Engineering and Testing RNA-circuits in Cell-Free Systems

TECHNISCHE UNIVERSITÄT DARMSTADT

Dissertation von François-Xavier Lehr

Erstgutachter: Prof. Dr. techn. Heinz Koeppl

Zweitgutachter: Prof. Dr. H. Ulrich Göringer

Vom Fachbereich Biologie
TECHNISCHE UNIVERSITÄT DARMSTADT

zur Erlangung des Grades
Doktor rerum naturalium (Dr. rer. nat.)
genehmigte Dissertation

Darmstadt February 2021

Lehr, François-Xavier : Engineering and testing RNA-circuits in cell-free systems,
Technische Universität Darmstadt,
Jahr der Veröffentlichung der Dissertationauf TUprints: 2021
Tag der mündlichen Prüfung: 30.04.2020

Ehrenwörtliche Erklärung:

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit entsprechend den Regeln guter wissenschaftlicher Praxis selbstständig und ohne unzulässige Hilfe Dritter angefertigt habe. Sämtliche aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie sämtliche von Anderen direkt oder indirekt übernommenen Daten, Techniken und Materialien sind als solche kenntlich gemacht. Die Arbeit wurde bisher bei keiner anderen Hochschule zu Prüfungszwecken eingereicht.

Ort,Datum	Unterschrift

Acknowledgements

I sincerely express my gratitude to my supervisors Prof. Heinz Koeppl and Prof. H. Ulrich Göringer, who gave me the opportunity to prove myself in the field of synthetic biology, and support me through the various challenges of experimental research. I also thank Prof. Chase Beisel and Prof. Viktor Stein for being on my PhD committee.

I would like to thank my colleagues from the Bioinspired Communication Systems Lab, who all contributed to create such a unique and exciting research environment: Dominik Linzner, Bastian Alt, Tim Prangemeier, Christian Wildner, Sikun Yang, Nikita Kruk, Kilian Heck, Lukas Köhs, Leo Bronstein, Nurgazy Sulaimanov, Jascha Diemer, Dr. Christiane Hübner, Klaus-Dieter Voss, Markus Baier, Christine Cramer, Adrian Šošić, Felix Reinhardt, Nikolai Engelmann. In particular, I would like to thank Maleen Hanst, Megan Baley, Werner Kleindiest, Faisal Islam, and Katharina Decker for their collaboration and precious help and knowledge in the wet-lab.

I am grateful for the technical and friendly help received from the neighbored labs: Pr. Johannes Kabisch, Pr. Beatrix Suess, Alexander Gräwe, Niels Schlichting, Thomas Zoll, Silke Hackenschmidt, Marc Vogel, Florian Gröher (highly skilled for 3D-printing!), and Cristina Bofill-Bosch.

Last but not least, special thanks go to the friends who have been on my side throughout those years. John, for throwing clods of butter into the abyss; Sarah, for her fierce meowing; Diane, feta-maniac, for competing with my hypochondria; Mathieu, forever bamboleo; Feli, for lightening up everything, Mariana and Georgio, the Manula and Manuli of Science, Genia, and Adrien, whom his chronic gros-lardisme only equals his unfailing friendship. I also thank my parents, as well as the Krak's inhabitants: Anne, Chris and Nathan.

And finally, a very special Dankeschön to my stubborn Zauberin!

Abstract

RNA molecules lie at the heart of living organisms where they are associated with most of the cellular processes. They have recently emerged as one of the most promising elements for developing programmable genetic regulatory systems. RNA regulators have been shown to offer great advantages to harness the power of synthetic biology. Versatility of functions, predictability of design, and light metabolic cost have turned RNA-based devices into components of primordial importance for therapeutic, diagnostic and biotechnological applications. However, advanced tasks require the use of sequential logic circuits that embed many constituents in the same system. Combining RNA-parts into more complex circuits remains experimentally challenging and difficult to predict. Contrary to protein-based networks, little work has been performed regarding the integration of RNA components to multi-level regulated circuits.

In the first part of this thesis, combinations of variety of small transcriptional activator RNAs (STARs) and toehold switches were built into highly effective AND-gates. To characterise the components and their dynamic range, an *Escherichia coli* (*E. coli*) cell-free transcription-translation (TX-TL) system dispensed via nanoliter droplets was used. Cell-free systems, which constitute an open environment, have removed many of the complexities linked to the traditional use of living cells and have led to exciting opportunities for the rational design of genetic circuits. A modelling framework based on ordinary differential equations (ODEs), where parameters were inferred via parallel tempering, was established to analyse the expression construct in a qualitative and quantitative manner. Based on this analysis, nine additional AND-gates were built and tested *in vitro*. The functionality of the gates was found to be highly dependent on the concentration of the activating RNA for either the STAR or the toehold switch. All gates were successfully implemented *in vivo*, displaying a dynamic range comparable to the level of protein circuits. Subsequent spacer screening experiment enabled isolation of a gate mutant with dynamic range up to 1087 fold change, paving the way towards multi-layered devices where tight OFF-stages are required for efficient computation. Expanding the repertoire of RNA regulatory parts with efficient inhibitors would complete the set of logic operations necessary for the building of dynamic

circuits, such as memory devices or oscillators. The TX-TL system was functionalized with pre-expressed dSpyCas9, a mutated version of Cas9 without endonuclease activity. Four functional small guide RNAs (sgRNAs) targeting the sfGFP reporter were engineered and characterized, all resulting in high repression efficiency. A three-inputs logic circuit containing toehold, STAR and sgRNA was successfully co-expressed, validating the orthogonality of NOT and AND gates based solely on RNA-based regulation. In order to minimize interactions which could arise from RNA-circuit of increasing complexity, the TX-TL system was functionalized with a second protein, the Csy4 endoribonuclease, which selectively binds and cuts a small RNA hairpin. Normalization of gene expression from various untranslated region contexts and enhanced processing of three-inputs small RNA operon were demonstrated via the use of Csy4. Finally, characterizing complex RNA-based circuits requires techniques that resolves dynamics. To overcome the batch-format limitations inherent to TX-TL systems, a microfluidic nanoliter-scaled reactor was implemented, enabling synthesis rates to stay constant over time. Dynamic control of RNA circuitry was demonstrated by modulating the concentration of ligands, reversing the gene state through the conformational change of riboswitches.

This thesis shows the potential of a rapid prototyping approach for RNA circuit design in TX-TL systems combined with a predicting model framework. Taken together, the characterization of a variety of RNA-parts : activators, repressors, or controllers culminating into logic modules; and augmented cell-extracts; form a complete RNA-toolbox for cell-free systems. The leveraging of this unique prototyping platform will ultimately enable the engineering and the study of highly dynamical RNA-circuits *in vitro*.

Zusammenfassung

RNA-Moleküle bilden ein Kernstück lebender Organismen, wo sie mit den meisten zellulären Prozessen in Verbindung gebracht werden. Sie sind in jüngerer Zeit hervorgetreten als eines der vielversprechendsten Elemente für die Entwicklung programmierbarer genetischer Regulationssysteme. RNA-Regulatoren bieten große Vorteile dabei, die Kraft der synthetischen Biologie zu nutzen. Durch die Vielseitigkeit ihrer Funktionen, der Vorhersagbarkeit beim Design und ihrer geringen Stoffwechselkosten sind RNA-basierte Elemente von grundlegender Bedeutung für therapeutische, diagnostische und biotechnologische Anwendungen. Fortgeschrittene Aufgaben erfordern jedoch die Verwendung von sequentiellen Logikschaltungen, die viele Bestandteile in dasselbe System einbetten. Das Kombinieren von RNA-Komponenten zu komplexeren Schaltkreisen bleibt experimentell herausfordernd und schwierig vorherzusagen. Im Gegensatz zu proteinbasierten Netzwerken wurden nur wenige Arbeiten zur Integration von RNA-Komponenten in mehrstufig geregelte Schaltkreise durchgeführt. Im ersten Teil dieser Arbeit wurden Kombinationen verschiedener kleiner Transkriptionsaktivator-RNAs, *small transcriptional activator RNAs* (*STARs*), und *Toehold*-Schalter in hochwirksame AND-Gatter eingebaut. Zur Charakterisierung der Komponenten und ihres Dynamikbereichs wurde ein zellfreies *Escherichia coli* (*E. coli*) -Transkriptions-Translations-System (TX-TL) verwendet, das über Nanolitertröpfchen verteilt wurde. Zellfreie Systeme, die eine offene Umgebung darstellen, haben viele der Komplexitäten, die mit der traditionellen Verwendung lebender Zellen verbunden sind, beseitigt und aufregende Möglichkeiten für die rationale Gestaltung genetischer Schaltkreise eröffnet. Zur qualitativen und quantitativen Analyse des Expressionskonstrukts wurde ein auf gewöhnlichen Differentialgleichungen, *ordinary differential equations* (*ODEs*), basierendes Modellierungsframework erstellt, dessen Parameter durch parallel tempering abgeleitet wurden. Basierend auf dieser Analyse wurden neun zusätzliche AND-Gatter gebaut und *in vitro* getestet. Es wurde festgestellt, dass die Funktionalität der Gatter stark von der Konzentration der aktivierenden RNA für den *STAR*- beziehungsweise den *Toehold*-Schalter abhängt. Alle Gatter wurden erfolgreich *in vivo* implementiert und wiesen einen Dynamikbereich auf, der mit dem

von Proteinkreisläufen vergleichbar ist. Ein anschließendes *Spacer-Screening-Experiment* ermöglichte die Isolierung einer Gatter-Mutante mit einem Dynamikbereich von bis zu 1087-facher Veränderung, womit der Weg geebnet wurde für mehrschichtige Bauelemente, bei denen enge *OFF-Level* für eine effiziente Schaltung erforderlich sind. Die Erweiterung des Repertoires der RNA-Regulatoren um wirksame Inhibitoren würde die für den Aufbau dynamischer Schaltkreise, wie Speicherbausteine oder Oszillatoren, erforderlichen logischen Operationen vervollständigen. Das TX-TL-System wurde mit vorexprimiertem dSpyCas9, einer mutierten Version von Cas9 ohne Endonukleaseaktivität, funktionalisiert. Vier funktionelle *Small-Guide-RNAs* (*sgRNAs*), die auf den *sfGFP-Reporter* abzielen, wurden konstruiert und charakterisiert und wiesen eine hohe Repressionseffizienz auf. Eine Logikschaltung mit drei Eingängen, die *Toehold, STAR* und *sgRNA* enthielt, wurde erfolgreich co-exprimiert und validierte die Orthogonalität von NOT- und AND-Gattern ausschließlich auf der Grundlage der RNA-basierten Regulation. Um Wechselwirkungen zu minimieren, die durch einen immer komplexer werdenden RNA-Schaltkreis entstehen könnten, wurde das TX-TL-System mit einem zweiten Protein, der Csy4-Endoribonuklease, funktionalisiert, welches selektiv eine kleine RNA-Haarnadel bindet und schneidet. Die Normierung der Genexpression verschiedener nicht translatierter Regionskontexte und die verbesserte Verarbeitung von kleinen RNA-Operons mit drei Eingängen wurde mithilfe von Csy4 demonstriert. Zuletzt erfordert die Charakterisierung komplexer RNA-basierter Schaltkreise Techniken um Dynamiken zu erfassen. Um die mit TX-TL-Systemen verbundenen Einschränkungen des Batch-Formats zu überwinden, wurde ein mikrofluidischer Nanoliter-Reaktor implementiert, der es ermöglicht, die Syntheseraten über die Zeit konstant zu halten. Die dynamische Kontrolle der RNA-Schaltkreise wurde durch Modulation der Ligandenkonzentration und Umkehrung des Genzustands durch Konformationsänderung von Riboschaltern demonstriert. Diese Arbeit zeigt das Potenzial eines *Rapid-Prototyping*-Ansatzes für das Design von RNA-Schaltkreisen in TX-TL-Systemen in Kombination mit einem Modellierungsframework. Zusammengefasst bilden die Charakterisierung einer Vielzahl von RNA-Teilen wie Aktivatoren, Repressoren oder Controllern, die in Logikmodulen gipfeln, sowie funktionserweiterte Zellextrakte, eine komplette *RNA-Toolbox* für zellfreie Systeme. Die Nutzung dieser einzigartigen Prototyping-Plattform wird es letztendlich ermöglichen, hochdynamische RNA-Schaltkreise *in vitro* zu konstruieren und zu untersuchen.

Table of contents

List of figures

List of tables

Nomenclature

Acronyms / Abbreviations

RBS Ribosome Binding Site

CRISPR Clustered Regularly Interspaced Short Palindromic Repeats

PAM Rrotospacer Adjacent Motif

crRNA CRISPR RNA

TX-TL Transcription-Translation

MCMC Markov chain Monte Carlo

GFP Green Fluorescent Protein

sfGFP Superfolder Green Fluorescent Protein

RNA Ribonucleic Acid

RNAP RNA Polymerase

MGA Malachite Green Aptamer

miRNA MicroRNA

MITOMI Mechanically Induced Trapping of Molecular Interactions

UTR Untranslated Region

3-PGA 3-Phosphoglyceric Acid

AA Amino Acid

ADP Adenosine Diphosphate

ATP Adenosine Triphosphate

Bp Base pair

DFHBI 3,5-DiFluoro-4-HydroxyBenzylidene Imidazolinone

DNA Deoxyribonucleic Acid

GUI Graphical User Interface

Kb Kilo base pairs

MFE Minimum Free Energy

mRNA Messenger RNA

NTP Nucleoside Triphosphate

OD Optical Density

ORI Origin of Replication

PCR Polymerase Chain Reaction

PDMS Polydimethylsiloxane

PEP Phosphoenolpyruvate

Pi Inorganic Phosphate

PURE Protein synthesis Using Recombinant Elements

PWM Pulse Width Modulation

SNR Signal to Noise Ratio

TCA Tricarboxylic Acid Cycle

TFs Transcription Factors

TPP Thiamine Pyrophosphate

tRNA Transfer RNA

W/o Water-in-oil

CFPS Cell-Free Protein Synthesis

RFP Red Fluorescent Protein

sgRNA Single Guide RNA

sRNA Small RNA

STAR Small Transcriptional Activator RNA

TFs Transcription Factors

tracrRNA Trans-activating RNA

Chapter 1

Introduction

1.1 Features of the Synthetic Biology Approach

The term 'Synthetic Biology' designates a field of research whose establishment in the scientific community has been reported in the early 2000s. It has been designed as the '21st century's fastest growing field of research' and has led to many promises in terms of technology output as well as basic fundamental research [24]. Synthetic biology is not a new discipline per se but rather built upon the more traditional fields of genetic engineering and molecular biology [18]. However, those fields were usually limited to editing one or a couple of genes. Synthetic biology aims to use genes and proteins as building blocks to enable the creation of new functional modules, and ultimately brought together to make very significant changes in living cells [23]. Performing those heavy modifications has been made possible by the continuous emergence of novel techniques that has shaped the field of genetic for the last half century [115]. The first milestone happened with the discovery of restriction enzymes, and their use for cutting DNA at specific sequence patterns, which led thereafter to the insertion of the insulin human gene into bacterial cells. Since the first restriction enzyme, HindII, discovered by Hamilton Smith in 1968, many others have follow, as well as other breakthrough technologies which are emerging since the late 1990's at an accelerated pace [93]. Affordable DNA synthesis and sequencing, high-throughput methods, and novel cloning strategies enabling the tuning of genes at a large scale have all contributed to the emergence of synthetic biology as described nowadays.

Those methods have all contributed to the emergence of numerous applications, whose some of them are already commercially available or near-market production. They have shown to cover a very wide and diverse area of functions. Bio-sensing, stem-cells repro-

gramming, point-of-care testing, engineered-cell therapies or bio-production of recombinant enzymes are a small sample list of what the field has to offer [56]. In 2005, Levskaya et al. [68] implemented a set of memory devices into bacteria for creating living photograph based on light-responsive toggle switches producing pigment. Paddon and Keasling [89] successfully implemented all the chemical pre-cursors to the antimalarial drug artemisinin into yeast to produce the drug at lower cost in a most reliable manner. Multiple other enzyme engineering projects are now commercially available or near market production, such as bio-diesel, bioisoprene (for rubber), or resveratrol [131, 137]. In 2015, the first complete synthesis and assembly of a minimal bacterial genome containing 473 genes has been successfully implemented, marking a major milestone for synthetic biology and opening promising possibilities to tackle a multitude of current research problems [52]. This year, less than 20 years after the first sequencing of the whole *E. coli* genome, Fredens et al. [36] demonstrated the full synthesis of the entire *E. coli* genome, recoding 18,214 codons to create an organism with a 61-codon genome, enabling the deletion of a previously essential transfer RNA. Finally, investigation of genetic programming through synthetic circuits engineering promise new treatments for incurable or complicated disease treatments. The creation of gene circuits able to trigger their components in a sequential manner, with diverse expression patterns such as oscillations, or with sophisticated time-dependent transcription profiles, bring hopes to program cells with therapeutic behaviors [127].

1.1.1 The Limits of the Engineering Approach

The engineering paradigm brings together most of those projects under the synthetic biology banner. Several authors have highlighted the importance of the engineering approach for synthetic biology. An example statement is: "Synthetic biology ... applies engineering formalism to design and construct new biological parts, devices, and systems for novel functions or life forms that do not exist in nature" [37]. The idea behind such statement is to bring the full methodology usually reserved to the classical engineering field (mechanical, information technology, etc...) to the biological branch. It traditionally involves a cycle of several steps which can be iterated to improve the overall performance of the final product [25]. Those steps are:

1. An initial design phase using predictive modeling
2. Optimizing designs based on standard components simulations
3. Hierarchical design

4. Fabrication using reliable and robust equipment

5. Testing of the finished assemblies using standardised protocol and devices

This protocol would present several advantages when transferred to synthetic biology. In particular, the use of standardized and well-characterised components would enable the fabrication of module libraries, which would, therefore, be themselves standardised for further use in unrelated or related projects. Similar to what electronic components are today to the television or computers, synthetic biology could take advantage of pre-written libraries of DNA circuits. The first major attempt in this direction has been the establishment of the Registry of Standard Biological Parts based on well-characterised 'Biobricks' comprised of elementary gene circuit such as promoters, fluorescent reporters and terminators [112]. This idea of using standard parts would confer major advantages to the field of biology. First, it enables the prototyping of simplified designs where each part has its own described characteristics and response behaviors to a sample of standardised conditions. Second, it would spare a tremendous amount of time and money as the bio-engineers could select a standard device or module in a numerous number of different projects, without having to re-invent and carefully characterise again the basic parts of their product. Third, the components of the system could easily be swapped in a plug-and-play configuration, so if a new better-performing part is released, any current working device could be updated without having to re-engineer the whole assembly [74]. Those advantages are well-established for the field of traditional engineering where the properties of modules such as electronic transistors or IT software libraries are well-known.

However, it has been pointed out that actually only a very small fraction of successful synthetic biology projects, as for instance the ones mentioned above, have demonstrated the use of the engineering pipeline. Davies [25] wrote in a recent opinion review that nine out of the ten most prominent synthetic biology projects have not been using standard components or predictive models. In other terms, their projects, despite being marketable, do not exhibit the classical features expected in a traditional electrical engineering enterprise, such as modularity of the components along with their characterisation in defined standardised conditions. Most of the projects resulted from a 'Design-Test-build' paradigm, where optimisation were mainly made by testing multiple iterating versions of the wet-ware constructs without any input coming from computer models [94]. Consequently, randomisation and selection of the best circuits is the current production pipeline of many synthetic biology projects.

This major discrepancy between the definition of synthetic biology and how it is usually performed emerges generally from the system components. Engineering genetic circuits

are usually derived from natural molecules, based on nucleic acid sequences encoding for diverse type of regulators or enzymes. Those sequences are still very poorly understood. For instance, it is now reasonable to predict the yield of a protein product from a given nucleic acid sequence, by taking into account known parameters such as the promoter or ribosome binding site (RBS) strengths. However, many unknown parameters such as sequence genomic context or unknown intragenic motifs could alter those predictions, requiring the need for these sequences to be experimentally verified [8]. Moreover, expressing synthetic circuits or producing recombinant proteins typically involves host cells, for which their whole understanding is still a very rough draft compared to the extent of their complexity [98]. The living host, which can range from bacteria to human cells, possess a very dense amount of endogenous molecules, for which their functionality is not always assessed. Even when their roles and structures are characterised, there is currently no convincing or systemic solution to predict potential cross-talks between the implemented gene circuit and the host metabolic network.

In order to circumvent most of the limitations imposed by cell-based processes, bio-engineers have recently turned to alternative solutions where they could synthesise proteins and implement genetic networks in an environment decoupled from the host complexity.

1.2 Cell-Free Systems for Synthetic Biology

Before the massive increase of popularity that cell-free transcription-translation systems (TX-TL) have known during the last decade, early cell-free protein synthesis (CFPS) experiments were reported back in the late 60s. One of the most iconic early example was its use to crack the first codon of the genetic code in 1961 by Nirenberg and Matthei [84]. Cell-free systems offer advantages closer to engineer objectives compared to the drawbacks of cell-based processes described in the previous section. Cell-free systems offer an open reaction environment for biotechnology where the easy addition or removal of reactants increases dramatically its flexibility compared to *in vivo* systems [45].

1.2.1 PURE Systems vs. Extract-based Lysates

The two main advantages of using TX-TL systems are linked to the decoupling of the implemented circuit from their host. First, as the complexity of the biological circuit increases,

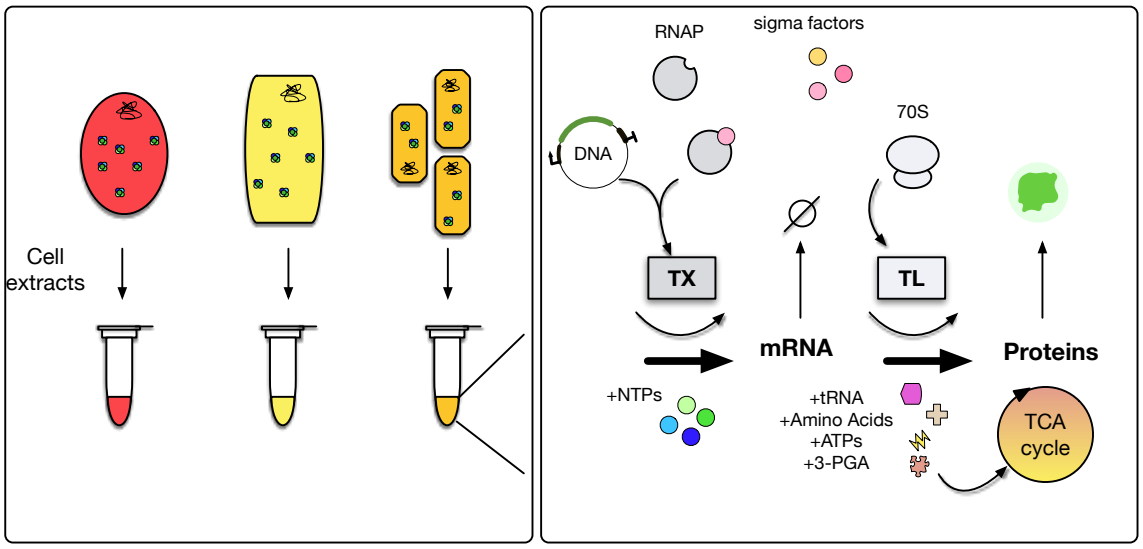


Fig. 1.1 Overview of lysate-based TX–TL systems. Cell-free crude extracts can be obtained from the lysates of various organisms (left). TX and TL processes require the supplementation of additional NTPs and substrates (ATP, 3-PGA, tRNA and amino acids) in order to sustain the reactions for an extended period of time (right).

the resource competition increases between the growth of the host and the synthesis of the circuitry components [69]. The resources become scarce, and a trade-off has to be found by the cell which will ultimately limit the synthetic circuit output [110]. Hence, in TX-TL systems, the cell viability is not taken into account anymore and all available resources can be redirected towards the circuit usage.

Second, as discussed earlier, the implementation of genetic networks into living cells may create unwanted interactions between the endogenous pathways and the circuit components [80]. Cell-free extracts do not possess any part of the host DNA anymore, reducing therefore the complexity of the number of possible interactions. This enables in turn the development of quantitative models describing more accurately the implemented networks.

Engineers can turn towards the use of two main categories of cell-free systems. The first one, considered as a "top-down" approach, is based on crude extracts from various organisms ranging from microbial to eukaryotic cell lines (Fig. 1.1). They are usually obtained by centrifuging the cell-lysate to remove the unwanted parts such as DNA, debris and unlysed cells while keeping intact the cell's machinery. The most popular by far is the one based on *E. coli* extract which possesses numerous advantages such as low-cost preparation, very high-batch yields and a wide range of applications from diagnostic platforms to antibodies production [81]. Other extracts based on multi-cellular organisms, e.g. insect, HeLa, or *Leishmania tarentolae* cells [33, 61, 129], have usually the disadvantage of having low batch

yield with long and difficult extract preparation. They possess, however, specific applications such as the production of complex proteins with post-translational modifications [78, 88].

The second category, considered as a "bottom-up" approach, is a cell-free system described as protein synthesis using purified recombinant protein elements (PURE) [104]. In the PURE system, each molecular component required for the transcriptional and translational processes in *E. coli* is independently purified and subsequently assembled together. This method offers a better control over the bio-chemical composition of TX-TL system than their extract-based counter-parts and can therefore be seen as an ideal environment to quantitatively describe the ensemble of system species. However, the PURE system is significantly more costly to produce because of the numerous purification steps needed to obtain the isolated components. Recently, Lavickova and Maerkl [64] developed a 'one-pot' method which considerably decreases both time and cost of the PURE production. However, the difficulty to use the endogenous host metabolism and proteome in PURE systems compared to lysate-based systems may limit the functionality of synthetic circuits.

CFPS technologies have enabled the bio-production of diverse recombinant proteins, with scales ranging from an Eppendorf tube to a reactor of 10^6 L, with yield reaching up to 9.7 g/L of Green Fluorescent Protein (GFP) in wheat germ extract [31]. In this thesis, the focus is made on the use of TX-TL reactions towards the prototyping and implementation of novel genetic networks. Hence, optimising the yield of the cell-free reactions will not be the main focus but plays an important role in enabling the co-transcription and co-translation of multiple components with concentrations high enough to be detected with fluorescence plate-reader or microscope assays.

1.2.2 *E. coli*-based TX-TL Systems

E. coli cell-free system is the platform that is most often used for *in vitro* synthetic biology applications [124]. It demonstrated potential for prototyping and testing novel genetic parts, allowing rapid characterisation of novel regulatory elements for further *in vivo* implementations. Several aspects of *E. coli* cell-free systems make them particularly suitable for circuit prototyping [85]. A couple of toolboxes of regulatory elements have been recently published for this extract, for which dozen parts of various transcription factors, promoters, or fluorescent reporters have been thoroughly characterised [106, 39]. Moreover, the catalogue of parts is still increasing, with for instance the recent addition of the CRISPR/Cas9 system or the evaluation of multiple ribo-regulators [72, 101]. A main concern related to the use of

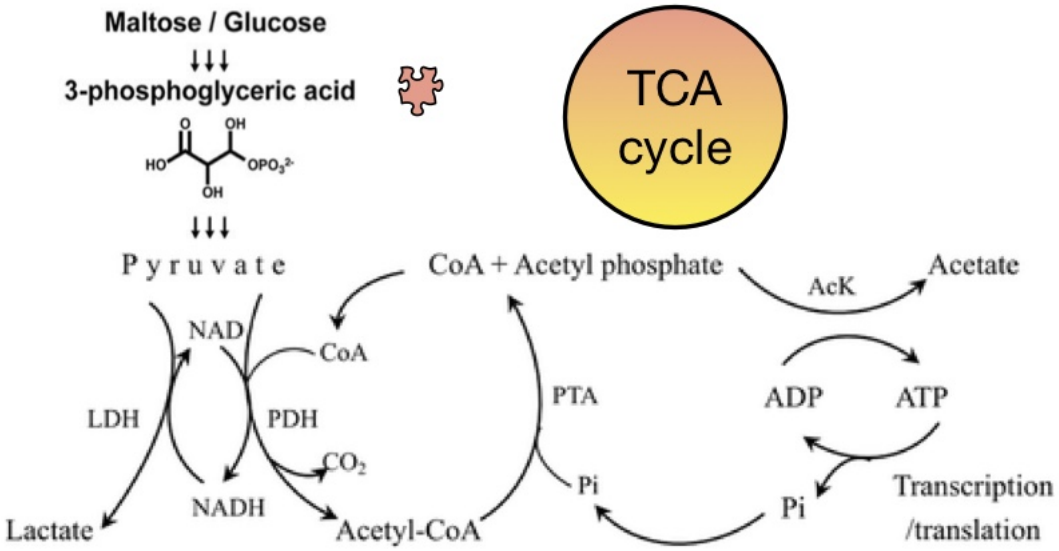


Fig. 1.2 Scheme of the TCA cycle for ATP regeneration in TX-TL systems. The TCA cycle is initiated with the addition of 3-PGA. Abbreviations: PDH, PTA, AcK, LDH stand for pyruvate dehydrogenase, phosphotransacetylase, acetate kinase, and lactic dehydrogenase, respectively. Figure adapted from [70].

cell-free systems is that the duration of protein synthesis is usually limited by the available energy supply in the reaction, since transcription and translation are both ATP-dependant processes. Fortunately, *E. coli* TX-TL systems have been the main subject of research regarding the extension of CFPS lifetime [70]. Enhanced energy regeneration systems have been developed along the years. The first-generation used expensive phosphoenol pyruvate (PEP) or pyruvate as an energy source in order to regenerate adenosine triphosphate (ATP) from the phosphorylation of adenosine diphosphate (ADP) [57]. The lifetime of protein synthesis was then extended from 2 hours up to 10 hours with current regeneration systems based on the use of glycolytic intermediate and additions of cofactors such as coenzyme A (coA) and nicotinamide adenine dinucleotide (NAD) to regenerate ATP from the pyruvate by-product. In this work, the energy regeneration system is based on 3-Phosphoglycerate (3-PGA) supplementation, as described from [73] (see Fig. 1.2).

Those elements taken together have contributed to the emergence of *E. coli* cell-free system as a leading expression platform with quick and robust preparation protocol and the most complete system for investigating CFPS.

