge. In chapter 3, we report the implementation of the TMS switch with small photo-switchable signalling molecules as inputs. The signalling molecules consist of two aminoglycoside molecules that are connected via an azobenzene unit. The light responsiveness of our system originates from the photo-switchable noncovalent interactions between the signalling molecule and the TMS switch, leading to the demonstration of photochemically controlled expression of two different genes. We believe that this modular design will provide a powerful platform for controlling the expression of other functional proteins with high spatiotemporal resolution employing light as a stimulus. Since most current bacterial switches are RNA structures, they cannot respond to inputs that bind to ssDNA. Hence, in chapter 4, we have overcome this challenge by developing the first trans encoded ssDNA switch. This switch again relies on mimicking a tRNA architecture and is
based on multi-modular gene control. This ssDNA switch is highly efficient and strictly regulated by the input with up to 150-fold change in gene expression levels. Moreover, it can accommodate small molecules and RNA as inputs. This approach broadens the application of switches in synthetic biology by making the system more sophisticated and flexible regarding the type of inputs. Fluorescence-based probes provide a unique insight into the otherwise obscure internal processes of living cells. However, the design of probes selective for novel analytes is notoriously challenging, especially where substrate-binding proteins are not available. In chapter 5, we present an entirely genetically encoded RNA-based sensor with a modular design feature that allows straightforward repurposing towards new analytes. We obtain a novel aptamer for GFP by a SELEX-type evolution directly in *E. coli*, which we conjugate to a second aptamer yielding a genetically encoded sensor. The sensor does not require exogenous ligands and is instead based on turn-on fluorescence as analyte-binding releases GFP. We validated the method versus biologically relevant small molecules and proteins in *Escherichia coli* by simple RNA-aptamer replacement. This method has the potential to be a standard tool for analyte detection in cells.

References