1.3 Microfluidics for the Study of Gene-expression

The expression of cell-free reactions can take place in multiple environments. The traditional approach consists of using a batch-tube format. Even though it is the most simple and accessible method to express cell-free reactions, it remains limited by several aspects. The limitations associated to the batch-format are detailed in the 4th chapter. Microfluidics gained popularity during the last decade among the biotechnology laboratories by bringing further benefits, such as low-cost fabrication and reduction of the sample volume. Microfluidic technologies allow the fabrication of cell-like compartments, in order to mimic and execute cellular functions in a cell-free environment, such as homeostasis (Fig. 1.3). TX-TL systems can be encapsulated in defined size compartments which can be programmed to perform certain tasks, such as mixing or merging, offering great opportunities for approaching the reactions with a control-theory perspective [28]. Indeed, parameter's environments can be tuned during the experiment, enabling online monitoring of the reactions and direct feedback over various optimisation critera, such as resource burden or network modularity. Niederholtmeyer et al. [82] were the first to propose a functional implementation of a microfluidic nano-reactor which could sustain TX-TL processes over long period of time (Fig. 1.3 A). The reactor was successfully used to implement oscillators of increasing complexity as well as to regulate the ON/OFF gene state through the modulation of an inhibitor concentration. Mechanically induced trapping of molecular interactions (MITOMI) devices are a great example of what high-throughput compartment-based microfluidics are able to achieve. Up to hundreds of parallel chambers can be used on the same device (k-MITOMI) to simultaneously perform the kinetic analysis of protein-ligand interactions (Fig. 1.3 B). It has been used for the determination of binding affinities of transcription factors (TFs), of protein interactions or even of small pharmacological molecules inhibitors [40]. This is a great improvement for high-throughput, low-cost diagnostics, compared to the techniques commonly used to determine binding affinities such as forced-based or optical methods. Additionally, having a good estimate over TFs parameters has been pointed out as a central element in establishing robust model for synthetic circuitry. PDMS-based micro-compartments can also be designed for a wide range of sizes and shapes. Karig et al. [55] have shown that femtoliter-scaled compartment can be made to encapsulate TX-TL systems, mimicking reaction volumes corresponding to volumes ranging from a single *E. coli* to a yeast cell (Fig. 1.3 C). Those volumes enable the observation of stochastic effects linked to the limited number of molecules confined in the small volume compartments. The resulting fluorescent traces can be used to infer intrinsic-noise properties of the TX-TL

system, otherwise more complicated to model in living systems because of the multiple endogenous noise sources.

Micro-compartments can also be arranged in spacial conformations in order to explore spatiotemporal patterns encoded by the TX-TL systems. Tayar et al. [122] created DNA-patterned micro-compartments arranged in a 2D array, enabling front propagation of a circuit-encoded signal through multiple cascaded chambers (Fig. 1.3 D). This design allowed the study of reaction-diffusion patterns, enabling the creation of artificial minimal systems in which one can recreate physical phenomena found in living cells such as self-organisation or long-range communication. Finally, microfluidics possess an alternative to the fabrication of micro-compartments with the generation of water-in-oil (w/o) droplets to encapsulate TX-TL systems. One main advantage over PDMS-based chambers is their ability to be generated in a high-throughput manner with a dynamical content. For instance, Hori et al. [47] successfully screened for the optimal concentration parameters of DNA components of an incoherent forward loop motif. They used three modulated inputs to generate a landscape mapping of the circuit yield in a single experiment (Fig. 1.3 E). High-throughput droplet platforms have the potential to close the gap between the use of *in vitro* and *in vivo* systems for screening purpose. Indeed, commercial devices such as Fluorescence-activated cell sorting (FACS) are already available to sort single cells in encapsulated droplets.

1.4 The Roles and Functions of RNA Molecules

Among the components of synthetic biology toolbox presented over the last sections, RNA molecules have emerged as one of their most promising elements. They are able to modulate almost all cellular processes, including transcription, translation, and mRNA degradation or splicing [11, 15] (Fig. 1.4). They have been shown to occupy essential functions in physiological and pathological conditions. Number of pathology or genetic disorders are treated based on RNA-targeted molecules, leading to an increasing wave of new therapeutic devices, drugs and tools. This in turn has led to the engineering of new regulatory RNAs, either derived from natural molecules or based on novel mechanisms, with a great diversity of functions and targets. For instance, microRNAs (miRNAs) possess functionalised secondary structures which upon special conformation, can mediate protein-binding and promote diverse downstream activities, such as degradation or maturation [38] (Fig. 1.4 E). RNA molecules can also organise themselves in large complex with other proteins, called

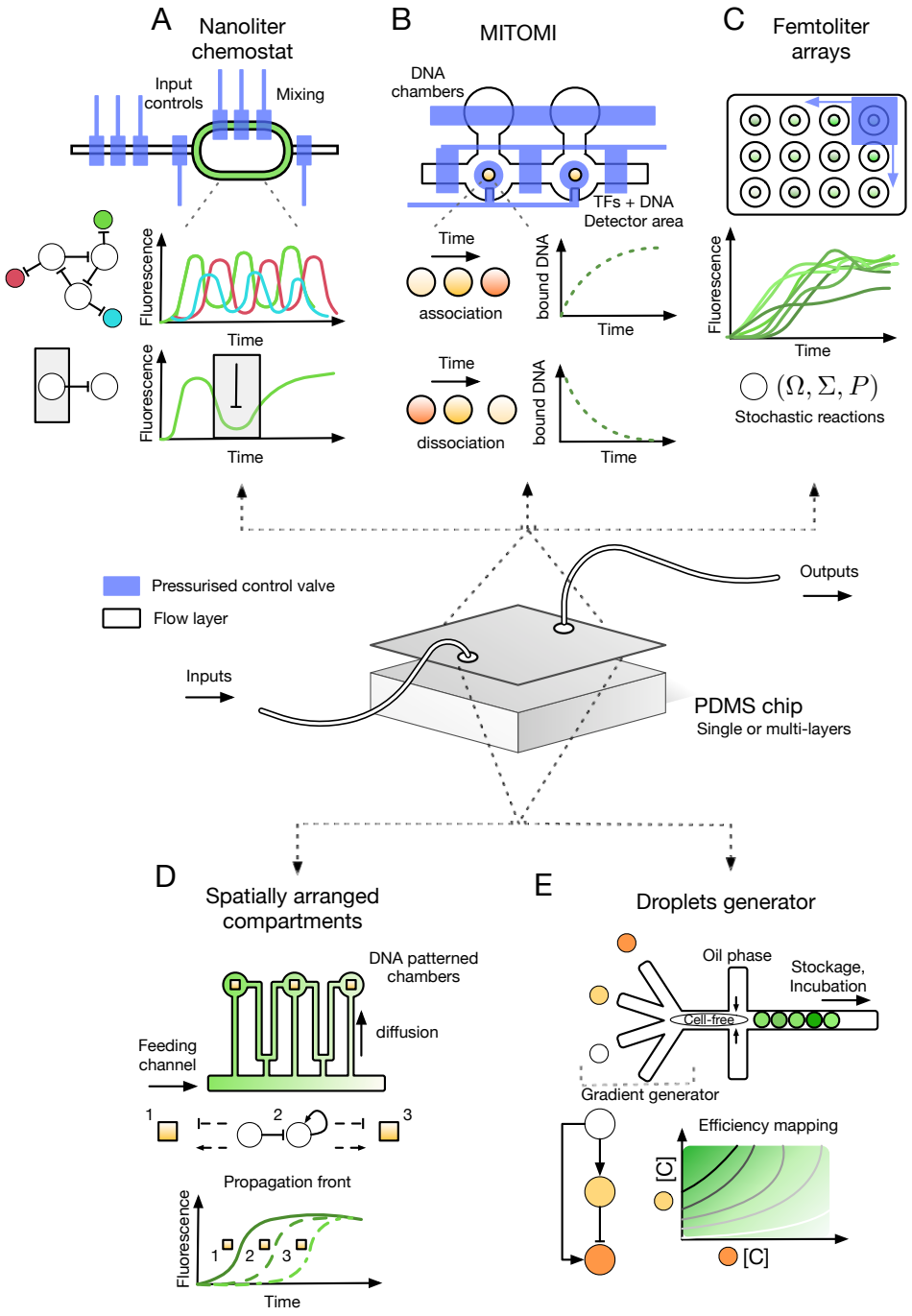


Fig. 1.3 Microfluidic platforms for studying TX-TL systems at the nano-liter scale. A) Nanoreactor enables continuous expression of TX-TL reactions at steady-state for the expression of dynamical circuits such as a three-nodes oscillator. B) MITOMI device allows the screening and measurements of binding parameters between DNA and transcription factors. Examples of association and dissociation curves are shown. C) Femtoliter arrays enclosing TX-TL systems generate stochastic time-course traces in a high-throughput manner. D) DNA brushes patterned in small compartments allow the study of diffusion patterns and front-propogation of gene expression. E) Microdroplets generated in a high-throughput manner allow the screening of component parameters of TX-TL systems, such as the concentrations of DNA parts.

quaternary assemblies, such as ribosomes and spliceosomes (Fig. 1.4 D). These regulators have been successfully used to address current problems in biotechnology, such as challenges in metabolic engineering [116] or therapeutic advances [59, 132]

1.4.1 Advantages of RNA-based Devices

Regulatory RNAs possess many features making them particularly suitable for genetic circuit design with advantages over the traditional use of TFs. One advantage is that RNA molecules are able to propagate information solely at the transcriptional level, without the need for intermediate proteins, which reduces the metabolic burden of a cell [54] and increases signal propagation speed [121]. In other terms, RNA-circuits of the same complexity as the ones based on TFs could perform equal computational tasks while sparing precious metabolic resources. Some of the RNA devices can be designed according to simple base-pairing rules either between RNA:RNA or DNA:RNA interactions. This considerably facilitates the adoption of a *de novo* engineering approach, compared to functional *in silico* protein design, which is a very challenging task. A growing number of algorithms, mainly based on free-energy minimisation, have shown the ability to compute secondary structures of RNA strands and link them to their regulatory functions. With a better understanding of this relationship between structures and functions, it also became recently possible to generate RNA sequences according to structural constraints [62, 38]. For instance, Zadeh et al. [136] have shown the successful use of the computational package Nupack to design and integrate a multi-co-localised 6-input OR device with toehold switches. However, inferring tertiary structures of RNA molecules still remains a challenge. Finally, similar to protein-based systems, RNA-devices are now able to regulate gene expression with very high dynamic ranges and great specificity for their targets. Conversely, some of the recent engineered RNA regulators are available in libraries of orthogonal parts or design rules for creating new ones [44, 16].

1.4.2 A Variety of RNA Regulators

Different RNA regulatory devices like riboswitches [100], ribozymes [102], small transcriptional activators (STARs) [13], toehold switches [44] as well as CRISPR/Cas based systems are available for building synthetic logic circuits. Lee et al. have recently assembled logic gates in *E. coli* based on the combination of engineered CRISPR/antisense RNA (asRNA) and

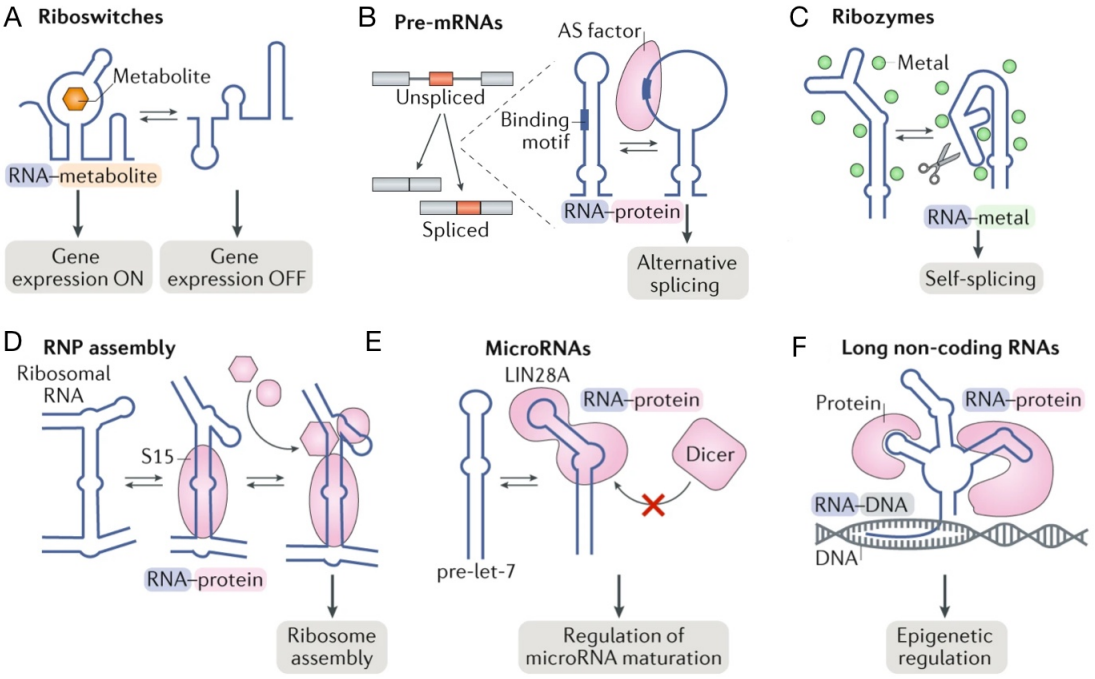


Fig. 1.4 RNA molecules display a variety of functions. Figure adapted from [38]. A) Riboswitches modulate the gene expression by conformation change of their tertiary structure upon ligand-binding. B) Alternative splicing protein factors bind and process RNA single strand motifs. C) Ribozymes are capable of catalysing self-cleaving reactions. D) Ribosomal RNA can promote the ordered assembly of the ribosome complex. E) MicroRNAs can bind to specific proteins and influence downstream activities. F) Long non-coding RNAs possess many cellular functions such as promoting protein interactions or epigenetic regulation.

STARs/asRNA [66, 65]. Shen et al. [102] have built layered AND-gates out of engineered sensor RNA, offering modularity at the post-transcriptional level.

Transcriptional Regulators

Transcription termination events are mainly due to rho-independant transcriptional terminators in *E. coli*. Briefly, a stable RNA hairpin loop motif of various lengths is followed downstream by a U-tract sequence comprised of 6 to 9 nucleotides. The U-stretch causes the polymerase to pause during transcription elongation and the RNA hairpin destabilises the DNA-polymerase complex resulting in its dislocation. Aside from the newly synthetic terminators engineered for *E. coli* based on those critera [17], a set of engineered devices have used the termination events as a mechanism of gene regulation. Activating or inhibiting transcription can be achieved by designing antisense small RNAs (sRNAs) that can respec-

tively stabilise or destabilise weak/strong terminator hairpin [96, 13]. Detailed mechanism of transcriptional activation by the STARs is given in chapter 2.

Translational Riboregulators

sRNAs can also be engineered to effectively regulate the access of the cell's ribosomal machinery to the mRNA. Regulatory sequences inserted in the 5' UTR of the targeted gene can occlude the RBS region and switch the genetic device in an OFF state. Upon binding with the cognate sRNA, structural change occurs and the RBS region is made accessible by the ribosomes. Various design rules are available for either riboregulators or toehold switches, for which the RBS sequence is maintained in an unpaired loop and therefore unconstrained from any sequence design [44, 53]. Repression by sRNAs can also be engineered in a similar way by simply designing a cognate binding sequence to the RBS and start codon regions, thus inhibiting translation initiation and elongation processes [9].

Riboswitches

Riboswitches can either act as transcription or translational regulators. They are found usually in the untranslated region of the regulated gene. They are made of two main domains: a binding-ligand aptamer, and an expression platform whose conformation changes upon ligand induction (Fig. 1.4 A). This expression platform triggers or represses gene expression through a variety of mechanisms. They can inhibit the access to critical parts of the mRNA, such as the RBS or eukaryotic 5' cap, promoting or disrupting the formation of intrinsic terminators in the 5' UTR region. Different riboswitches sequences have been discovered or engineered which can sense identical ligands, but act upon different regulation process, e.g. the theophylline riboswitch which can be found either as a translational or transcriptional regulator [130]. Most of the existing riboswitches were found in natural organisms, binding to a wide range of small molecules, vitamins and secondary metabolites. Engineering synthetic riboswitches *de novo* from computational methods is still a challenge, but *in vitro* selection methods such as systematic evolution of ligands by exponential enrichment have led to the identification of considerable numbers of novel ligand-binding aptamers [51].

CRISPR–Cas systems

The CRISPR systems are part of antiviral immune systems for counteracting foreign DNA or RNA that are present in a large percentage of bacteria and archaea species. Multiple

CRISPR–Cas systems have been found and categorised based on the configuration of the Cas protein. Of particular interest to gene editing, Cas9 or Cas12a effects a double strand break of a specific DNA [133]. Their specificity is defined by CRISPR RNA (crRNA) which forms a complex with trans-activating CRISPR RNA (tracrRNA) for guiding the Cas endonucleases to their cleavage target. Beyond gene editing, the inactive CRISPR–Cas systems have been reprogrammed for repressing gene expression at specific location. By fusing them to other domains such as TFs, they have also been used to activate gene expression. Cas-based technology has shown to be highly orthogonal due to the design of the single guide RNA (sgRNA), specific of the gene target (sgRNA is a engineered version of the crRNA and the tracrRNA expressed together). Extended functionalities have been shown, by extending the crRNA sequence to recruit other effector proteins or by fusing directly the Cas protein with the effector of interest.

1.4.3 RNA-based Circuitry

The amount of available RNA parts is therefore in continuous increase, with a catalogue of hundreds of RNA-based components harboring diverse dynamic ranges, mechanisms of regulation, complexities of the folded structure, or even different degrees of cross-species portability [13]. RNA-based circuits built exclusively on those components have led to the implementation of logic gates, feedback control units, regulatory cascades, or feedforward control strategies [42]. Successful examples are available for most of those components: ribosomal circuits based on orthogonal ribosomes with tuned specifities, sRNA-based circuits mediating complex logic computations, riboswitches combination used as band-pass filters, or layered complex circuits regulated with multiple small guide RNAs (sgRNAs) and dCas9 [62]. Many of these RNA-based circuitry have found applications in the biotechnological and medical fields. For instance, the use of RNA-scaffolds have shown to increase hydrogen output up to 48-times and post-transcriptional classifier has been implemented to selectively kill cancer cells expressing various levels of different miRNAs [60].

1.5 Aim of this Study

Novel RNA devices with various functions and high dynamic range have been characterised over the last few years. They have been shown to regulate specific processes of the cell machinery, e.g. transcription or translation. The previously given examples show the great potential of RNA-based devices, but also highlight some of the challenges in the assembly

of highly complex synthetic circuits required for more advanced projects in synthetic biology, such as establishing efficient strategies for gene and cell-engineered therapy. Despite numerous current research projects demonstrating a tremendous potential in RNA-based circuits, an important limitation remains in the lack of extensive control over the circuit expression parameters. Indeed, the control of the expression strength and dynamical timing of RNA-circuitry ,independently from the cellular context, remains a challenging task [60]. For example, over-expression of the targeted output may be harmful with unwanted side-effects such as toxicity under excessive activation of engineered cells. Another problem is that advanced applications reviewed in section 1.1 require a variety of well characterised and orthogonal devices. One way to address these challenges is to use RNA-based circuits able to regulate gene expression at different levels. Recent studies have shown the efficient integration of multiple RNA-activators in the same circuit. For instance, toehold switches have been engineered into a layered 4-input AND-gate mediated by TFs [44], or integrated into a co-localised 6-input OR device [43]. Similarly, STARs have been used in tandem to build logic AND-gates [15]. However, little work has been done yet with respect to the integration of heterogeneous RNA regulators STARs and toehold switches in the same circuit. In the 2nd chapter of this work, I present the implementation of advanced logical circuits based on both translational and transcriptional activators in a cell-free environment. As explained in section 1.2, cell-free platforms offer an ideal environment to overcome the current limitations of genetic network implementation in living cells, since every component of the circuit can be calibrated to the desired level of activation.

Multiplexed RNA-based circuits suffer from another limitation: a low dynamic range, which is however a critical feature when tight regulation is needed. Indeed, as the complexity of the circuits increases and involves several regulating RNA species, cross-talk effects can occur within the circuit. This can either be due to co-folding issues related to high sequence similarity between RNA components, or impaired functionality due to perturbation of their tertiary structures in response to various cellular conditions. In the 3rd chapter, TX-TL systems functionalised with two different proteins are analysed and enable the leverage of the logic gates from the 2nd chapter with improved functionality and efficiency. The first one consists of the catalytically "dead" endonuclease Cas9, which can still recognise and bind to DNA, and therefore be used as an inhibitor of gene expression by blocking the transcription. This aims to implement novel Boolean operations at the transcriptional level, and complete the toolbox based on RNA-parts for TX-TL systems. The second protein is Csy4, an endoribonuclease which recognises, binds, and ultimately cleaves a short RNA

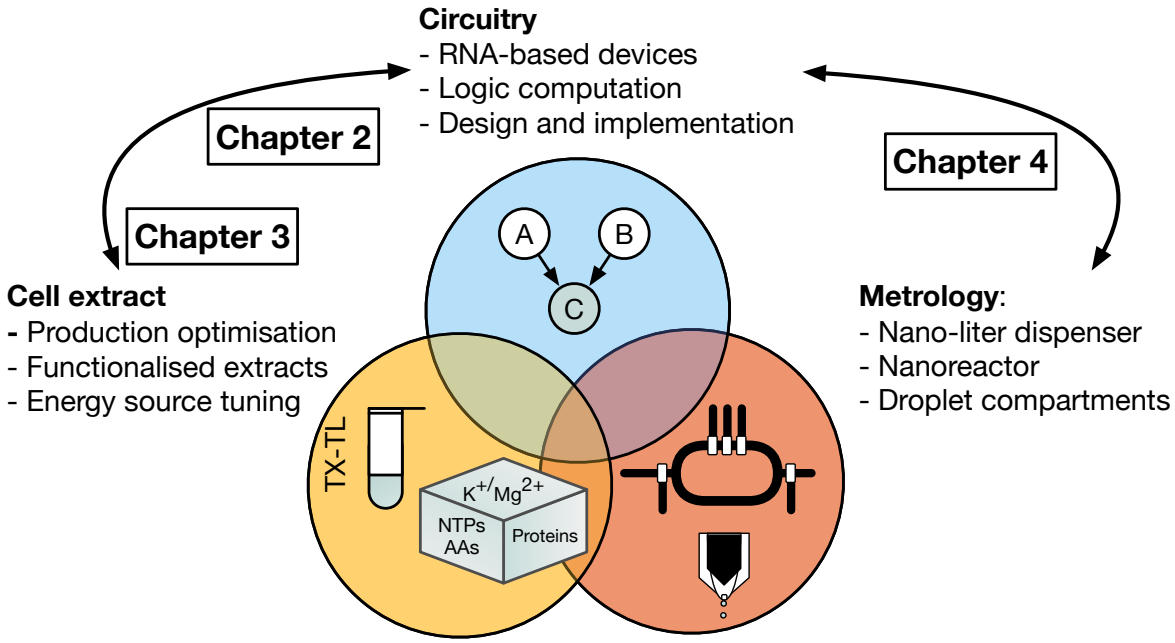


Fig. 1.5 General outline of the research subjects presented in this thesis. It can be divided into three main parts represented by, the cell-free extract production, the RNA-based circuitry prototyping, and the microfluidic metrology.

hairpin. Csy4 processing could serve multiple goals for synthetic biology engineering. It could normalise the expression level of heterogeneous DNA constructs, as well as insulating RNA components composed of multiple parts. So far, the co-expression of Csy4 in TX-TL systems has not been studied and I anticipate it may constitute a useful tool in the context of RNA-based prototyping. Finally, the 4th chapter is devoted to the implementation of some of the RNA-toolbox elements through two different microfluidic platforms. I will explore how the use of compartments with sizes ranging from the femtoliter to the nanoliter scales could be used for the study of RNA-based circuits in TX-TL systems. I will especially focus on the implementation of the nanoreactor which provides a technological solution to mimic cellular homeostasis in cell-free conditions. Taken together, this thesis aims to generate all the requisites parts necessary for the construction of advanced dynamic RNA-based networks, such as oscillators or memory devices. Their current implementations based solely on transcription regulation remain so far at the embryological stage.

Chapter 2

Engineering AND-logic Gates Based on Heterogeneous RNA Regulators

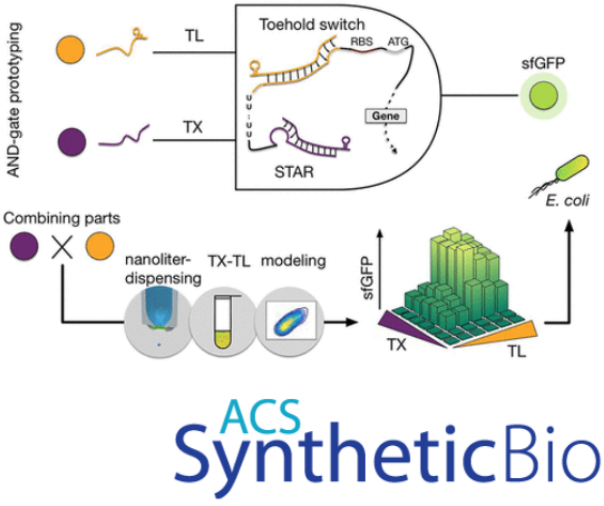


Fig. 2.1 Cover art made for the publication in ACS Synthetic Biology, "Cell-Free Prototyping of AND-Logic Gates Based on Heterogeneous RNA Activators". [67]

In this chapter, I am presenting the prototyping and implementation of multiple Boolean logic circuits in a cell-free system environment. Many synthetic biological circuits have been based on Boolean regulation, but this work represents the first attempt - to the best of my knowledge - to combine *de novo* engineered RNA-regulators to control both transcription and translation in a highly effective manner. This work has led to a publication in ACS Synthetic Biology [67] and several research projects have been started on the basis of this work.

2.1 RNA-based Regulators

Synthetic gene constructs regulated by either STARs or toehold switches have shown unprecedented regulatory performance in diverse contexts of transcriptional and post-transcriptional circuit regulation. Similar to the efficiency of protein-based systems, STARs and toehold switches regulate gene expression with very high dynamic ranges and great specificity for their targets. They are either available as libraries of orthogonal parts or with design rules for creating switches *de novo* [44, 16]. Additional reasons have led me to chose the STARs and toehold switches. Firstly, they are able to control transcription and translation independently, delivering a promising option for applying multiplexed regulation to virtually any gene (in *E. coli* at least). Secondly, using such a multi-level regulator system has been shown to increase the dynamic range of gene expression [128] by reducing dramatically the OFF-level of a gene otherwise leaky either on the transcriptional or translational level. Thirdly, both systems have been thoroughly characterised with fine-tuned sequences or coupled with various synthetic parts, turning them into candidates of choice for integrated circuits [77, 125]. Finally, because both switching mechanisms are based on Watson-Crick base pairing, gate design could be aided by the growing number of algorithms for computing secondary structures or solve inverse structure problems from the sequences of nucleotides given as an input [35, 41].

2.1.1 Characterisation of Single Activators in TX-TL

Mechanisms and implementations

In order to build multilevel RNA-regulated AND-gates based on the STARs and toehold switches, the first step was to test their functionality as independent activator unit in TX-TL. STARs are small RNAs designed to form RNA-terminator hairpins. Synthetic STAR antisense sequences can be co-expressed to disrupt the formation of these intrinsic terminators placed upstream of the 5' UTR of the regulated gene - called in this thesis - targets. In the absence of the anti-terminator STAR activator, the STAR-target leads to an early rho-independent termination of the regulated transcript (see Fig. 2.2 A. The toehold system controls gene expression at the translational level and also contains a small antisense RNA which is called trigger. In the absence of the trigger, the toehold blocks the RBS. The trigger invasion in the toehold region leads to a conformation where the RBS can be accessed by the ribosome and translation can thereby occurs (see Fig. 2.3 A). Thus both devices become active in the presence of an interacting antisense RNA, either trigger or STAR activator.

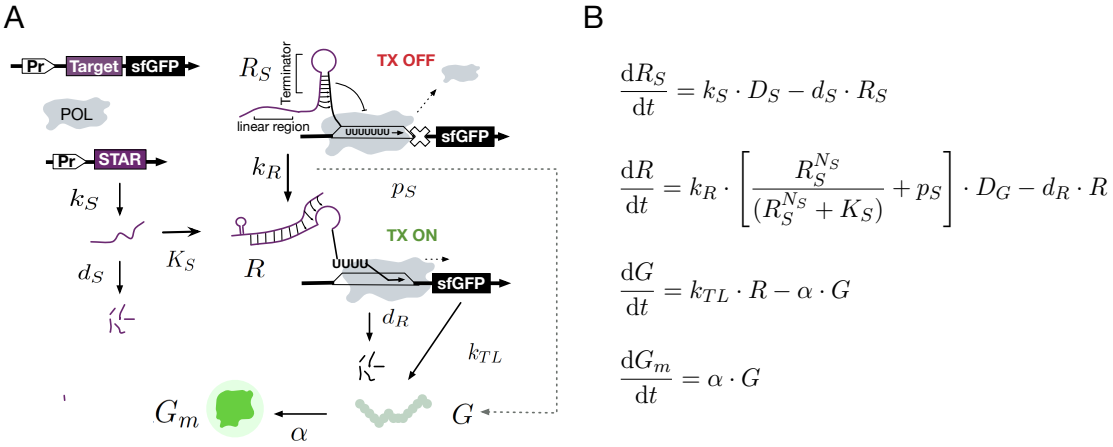


Fig. 2.2 Small Transcriptional Activator RNAs regulate gene expression at the transcriptional level. A) Schematic of the STARs mechanism. B) ODEs model for the single STAR activator with species labels corresponding to the one written on the schematic.

For the experimental implementation, I used a two-plasmid-system similar as the one used by Chappell et al. [13]. One plasmid contained the small activating RNA, which is driven by the strong constitutive J23119 promoter (Anderson collection). The other plasmid codes for the super folder GFP (sfGFP) [92] reporter gene, which is also driven by the J23119 promoter. Either the STAR target or the toehold switch sequence is placed upstream of the reporter gene start codon to control its expression. In both plasmids, the t500 terminator is used to terminate transcription. For the *in vivo* experiments, the activator cassettes were combined in the same plasmid. The parts and plasmid constructs are available in the material chapter, in section 6.19 and 6.20.

Characterisation in TX-TL

The expression of the most efficient STAR from the original publication (STAR AD1) [13] together with one of the top-performing engineered toehold switches (toehold 2) [44] were initially characterised. Dose response curves were generated for each part using an *E. coli* cell-free TX-TL system by varying the concentration of the activator plasmid. I used a final reaction volume of 2 µl by dispensing cell extract, energy buffer and plasmid DNA with the help of a liquid handling robot (I-Dot). The reporter plasmid concentration was 2 nM, and

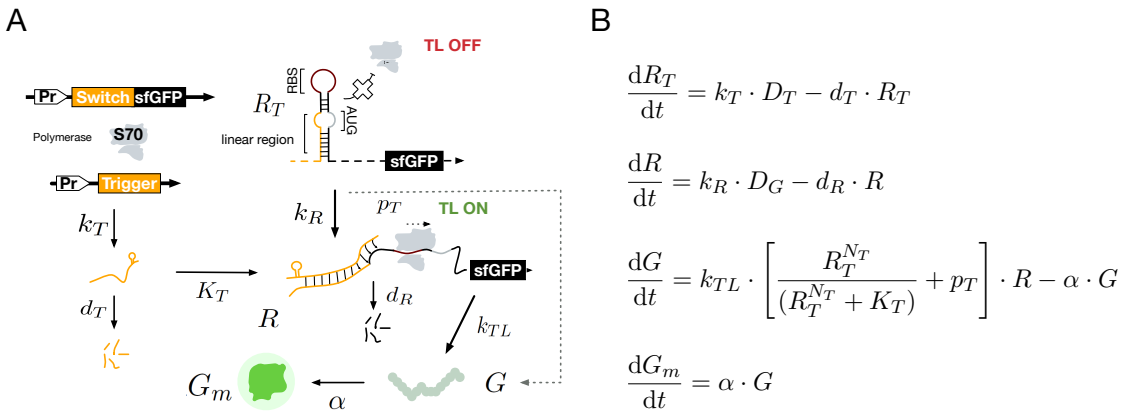


Fig. 2.3 Toehold switches regulate gene expression at the translational level. A) Schematic of the toehold switch mechanism, with and without trigger activator. B) ODEs model for the single toehold activator with species labels corresponding to the one written on the schematic.

I varied the concentration of activator plasmid (trigger 2 or STAR AD1) between 0-22 nM. Each DNA sample was independently calibrated beforehand in the I-Dot to insure that the dispensed volume properly fits to the dose-response curve. The results of the endpoint measurements are presented in Fig. 2.5 B and individual trace kinetics are presented in Fig. 2.4.

The TX-regulation-system and the TL-regulation-system reached saturation at about 20 nM of STAR plasmid and 7 nM of trigger plasmid, respectively. Similar to previous *in vivo* data [44, 13], the toehold 2 has a three times higher dynamic range than the STAR AD1. Rather than due to impaired functionality, those lower dynamic ranges could be attributed to the TX-TL system. Previous results have highlighted the decrease in dynamic ranges for RNA devices when transitioning from *in vivo* to *in vitro* [72]. This is likely due to the batch-reaction format, where the synthesised protein accumulates in the OFF-state without getting diluted as in a dividing cell. The dose-response relationships follow either a linear response for STAR AD1 or a fast saturation curve for toehold 2 activation. This difference may be explained by the longer time (mRNA life-time) in which the trigger RNA can bind and invade the stem of the toehold switch compared to the transient event characterising the STAR termination [77]. To optimise the sfGFP read-out (Suppl. Fig. 2.6), I tried to increase the reporter plasmid concentration to 6 nM. This led to a significant increase in translated sfGFP and therefore strength of the fluorescent signal. However, the concentration of activator plasmid needed to reach the plateau tripled from 7 nM to 20 nM, suggesting that

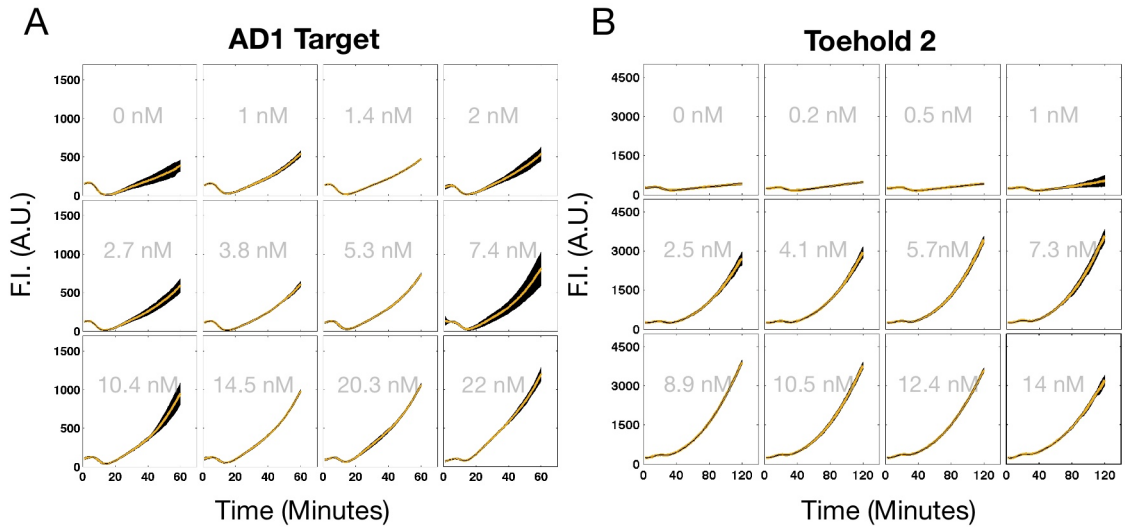


Fig. 2.4 Kinetic characterisation of toehold 2 and target AD1 in presence of 2 nM of reporter template, either toehold2-sfGFP or targetAD1-sfGFP. Yellow lines correspond to the averaged experimentally measured fluorescence. Black shaded regions represent the experimentally measured standard deviation. Concentrations written in gray in the background correspond to the activator plasmid concentration, either trigger 2 or STAR AD1 activator plasmid.

RNA polymerases were not saturated yet with 2 nM of DNA template. Because such high DNA concentrations caused issues with dispensing and to calibrate the samples in the I-dot adequately, I kept the reporter plasmid concentration set at 2 nM in the following experiments. The dynamic range of around 20-fold was unaffected by the different concentration levels of the reporter plasmid.

2.1.2 Mathematical Modeling of Single Components

Cell-free system reactions in bulk-sized format have the advantage to offer a deterministic framework based on the chemical species present in the system. Many examples have shown that traditional ODE models have led to convincing results to describe and predict synthetic circuits in a cell-free environment [79, 49]. Based on the TX-TL dose-response kinetics for the STAR AD1 target and the toehold 2 shown in section 2.1.1, coarse-grained parameterised ODE-models were built for both individual activating systems.

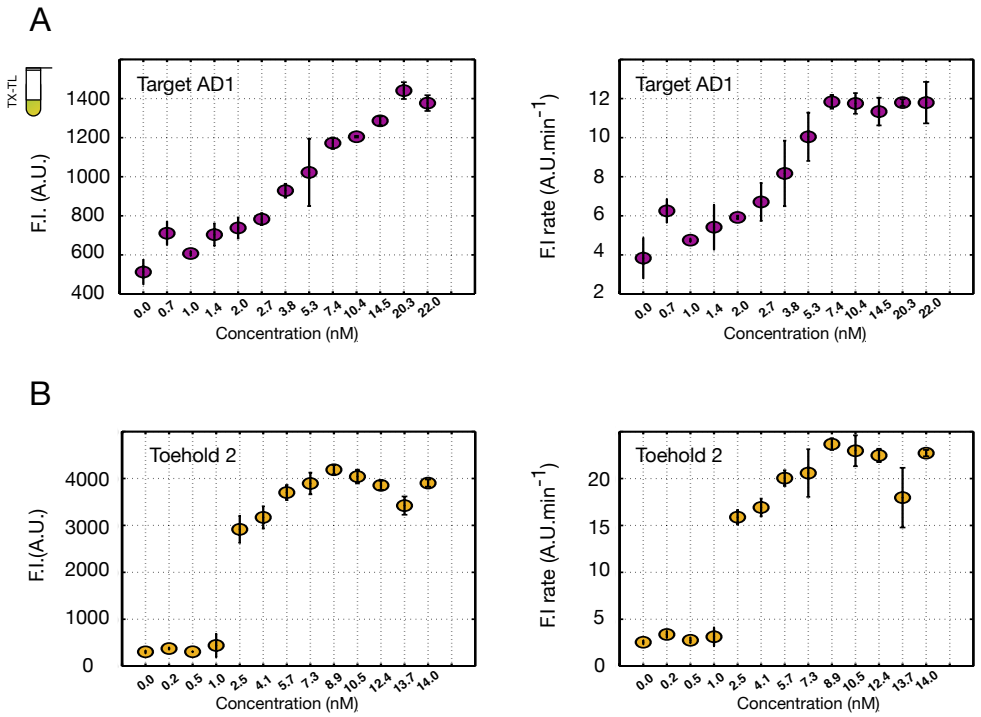


Fig. 2.5 Dose response curves of TX-TL reactions with 2 nM of regulated reporter plasmid and titrated trigger 2 or STAR AD1 activator. A) Endpoint fluorescence measurements and B) sfGFP production rates are presented. Rates were determined by computing the slope in the linear regime of each time-course trace. Errors bars represent the standard deviation over three independent experiments.

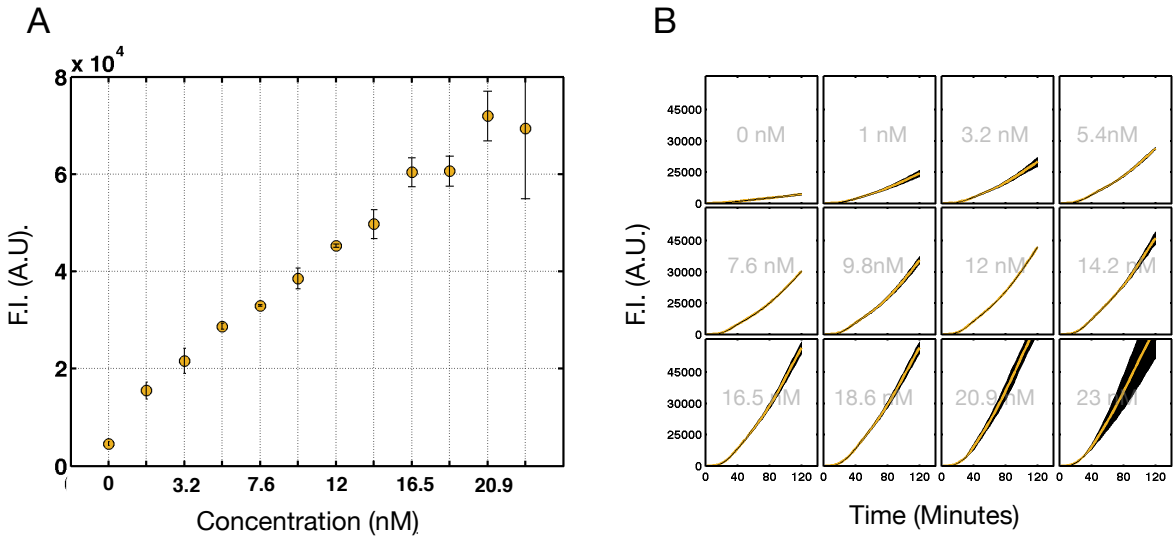


Fig. 2.6 Kinetic characterisation of toehold 2 switch in presence of 6 nM of toehold2-sfGFP template. Yellow lines correspond to the averaged experimentally measured fluorescence. Black shaded regions represent the experimentally measured standard deviation. Concentrations written in gray correspond to the activator plasmid concentration, i.e trigger 2 encoding plasmid.

Table 2.1 Description of species and kinetic parameters used for the modeling of TX and TL-regulated components.

Model species	Definition
R_S, R_T	RNA activators for STAR and trigger
R	sfGFP reporter mRNA
D_s, D_T	DNA template for STAR and trigger activators
D_G	sfGFP reporter DNA template
D_R	mRNA of the sfGFP reporter
G	Immature sfGFP
G_m	Mature sfGFP
Model parameters	**Definition**
k_s, k_T	Transcription rate of the STAR and trigger activators
k_R	Transcription rate of the sfGFP reporter
k_{TL}	Translation rate of the sfGFP reporter
d_s, d_T	Degradation rate of the STAR and trigger activators
d_R	Degradation rate of the sfGFP reporter
α	Maturation rate of the sfGFP reporter
N_S	Hill coefficient of the target-STAR binding
N_T	Hill coefficient of the toehold-trigger binding
K_S	Dissociation constant of the target-STAR binding
K_T	Dissociation constant of the toehold-trigger binding
p_S, p_T	Activation probability of the STAR and trigger

The ultimate goal was to infer and use the parameters from those individual activator models to predict the performance of the assembled AND-gate.

For each model, a minimal number of parameters and species were used to account for the unidentifiability of the parameters because the sfGFP was the only species being measured. The models were mainly based on mass-action kinetics except for the TX- and TL-regulation by the STAR and trigger activators, for which Hill-kinetics was employed to capture the corresponding dose-response shapes. The detailed systems of ODEs can be found in the Fig. 2.2 B for the STAR activator and in Fig. 2.3 B for the toehold switch. Table 2.1 describes the list of species and parameters used for those models.

Once the basic set of equations for ODEs were set up, the multiple parameters have to be fitted to the experimental results. Taking rates directly from the literature and applying them in our system was not suitable as a huge variance exists between cell-extracts, constructs assembly, and even laboratories with similar conditions performing CFPS [22]. Additionally,

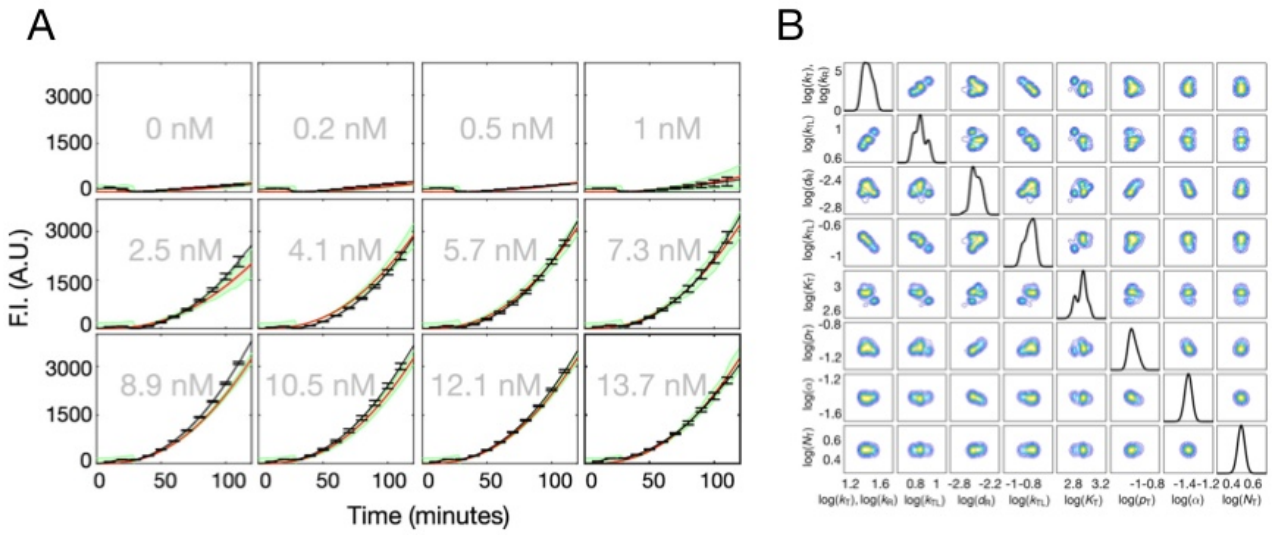
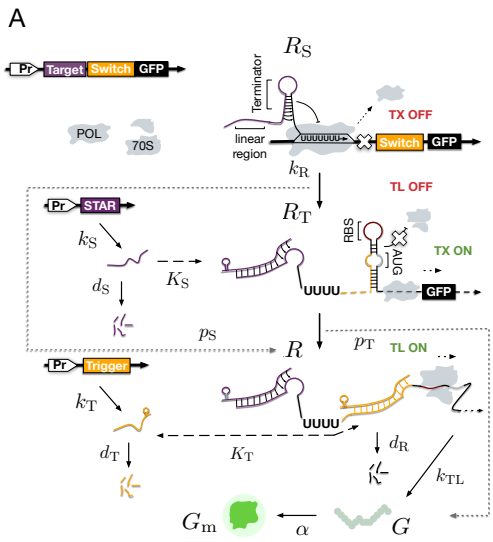


Fig. 2.7 Calibration of single activator component. A) posterior predictive distribution of the toehold-trigger 2 system for 2 nM input plasmid concentration. Input concentrations of trigger activator plasmid are written in grey; black: experimental data; red: median of parameterised model according to identified parameter posterior distribution displayed in (B); green: observation model corresponding to experimentally measured standard deviation and assumption of normally distributed observation errors, 95%-quantil and 5%-quantil. B) parameter posterior distributions of the toehold-trigger 2 system based on 150 000 samples in the logarithmic space. Univariate distributions are on the diagonal, bivariate distributions on the off-diagonal. The blue to yellow scale indicate the low to high density values. Details of the implementation can be found in section 6.1.5.

some of them were not available (like for the dissociation constant of the activators). The kinetic model curves of the mature sfGFP of the two single parts had therefore to be calibrated to the experimental data. A mathematician researcher from the BCS group, Maleen Hanst, collaborated on this project for implementing a Bayesian inference method with Markov chain Monte Carlo sampling [50]. To improve the convergence speed compared to standard approaches, she performed parallel tempering, a modification of the Metropolis-Hastings algorithm (for details see section 6.1.5). STAR AD1 and toehold 2 were calibrated to the experimental dose-response corresponding to different input concentrations of activator plasmid. The resulting parameter posterior distributions and the corresponding model kinetics for toehold2-trigger2 systems are shown in Fig. 2.7. An appropriate set of parameters posteriors were therefore found for all concentrations to closely capture the dose-response kinetics. This is not only useful to determine the best fitting set of posterior parameters, but also gives some clues about the general system behavior. For instance, for the toehold 2 system, a positive correlation is observed between the transcription, k_{T}, and degradation, d_{T}, rate of the trigger RNA, while the transcription, k_{R}, and translation, k_{TL}, rate of the reporter are negatively correlated. This matches expectations, since the higher the TX-rate of



B

$$\frac{dR_S}{dt} = k_S \cdot D_S - d_S \cdot R_S$$

$$\frac{dR_T}{dt} = k_T \cdot D_T - d_T \cdot R_T$$

$$\frac{dR}{dt} = k_R \cdot \left[\frac{R_S^{N_S}}{(R_S^{N_S} + K_S)} + p_S \right] \cdot D_G - d_R \cdot R$$

$$\frac{dG}{dt} = k_{TL} \cdot \left[\frac{R_T^{N_T}}{(R_T^{N_T} + K_T)} + p_T \right] \cdot R - \alpha \cdot G$$

$$\frac{dG_m}{dt} = \alpha \cdot G$$

Fig. 2.8 Schema of the computational RNA-based AND logic gates based on transcriptional (TX) and translational (TL) control. A) TX is regulated through the presence or absence of a STAR activator. TL is regulated through the binding of the trigger RNA to the linear toehold region. Both types of small RNA are required to activate the gene expression. B) Parameterised ODE-model of the AND-gate as a combination of the single part models. Details regarding the parameters, species, and single part models can be found in the table 2.1 and section 6.1.5.

the reporter mRNA is, the lower its TL-rate should be to reach the measured fluorescence. Finally, the model was able to capture the steep increase between 1 nM and 4.1 nM input concentration of activator plasmid (see Fig. 2.8). This is possible due to the Hill-kinetics equation used to model the regulation.

2.1.3 A Prototype Gate for Multi-level Regulation

Mathematical Modeling

After demonstrating the proper function of the single activators STAR AD1 and toehold 2 in TX-TL system and building an ODE framework containing fitting parameters for a wide range of concentrations, I wanted to assess the ability of the independent models to be joined in order to predict the behavior of Boolean RNA-based circuits. Therefore, an ODE model of the AND-Gate controlling the sfGFP production based on both STAR and toehold was

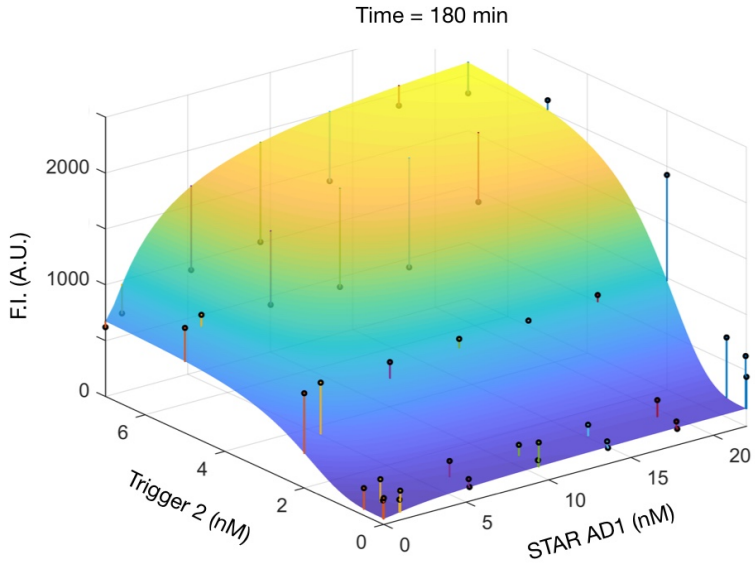


Fig. 2.9 Computational dose-response of AND-gate based on STAR AD1 and toehold 2 components after 180 minutes of expression. Experimental data are represented by the black dots, and the variance represents the standard deviation over three replicates.

built from single part models and their inferred parameters (maximum a posteriori values). In this new AND-gate model, the transcription process was regulated by the STAR and the translation by the toehold switch. The parameters associated with the sfGFP mRNA such as maturation time α and k_R were therefore averaged over the inferred parameters from the single component models.

The model prediction of the gate can be seen in Fig. 2.9. The plateau or maximal activation threshold was reached with approximately three-fold more STAR activators than trigger. In particular, the model predicted a high leakage for 0 nM STAR AD1 plasmid input concentration.

Experimental Validation

The next step was to implement the AD1 target-toehold 2 gate in TX-TL to verify the prediction of the *in silico* model. The STAR AD1 target and the toehold 2 switch were placed in tandem upstream of the sfGFP reporter gene without further re-engineering (see Table 6.17). Guided by the model, a combination of 42 concentrations was tested (six

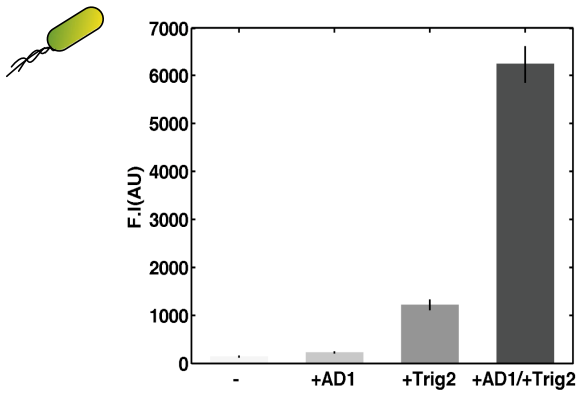


Fig. 2.10 *In vivo* characterisation of the prototype gate based on STAR AD1 and toehold 2. Fluorescence endpoint measurements of the gate in top10 *E. coli* cells in presence (+) or absence (-) of the trigger/STAR DNA-encoded activators.

concentrations for the trigger 2 and seven for the STAR AD1). The predictions of the AND-gate model captured the qualitative trend of the experimental characterisation (Fig. 2.9). The TX-TL endpoints at 180 minutes fit to the predicted values of the sfGFP leakage along the STAR concentration gradient when no trigger is present. Some qualitative disagreements are observed at intermediate concentrations of activators, i.e. on the surface slope.

Predicting transcriptional leakage *in vitro* could also provide insights into the *in vivo* behaviour. Therefore, *E. coli* cells were transformed with the same reporter plasmid as used for the TX-TL measurements. This plasmid contains the p15A origin of replication, which leads to a low plasmid concentration inside the cell. For the activators, a plasmid with the ColE1 origin of replication was used, which leads to a high copy number inside the cell. I thought this configuration would mimic the conditions studied in TX-TL, since it reflects the lower amount of DNA reporter used there. It would be interesting in future works to swap the constitutive promoters with an inducible system[106], and reproduce *in vivo*, the dose-response shape of each activator *in vivo*. To analyse the behaviour of the AND-gate, the *E. coli* cells were co-transformed with an activator plasmid containing either the STAR, the trigger or both activators on the same plasmid. The obtained endpoint fluorescence measurements show a similar behaviour of the AND-gate *in vivo* and *in vitro* (Fig. 2.10 A). A slight activation of the gate when only trigger 2 is present was observed, i.e., the transcriptional leakage observed *in silico* and *in vitro* occurred as well in living cells. This indicates that the use of TX-TL as a prototyping platform is not only beneficial for screening various circuit implementations, but can also be used to analyse more finely the effects of single part behaviour.

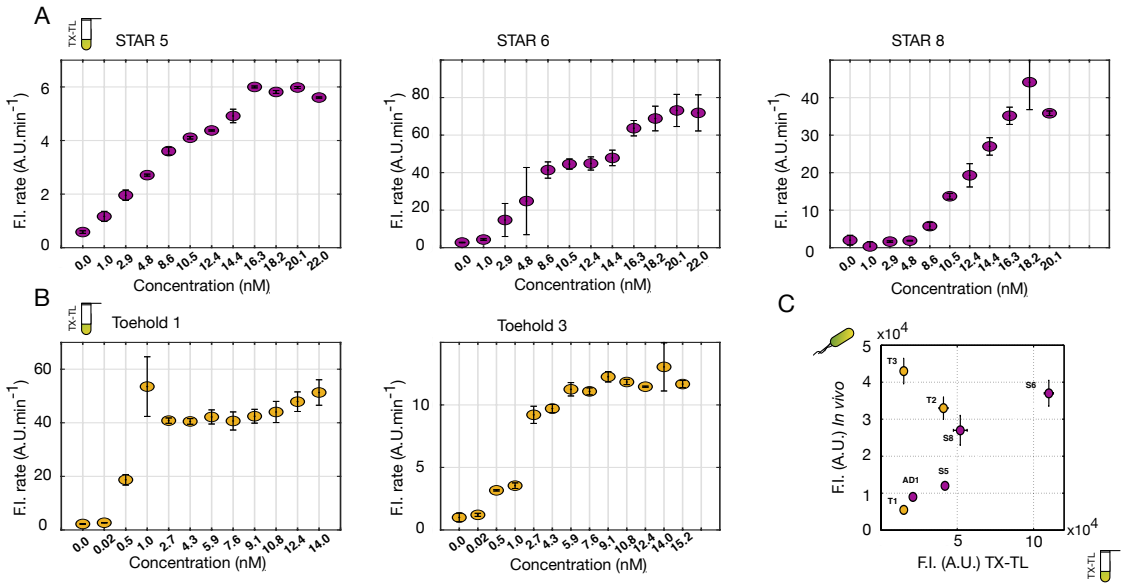


Fig. 2.11 STARs and toehold switches characterisation for five additional single components. A) Dose-response curves of sfGFP production rates for STARs 5, 6, 8 in TX-TL. B) Dose-response curves of sfGFP production rates for toehold 1 and 3. DNA-encoding STAR or trigger activators are titrated in presence of 2 nM of regulated sfGFP plasmid. Rates were obtained by calculating the slope in the linear regime of each time-course trace. C) Corresponding relative fluorescence intensities between *in vivo* and *in vitro* for seven single parts. Errors bars represent the standard deviation of three independent experiments, for *in vitro* and *in vivo* data.

2.2 Optimising RNA-based Logic Gates

2.2.1 Additional Parts Characterisation

The prototype gate based on STAR AD1 and toehold 2 showed potential for proper functionality in TX-TL and *in vivo*, I sought next to create a collection of gates made of various STAR and toehold parts. Multiple studies have shown that combining single components may result in unwanted co-folding or context effects, challenging the implementation results of assembled RNA-parts [71, 48]. However, STAR and toehold components are available as orthogonal matrices of numerous parts, and are still under active investigation for further improvements. [77, 20]. As the relatively high OFF-state of the STAR AD1 component was

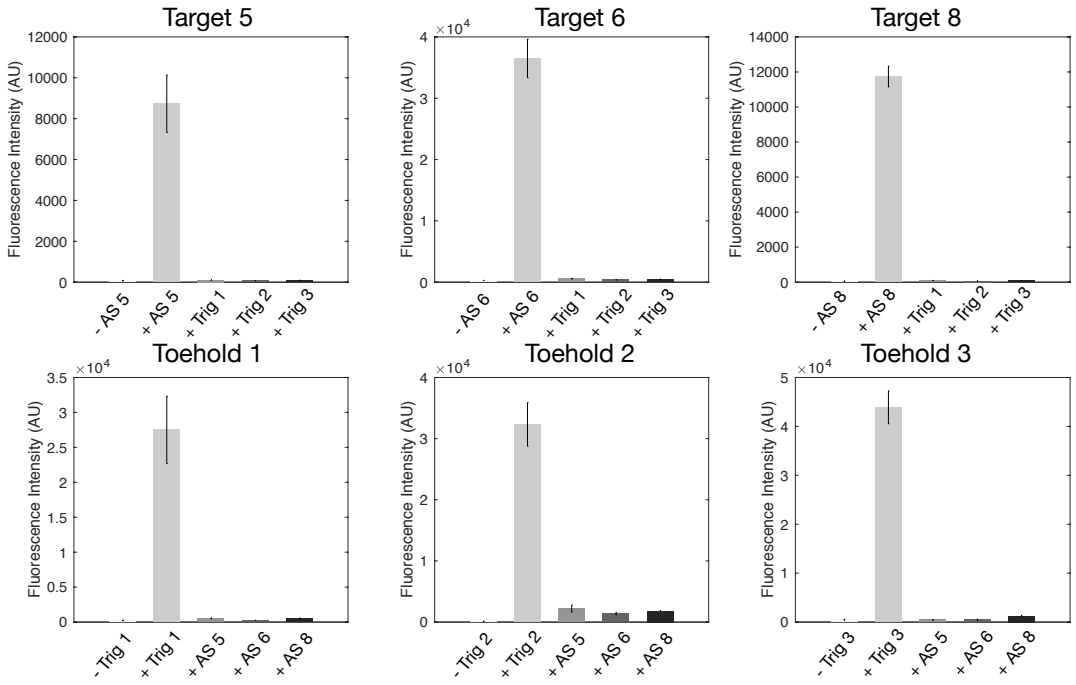


Fig. 2.12 *In vivo* measurements of endpoint fluorescence for each single part, either toehold (1, 2, 3) or STAR target (5, 6, 8), in presence (+) or absence (-) of their cognate or orthogonal trigger or STAR antisense. Error bars correspond to the standard deviation of three independent measurements.

identified as the source of the consequential translational leakage in the previous section, I selected three newly engineered STARs with lowered OFF-levels (STAR 5, STAR 6, STAR 8). In addition, I chose two other high performing toehold switches (toehold 1 and toehold 3) in order to increase the diversity of the gate library. The context-effect of the toehold sequences could therefore be studied as well, as the diversity between them is far greater than the one among the STAR library. Indeed, one current disadvantage of the STARs, at least for the best performing ones, is that they are built upon the same terminator sequence, which could be a potential source of cross-talk in multi-layer circuits. Those five RNA single activators were tested under the same conditions in TX-TL as the initial parts AD1 and toehold 2. The corresponding dose-response curves of the sfGFP production rates are plotted in Fig. 2.11. Like previously seen with AD1 STAR and toehold 2, a similar dose-response behavior is observed i.e. less trigger plasmid is needed to activate the toehold than STAR plasmid is needed to activate the STAR target. sfGFP production rate is maximal for 3 nM to 5 nM of trigger-encoding plasmid and for 15 nM to 20 nM of STAR-encoding plasmid. Importantly, the OFF-level in the absence of the activator STAR DNA (5,6,8) was lower than the one from STAR AD1, suggesting those parts will likely have a better transcriptional tightness

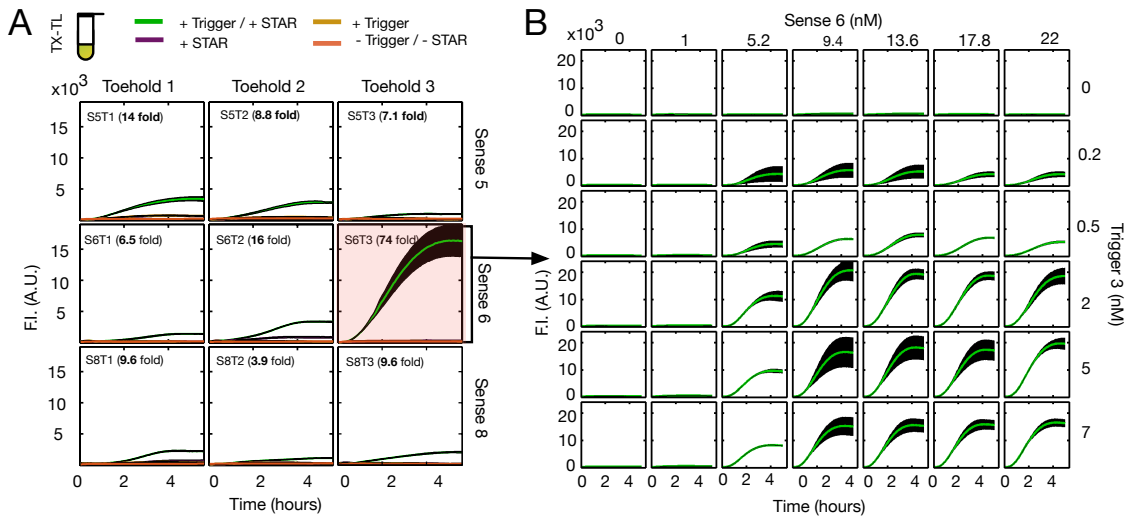


Fig. 2.13 Characterisation of logic AND-gates with TX-TL time course reactions. A) Pre-screening in TX-TL reactions of RNA-based gates composed of a cross design of three toeholds (1, 2, 3) and three STARs (5, 6, 8), for nine gates. The black shaded regions correspond to the standard deviation over three independent reactions. Colored lines correspond to the average measurement over those replicates, according to the presence of either trigger-encoding DNA or STAR encoding DNA or both. The fold-range is computed as the difference between the active state (both trigger and STAR are present) divided by the inactive state (trigger and STAR are not present), corrected by the background value of the TX-TL mix. B) Concentration matrix of TX-TL time course reactions for the sense 6-toehold 3-sfGFP gate (S6T3). Trigger and STAR encoding plasmids were simultaneously titrated from 0 nM to 7 nM and from 0 nM to 22 nM, respectively.

when combining into AND-gates. The dynamic range increased as well, in particular for the STAR 6 construct.

An other important criteria of efficient logic devices is the requirement of orthogonal parts. In the case of the RNA-based AND-gates, it means that the single components can not activate a different cognate sequence. This was verified *in vivo* whether the STAR activates the toehold or if the trigger is able to activate the STAR target. This was not the case for any of the tested combinations (Fig. 2.12). The toehold 2, however, had a slighter diminished dynamic range when any of the STAR were present, suggesting some minor activation level of the toehold strand by the STAR sequences.

2.2.2 Gates Library in TX-TL System

Based on the single component parts from the previous section, nine different gates have been engineered by crossing each STAR target with each toehold switch. This was done in two distinct stages. First, a screening to detect the gates with the best dynamic ranges was performed. I described this step as a screening phase even though it required the assembly and sequencing check of each independent plasmid construct, an approach which is not compatible with a high-throughput design. The constructs could be as well expressed from linear template by the addition of GamS (see chapter 4). This could be a way to bring this screening closer to a higher throughput methodology. The limitation of cell-free reactions in batch-format for screening of high dimensions parameters needs to be noticed as well. Each construct needed to be tested in a separate well, which is a disadvantage compared to the possibilities offered by *in vivo* screening (see also Section 2.2.4).

As the prototype AD1-toehold 2 gate performed best at 15 nM of STAR and 5 nM of trigger, I chose the same concentrations of activators for testing the nine additional gates in TX-TL which are :

- S5T1: Sense 5 - Toehold 1
- S5T2: Sense 5 - Toehold 2
- S5T3: Sense 5 - Toehold 3
- S6T1: Sense 6 - Toehold 1
- S6T2: Sense 6 - Toehold 2
- S6T3: Sense 6 - Toehold 3
- S8T1: Sense 8 - Toehold 1
- S8T2: Sense 8 - Toehold 2
- S8T3: Sense 8 - Toehold 3

In a second stage, I performed a detailed kinetics analysis of the selected gate(s). Fig. 2.13 A displays the resulting kinetics over a period of 5 hours of TX-TL expression. For every gate, the plateau is reached between 3 and 4 hours. All gates display Boolean functionality regarding to the trigger and STAR activator. However, they have various efficiencies ranging from around 4-fold for the least efficient (S5T2) to 74-fold for the most efficient (S6T3). Remarkably, S6T3 presents an absolute fluorescence intensity and a dynamic range almost one order of magnitude above the pool. Most of the gates display a symmetrical leakage when only trigger or STAR is expressed, except for S6T2, which shows higher leakage when only STAR is present. This last result can be seen as well as the level of the

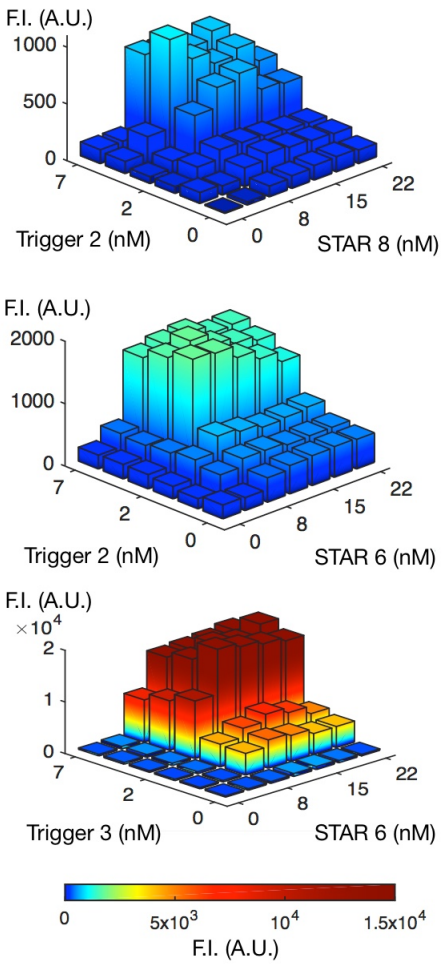


Fig. 2.14 Dose-response comparison of endpoint fluorescence measurements of three different AND-gates after 4 hours of expression in TX-TL system; from top to bottom: sense 8-toehold 2, sense 6-toehold 2, and sense 6-toehold 3.

single components in figure 2.12 where toehold 2 show some slight leakiness when STAR 6 was present.

Despite this last fact, correlation between gate performance and part composition has not been clearly identified. My assumption was that for all constructs, the final fluorescence level could only be as high as the least efficient component of the gate, i.e. the STAR target or toehold switch. In other terms, the level of activated mRNAs from the gate constructs could not exceed the level from the single STAR or toehold switch composing this gate.

To a certain extent, it can be observed for most of the gates. For instance, the gate S5T3 has the lowest fluorescence value – and is based on the two single components with the least efficient sfGFP production rate (target 5 and toehold 3, cf Fig. 2.11). However, the

S6T3 gate is a remarkable counter-example as the high yield of target 6 is not diminished by the presence of the toehold 3 component. In order to look into more details about the functioning of this gate, I performed extensive kinetic measurements for S6T3 with similar concentration conditions as for the initial gate AD1-toehold 2 (see figure. 2.13 B). Initial screening concentrations were close to the optimal expression level, which were found to be 2 nM of trigger and around 9 nM of STAR. To assess the qualitative disparity among the gates' dose-responses, I performed kinetic measurements for two additional screened gates. The results are shown in Fig. 2.14. Compared to S6T3, the fluorescence levels were either intermediate (S6T2) or very low (S8T2). Overall, the qualitative shape remained similar. The initial screening concentrations were again near the maximum sfGFP-expression. A decrease in sfGFP fluorescence is observed when both activators are present in high concentration. This may be due to the toxicity effects reported for high DNA concentration in TX-TL [7]. Resource competition is critical in TX-TL where the energy supply is not replenished. A relevant experiment to test this hypothesis would be to run the gates in the nano-reactor, where resources can be replenished to overcome fast energy consumption by the activator encoding plasmids (see chapter 4).

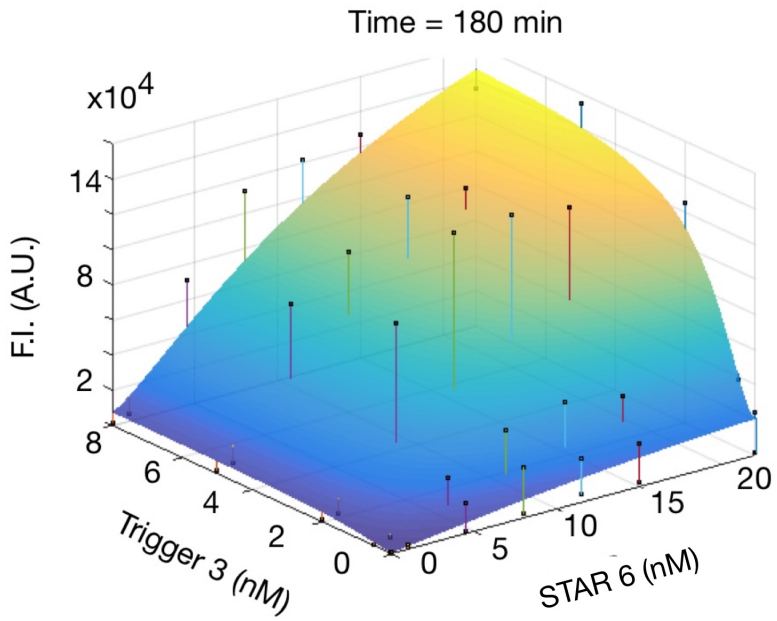


Fig. 2.15 Computational dose-response of the S6T3 gate after 180 minutes of expression in TX-TL systems. Experimental data are represented by the black dots, and the variance represents the standard deviation over three replicates. The model data was rescaled by the factor 80 to visualize the qualitative comparison.

Monitoring the transcription process would add some precious insights to enhance our understanding of the system's state. One advantage of the RNA-based regulatory system presented here is that the sfGFP 3' untranslated region (UTR) is not occupied by any regulating sequence and can therefore be functionalised with a fluorescent reporter aptamer [58]. Malachite Green Aptamer (MGA) can specifically bind to malachite green (MG) dye and induces it to emit far-red fluorescence signals. I used the MGA construct containing 4 tandem repeats (4x MGA) derived from the work of Yerramilli and Kim [134], which show that placing tandem of MGAs increase the signal to noise ratio (SNR) in living cells. The 4x MGA array was cloned into the 3' UTR of the the S6T3-sfGFP gate. Fig. 2.18 displays the kinetics results of the simultaneous transcription and translation processes. The resulting data are in agreement to what is expected from the mechanistic gate system. MGA signal remains at an undetectable level when the gate is not induced by the STAR or the trigger, or induced with the trigger alone. The transcription signal increases quickly in the two first hours when the gate is induced by the STAR alone, but no translated sfGFP is observed. When both STAR 6 and trigger 3 are present, the sfGFP increases until reaching a plateau around 5 hours, as observed previously. After reaching a production peak 1 hour after the start of the reaction, the mRNA level decreases in the TX-TL system as the degradation rate of the mRNA exceeds its production rate. Those results are in agreement with similar expressed constructs reported in the literature [79]. Finally, the production rate of the gate mRNA may be stronger when the two activators are present compared to the STAR alone, which would suggest a synergistic effect of those activating sRNAs, but the standard deviation of the mRNA signals are not different enough to fully support this hypothesis. A deeper characterisation about the MGA implementation in cell-free systems can be found in the next chapter.

Applying the ODE modeling framework to some additional gates, similar results as for the prototype gate were obtained, i.e., rough qualitative agreements which are helpful to guide the users to predict the dose-response concentrations. However, the quantitative agreement was mostly not achieved, especially for the S6T3 gate where a clear synergistic effect takes place that cannot be captured by the composition of the individual models (see Fig. 2.15).

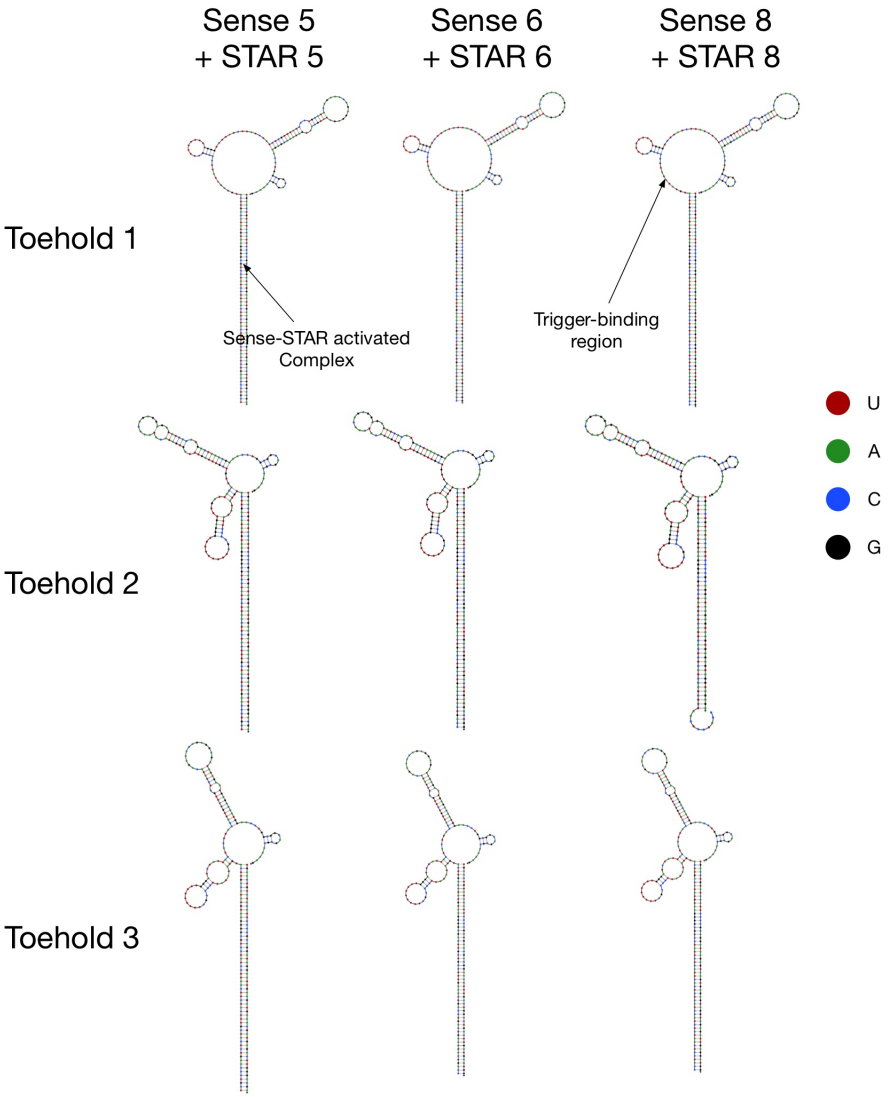


Fig. 2.16 MFE secondary structures computed with NUPACK online webserver. The nine AND-gates are shown in presence of their cognate STAR binding activator. The analysis tool was used with the following parameters: 37 °C, 0.0 M Mg^{2+}, 1.0 M Na^{+}, two RNA strands. The MFE structure is the structure of the ordered complex (Gate + STAR activator) which has the minimal total free energy. Details of the the equilibrium base-pairing properties computed by NUPACK are explained in [136].

Finally, I investigated if *in silico* folding simulations could explain the difference of the sfGFP production between S6T3 and the other gates. Free energies of the RBS regions were computed for a range of sub-optimal structures (Fig. 2.17), showing comparable results for gates with the same toehold part. Minimum Free Energy (MFE) structures of the gates complexed with their STAR cognates are displayed in Fig. 2.16. One can observe independent folding of the STAR and toehold secondary structures. The toehold component was not disrupted by the upstream spacer and STAR target terminator. The RBS-calculator tool

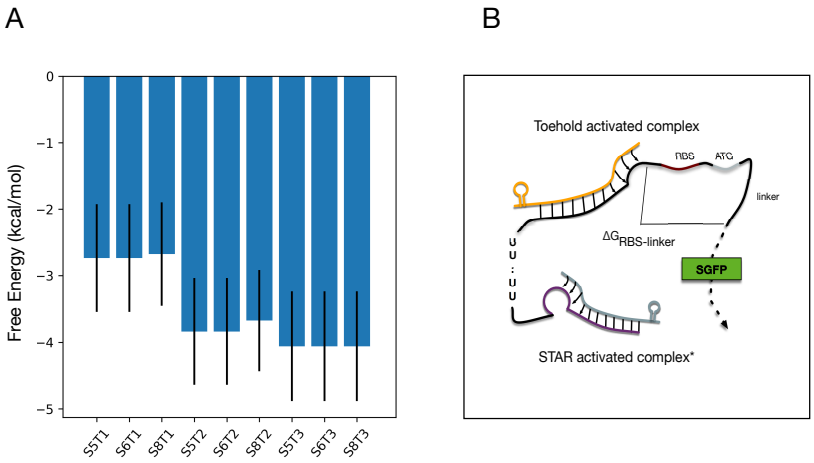


Fig. 2.17 Free Energy of the complexes RBS-linker regions. A) Mean Free Energies of the RBS-linker regions for a range of suboptimal structures. All suboptimal structures in an energy gap of $\Delta G = 2.5$ kcal/mol have been computed with NUPACK's `subopt` method. The means are calculated over the remaining suboptimal structures. Error bars: one standard deviation. B) Schema of the RBS-linker region for which the free energies are shown.

[99] was used to analyse out-of-frame upstream start codons, but no difference between gates sharing the same toehold sequences was detected. I suspect that the exceptional S6T3 performance is due to the co-folding effect of the gate sequence on the sfGFP mRNA, which either affects the translation process or the mRNA stability [29].

2.2.3 Dose-response Matrices Comparison

To further assess the behaviour of S6T3, I compared the traces corresponding to the most efficient concentrations (9.4 nM of STAR/2 nM of trigger) with a positive control that contains only a very strong RBS in its 5' UTR. As shown in Fig. 2.19 A, fluorescence intensity between the two constructs shows a similar order of magnitude, even though the plateau phase of S6T3 is reached one hour earlier. We computed the time derivative of the linear phase (Fig. 2.19 B) and did not observe a significant difference of the sfGFP production rate. This indicates that the translation efficiency of the S6T3 construct is near its optimum. This suggests that the observed earlier plateau is likely an issue of energy limitation related to the transcriptional burden associated with the presence of plasmids encoding the activators' RNAs. However, the depletion effect is small as both fluorescent values end up on the same order of magnitude. We found this result to be in agreement with previous experiments

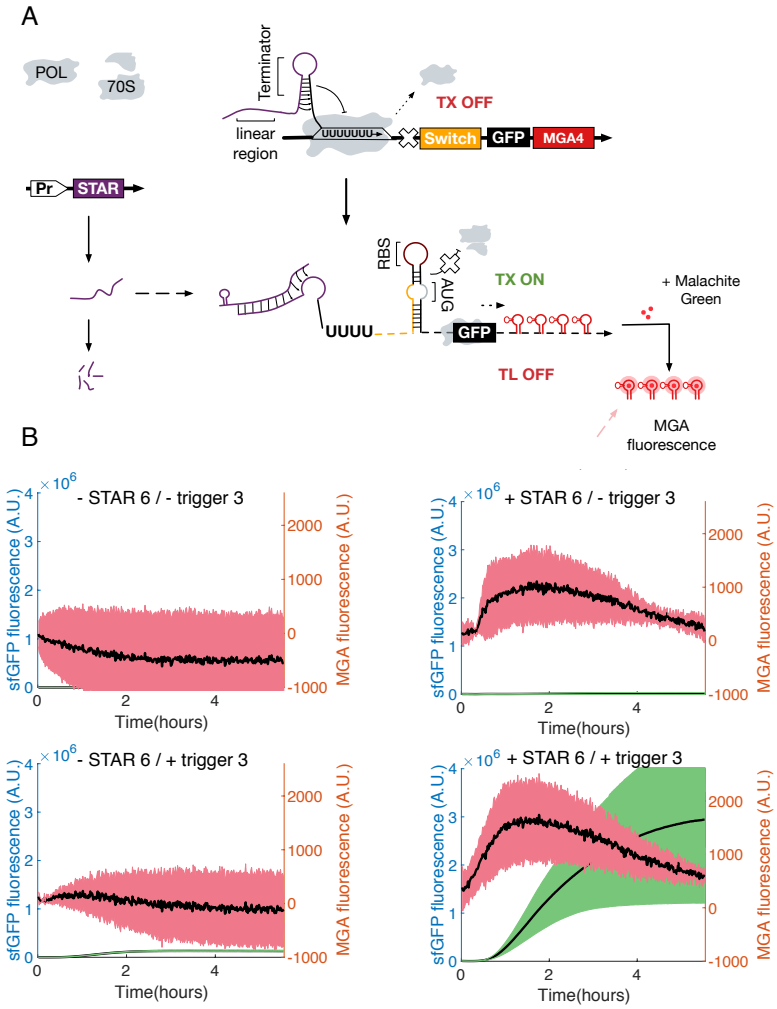


Fig. 2.18 Transcription and translation monitoring of the S6T3 gate. A) Scheme of the MGA-4x tandem located in the 3' UTR of the gate construct. The tandem of MGAs is transcribed upon the activation of the transcription by the STAR activator. B) TX-TL time-course reactions containing 2 nM of S6T3-sfGFP plasmids, in presence or absence of STAR, trigger or both activators. Green area represent the sfGFP fluorescence and red area, the MGA fluorescence level. Black lines represent the mean of the replicates, while the shaded regions represent standard deviation of those replicates (n = 3).

and model simulations of the influence of non-coding RNA transcription on the translation efficiency in TX-TL [109]. I then replaced sfGFP with dTomato, a red fluorescent protein, to compare the behaviour of the gate with a different downstream coding sequence. Similar to sfGFP, the order of magnitude of the fluorescence is comparable after 15 hours. However, the production rate of dTomato regulated by S6T3 is about half of the unregulated one. The dTomato sequence greatly varies from sfGFP (43.9% similarity score), and it has been shown

that the folding of the mRNA as well as the codon context can influence protein expression [21].

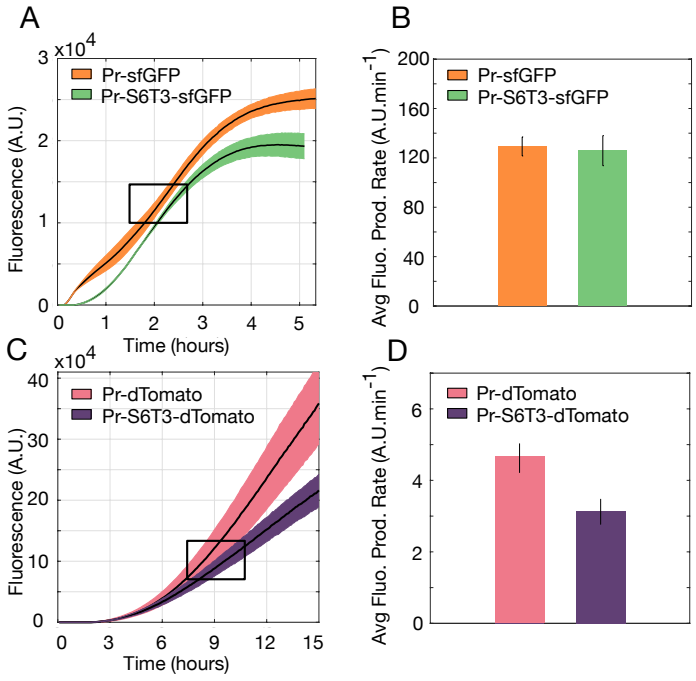


Fig. 2.19 Comparisons between the S6T3 gate and unregulated fluorescent reporters. A) and C) Fluorescence time courses of TX-TL reactions containing either 2 nM of DNA encoding sfGFP/dTomato with no RNA regulation, or with 2 nM of S6T3 gate in presence of 5 nM of trigger 3 and 15 nM of STAR 6 plasmids. Black lines represent the mean of the replicates, while the shaded regions represent standard deviation of those replicates (n=3). B) and D): Average sfGFP/dTomato production rates computed according to the data in the black boxed region from A and C, respectively.

Co-expression of protein in TX-TL seems to impose a heavier burden on resources consumption compared to the co-expression of non-coding RNAs. For instance, Shin et al. [106] reported a decrease of about 12-fold between the sfGFP production controlled by σ^{70} and an AND-gate based on the enhancer protein Nrtc and σ^{54}. Furthermore, STAR and trigger activators are considerably faster than protein activators, which require longer transcription, as well as supplementary steps (translation and maturation). This observation agrees with our expectation to obtain quick logic modules, notably faster than their protein-based counterparts [19]. The time delay of the circuit, defined as the time necessary to observe a detectable sfGFP signal, was about 15 minutes in this experiment. The response time of the P_{54}-based AND-gate was found to be around 30 minutes in TX-TL, therefore

twice as long as the time required by S6T3. Takahashi et al. [120] have found even more striking evidence for faster signal propagation with about 5 minutes per step of a RNA-based cascade in TX-TL against 140 minutes as measured *in vivo* [46].

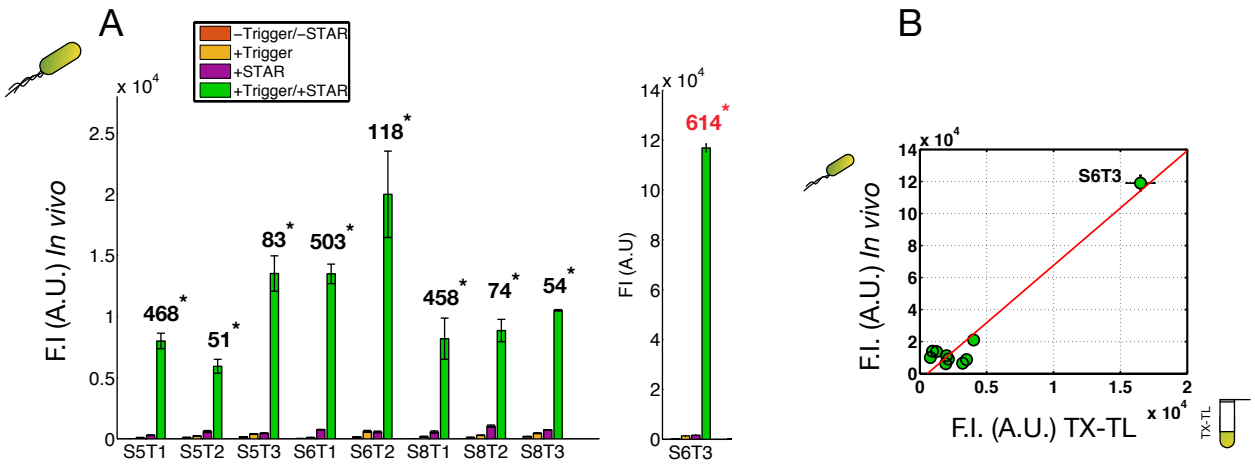


Fig. 2.20 Implementation of the RNA-based gates *in vivo*. A) Fluorescent measurements of the nine newly combined AND-gates in *E. coli*. The different colors indicate controls without any activators, with only STAR or only trigger. Green bars represent experimental measurements in presence of both STAR and trigger activators. Numbers with * indicate the fold range computed by the ratio between the activated gate and the signal without any activators. Error bars correspond to the standard deviation of three independent measurements. B) Corresponding relative fluorescence intensities between *in vivo* and *in vitro* for ten RNA-based gates.

2.2.4 *In vivo* Implementation

Performance Comparison

Fast prototyping of synthetic circuits *in vitro* often aims to transfer the selected candidates into living organisms. *E. coli* TOP10 cells were transformed with two plasmids, one containing the AND-gate controlling the sfGFP-reporter and the other containing one or both activators. Endpoint fluorescent measurements are shown in Fig. 2.20 A. The nine gates show a very low OFF-state in the absence of either STAR RNA or trigger RNA. In the presence of one RNA activator the sfGFP fluorescence increases only between 3- to 10-fold. If both activating RNAs are present, the reporter fluorescence goes up by more than 50-fold. Comparing the *in vivo* to the *in vitro* measurements, it is possible to see that both measurements are

in good agreement with each other (Fig. 2.20 B). The S6T3 gate has by far the highest ON-state in both the TX-TL and the *E. coli* measurements (614-fold). Those results add a supplementary evidence to the list of demonstrated correlation between cell-free-based and *in vivo* measurements, ranging from the characterisation of single parts [14, 101] to complex dynamic gene circuits [83].

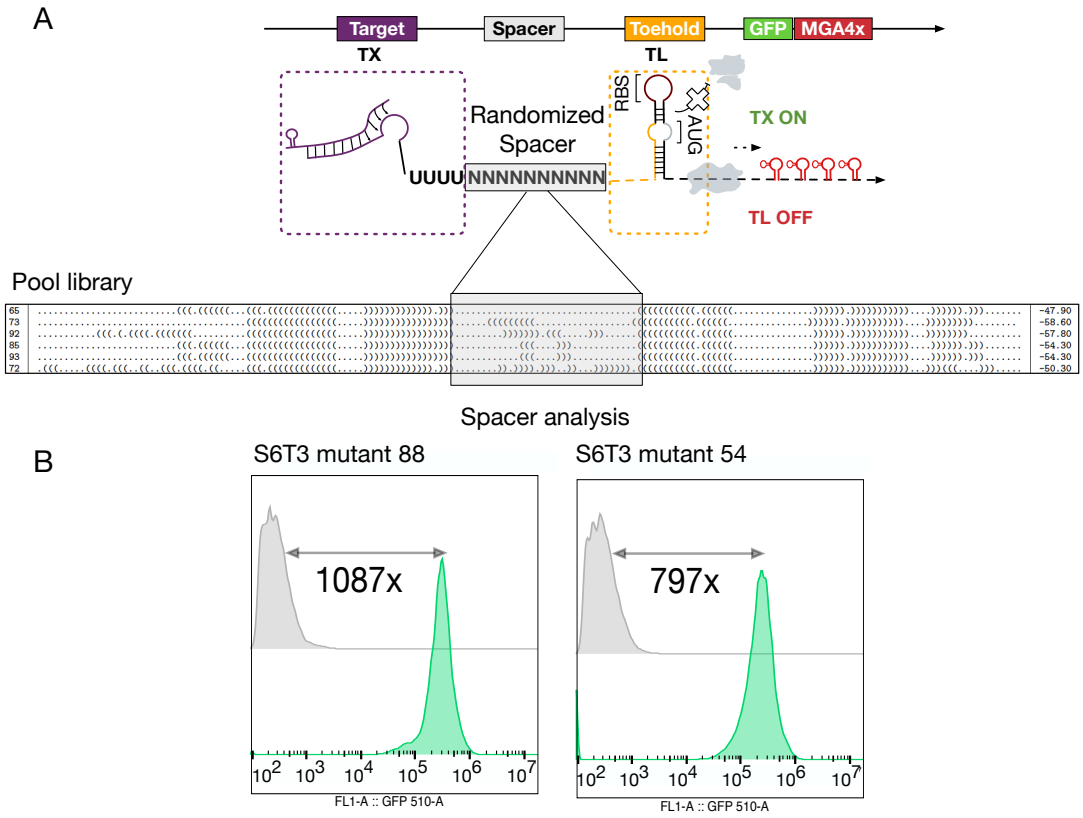


Fig. 2.21 Randomised spacer in S6T3 led to mutants with enhanced dynamic ranges. A) Design of the randomised spacer between the target STAR and the toehold regions. Examples of sequenced mutants are shown in the frame with corresponding secondary structures computed with the RNAfold package. B) Flow cytometry analysis of the two selected mutants, $S6T3-54$ and $S6T3-88$ with enhanced dynamic ranges. Fold-change is computed as the normalised ratio of the activated population median (green area) with the unactivated population median (grey area).

Spacer Analysis

A remaining question which was not investigated through the previous designs presented in this chapter is whether the spacer region between the two regulatory RNA elements

has an influence on the genetic expression of the combined gate. I have shown in section 2.2.2 that traditional secondary structure algorithms do not predict any binding of critical sequences (STAR terminator or toehold stem) to the spacer region. However, some of the gates presented some internal base-pairing between the spacer and the trigger-binding region. I wanted to assess if this sequence region critically influences the gate efficiency. I chose the best performing S6T3 gate to study this question since it possesses the highest dynamic range, increasing the chances of observing sequence-dependant variations. A library based on the S6T3-sfGFP template was generated by inserting ten randomised nucleotides between the uradine stretch of the STAR terminator and the the beginning of the trigger-binding region of the toehold switch. The pool was analysed in presence and absence of the co-activators in *E. coli* with FACS measurements. 100 sequences were sorted out and selected after gating on the criteria to belong either to the highest, medium or lowest fluorescence pool. Sequencing analysis was then performed and secondary structure computed for each of them. Two of the most occurring sequences (54 and 88) in the high pool and with the lowest occurrences in the low and medium pool were re-transformed and analysed with FACS (figure 2.21). The overall dynamic ranges of those gates increased compared to the wild-type S6T3 (797- and 1087-fold). This could be attributed to the lowered OFF-state while the ON-state level remained similar to the wild-type S6T3. The presence of the spacer #54 and #88 seem to have contributed to the independent processing of each gate component. This is supported by the secondary structures where nucleotides of the spacer region do not form any base-pairs with neither the STAR or toehold regions.

Chapter 3

Functionalising Cell-Free Systems with CRISPR-Associated Proteins

Preparation of *E.coli*-based cell-free extracts (CFEs) has been thoroughly described in the literature and many parameters of the protocol can be finely tuned to achieve different yields according to the gene expression context. For instance, expressing genes from a T7 promoter has been shown to work more efficiently in lysates in which the run-off reaction has been either skipped or shortened. The run-off reaction designates the step where the lysate is incubated at 37 °C to free the mRNA from the processing ribosomes. Inversely, expressing genetic constructs from an endogenous *E. coli* promoter requires run-off reaction lasting around 80 minutes. Sometimes, minor differences in the protocol steps have been linked to increased yield variances between batches and can therefore be a source of concern for achieving reproducible results. One part of the work presented in this chapter was to leverage the protocol steps found in the literature in order to achieve robust and reproducible cell-extract production in our laboratory. The second part focuses on the production of recombinant proteins in the cells, bringing specific functionalities to the resulting cell-extract. So far, only few reports demonstrated the ability to pre-expressed certain proteins or precursors while maintaining a high TX-TL activity in the produced lysate. In this thesis chapter, I studied various parameters of the cell-extract protocol as well as the expression vector in the chosen *E. coli* strain, and established a clarifying methodology for producing functionalised cell-free systems with two different CRISPR-associated proteins, dSpyCas9 endonuclease and Csy4 endoribonuclease.

3.1 A Wide Range of RNA-guided Nucleases

CRISPR-associated systems have become established tools in biological engineering. They were initially characterised as adaptive immune system that protects prokaryotes from foreign genetic materials, such as viruses. Indeed, a vast majority of archaea and half of bacteria harbor one or several CRISPR loci which are highly repetitive genomic regions composed of a series of nucleotide repeats separated by unique spacer sequences [2]. CRISPR loci can often be found co-expressed and co-adjacent to CRISPR-associated (Cas) genes, which are essential to the immune response. The unique spacers share partial genetic matches with bacteriophage genomes and plasmids. The CRISPR loci can thus be seen as a chronological record of the encountered viruses by the cell and its ancestors. CRISPR loci possess a large difference of repeat numbers: from 2 up to 250 and can be found either as a single CRISPR locus in many prokaryotes, but some can have dozen more [113]. Repeat sequences contain typically between 28 and 40 nucleotides in length, and spacer sequences go up to 72 nucleotides [2].

The CRISPR-Cas mediated immunity can be divided into three main steps [97]. The first step leads to the integration of newly encountered spacers in the CRISPR locus. During the second step, the system is translating the necessary Cas genes and transcribing the CRISPR array into a long precursor CRISPR RNA (pre-crRNA) which is processed into mature crRNA by Cas proteins and co-factors. The final step is the targeting of the matching nucleic acid sequence and its destruction by the complex formed by crRNA and Cas proteins.

CRISPR-associated systems have been re-engineered for many applications ranging from circuit regulation to genome engineering. Targeting CRISPR-Cas nucleases to cleave or bind to a chosen DNA sequence involve solely the co-expression of a non-coding sgRNA, which includes a spacer sequence needed for base-pairing with the sequence of interest [76]. Only few constraints exist regarding the design of the spacer: the main one requires the spacer to be flanked with a pre-determined protospacer-adjacent motif (PAM). Many different Cas proteins have been characterised and new ones are still discovered up to now. They are categorised according to the structure and function of Cas protein. The CRISPR-Cas systems are categorized into two classes (class I, class II), and subdivided into six types (type I–VI). Class I includes type I, III, and IV, and class II includes type II, V, and VI. Type I, II, and V systems are able to target and cleave DNA, the newly found type VI can target and edit RNA, and type III can edit both DNA and RNA. These types are further sub-divided in 30 subtypes with specific activities such as cleavage or degradation of the target, diverse PAM sequences recognition, processing of their own sgRNA arrays, and different levels of off-targets cleavage [76].

The method proposed in this chapter enables the enhancement of TX-TL systems with two pre-expressed CRISPR-associated nucleases: Csy4 or dCas9. They are among the most common Cas nucleases used in the circuit engineering field, but this method could be easily extended to any CRISPR-associated systems. Many CRISPR-Cas systems are not yet fully characterized and/or adapted for use in biotechnology, and TX-TL platforms may be a platform of choice for characterising and screening their associated components, as recently demonstrated by Wandera et al. [126].

3.2 Preparation of Cell-free Extracts

3.2.1 *E. coli* "Wild-Type" Extract Preparation

The general CFE protocol used in this work is based on the BL21 *E.coli* Rosetta™ 2 strain and is closely related to the protocol described from Silverman et al. [111] with some slight adaptations. More details about the preparation are available in the method chapter (section 6.1.1). The dialysis step is also omitted as it did not improve the yield of the CFE (data not shown). Table 3.1 describes the different CFE protocols developed during the last decade as well as the first one originally reported from Zubay [138]. The lysis and post-lysis are the steps with the high variability of parameters. Interestingly, it has been shown that the promoter-driven systems used to express genetic constructs in the CFE are step-dependent and strain-dependent. For instance, expressing a gene driven by a T7 bacteriophage promoter requires skipping the run-off reaction [63] for the BL21 strain. In this thesis, I focused on expressing RNA-circuits driven by the endogenous E. coli RNAP, and the protocol parameters for the lysis and post-lysis steps have been chosen accordingly.

To report on the yield of the CFE, a standardised control plasmid is used. It consists of a strong σ_{70} promoter followed by a 5' UTR containing a stability hairpin, a strong RBS, a gene coding for sfGFP, the MGA 4x tandem, and the T500 terminator (Fig. 3.2 A).

Lysis step

The sonication energy and the corresponding proceeded volume (in which the cell's pellet is resuspended) have been shown to be critical parameters for the yield of the CFE. Two specific volumes were used for the production of different extracts used throughout this thesis, depending on the initial amount of cultivated cells. 1 mL and 0.7 mL were used with an optimal sonication energy of 700 J and 500 J. This step is likely the one prone to most of the

Table 3.1 Summary of critical step differences between CFE preparation protocols. The red cells outline the steps changed compared to the most recent published protocol from Silverman et al. [111].

CFE protocol	Zubay, 1973	Sun et al, 2013	Kwon et al, 2015	Silverman et al, 2019	This thesis
starter culture	one 16-h	two 8-h	one 16-h	one 16-h	one 16-h
cuture media	2x YT	2x YT + P	2x YT + P + G	2x YT + P	2x YT + P
lysis method	French press	bead-beating	sonication	sonication	sonication
centri-fugation	300000 g 10min	12000 g 10 min	12000 g 10 min	12000 g 10 min	12000 g 10 min
run-off reaction	80 min	80 min	none for BL21	80 min	80 min
dialysis	18 h	3 h	none	3 h	none
flash freeze	none	none	before lysis	before lysis	none

yield variation between batches due to the human-handling insertion of the sonication probe in the resuspended pellet. To improve the consistency of the lysed CFE, an homogeniser could be used at this step to process large cultures, remove insoluble components and diverse cells debris [63]. A cooling time of 10 seconds for every 10 seconds of sonication were used in all cases. This parameter has previously been shown to not have an impact on the final yield.

Effect of the cells harvest time on the CFE activity

The machinery composition of the cells is affected by the time of harvest, which ultimately influences the yield of the CFE [63]. CFEs have shown functionality for a wide range of cells harvest times, from the exponential growth phase up to the saturation phase. The TX-TL enzymes are harvested in a more active state during the early exponential phase, but the increased cell mass obtained with higher OD_{600} compensates for the loss of enzymes activity [34]. In this work, I reproduce the results from Kwon and Jewett [63] with the standardised plasmid reporting for both mRNA transcription and protein production. *E. coli* BL21 rosetta 2 cells were grown to OD_{600} 1.9, 2.5, 3.5, 4, 4.5, 5.2, covering the range of

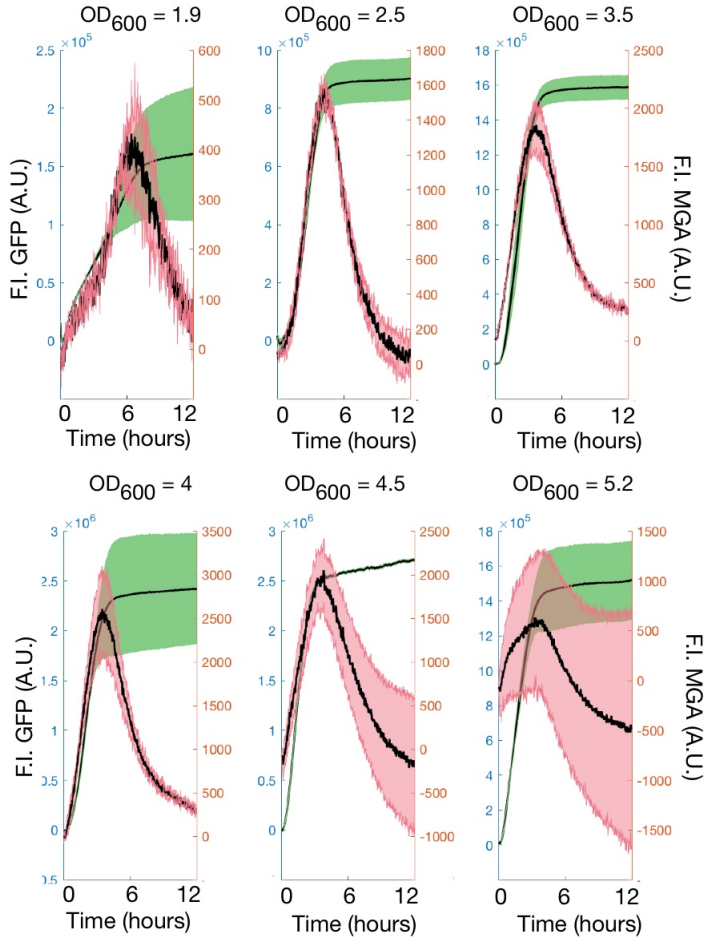


Fig. 3.1 TX-TL time-course reactions in CFEs produced at different times of cells harvest. Expression of 5 nM of plasmid Pr-sfGFP-MGA4x-T500. sfGFP fluorescence has a green shaded green region corresponding to the standard deviation of 3 replicates. MGA fluorescence has a shaded red region corresponding to the standard deviation of 3 replicates. Average of time traces are denoted in black.

mid to late exponential growth. For each OD_{600}, an independent crude lysate was processed and supplemented with energy buffer (see methods for a detailed explanation of the energy buffer preparation). The CFE transcriptional and translational activities were constant over all OD_{600} (Fig. 3.1). The best yields were obtained for ODs_{600} 3.5 and 4.2, and the lowest yied was for OD_{600} 2.1, confirming the trend published previously by Kwon and Jewett [63]. The mRNA level showed a much greater variation for any OD_{600}, with a maximum at OD_{600} 3.5 and a minimum for OD_{600} 2.1. Interestingly, mRNA signal reaches its maximum intensity level later in the lowest OD_{600} (around 6 hours) and earlier in the other ODs_{600} (between 4 and 5 hours) . The variation in the qualitative and quantitative states of the *E.coli* machinery, i.e polymerases, ribosomes or RNAses, resulting from the harvest at different OD_{600}, could explain those kinetic differences.

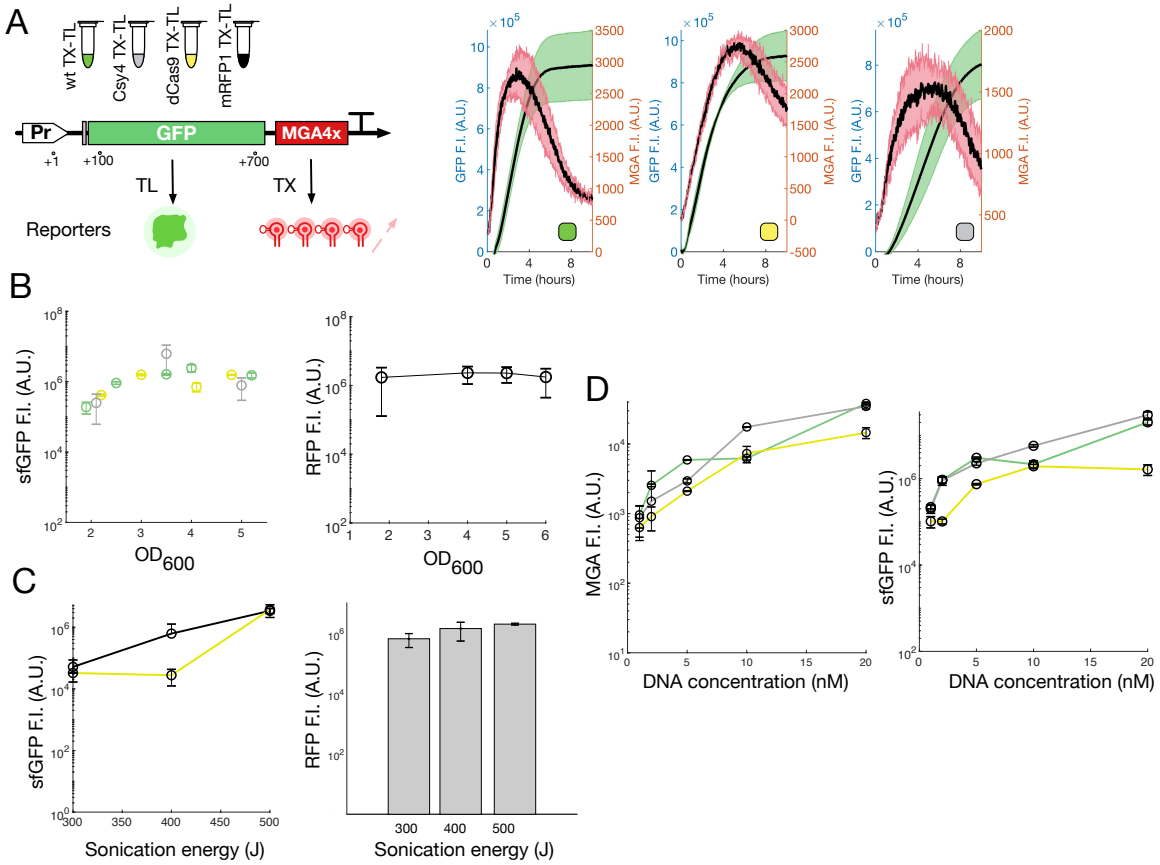


Fig. 3.2 Influence of the pre-expression of proteins in BL21 rosetta™ 2 cells on the resulting CFE efficiency. A) The standard plasmid reporter Pr-sfGFP-MGA4x enables simultaneous monitoring of the transcription and translation production. Four different CFEs are considered here: the wild-type (WT), and three CFEs pre-expressing Csy4 (grey color), dSpyCas9 (yellow), or mRFP1 (black). Examples of TX-TL time-course reactions for the WT, Cys4 and dCas9 CFEs expressing 1 nM of reporter plasmid. B) sfGFP (left) and mRFP1 (right) yields of the CFEs for different ODs_{600} (1 nM of reporter plasmid). C) Influence of the sonication energy on the CFE performance for mRFP1 and dCas9 CFEs (1 nM of reporter plasmid). D) Pr-sfGFP-MGA4x plasmid titration in three CFEs, with sfGFP fluorescence endpoints on the left and MGA fluorescence endpoints on the right.

DNA titration

Every CFE batch needs to be tested with a range of concentrations of the reporter plasmid. Indeed, the quantity and quality of the lysate machinery can variate considerably between CFE batches, which ultimatelty influences the DNA concentration boundaries suitable for circuit prototyping. (Fig. 3.3 A).

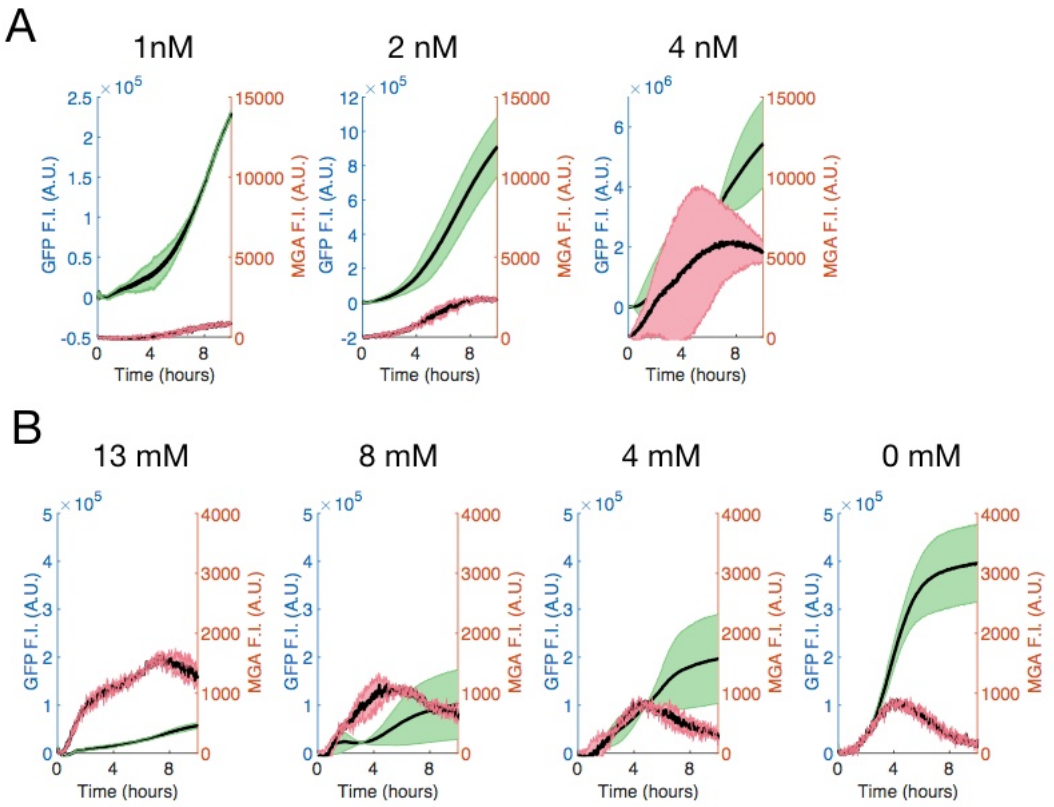


Fig. 3.3 DNA and maltose concentrations influence transcription and translation performance in CFE. A) TX-TL time-course reactions for 1, 2 and 4 nM of Pr-sfGFP-MGA4x-T500 plasmid. B) TX-TL time-course reactions for 0, 4, 8 and 12 mM of maltose. sfGFP has a shaded green region corresponding to the standard deviation of 3 replicates. MGA fluorescence has a shaded red region corresponding to the standard deviation of 3 replicates. Mean of fluorescent traces are denoted in black.

ATP regeneration system

In this work, the CFEs are supplemented with an energy regeneration system based on 3-phosphoglycerate (3-PGA). 3-PGA increases ATP regeneration as a precursor of the tricarboxylic acid cycle (TCA) [119]. Recent update of the protocol suggests the use of the maltose metabolic pathway for recycling inorganic phosphate (Pi), by-products of the protein synthesis, through the glycolytic pathway, which could be easily implemented due to the endogenous enzymes present in the crude extract [10]. In accordance with the literature, maltose was titrated between 0 and 13 mM in the CFE and resulting fluorescent time-course traces for translation and transcription are displayed in Fig. 3.3 B. Contrary to what was stated by Caschera and Noireaux [10], the protein yield decreases with the increase of maltose concentration in our CFE. However, the mRNA production level increases linearly with the

maltose concentration. These counter-intuitive results could indicate a complex inter-play of resource usage in the CFE. The maltose may shift the resource usage towards the RNAPs at the expense of the ribosomes. In a context of RNA-circuit prototyping in a cell-free environment, it is interesting to notice that TX activity could be enhanced through a simple maltose supplementation.

3.2.2 Pre-Expressing Csy4 and dSpyCas9 Proteins

Characterisation of the augmented CFEs

The CFE protocol was modified by transforming the BL21 Rosetta™ 2 strain with a second, high copy, ampicillin-resistant plasmid. Every step of the protocol was the same as the wild-type (WT), except that the cells had to be cultivated with both ampicilin and chloramphenicol antibiotics in the medium. The growth kinetics for the WT, Csy4 and dSpyCas9 were obtained by monitoring OD_{600}. The second plasmid expressing the recombining protein did not have a dramatic impact on the growth kinetics or doubling time (data not shown). Both Csy4 and dSpyCas9 CFEs were tested for different ODs_{600}, similarly to the experiment performed in the last section for the WT strain. The sfGFP endpoint measurements for 1 nM of control plasmid are displayed in Fig. 3.2 B and show a robust consistency between CFEs. The optimum yield is reached between OD_{600} 3 and 4 and the lowest yield are found between OD_{600} 1.8 and 2.5, consistent with previous experiments performed by Kwon and Jewett [63]. Kinetic fluorescent traces for the three CFEs in presence of 1 nM of control plasmid are shown in Fig. 3.2 A. Kinetics and yields were similar in those three systems, with sfGFP and MGA levels reaching similar values after 10 hours of expression. The mRNA signal reaches its highest level between 3 and 5 hours of expression. The Csy4 CFE exhibits, however, a slight delayed expression, but this variation of kinetic speed is likely due to the high variability of CFE batches, a known problem in the CFE field [105]. The reporter plasmid was titrated between 1 to 20 nM in the three CFEs. Similar saturation curves were found in the three systems for both mRNA and protein expression (Fig. 3.2 D) with highest yields found at 20 nM of plasmid, one order of magnitude higher than the yields obtained with 1 nM. Calibration curves were performed in order to quantify both mRNA and protein levels expressed by the CFEs. For the highest concentration of Pr-sfGFP plasmid, yields between 1 and 10 µM of sfGFP are reached in the three extracts, showing similar performance than the CFE protocol from Silverman et al. [111]. Correspondingly, mRNA level was found to reach 250 nM to 900 nM (corresponding calibrations can be found in Fig. 6.1).

Quantification of the pre-expressed protein

In order to quantify the concentration of pre-expressed proteins obtained in the final lysate, either for Csy4 or dSpyCas9, I produced a fourth CFE based on the constitutive production of mRFP1 with the J23108 promoter. sfGFP expression in the mRFP1 CFE follows a similar trend than the three other CFEs, indicating the correct functionality of the CFE. A first experiment in which the lysates were processed at various ODs_{600} did not show any significant variation of the mRFP1 level (Fig. 3.2 B). Sonication energies were varied between 300 and 500 J in a second experiment. Lysing the cells with different energies could indeed lead to different concentration of proteins in the lysate. An increase in sfGFP yield was observed with the increase of sonication energy for the mRFP1 and dSpyCas9 CFEs, similarly as previoulsy decribed by Kwon and Jewett [63]. However, the sonication energy was not critically influential on the level of mRFP1 (Fig. 3.2 C). Finally, a calibration curve was measured for mRFP1 fluorescence (See Fig. 6.1) and the mRFP1 quantity was found to be in the range of 30-50 nM in the final CFE. However, the quantity of dSpyCas9 and Csy4 cannot be directly inferred from those data, but it may give a rough estimate. For instance, I would expect dSpyCas9 to be several times less abundant since its coding gene size is about 4236 bps versus only 672 bps for mRFP1. However, Csy4 has a very similar coding size as mRFP1 with 561 bps and is expected to be in a similar concentration in the final CFE.

3.3 Assessing the Functionality of the CRISPR-associated Proteins

After having shown that the enhanced Csy4 and dSpyCas9 CFEs have a similar efficiency as their WT counterparts, the next step is to characterise the qualitative effects of those nucleases for the prototyping of synthetic circuits in TX-TL.

3.3.1 Csy4 Endoribonuclease as a Tool for RNA Processing

Csy4 or (Cas9f) is a endoribonuclease found in *Pseudomonas aeruginosa*, which recognises and cleaves a specific 28-nt long RNA sequence (Fig. 3.5 A). It was found to help the

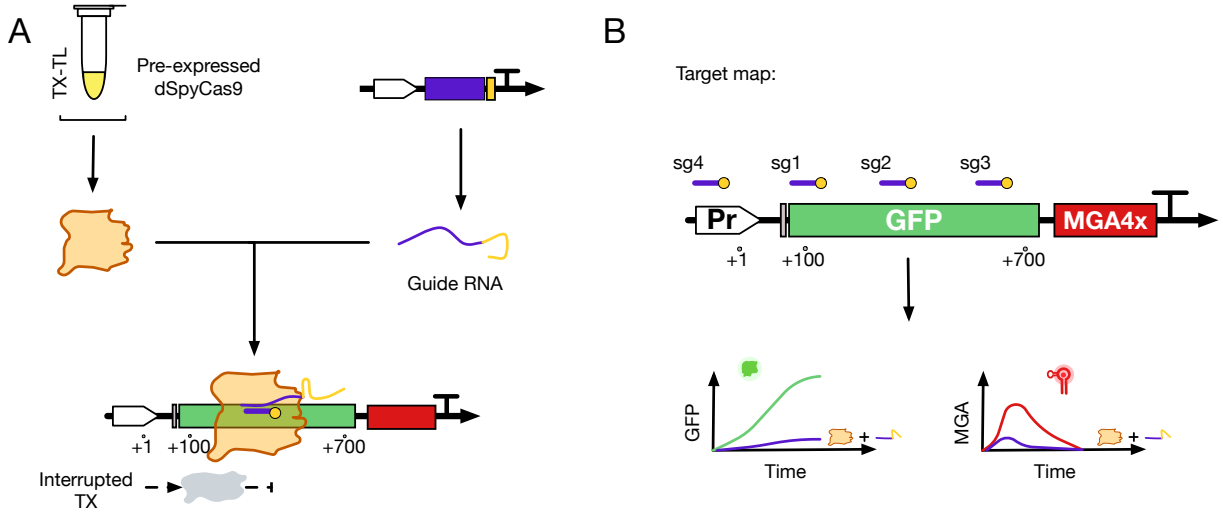


Fig. 3.4 Scheme mechanism of dSpyCas9. A) dSpyCas9 prepackaged in TX-TL. A plasmid expresses a small guide RNA, which after assembly with dSpyCas9, enables the blockage of the transcription elongation at the targeted location. B) The four different sgRNAs designed in this work target diverse locations on the reporter plasmid. They target regions from the endogenous promoter sequence up to the 3' end of the GFP. The two schemes indicate the expected behaviors for the mRNA and protein outputs, in presence (blue) or absence (green or red) of dSpyCas9.

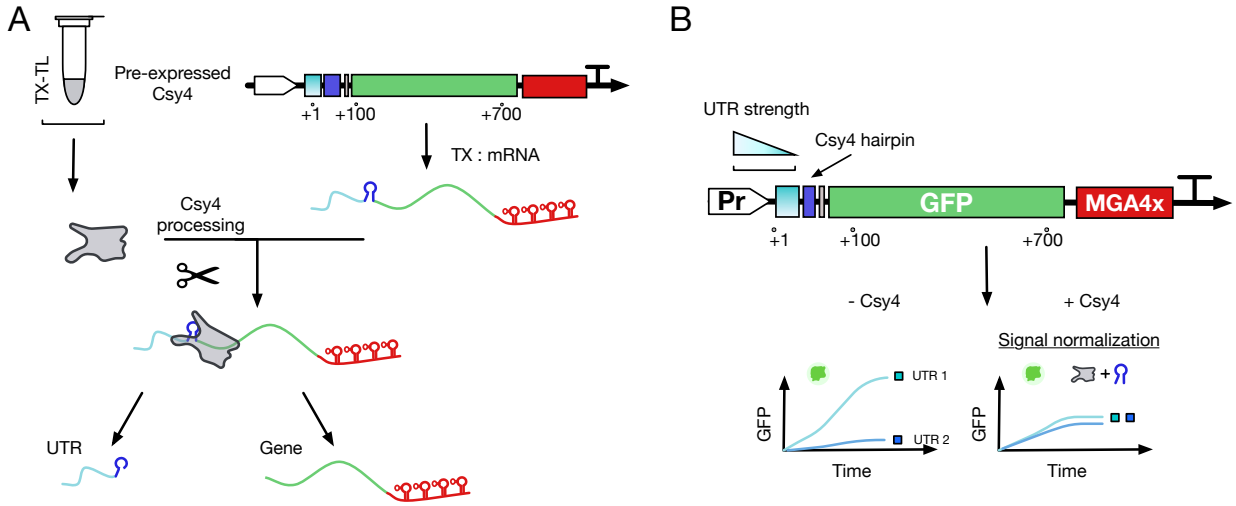


Fig. 3.5 Scheme mechanism of the endoribonuclease Csy4. A) Csy4 prepackaged in TX-TL system. Example of DNA construct expressing sfGFP with a 5' UTR and a Csy4 cleavage hairpin upstream to the start codon. Upon cleavage by the Csy4 endoribonuclease, the 5' UTR mRNA and the sfGFP coding mRNA are separated. B) Design experiment where 5' UTRs of various strengths are tested in presence (right curves) or absence of Csy4 protein (left curves).

transcribed processing of the long repetive elements belonging to the CRISPR locus. Recently, Qi et al. [96] demonstrated the use of Csy4 endoribonuclease in bacteria for improving the processing of synthetic regulators in a variety of circuit implementations. In order to test the efficiency of Csy4 as a RNA-processing tool in our TX-TL platform, a similar context experiment as the one described by Qi et al. [96] was performed. Briefly, various 5' UTR sequences were cloned upstream to the sfGFP, replacing the original 5' UTR used in the standarised reporter plasmid. These UTRs have been characterised with a range of translational strengths, leading to variable amount of translated proteins. However, after insertion of the Csy4 hairpin upstream to the gene, the downstream 5' UTR should lose its influence on the translational efficiency upon cleavage by the Csy4 endoribonuclease. Diverse UTRs processed by Csy4 should therefore lead to a similar, normalised GFP signal (Fig. 3.5 B).

5' UTR regions can be cleaved with Csy4 in TX-TL

In order to assess their respective strengths, the six chosen 5' UTRs were tested in the WT CFE at diverse ODs_{600} (Fig. 3.6 B). The 5' UTR sequences were chosen from the work of Evfratov et al. [32]. The data show a wide range of sfGFP yields, and follow a consistent trend between CFEs of different ODs_{600}. Next, I selected two UTRs elements harboring a large strength difference (Ion-UTR and RpvI-UTR) and performed kinetics with a set of control conditions (Fig. 3.6 A). First, the difference in sfGFP yield was verified in the WT CFE with 1 nM of constructs containing the Csy4 hairpin. 1 nM of constructs without Csy4 hairpin were also tested in the Csy4 CFE. A large difference in sfGFP yield was observed in this case as well. Finally, the constructs containing the Csy4 hairpin were tested in the Csy4 CFE, leading to a similar yield in sfGFP for the both UTR constructs. Those results strongly suggest the effective processing of the 5' UTRs by Csy4 in TX-TL. Interestingly, in these experiments, the mRNA output was found to exhibit a similar pattern as well, even though no effects related to transcription normalisation was excepted from the Csy4 endoribonuclease. However, the use of a DNA template at low concentration (1 nM) leads to a weak and noisy MGA signal, which could easily confound this hypothesis.

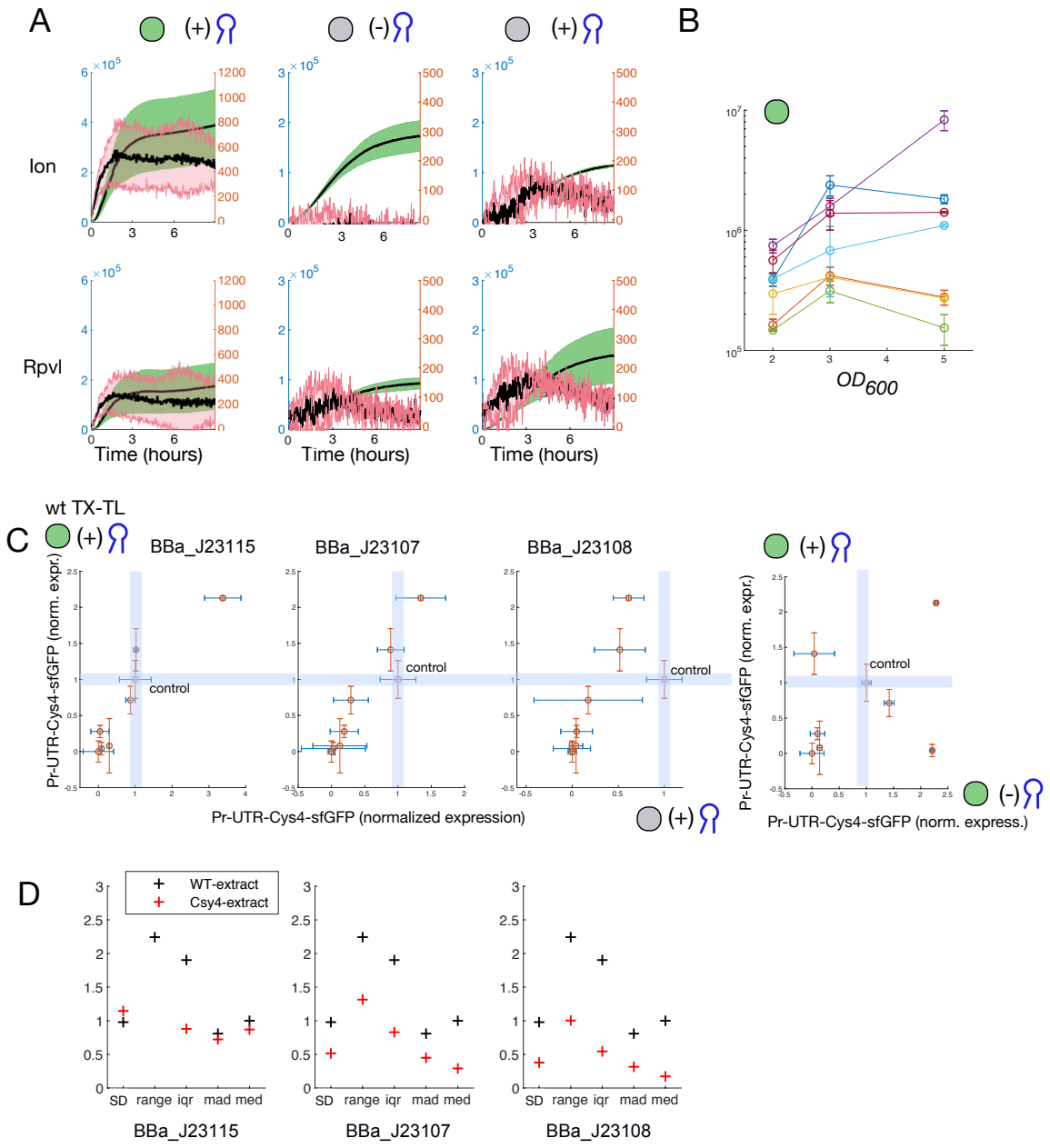


Fig. 3.6 Csy4 can process mRNA in TX-TL systems. Green circle = WT CFE, grey circle = Csy4 CFE. A) TX-TL time-course reactions of sfGFP and MGA for two different UTR constructs, expressed either in WT CFE or Csy4 CFE, with and without the Csy4 processing hairpin. B) sfGFP endpoint measurements of 7 different UTRs in WT CFE from various ODs_{600}. C) Normalised endpoint measurements for 1 nM of each UTR plasmid in WT and Csy4 CFEs. Three different Csy4 CFEs are compared, varying in the strength of the Csy4 promoter used. All data are normalised to 1 nM of the reporter plasmid. D) Measures of dispersion applied to experiments presented in C). The dispersion statistics are computed for each CFE experiment over all UTRs data (n = 18 in each experiment, with 6 different UTRs and 3 technical replicates for each). SD = standard deviation, iqr = interquartile range, mad = median absolute deviation, med = median.

Csy4 efficiency scales up linearly with the strength of the promoter

In order to confirm the results obtained previously, all UTRs constructs with Csy4 hairpin were compared as endpoint fluorescent measurements in both WT and Csy4 CFEs. Values were normalised to the reporter control plasmid which does not contain a Csy4 hairpin (Fig. 3.6 C). For the Csy4 CFE produced with the promoter J123108 (strength = 0.5), the UTRs constructs in WT CFE displayed diverse yields between 0.1 and 2.2 fold of the control construct, as expected from this short library of 5' UTRs. When expressed in the Cys4 CFE, the yield of every UTR construct was narrowed down to 0.1 to 0.9 fold of the control construct. As a complementary negative control, the test of all UTRs in WT CFE with and without Csy4 hairpins led to a bigger dispersion of yields in both cases (between 0.1 and 2.25 fold of the control construct), validating the effect of Csy4 processing in the Csy4 CFE. Two additional Csy4 CFEs were made with different constitutive promoter strengths. J23115 and J23107 correspond respectively to a relative strength of 0.15 and 0.34 compared to the reference J23119. The results are displayed in Fig. 3.6 C. The UTRs containing Csy4 hairpin expressed a narrower range of sfGFP yields in the Csy4 CFE than the WT CFE. However, the processing seem to not be as efficient as in the Csy4 CFE with the stronger J23108 promoter. In order to summarise the results, a set of five different statistics of dispersion are plotted in Fig. 3.6 D for each CFE. The standard deviation, the range, the interquartile range, the median absolute deviation and the median have reduced values in the three Csy4 CFEs compared to the WT CFE. Remarkably, the values of the selected dispersion measures decrease along with the increase of the promoter strength. This likely indicates a greater processing ability when a higher concentration of Csy4 protein is present in the CFE. It also suggests that Csy4 may be saturated by mRNA and indicates that the processing efficiency may decrease with an increase of mRNA production.

3.3.2 Characterisation of dSpyCas9 Activity in CFE

In order to assess the pre-expressed dSpyCas9 activity in the produced CFE, four different sgRNAs were designed with the help of the benchling CRISPR tool. Briefly, the 4 sgRNAs were selected to maximize the ON-target score and minimise the OFF-target score according to the model scoring guide by Doench et al. [27]. Three sgRNAs bind in the non-target strand of the sfGFP coding sequence (5', middle and 3' of the gene, denoted as sg-1,-2,-3) while the fourth sgRNA binds in the J23119 promoter sequence (sg-4). In presence of the sgRNAs,

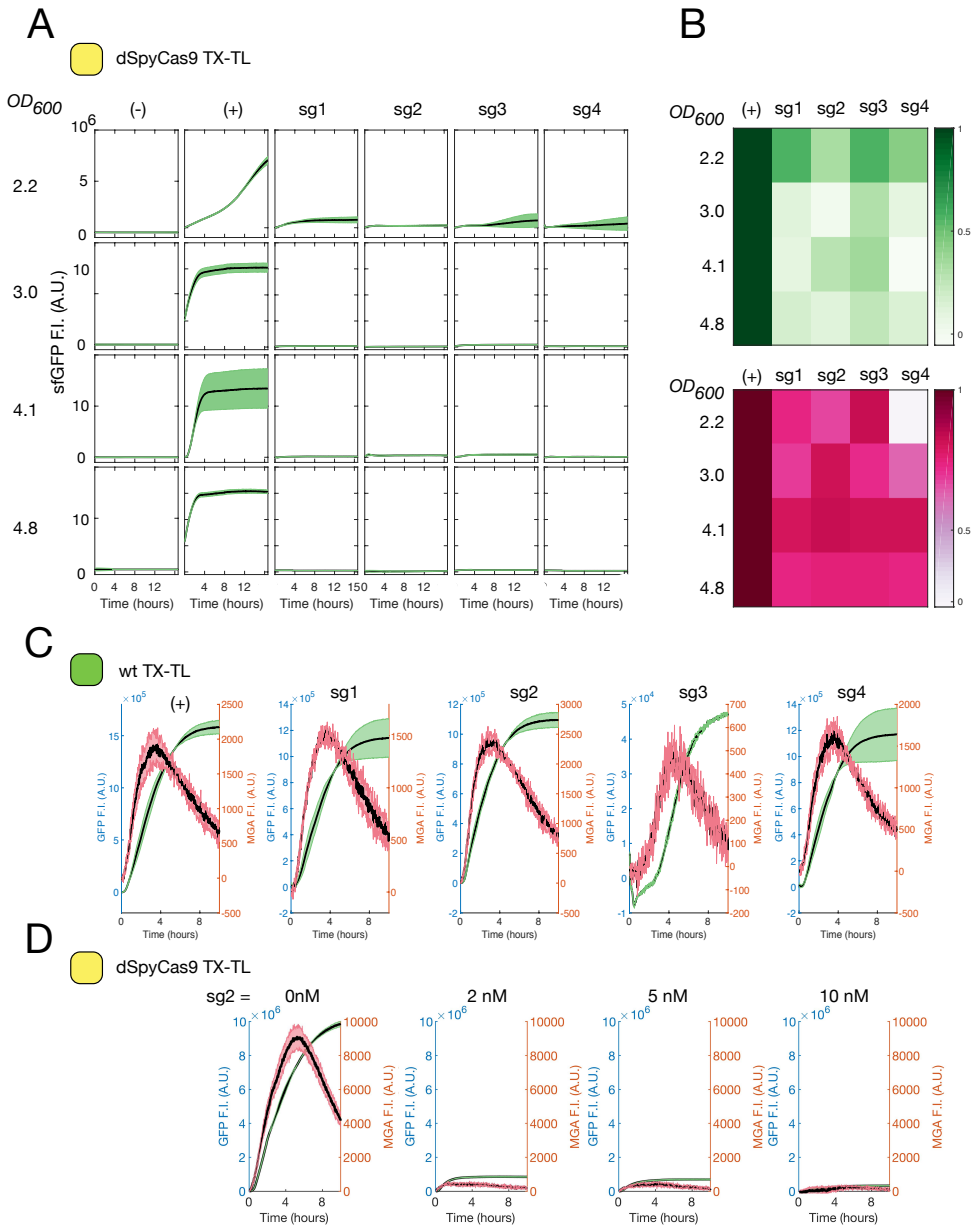


Fig. 3.7 dSpyCas9 is an efficient tool for gene repression in TX-TL systems. A) sfGFP time-course traces of 1 nM of reporter plasmid in presence of 4 different sgRNAs encoding plasmid (10 nM), expressed in dCas9 CFEs made from different ODs_{600}. The positive control (+) has been performed in presence of 10 nm of a plasmid coding for a decoy sRNA. B) Heatmaps summarising the corresponding endpoint fluorescences for sfGFP (green) and maximum fluorescence reached by MGA (red). Colorbar scales are transformed into the log space for increased visibility (1 = no repression, 0 = total repression). C) sfGFP and MGA time-course traces performed in a WT CFE with the same concentrations of reporter and sgRNA plasmids. D) sfGFP and MGA time-course traces performed in a Cas9 CFE of 2 nM of reporter plasmid titrated with sg2 (2 to 10 nM). Black lines represent the mean of the replicates, while the shaded regions represent standard deviation of those replicates (n = 3).

the control plasmid should therefore not express the sfGFP since the transcription elongation is hindered by the complex dSpyCas9-sgRNA located on the DNA template (Fig. 3.4). Both sfGFP and MGA fluorescent signals are therefore expected to be lowered in presence of the sgRNA expressing plasmid. First, 10 nM of sgRNA encoding plasmids were tested in presence of 1 nM of the reporter plasmid in the WT CFE (Fig . 3.7 C). The results show similar qualitative and quantitative kinetic traces for MGA and sfGFP fluorescences between the control experiment and the four different sgRNAs. Since no dSpyCas9 is present in the WT CFE, no inhibition of gene expression is observed. In the following experiments, 10 nM of plasmid coding for a decoy RNA (AD1) is used along with the positive reporter plasmid to mimic the resources consumption used by the sgRNA plasmid. the TX-burden imposed by the sgRNA encoding plasmid could decrease the TL yield, as demonstrated by Siegal-gaskins et al. [109].

dSpyCas9 is active regardless of the CFE harvest time

Four dSpyCas9 CFEs made from different harvest time spanning from a range of mid to late exponential growth (OD_{600} = 2.2, 3.0, 4.1 and 4.8) were used to test the activity of the four constructed sgRNAs. Results are shown in Fig. 3.7 A. Compared to the positive control, all sfGFP kinetic traces show a dramatic decrease in fluorescence for all conditions, regardless of the OD_{600} or of the sgRNA. This reflects the result obtained earlier about the quantification of the mRFP1 protein through a similar range of ODs_{600}, i.e. dSpyCas9 is present in similar concentration in CFEs made from different ODs_{600}. However, the gene target is less repressed in CFE made from cells harvested at OD_{600} 2.2. This could be linked to the slower kinetic, characteristic of CFEs made from cells harvested at the early exponential phase, like observed for the WT CFE in Fig. 3.1. Heatmaps summarising the corresponding fluorescent endpoints for sfGFP and MGA are displayed in Fig. 3.7 B. Sg1, sg2 and sg4 display a similar sfGFP repressing strength, and sg3 is the less efficient sgRNA throughout each OD_{600}. This may be explained by the target of sgRNA 3, located towards the 3' end of the sfGFP coding sequence, enabling the transcription of a near-to-complete sfGFP mRNA. The heatmap of the MGA signal, corresponding to the maximum MGA fluorescence intensity recorded during the time-course reaction, does not show any obvious correlations between sgRNA targets or ODs_{600}. However, it is consistent with the observation that the transcription is repressed for all experiments, compared to the positive control. This lack of more detailed transcriptional information could come from the low level of gene expression, leading to a low signal-to-noise ratio (SNR) for the MGA signal.

3.3.3 Applications for RNA-based Circuitry

After having determined the efficiency of the Csy4 processing and characterized the inhibition strength of the complex dSpyCas9 with various sgRNAs, the next step was to merge those new synthetic tools for TX-TL system with the RNA-based circuits prototyped in the 2nd chapter.

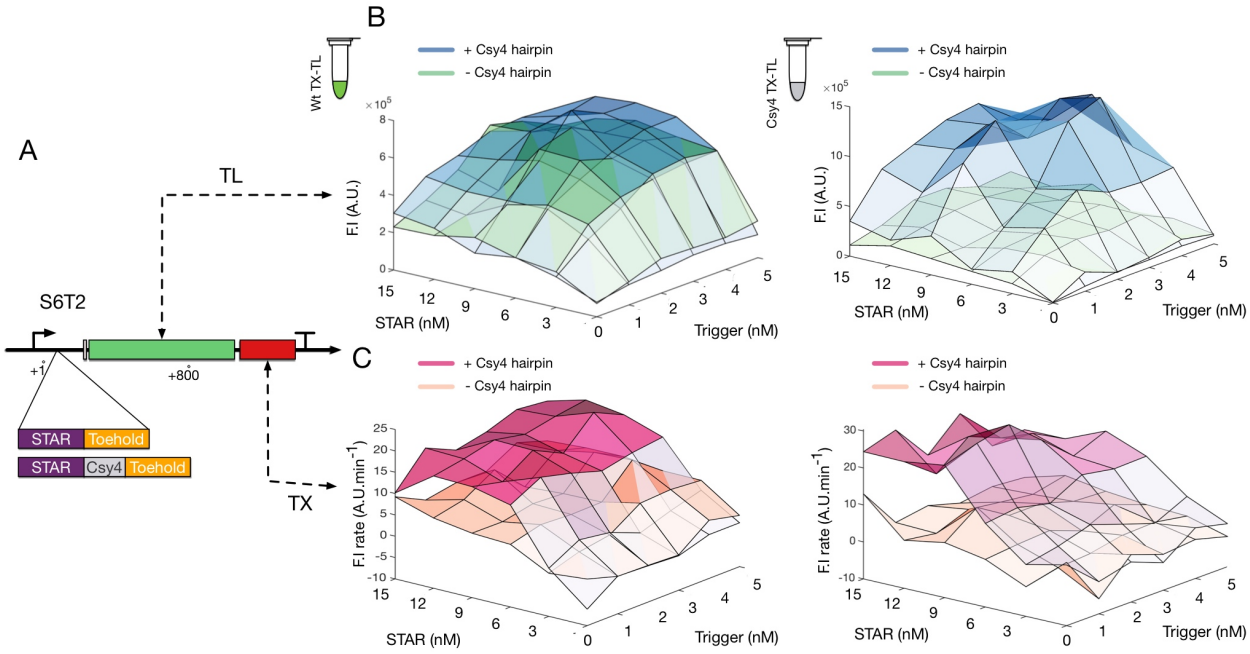


Fig. 3.8 Csy4 activity enhances the translational efficiency of the S6T2 AND-gate in TX-TL systems. A) Scheme of the S6T2-sfGFP-MGA4x construct, with and without the Csy4 hairpin between the sense 6 and toehold 2 components. B) sfGFP endpoint fluorescence measurements of the two gate variants in WT CFE (left) and in Csy4 CFE (right) with (blue) and without Csy4 hairpin (green). C) MGA fluorescence production rate of the two gates in WT CFE (left) and in Csy4 CFE (right) with (pink) and without (orange) Csy4 hairpin.

Csy4 enhances translation activity of S6T2 in cis-configuration

A direct application of the Csy4 ability to process RNA elements would be to test whether the system could enhance the RNA-based AND gates. I chose a low efficient gate, S6T2, from the screening procedure (see Fig. 2.13) and created a variant with the Csy4 hairpin inserted between the sense 6 and toehold 2 components (Fig.3.8 A). This configuration is

denoted as "cis-acting" since the three RNA components: sense, Csy4 hairpin and toehold, influence directly the TX-TL activities of the downstream gene. I hypothesised the Csy4 enhancement would be more obvious in a RNA-based gate with a low efficiency, since it is likely where RNA co-folding issues arise the most. First, the gates' dose responses titrated with STAR 6 and trigger 2 were compared in a WT CFE (Fig. 3.8 B). The dose-responses show similar landscapes for the sfGFP endpoints, indicating no difference of expression between the two constructs. This is a negative control used to verify if the Csy4 hairpin sequence alone could influence the gate, without the processing of the Csy4 protein. The two gate variants were then expressed in a Csy4 CFE with the same dose-response setup (Fig. 3.8 B). S6-cys4-T2 displays a higher dynamic range (around 10 fold increase) than S6T2 for the sfGFP fluorescence endpoints. Moreover, the fluorescence level remain similar when the two activators STAR 6 or trigger 2 are absent. This indicates that the enhanced dynamic range is likely not a Csy4 artefact due to the gate misfolding, which could result in high TX or TL leakage. Finally, Fig. 3.8 shows the MGA fluorescence rate of the corresponding experiments in WT and Csy4 CFEs. Interestingly, S6-csy4-T2 displays a higher TX rate than S6-T2 for concentrations of STAR above 6 nM in both CFEs, suggesting than Cys4 hairpin alone enables a better TX processing of the gate. The observed increased sfGFP level of S6-csy4-T2 can be attributed to the indirect enhancement of the TL process by the Csy4 endoribonuclease, since it was not observed in the WT CFE. This experiment also confirms the capacity of the sfGFP-MGA4x construct to monitor independently both mRNA and protein level for complex synthetic circuits. In particular, the decoupled TX and TL dose-responses obtained here confirm the independent activation of TX and TL processes by STAR and trigger (see section 2.2.2).

Blending of CFEs

Two constructs were assembled in order to combine the computational logic of the S6T3 gate and sg1-dCas9 system. Briefly, in the first construct, STAR6, trigger 3 and sg1 were assembled in tandem without spacer (I1). In the second construct (I2), a Csy4 hairpin was inserted between each sRNA sequence (Fig. 3.9 A). The combining logic of those 3 sRNAs should result into a NOT∧AND type of logic circuit. In other terms, STAR 6 and trigger 3 should act as the AND inputs of the gate, and the sgRNA should act as the logical NOT complement. Expressing the three inputs from the same genetic layer has multiple aims: First, it is important to be able to compress information encoding systems into the less possible space, as the complexity of the integrated circuits increase. Second, Csy4 processing ability could be useful to improve the processing of such synthetic sRNA operons. Three CFEs

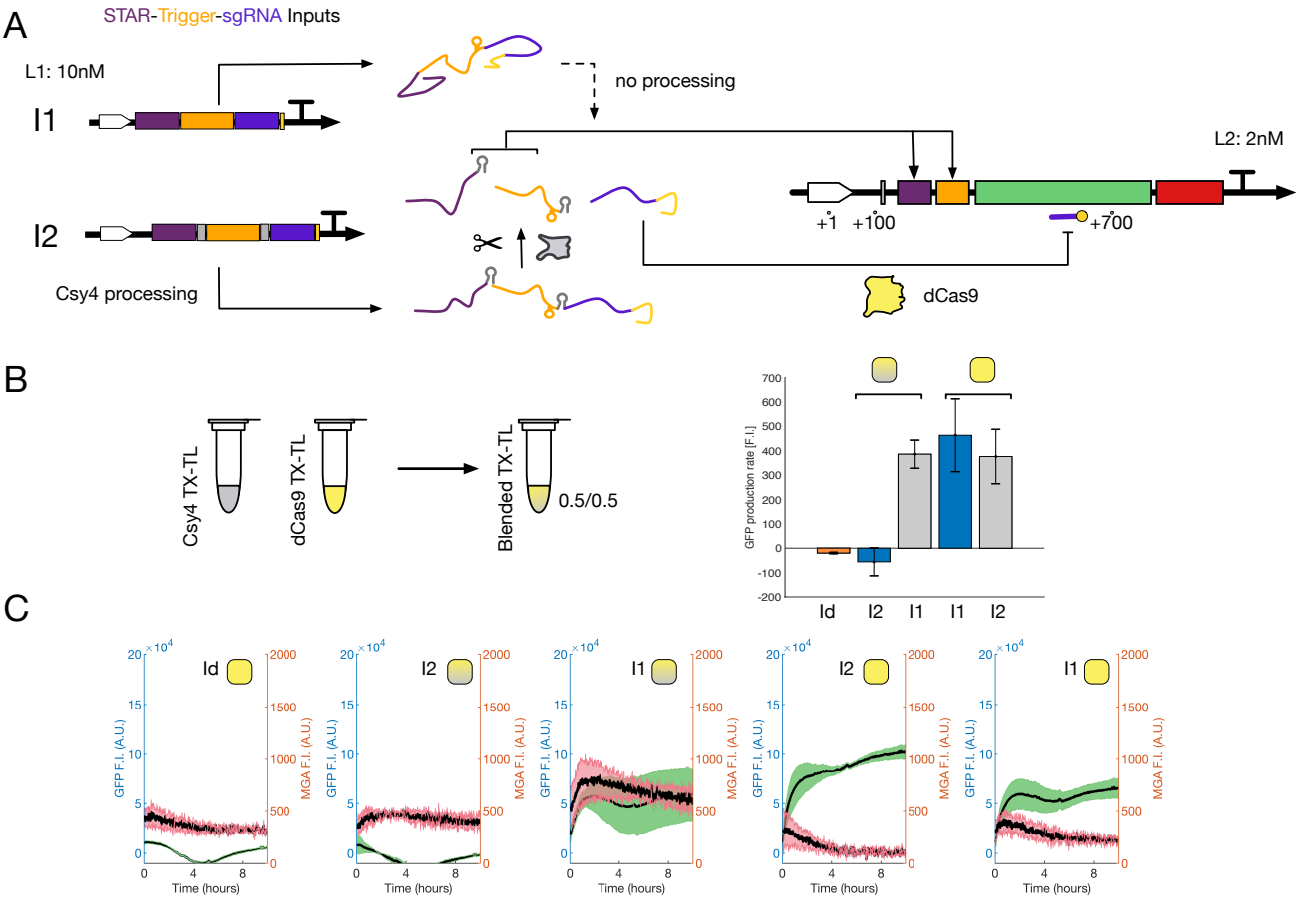


Fig. 3.9 Csy4 and dSpyCas9 CFEs can be efficiently combined to enhance the RNA-based toolbox in TX-TL systems. A) Scheme of the two different input DNA constructs, I1 and I2. I2 contains the Csy4 processing hairpin, hence once expressed in a Csy4 CFE, the inputs can be processed by the endoribonuclease. STAR 6 and trigger 3 activate the TX and TL of sfGFP, while the sgRNA 1 inhibits it. B) Csy4 (grey) and dSpyCas9 (yellow) CFEs mixed to form a "blended" CFE (shaded). Endpoint fluorescent measurements of sfGFP produced with 2 nM of S6T3 plasmid and 10 nM of either I1, I2 or ID (control) in different CFEs (colored squares). C) Corresponding detailed kinetic traces for sfGFP and MGA from the experiment shown in B). Black lines represent the mean of the replicates, while the shaded regions represent standard deviation of those replicates (n = 3).

were used to test these two constructs: Csy4, dCas9, and a blended CFE made from 0.5 part of Csy4 CFE and 0.5 part of dCas9 CFE (Fig. 3.9 B). In all CFE experiments, concentrations of I1 and I2 were kept at 10 nM and concentration of the gate reporter was set to 2 nM. Results are shown in Fig. 3.9 C and D. In the dSpyCas9 CFE, both I1 and I2 show slight activation, with similar sfGFP production rates (about 400 A.U. per minute). In the blended CFE, I1 triggers a similar activation level, likely due to the absence of the Csy4 hairpin. In presence of I2, no activation is observed, which demonstrates a complete inhibition of the gate. Additionally, expressing I2 in the blended CFE leads to a similar level of inhibition as

when expressing the gate in presence of a decoy sRNA (ID). Taken together, those results indicate that Csy4 and dSpyCas9 could be used in synergy when applied to RNA-based circuitry to decrease circuit leakage, increase the type of computational logic available, and increase encoding possibilities for the sRNA inputs.

Chapter 4

Beyond Cell-Free Batch Reactions: From Steady-State to Stochastic Experiments

4.1 Development of a Nanoliter-Scaled Chemostat

4.1.1 Limitations of the Batch Reaction Format

As shown in the previous chapters and from multiple published studies, genetic and metabolic networks can be implemented and studied in TX-TL systems. However, in most of the cases, the complexity of the circuits remains limited to sequential events i.e. genetic cascades, where the first layer of information activates or inhibits the other layers. It was the case for instance for the implementation of the RNA-based logic gates where two RNA-activators trigger the production of sfGFP (see 2nd chapter). Studies demonstrating the *in vitro* implementation of auto-catalytic circuits, either encoded as a negative or positive feedback, remain basic and limited to a couple of interacting species [48]. The implementation of more complex dynamical circuits is therefore limited by the batch-format in which the cell-free reactions are expressed. The batch environment of a test tube limits the output of the reaction and the gene synthesis rate decreases as energy precursors are consumed. The accumulation of by-products and the associated pH shift have also been shown to contribute to the slowing down of enzymatic reactions [82]. After a couple of hours, the chemical reactions are therefore at equilibrium in a batch-tube. Cell-free reaction times have been successfully extended with diverse strategies like the use of exchange devices of large scale, the semi-batches systems [114], or at the micrometer scale by expressing the TX-TL system inside vesicles with a phospholipid membrane in a feeding solution [86, 39].

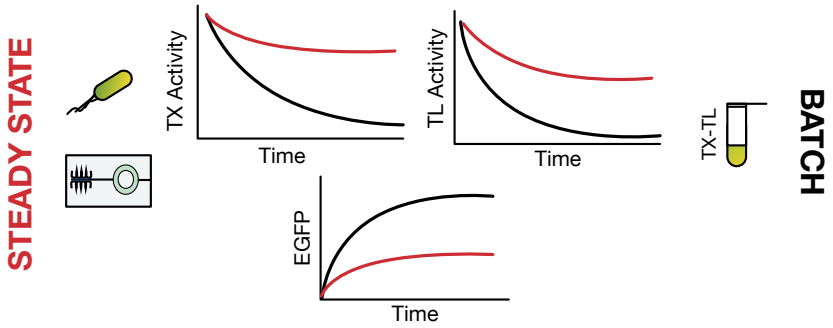


Fig. 4.1 Scheme of the transcription-translation rates of a cell-free reaction when performed in a batch-format (black lines) or at steady state in living cells or in the chemostat (red lines).

These examples have contributed to increase significantly the protein yield of TX-TL system, which is the main concern of many biotechnology production processes. However, those set-ups with continuous reaction conditions are not fully adequate for the prototype of dynamical synthetic circuits such as oscillators or toggle switches. These circuits require the degradation of intermediate inhibitors products to reset the state of the circuit nodes when several negative feedback loops are presents. The degradation pathways are extremely efficient *in vivo*, where part of *E.coli* metabolism is placed upon multiple feedback loops [75]. In TX-TL, the protein products accumulate in the batch tube since the proteases are usually in negligible amount and expressed proteins would require a specific degradation tag. Strategies to implement active degradation pathways such as the use of ClpXP to specifically target proteins harboring degradation tags have been investigated and offer promising solutions [39, 103]. It remains still challenging to implement large genetic networks since active degradation will be ultimately limited by the energy capability of the batch system and Clp are ATP-dependent proteases. Recently, Niederholtmeyer et al. [82] have developed a microfluidic device to overcome those limitations [83]. It consists of a 33-nL reactor operating at steady-state for an extended period of time where successive dilutions of the cell-free/DNA mixture occur in discrete steps. The first stage of this project was to implement a similar device at the BCS laboratory, from the device fabrication to the software operation.

4.1.2 Microfluidic Implementation

Design and Fabrication

For running the TX-TL reactions at steady state, the chip was designed with two layers of PDMS. The chip design is similar to previous published implementations, in particular the one published by Niederholtmeyer et al. [82]. However, I down-scaled the reactor volume

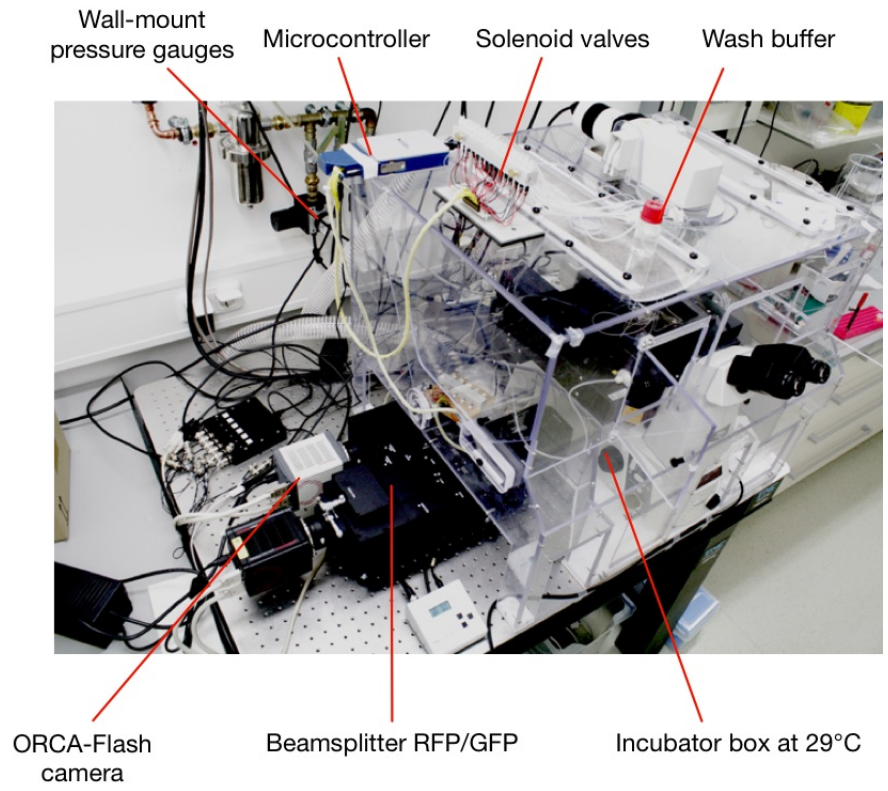


Fig. 4.2 Photography of the microscope set-up for microfluidic experiments. The solenoid valves are connected to the control layer of the chip. The computer controlling the microscope is not shown.

from 33 nl to 10 nl. The first layer contains the flow channels and the reactor where the TX-TL system and reactants are flushed with (flow layer). The second layer contains the control channels and Quake valves [5] connected to the solenoid valves (control layer).

The first version of the chip contained only one reactor. This single reactor design aimed to facilitate the set-up implementation and configuration, and to minimise the complications arising from having to manage parallel reactors. One main distinction of my design is that a pulse width modulation control (PWMC) module was incorporated between the inlets and the reactor core. This module will enable the generation of a range of desired concentration inputs by mixing the packages of different inputs at different frequencies [1].

In the single-core version of the chemostat (Fig. 4.3), four inlets are available to connect the fluids, i.e. DNA, buffer or TX-TL system. These inputs can be controlled independently by Quake valves, c.g. for opening or closing the corresponding channels. The bypass function enables the rapid flushing of the reactants in the channel when switching the inputs. The first peristaltic pump after the PWMC module enables to meter the reactants of different package size according to the flow rate and frequency patterns of the valves. A second peristaltic

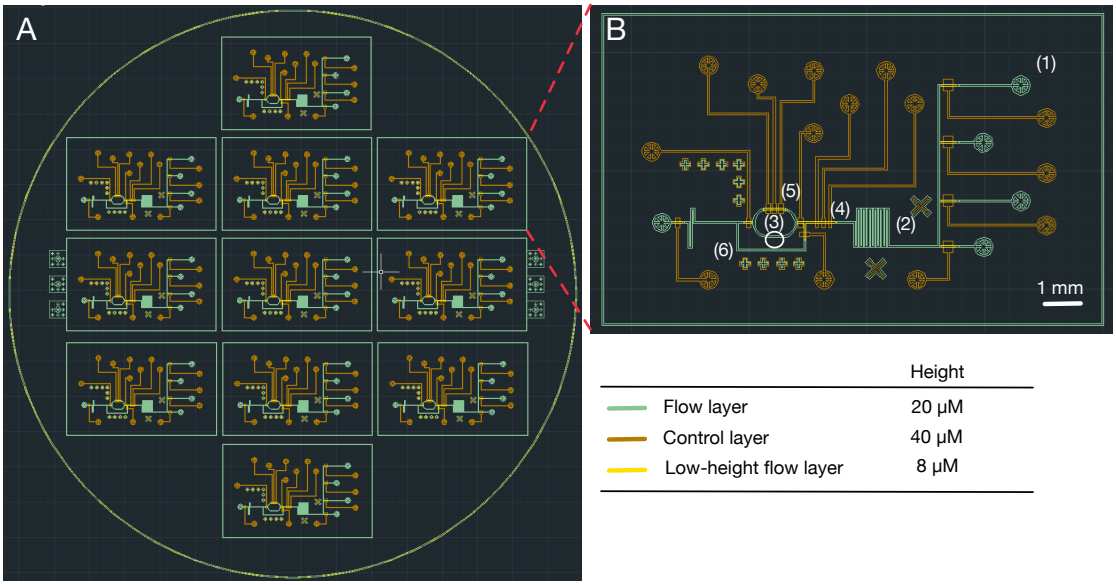


Fig. 4.3 AutoCAD design of the microfluidic chemostat. A) 10 parallel chips can be produced from the same wafer. B) Close-up of a single chip design. High-height flow channels are in green, control layer is in marroon, and low-height flow channels are in yellow. (1) 4 inlets connected to the flow channels, (2) PWMC module, (3) chemostat reactor and imaging position (white circle), (4) first peristaltic pump, (5) second peristaltic pump, (6) bypass channel.

pump is located inside the reactor and enables the proper mixing of the reactants. The small volume of the microfluidic reactor requires an active mixing strategy to be completely mixed within the time-scale of the TX-TL reaction. The reaction ring can be completely sealed from the external flow channels by closing its inlet and outlet valves.

Two different wafers were used for the fabrication of the flow and control layers. The fabrication protocols for each layer are described in Table 6.7. Briefly, I used standard photolitography techniques. The flow layer required the use of two different photoresists, as the channel heights needed to be lower beneath the valves to close quickly and robustly the flow channels. Finally, the PDMS chips were fabricated according to standard multilayer soft litography and the detailed corresponding methods can be found in section 6.1.4.

Microscopy Set-up and Chip Operation

The microscope set-up is described in Fig. 4.4 and the detailed list of the components are listed in Table 6.12 and Table 6.9. Shortly, the chip was first mounted on a metal plate and the incubator box was set up at 29 °C. The solenoid valves (15 independent peristaltic valves) were pressurised with air pressure ranging from 1.5 to 2 bar and were controlled by an internal lab-made Matlab script. They were connected to the chip and filled with water to completely

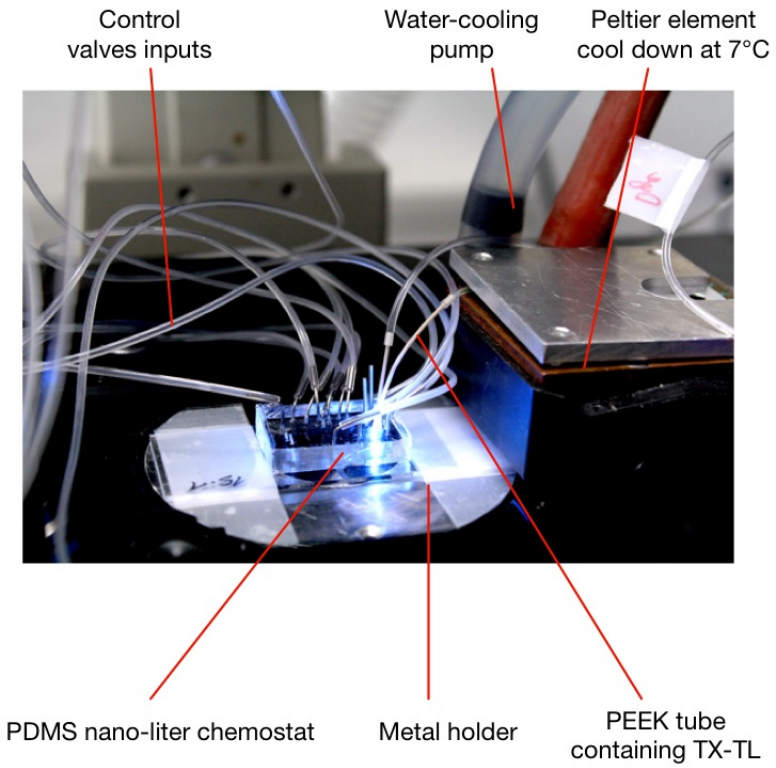


Fig. 4.4 Close-up photography of the PDMS-based microfluidic chemostat inside the microscope incubator chamber.

remove the air from the control channels, which could have otherwise leaked into the flow channels. The valves where then tested independently for each channel by an automatic script closing and opening them at 1 minute interval during 20 minutes to test their robustness and assess any possible leakage. The flow inlets were then connected to the different components (TE buffer, one or two inputs DNA, and TX-TL system) and pressurised with 0.1 to 0.4 bar depending on the experiment. I used a home-made cooling system to keep the TX-TL at 7 °C during the time of the experiment. A watercooling pump was used to cool down the heated side of a peltier element where the cooled TX-TL system was stored in a tygon tube. The tube was connected to the chip with a polyetheretherketone tube to minimise the amount of dead volume not cooled by the peltier element. A Graphical User Interface (GUI) was developed in Matlab to run the chip characterisation and perform the experiments. The software enables to program the successive dilutions of the TX-TL-DNA mixture contained in the reactor at discrete interval of time with a defined number of DNA and TX-TL packages. Those numbers were computed from the chip characterisation described in the section below. The loading of DNA and TX-TL system were operated through the use of the first peristaltic pump (see Fig. 4.5) and after each injection of fresh reactant, the second peristaltic pump

was used to mix the content of the reactor during 1 min. The duration of the experiment ranged from a couple of hours to 15 hours.

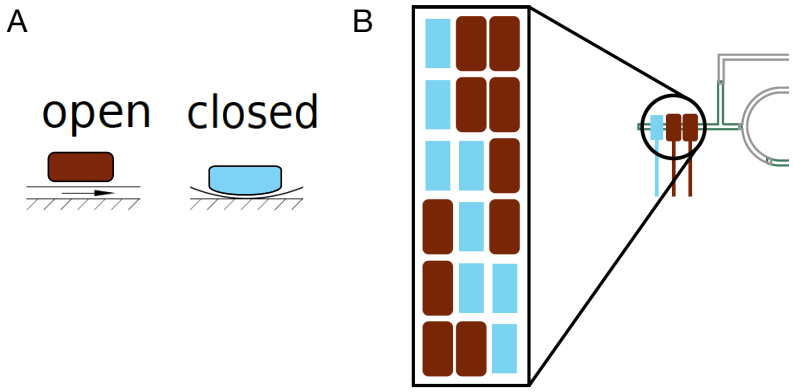


Fig. 4.5 Schema of a microfluidic peristaltic pump based on Quake valves. A) Mechanism of a single PDMS Quake valve. On the left, the valve is in its open configuration when it is not pressurised (red). On the right the valve is in its closed configuration, as pressurisation causes compressing of the flow channel underneath (blue). B) Actuation sequence of the peristaltic pumps used in the chemostat chip for mixing and reactants metering.

Characterisation of the microfluidic chemostat

Before each experiment, a calibration procedure was made with the chip to determine the dilution time t_d. This term can be found by finding the dilution rate λ which is unique to each chip. λ may vary because of differences in PDMS ratio, leading to diverse thickness and stiffness properties between batches of chips. Adjusting the flow rate contributes as well to the variance of λ between each experiment since the pressure gauge control was analog-based.

n represents a fixed number of pump cycle delivering a fixed amount of volume nV_d at each interval of time ΔT. For any liquid volume at initial concentration c_0, I get

$$V_T c_0 - nV_D c_0 = V_T c_1,$$

where each dilution step of n pump cycles gives a concentration ratio

$$\frac{c_1}{c_0} = 1 - n\frac{V_T}{V_D},$$

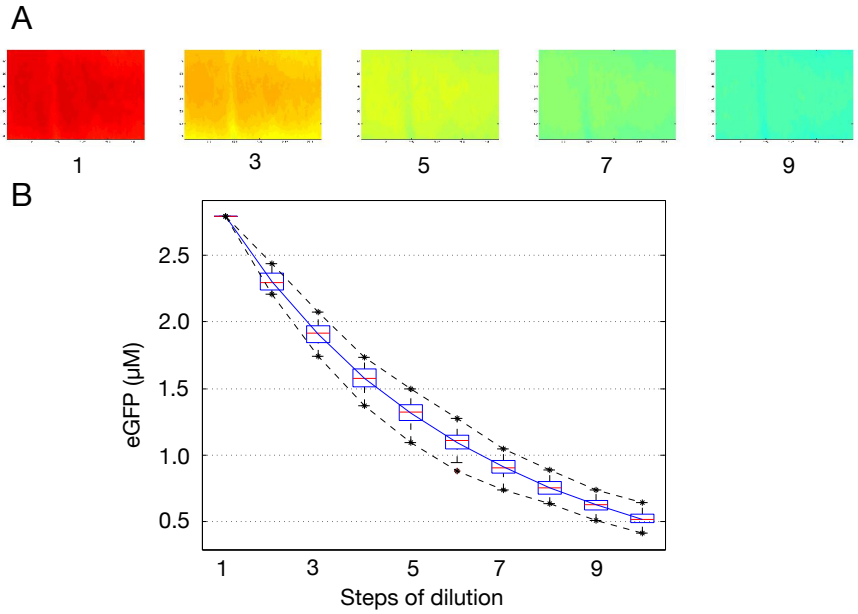


Fig. 4.6 Characterisation experiment of a single chemostat reactor to compute the dilution rate, λ. The number of cycle pumps was fixed at 10 and the successive values of eGFP fluorescence are plotted in figure B. The blue line represents the exponential fit of the data and the standard deviation are represented by the dashed black lines. A) Fluorescent image channel (colored) of the reactor at different dilution time steps (red = high GFP fluorescence, blue = low GFP fluorescence)

After N steps of n pump cycles each, I have

$$\frac{c_N}{c_0} = \left(1 - n\frac{V_T}{V_D}\right)^N,$$

with $\nu = 1 - n\frac{V_T}{V_D}$ being the fraction of the reactor exchanged at each step, I have

$$c_N = c_0 e^{Nln\nu} = c_0 e^{-\lambda t},$$

giving the dilution rate

$$\lambda = \frac{ln\nu}{\Delta T},$$

from which I can now define the dilution time which is the half-life corresponding to the dilution rate

$$t_d = \frac{ln(0.5)}{\lambda},$$

I can now performed the experimental characterisation and compute t_d from the resulting data. I used a purified solution of eGFP (3 µM) to initially fill the reactor ring and performed sequential dilution with Tris buffer. This was repeated for a different range of pump cycles ($n \in [3, 10]$) to get a robust estimation of the parameters. An example of resulting curve is given in Fig. 4.6. Finally, the following TX-TL experiments can have their dilution rate tuned by modulating n and ΔT, with the following formula :

$$t_d = \frac{ln(0.5)}{1 - n\frac{V_T}{V_D}}\Delta T.$$

Interestingly, the number of pump cycle required to exchange the same percentage of fraction volume v was found to be much lower in my design than the original device reported from Niederholtmeyer et al. [82]. The both designs possess Quake valves of the same dimensions (100 µM x 100 µM) but my implementation has narrower and shallower channels, and therefore a greater quantity of fraction volume is displaced when the valves are pressurised. The proper function of the input peristaltic pump is checked by plotting a package-responses relationship (see Fig. 4.7 B). The linear correlation between the number of inserted packages and the average fluorescence cropped from the reactor ring is an important criteria for inserting the correct fraction of different reactants. The second peristaltic pump inside the reactor is also checked by plotting the average fluorescence of the reactor through the mixing time. It should result in a stabilised fluorescence value after some rounds of mixing, as shown in Fig. 4.7 C, which shows a efficient mixing inside the reactor under 1 minute. Each damped oscillation corresponds to a complete turn of the fluorescent package getting progressively diluted in the surrounding non fluorescent buffer.

The final step is to proceed to the start of the experiment by setting up the GUI control software with dilution time computed from the values of ΔT and n previously found. The detailed procedure steps can be found in section 5.1.4. The typical ΔT was chosen according to the doubling time of living *E. coli* at 37 °C, i.e $\sim$ 20 minutes. It corresponds to a dilution rate of the reactor of 0.2 h^{-1}. A live-data analysis plugin from ImageJ was used to ensure that the TX-TL reaction entered a steady state after few hours and to detect any flow abnormalities.

4.1.3 Transcription and Translation at Steady-State

I hypothesised the nanoreactor would enable TX-TL synthesis at constant rate given that the decreasing rate of TX and TL activities would be compensated by the fresh intake of enzymes and energy molecules. For the first testing procedure expressing synthetic circuits in TX-TL,

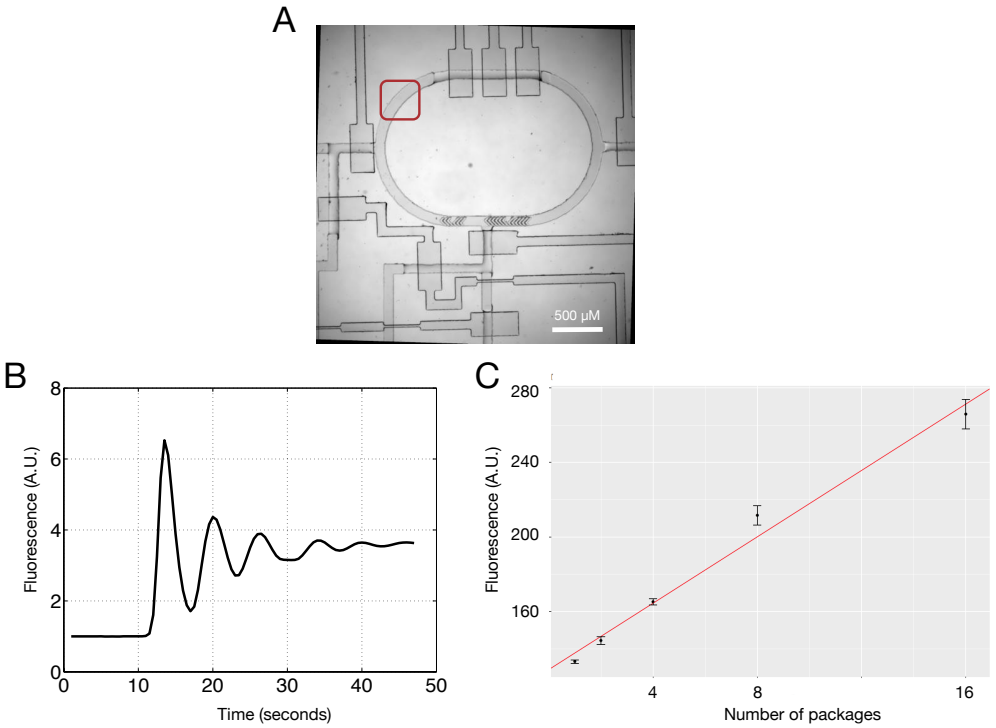


Fig. 4.7 Characterisation experiments of a single chemostat reactor. A) Bright field of the microfluidic reactor, the red rectangle area indicates the microscope field of view with the 10x lens for the fluorescence channel. B) Average fluorescence value of the area over time when active mixing is operated in the reactor. C) Linear dependency between the number of eGFP packages inserted in the reactor and the corresponding fluorescence values after 1 minute of active mixing.

I performed protein synthesis from a DNA construct expressing a deGFP reporter driven by a strong *E. coli* promoter [39]. deGFP represents a start codon-optimised version of the enhanced (e)GFP. In order to be compatible with a fast prototyping approach, I used linear DNA fragments from PCR amplification as template for gene expression [119] (Fig. 4.8 A). Since the complex RecBCD has been shown to degrade linear DNA in TX-TL, purified GamS was added to the TX-TL system. GamS is a nuclear inhibitor of the RecBCD complex by interacting with its DNA binding site [26]. The steady state of the TX-TL system, i.e. when the protein synthesis rate is replenished at the same rate as it is diluted, was reached after an average of 2 hours with a dilution time of 20 minutes. Increasing the dilution time to 27 minutes did not have an influence on the time needed to reach the steady state level (Fig. 4.8 B). The dynamic nature of the dilution time was shown by tuning the time interval ΔT from 20 to 40 minutes, after the TX-TL reaction reached the initial steady state. Overshooting of the fluorescence level was observed, corresponding to an increase of deGFP protein accumulation, which was progressively damped over the dilution steps until reaching a new steady state several hours later (Fig. 4.8 C).

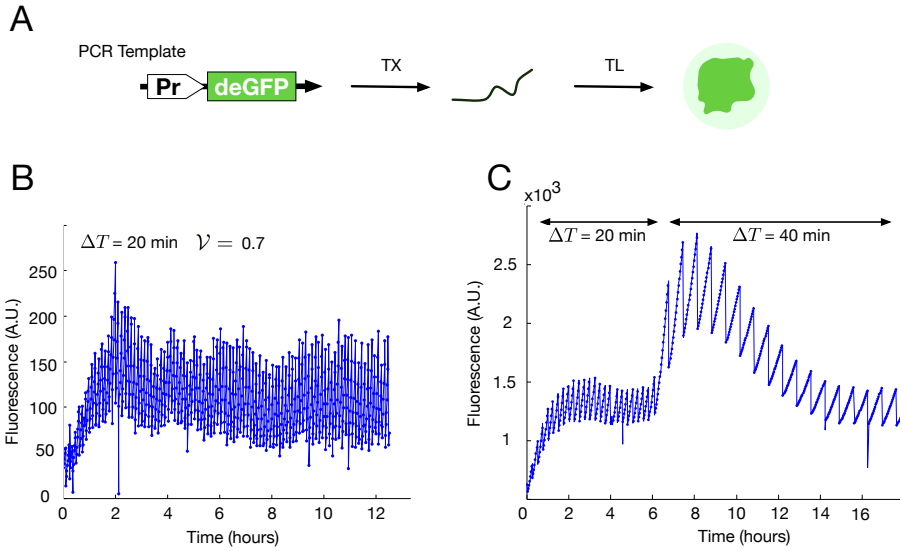


Fig. 4.8 TX-TL systems reach a steady state in the chemostat. A) Scheme of a simple linear DNA template used for the dilution time tuning experiments. The deGFP mRNA is trancribed from a PCR product downstream of a strong promoter, and is translated into a mature deGFP protein. B) Steady-state reaction of the corresponding reporter circuit with 7 nM of DNA template, with a dilution time of 27 minutes. C) Steady-state reaction of the corresponding reporter circuit with 7 nM of DNA template, with dilution time switching from 20 minutes to 55 minutes.

Monitoring the GFP signal is a robust way to follow the translation process, but it is important as well, in the context of the RNA-circuitry work presented in this thesis, to monitor the mRNA production. Therefore, in order to track the transcription production, I tested a DNA template containing the iSpinach reporter downstream of a T7 bacteriophage promoter [6]. Like the MGA aptamer, iSpinach is an aptamer which upon binding to its fluorogenic co-factor 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI), emits fluorescence in the GFP spectra (Fig. 4.9 A). Compared to its parent aptamers (Spinach, Spinach 2.0, Brocoli), iSpinach harbors a greater stability when expressed in *in vitro* environments such as TX-TL system. The corresponding experiment in the nano-ractor is presented in Fig. 4.9 B. We observed that the transcription reaches as well a steady-state, in a similar way that the previously monitored translational GFP activity. The iSpinach reporter is therefore compatible with the use of the chemostat/microscope set-up, which will be an important asset for the context of the future experiments based on RNA-based circuits.

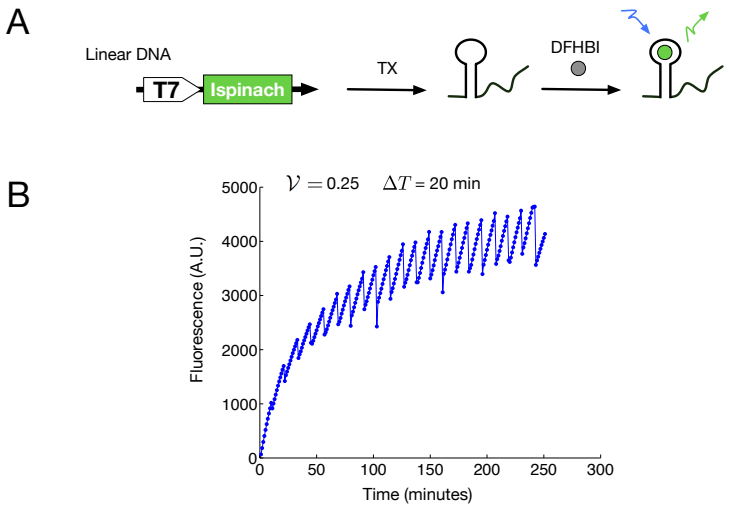


Fig. 4.9 TX system reaches a steady state in the chemostat. A) Scheme of a linear DNA template used for the experiment. iSpinach is trancribed from a PCR product downstream of a T7 promoter, and upon folding and binding to its cognate ligand, can be monitored in the GFP channel. B) Steady-state reaction with 20 nM of the corresponding reporter circuit (dilution time of 45 minutes.

4.1.4 Implementation of Riboswitch Regulators

The next step was to implement genetic circuits, which could be regulated post- or co-transcriptionally by tuning the concentration of a signal inducer in the nanoreactor. This would enable the control of the genetic network state by modulating the external inducer at various concentrations. As a first implementation, the thiamine pyrophosphate (TPP) riboswitch from *B. subtilis tenA* gene was cloned into the 5' UTR of the deGFP reporter construct. The TPP riboswitch enables the control of gene expression through a variety of mechanisms in different species [117]. The one used here acts as an OFF-switch in presence of the TPP molecule. Upon binding with TPP, the stem-loop of the riboswitch sequence forms a terminator hairpin preventing the host polymerase to extend. In the absence of TPP, the riboswitch structure forms an anti-terminator structure disrupting the terminator, and therefore enables the transcription processing. Control experiments were first performed in a batch-format to determine the activity of the TPP riboswitch in TX-TL systems. Interestingly, a high concentration of TPP ligand (above 500 µM) was found to reduce significantly the gene expression level in TX-TL systems (Fig. 4.10 B). To minimise its toxicity level in the following experiments, the TPP concentration has been set at 100 µM. The TPP riboswitch functioned as expected in the TX-TL system, with a yield about five-fold higher when expressed from a circular template compared to the linear PCR product (Fig. 4.10 C). The GFP expression was reduced by about three-fold in both cases when the TPP-ligand was present in the system (Fig. 4.10 C). It is interesting to notice that the dynamic range of the

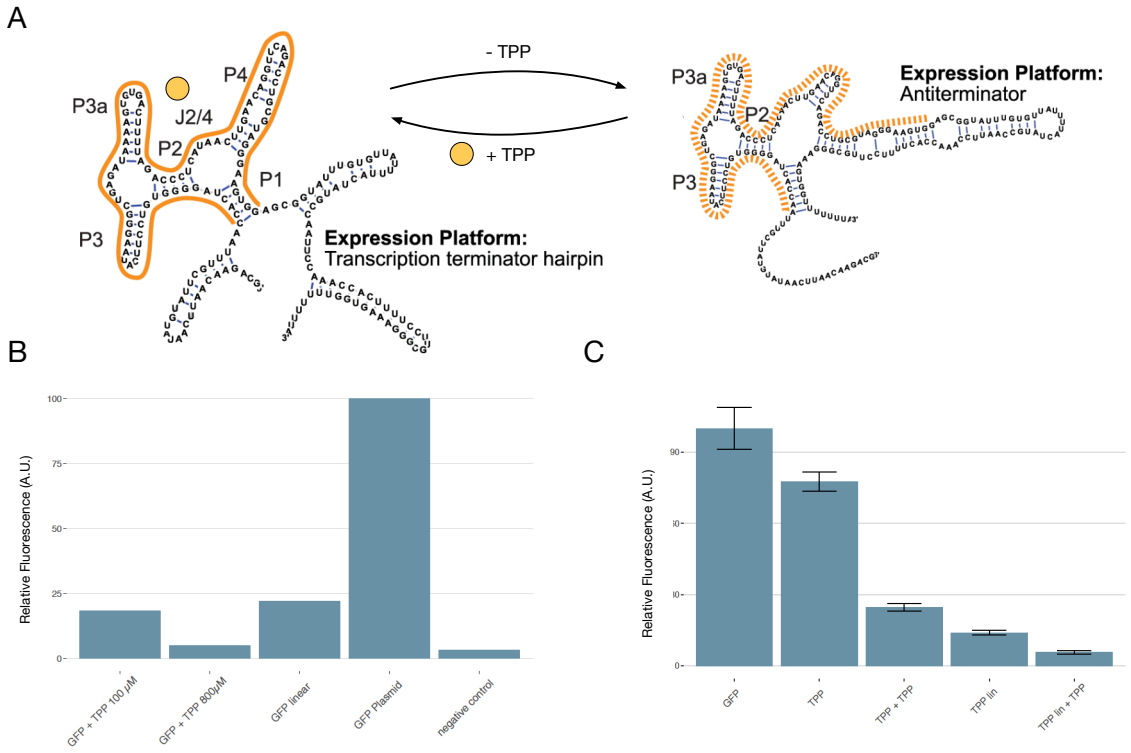


Fig. 4.10 TPP-riboswitch expression in batch-tube format. A) The two structures of the tenA thiamine pyrophosphate (TPP, orange circle) riboswitch are shown. Schema adaptated from Sudarsan et al. [117]. Aptamer domains are highlighted with orange straight or dashed lines. B) Endpoint fluorescence measurements to determine the toxicity level of TPP on TX-TL systems. C) Normalised endpoint fluorescence measurements of the Pr-TPP-GFP constructs, either linear or circular in presence or absence of TPP. Error bars represent the standard deviation over three replicates. Plasmid and linear DNA concentration were respectively set at 5 nM and 10 nM.

TPP riboswitch is similar with linear or circular template, even tough recent studies suggest that circular DNA might be a better template to mimic *in vivo* transcription rates from *E. coli* promoters [12].

Modulating the concentration of external inducer in a batch-format reaction can be cumbersome. The reaction tube consists indeed of a close environment where the tuning of the concentration parameters is difficult while proceeding to the measurements of TX-TL activity. The next step was to test if the TPP riboswitch circuit could be expressed in the nanoreactor. This would enable a direct control of the TPP concentration in the TX-TL environment, and consequently, of the gene expression. I used 5 nM of Pr-TPP-deGFP circular template and started the reaction when no TPP was present in the reactor (see Fig. 4.11 B). When the reaction reached a steady-state (after 200 minutes in this experiment), the input was switched to the same concentration of DNA template mixed with 400 µM

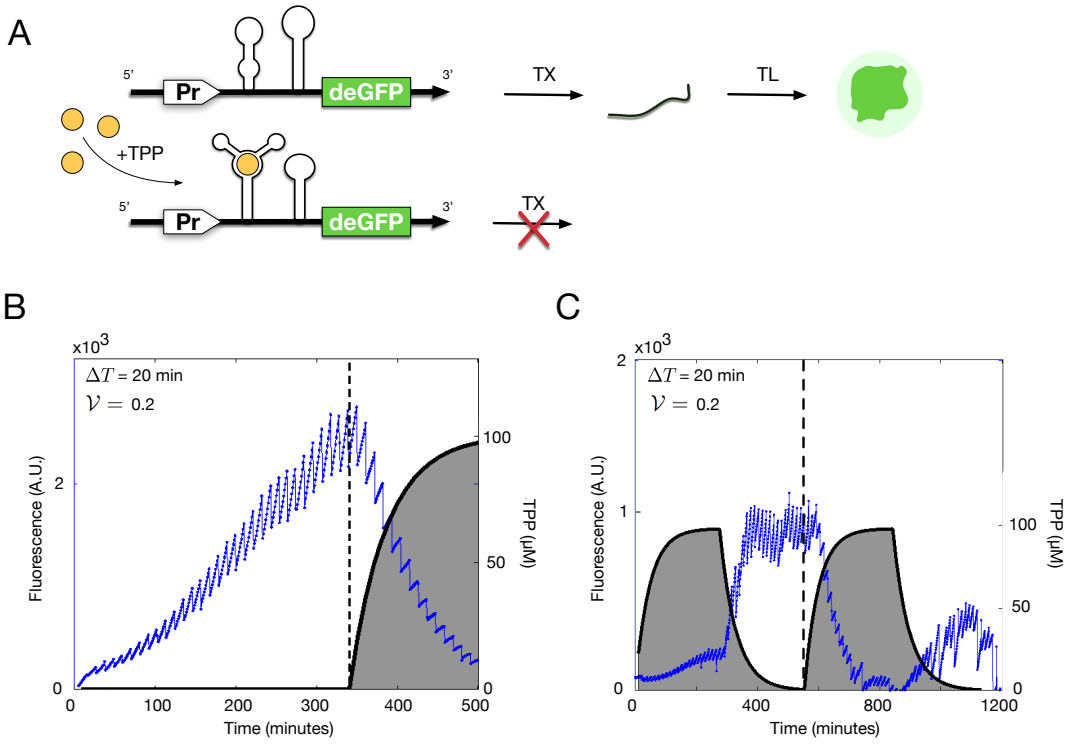


Fig. 4.11 TX-TL systems express riboswitch circuitry at steady state. A) Schematic of the TPP-riboswitch circuit. TPP presence promotes the terminator folding and prevents the transcription elongation. B) and C) Steady-state reactions of the corresponding reporter circuit with 5 nM of circular DNA template, with a dilution time of 35 minutes. Grey area indicates the concentration of TPP inside the reactor.

of TPP and loaded in the reactor with an identical number of packages (1 package of DNA/buffer/TPP for 3 packages of TX-TL). The concentration of TPP increased therefore with every additional dilution step up to almost 100 µM after 2 hours or 12 dilution steps. As expected, the deGFP production rate started to decrease a soon as after the second dilution step of TPP injection (about 20 µM of final concentration in the reactor). As the TPP concentration was increasing with each dilution step, the deGFP production rate was slowed down until reaching a stabilised level slightly above the background, corresponding to the OFF steady-state of the genetic circuit.

In order to test the limits of our microfluidic reactor, the same TPP-riboswitch circuit expressed in similar conditions was run in a longer experiment (20 hours) (Fig. 4.11 C). This time, the reaction started in presence of 100 µM of TPP, and the steady-state remained at a low level of fluorescence. After 300 minutes, the TPP input was switched off and its concentration in the reactor decreased according to an inverse exponential decay. The deGFP

production rate increased immediately after the second dilution step, reflecting the results obtained in Fig. 4.11 B. Once TPP was in negligible working concentration in the reactor ($<$ 1 μM), the TPP input was switched on again to assess the ability of the reactor to reverse the state of a genetic circuit multiple times. The deGFP rate decreased successfully, and came back to the initial OFF steady-state level. For the final step of the operating sequence, the TPP was switched off once again. We observed a return of deGFP expression, but this time the level was not as high as the first ON steady-state reached. This likely indicates a problem with one of the Quake valves, which can become stuck after a long time of operation. The cooling of our TX-TL system, kept at 7 °C throughout the experiment, may also need to be enhanced for such long-term measurement. Upgrading the set-up with a better performing cooling pump or peltier element could enhance the preservation of the TX-TL activities over time.

4.1.5 Future Use and Development

The nanoreactor was a time-consuming device to implement and characterise but it constitutes a relevant microfluidic device for the future of the TX-TL experiments presented in this thesis. The platform has been made accessible for different users by creating a GUI automatising the dilution operations on the chip and by standardising the numerous input channel connections. The ensemble of RNA-parts characterised in the 2nd and 3rd chapters will serve as a basis to build complex logical and dynamical circuits. The nanoreactor will enable the implementation of genetic circuits with multiple stable states triggered by a signaling molecule, such as the toggle switch or memory devices. Another promising application will be the engineering of RNA-based oscillators made of the previously characterised sgRNAs. The combined SpydCas9/Csy4 CFEs will serve as an ideal expression platform to prototype and test those multi-components circuits. So far, no oscillators based on RNA-parts have been reported in the literature, and I believe that a multi-disciplinary approach merging microfluidics, the RNA cell-free toolbox and computational modeling would be successful.

Furthermore, a new generation of microfluidic chips is currently being tested with eight independent parallel reactors. This will enable the test of a higher number of conditions simultaneously on the same chip. This is particularly relevant for complex genetic circuits made of numerous parts, since the contribution of each DNA part on the circuit performance could be assessed independently into each reactor with the same environment and set-up conditions. A prototype chip containing height parallel reactors is shown in Fig. 4.12. The design is similar as the single reactor chip, with the difference that pressure control lines were made thinner when crossing over non-targeted channels to not block the flow. The

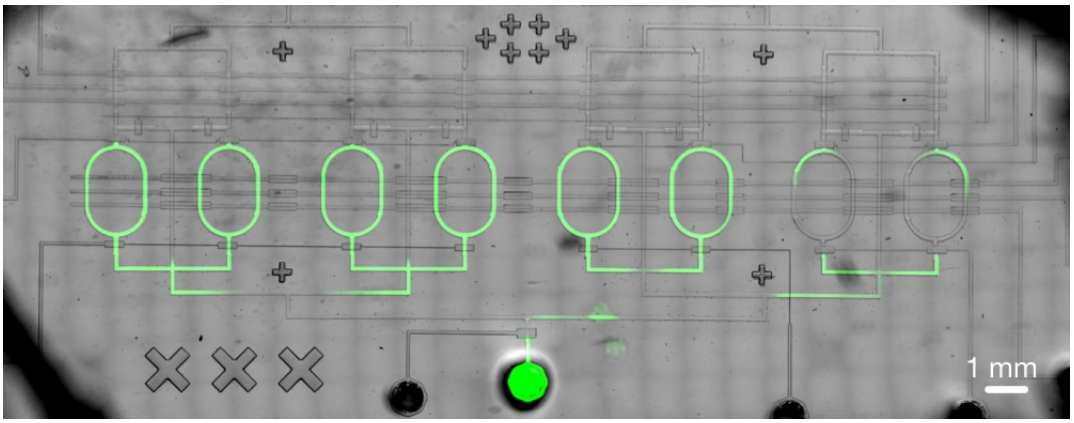


Fig. 4.12 Microscope image of the microfluidic chip containing 8 parallel reactors (Overlay of bright field and fluorescent channels). Most of the reactors are filled with FITC, while the two on the right have been diluted with several packages of water, demonstrating independent control of each pair of reactors. Scale bar in white indicates 1 mm.

number of inlets have been increased (from 4 to 6) to increase the possible number of control parameters.

4.2 Cell-free Expression in Small Volume Compartments

Expressing cell-free reactions at steady-state offers great opportunity for the design of genetic circuits as deterministic mathematical models of gene expression can be used to predict and optimise challenging implementations. Furthermore, those bulk cell-free reactions possess negligible stochastic effects, which is however an important element to consider when assessing the behavior of circuits expressed in single living cells [100]. A critical step is to bridge the gap between *in vivo* implementations of artificial genetic circuits (for example based on yeast or *E. coli* cells) and their *in vitro* counterparts. In this section, I will explore one novel microfluidic strategy for expressing cell-free systems in confined spaces at the femtoliter scale.

4.2.1 Femtoliter-Droplet On Demand

Context and Implementation

The work presented here took place in the context of a Compugene collaboration with the Hardt group (TU Darmstadt, department of Nano- and Mikro-fluidik). The microfluidic work

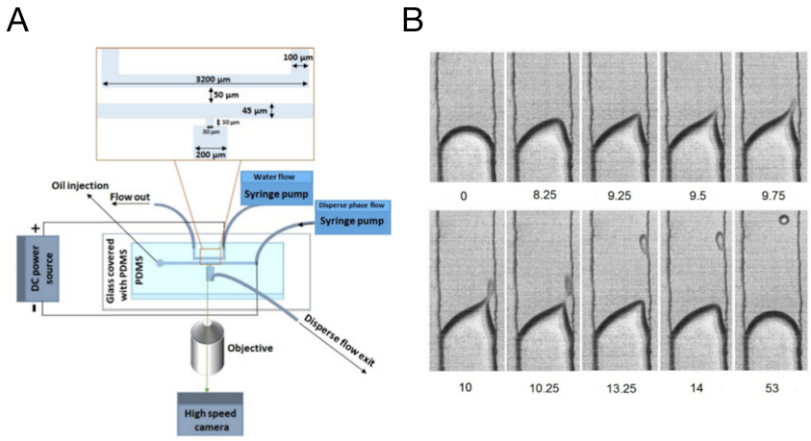


Fig. 4.13 A) Schema of the microfluidic set-up. B) Time-lapse images of the on-demand production of a water in oil droplet. Time in miliseconds. (400 V, pulse duration = 10 ms).

has been realised by Mostafa Shojaeian. This work has led to a publication in ACS analytical chemistry [107]. Some elements of the papers (i.e. figures) are reproduced here. Mostafa Shojaeian created a method allowing the generation on-demand of single femtoliter-size droplets in a reproducible manner. If those droplets can be used as reaction compartments, it would be an important potential tool in the field of synthetic biology for measuring the performance of circuit in cell-free in the light of their droplet-to-droplet variability.

The microfluidic platform is described in Fig. 4.13. Simply, the device consists of a constricted T-junction channel with a parallel and separated U-shaped channel. The U-channel is filled with water and used as an electrode when connected to the power supply. The continuous phase (oil) is loaded on the side of the T-channel and the disperse-phase liquid (such as TX-TL system) is introduced on the other inlet side. The disperse-phase liquid is then pressurised until partially displacing the continuous phase. A high voltage sequencer was used to create a voltage pulse of several hundred milliseconds and droplet formation was observed at the liquid/liquid interface close to the junction. Fabrication used standard lithography protocol and the methods and material are basically the same as from the nanoreactor device (see section 4.1.2).

The drop-on demand generator brings along two main advantages for TX-TL systems. First, it allows the creation of biological reaction compartments with a size comparable to that of yeast cells and only with one order of magnitude higher than the typical volume of an *E. coli* cell, two of the most frequently used cell types in synthetic biology. In other terms, a comparable number of biologically relevant molecules can be studied in similar sized compartments. That way, it becomes possible to mimic the stochastic behavior of biological cells. Second, the droplet generator platform is compatible with the crowding

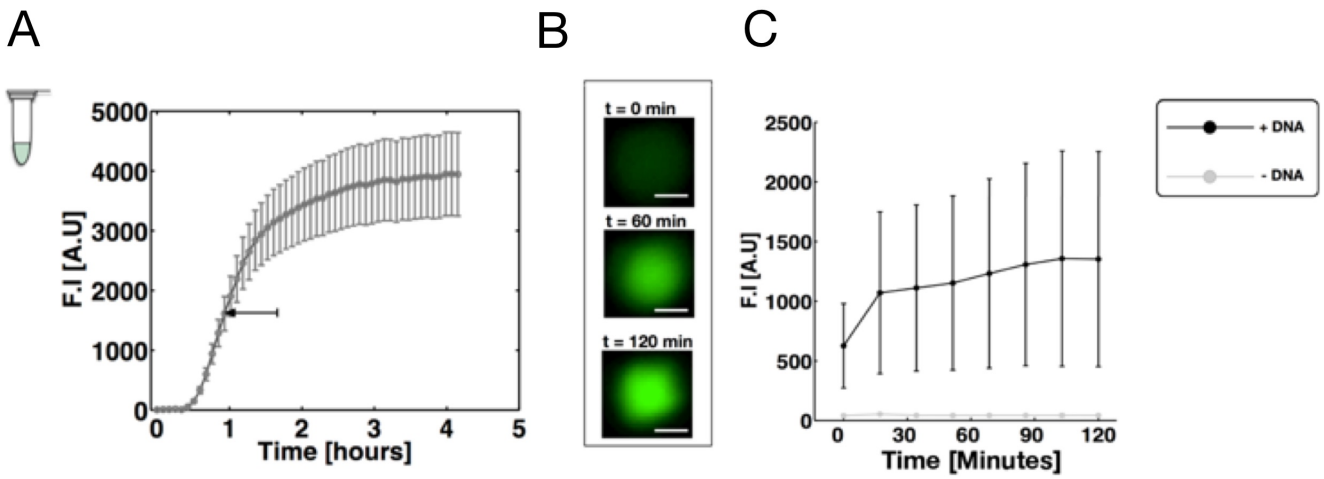


Fig. 4.14 TX-TL systems expressed in femtoliter w/o droplets. A) TX-TL expression of 5 nM of deGFP plasmid inside a batch-format tube. B) Fluorescent channel images of a droplet at different time points (0, 60, 120 minutes). Scale bar represents 2.5 µM. C) Averaged fluorescence time-course kinetics of n = 23 droplets containing 5 nM of deGFP template. Droplets without DNA template is represented by the grey line (n = 3). Errors bars are the standard deviations from the mean.

characteristic required for biological cells. The interior of living cells is densely packed with different molecular species, translated into a functional highly viscous environment. The platform presented here has demonstrated the compatibility with generating droplets having a viscosity more than 104-fold higher than that of water [107]. This will be highly useful for the characterisation of RNA-based circuits where little has been done yet to understand the localisation and bursting effect of mRNA on circuit prototyping [87].

Femtoliter-size droplets as biological compartments

In order to test if this platform is compatible with synthetic biology applications, I performed two of the most fundamental biochemical reactions inside the droplets: gene transcription and mRNA-translation. The CFE and reporter plasmid used were the same as in section 4.1.3 for the chemostat characterisation. The gene is first transcribed into mRNA and subsequently translated into protein. The synthesis of deGFP was analysed in a time-dependent manner, and the product of the reaction when performed at standard ("bulk")-conditions in a batch-format volume (10 µL) is quantified in Fig. 4.14. deGFP-fluorescence starts at roughly 30 min and follows a linear increase for 1 hour before to reach a plateau roughly at 3 h.

The TX-TL reaction mix was enclosed in the w/o droplets after 60 min of "bulk" reaction (arrow in Fig. 4.14 A). Droplets with diameters between 3 and 8 µM were tested, corre-

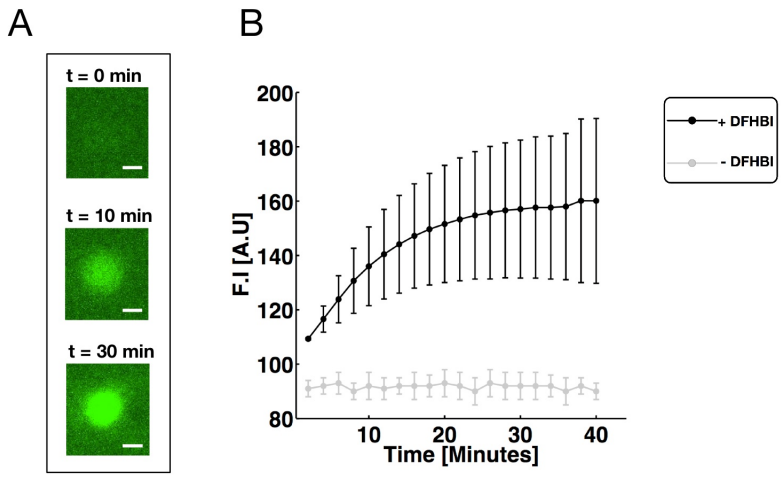


Fig. 4.15 *In vitro* transcription of the iSpinach aptamer inside femtoliter w/o droplets. A) Fluorescent channel images of one droplet at different time points. Scale bar represents 3 µM. B) Averaged fluorescence time-course kinetics of 9 droplets containing 20 nM of T7-iSpinach linear template (Black curve). Droplets without DNA template is represented by the grey line (n = 3). Errors bars are the standard deviations from the mean.

sponding to droplet volumes between 20 fL and 300 fL. Cropped images of the fluorescent channel show a 5.1 µm diameter droplet at three different time-points (Fig. 4.14 B). The time-course fluorescence signal of deGFP averaged over 23 droplets is shown in Fig. 4.14 C). While the droplets containing no DNA template do not show any detectable signal over the observed time frame, the fluorescence signal in the droplets increases over 2 h. An initial fast increase within the first 15 min is followed by a second slower increase thereafter. As expected and in accordance with the literature, the data show a high statistical variance, likely due to the stochastic distribution of the TX-TL-extract components, e.g. number of polymerases and ribosomes, in the small reaction volumes. In that sense, the droplets seem to mimic the stochastic behavior of gene expression in single cells. This platform would be therefore suitable for the studying of the extrinsic and intrinsic noise distribution among droplet populations.

As an additional experiment to expand the platform possibilities, I tested if low-molecular mass ligands can diffuse across the oil/water interface of the droplets in order to bind to a cognate receptor molecule and triggers a biochemical switch.

For this, *in vitro* transcription reaction was set up within the droplets to synthesise the "light-up" RNA aptamer iSpinach (see section 4.1.3 for more details). The fluorophore was solubilised in the oil phase and thereafter injected into the microfluidic device. About 2–3 minutes after droplet formation, the droplets started to fluoresce, reaching a plateau after

only 30 min. Those data are compatible with the previously published data recording both the deGFP and mRNA level in TX-TL systems (see chapter 3), where usually the maximum mRNA production rate is reached earlier than the maximum protein production rate. Cropped images of the fluorescent channel of a 5.9 μm diameter droplet at different time points are shown in Fig. 4.2.1 A. Fig. B displays the averaged time-course fluorescence intensity from a sample of nine droplets. As before, the strong electric field required for droplet formation did not adversely disrupt the bioactive structure of the biomolecules, in that case the polyanionic RNA molecules. Similar to the case of the TX-TL system, this reinforces the compatibility of the droplets produced using the method described above with biomolecular reactions. Demonstrating the proper folding ability of mRNA in those droplets is important for testing synthetic circuits regulated by RNA-based devices whose functionality relies on the folding of secondary and tertiary structures.

Chapter 5

Conclusion and Outlook

Characterisation of several logic AND-gates built on multi-level RNA activators was performed. To the best of my knowledge, this is the first study which combined the STAR and toehold regulators, two highly efficient RNA activators. TX-TL system was used together with a robotic liquid handler (I-Dot), which enabled to perform precise, fast and down-scaled parallel reactions for testing a large range of concentrations. From the resulting dose-response matrices, the saturation concentrations which maximise the dynamic range of the AND-gates were determined. A similar qualitative behaviour was observed for the gates as for the single parts; the transcriptional activator STAR needed threefold the amount of the translational activator trigger to reach saturation. This agreement between single parts and gates shows the potential of the STAR-toehold system for synthetic circuit design where modularity of parts is desirable. Establishing the I-Dot framework for TX-TL system was also the first step towards a pipeline automation for circuit-prototyping with the Compugene robotic platform. The next critical step will be to merge this framework with sequential optimal experimental design for testing and prototyping RNA-based circuitry. An iterative feed-forward loop could be implemented between the I-Dot and the fluorescence plate reader. This would enable the automatic screening for optimal concentrations of DNA parts or inducers in TX-TL systems. As an example, it would be possible to automatically determine the concentrations maximising the highest dynamic range for several different AND-gates, and combine them thereafter based on those concentrations.

Experimentally, a remarkable AND-gate based on the STAR 6 and toehold 3 parts was identified. Its production rate was as high as for the unregulated sfGFP, showing a great potential for metabolic engineering. I suspect that local co-folding of the RNA regulators hinders the translational process of the remaining gates. This hypothesis is supported by the

dramatic enhancement of the AND-gate with the Csy4 processing hairpin between the STAR 6 and toehold 2 components. In future work, it may be worthwhile to perform more advanced secondary structure analysis and combine it with the ODE population model to obtain better quantitative insights. However, Angenent-Mari et al. [4] have highlighted the weakness of state-of-the-art thermodynamics parameters and kinetic models for predicting the efficiency of toehold switches. Instead, they found that deep neural networks outperformed those features in engineering and understanding of RNA-based switches. It would be interesting to apply this novel computational methodology to the RNA-based gates in order to confirm this trend. Moreover, Yim et al. [135] have demonstrated the use of next-generation sequencing in TX-TL systems, which could be a solution for the generation of a high volume of data, required for neural networks analysis.

Finally, TX-TL based diagnostic platforms with toehold switches have been recently used for the rapid and low-cost detection of specific RNA genome sequences [90, 91]. As the STAR activators can be computationally designed [16], coupling the toehold sensors with a STAR module could increase the flexibility and sensitivity of the current bio-molecular diagnostic tools. Since the current ones are mainly based on the toehold switches, the users have typically a waiting time over a couple of hours (due to the translation and protein maturation times), which could be scaled down to the order of minute with a STAR-based sensor. The transcription process could be read out with an aptamer, like the tandem of MGAs used in this thesis. Recently, Alam et al. [3] have applied this approach to a real world problem by designing a cell-free biosensing platform that uses Spinach aptamers activated by ligand induction to detect harmful contaminants in aqueous samples.

Finally, extending the catalogue of RNA-based components to the TX-TL toolbox will enable the engineering of additional Boolean functions. In particular, the sgRNA repressors allow the construction of NAND gates, which are universal gates. In other terms, any other gate can be represented as a combination of NAND gates, which has important implications for the creation of advanced RNA-based circuits. A current limitation to this paradigm is, however, the limited number of orthogonal STAR components available. As the current library is mainly based on the same terminator hairpin, only a few STARs can be effectively combined without crosstalks. In theory, there is nothing preventing the engineering of novel STAR devices based on a set of artificial terminators [17], as this would enable a greater library diversity.

All gates were successfully transferred from *in vitro* to *in vivo*, demonstrating the versatility of TX-TL systems for the fast prototyping of RNA-based circuits. Additional studies to transfer the dose-response matrices from the TX-TL systems to an *in vivo* counterpart (for

instance with inducible promoters) may help engineer more balanced and efficient circuits in living cells. As both dSpyCas9 and Csy4 systems have been first established *in vivo*, there is no obvious limitations to their transfer in living cells. The cloning of advanced circuits composed of numerous sRNAs can, however, become increasingly challenging. The microfluidic chemostat will be a valuable instrument to measure, test, and rewire these complex circuits in an accessible way. It will be particularly useful for the dynamic characterisation of RNA-based circuits in TX-TL systems, since the circuit state can be switched under microfluidic control.

Chapter 6

Material and Methods

6.1 Methods

6.1.1 Cell-free Extract Preparation

Lysate preparation

The 3rd chapter of this thesis investigates parameter variations related to the CFE protocol. In this subsection, only the most efficient protocol for the production of cell-free system is presented.
E. coli BL21 Rosetta 2 were streaked overnight on agar plate containing chloramphenicol (and ampicilin, for the production of functionalised extracts). One colony was picked and inoculated overnight in 50 mL 2xYT supplemented with antibiotic(s) for growth at 37°C. After a minimum of 15 hours, 10 to 20 mL of the stationary culture was used to inoculate 200 mL to 1 L of 2xYT + P media (16 g/L tryptone, 10 g/L yeast extract, 5 g/L sodium chloride, 7 g/L potassium phosphate dibasic, 3 g/L potassium phosphate monobasic) in a 1 L to 5 L baffled flask. For most of the experiments not related to the study of optical density (OD_{600}), cells were grown at 37 °C and 200 rpm to 3.5 $\pm$ 0.1 OD_{600}. Centrifuge bottles were filled up to 300 mL and centrifuged for 10 minutes at 4000 g at 4 °C and supernatants was discarded. The pellets were next washed three times with 25 mL buffer S30A (50 mM Tris-base, 14 mM magnesium glutamate, 60 mM potassium glutamate, 2 mM DTT, brought to pH 7.7 with acetic acid). Each step of washing was separated with a centrifugication step at 4000 g at 4 °C for 10 minutes. A fourth centrifuging step at 3000 g at 4°C for 10 minutes enabled to remove the remaining traces of buffer.

The pellets were then resuspended in 1 mL of Buffer S30A per gram of dry pellet. 700 µL of the suspension was splitted in the necessary number of 2 mL eppendorf tubes. The pellet suspensions were then lysed with a sonicator (QSonica Q125 with a 3.175 mm diameter probe, 50% amplitude, 20 kHz, and 10 secondes ON/OFF pulses). Each sample was sonicated until reaching 500 J. For CFEs corresponding to the experiments detailed in the 3rd chapter, 1 mM of DTT was added to each crude lysate immediately after sonication. The cell lysate was then centrifuged for 10 minutes at 4 °C and 12000 g. The supernatant was removed and placed into an incubator set up at 37 °C and 200 rpm for 80 minutes. After this run-off reaction, the supernatant was taken and centrifuged again for 10 minutes at 4 °C and 12000 g. For cell-free extract preparation including a dialysis step, the extract was dialysed into a 10K MWCO Slide-a-Lyzer dyalisis cassette put in a 1 L buffer S30B (14 mM magnesium glutamate, 60 mM potassium glutamate, 1 mM DTT, bought to pH 8.2 with 2 M Tris-base) refrigerated at 4°C. A final centrifugation step for 10 minutes at 4 °C and 12000 g was performed and the supernatant was aliquoted as 40 µL samples into PCR tubes. The aliquots were snap-frozen into liquid nitrogen and stocked at -80 °C.

Energy buffer preparation

Energy buffer for the cell-free system was prepared according to table 6.8. The master-mix was aliquoted as 45 µL samples into PCR tubes. The aliquots were snap-frozen into liquid nitrogen and stocked at -80 °C. The ratio between cell-extract and energy buffer were kept to 33:42. The standard concentration of magnesium glutamate and potassium glutamate were kept respectively at 10 nM and 140 nM. Each newly produced cell-extract was titrated with magnesium glutamate in the range of 10-20 nM. Some experiments involved the titration of potassium glutamate between 80 nM and 200 nM.

Cell-free experiment

The final cell-free mixture is composed of: 10-20 mM magnesium glutamate; 130 mM potassium glutamate; 1.5 mM ATP; 1.1 mM each of GTP, UTP, and CTP; 0.034 mg/mL folinic acid; 0.171 mg/mL tRNA; 2 mM amino acids; 30 mM 3-PGA; 0.33 mM NAD; 0.27 mM CoA; 1.5 mM DTT; 1.5 mM spermidine; 57 mM HEPES, 0.75 mM cAMP, 34% CFE per volume, plasmid DNA and Milli-Q water. Each experiment was prepared in a final volume of 20 µl and 3 replicates of 5 µl were dispensed into a low volume 384 well microplate. For experiments involving MGA RNA read-out, the plates were chilled beforehand into the -20 °C freezer to slow the start of the kinetic reactions. Once all reactions were pipetted, a transparent

seal was applied to prevent evaporation of the reactions. Finally, the microplate was vortexed and centrifuged again before being ready for either kinetic or endpoint measurements. The plate was read with a ClarioSTAR plus or PHERAstar FSX set-up at 29 °C for 10 hours. For endpoint measurements, the plate was let overnight (16 hours) in the incubator at 29 °C. The GFP signal was measured every 2 minutes with the direct optic bottom module FI 485/520 (Ex/Em), adjusting gain to 120 with 20 flashes per well and z-height optimised according to each experiment. For recording multiple fluorescence channels, the ClarioSTAR plus was used with the following (Ex/Em) optic parameters: GFP (485/520), RFP (530/585), MGA (610/660).

Calibration curves for cell-free reactions

To quantify the production of mRNA and protein in TX-TL, calibration curves were obtained by titrating purified mRNA and proteins at various concentrations with the same plate reader parameter conditions. sfGFP calibration was obtained by titrating eGFP, and mRFP1/dtTomato from purified dtTomato (Takarabio). For the mRNA quantification, purified sfGFP-MGA4x mRNA was used, obtained from a T7 *in vitro* transcription. The mRNA was heated at 80 °C for 5 minutes and then cooled down at RT for 15 minutes before measurement.

6.1.2 Purification of Biological Molecules

RNA purification

For experiments involving pre-expressed mRNAs, the RNAs were transcribed from PCR-amplified templates, all containing three 5-terminal guanosyl residues to facilitate transcription in vitro using T7 RNA polymerase. For this, two oligonucleotides were designed with an overlap of 35 bp containing the T7 promoter sequence and amplified using Q5 High-Fidelity DNA polymerase (NEB) according to the supplier's instructions. After ethanol precipitation, the DNA template was used for *in vitro* transcription with T7 RNA polymerase (NEB) as reported by the manufacturer. The RNA was gel-purified and molarity was determined by spectrophotometric measurements using NanoDrop 1000 Spectrophotometer (Thermo Scientific).

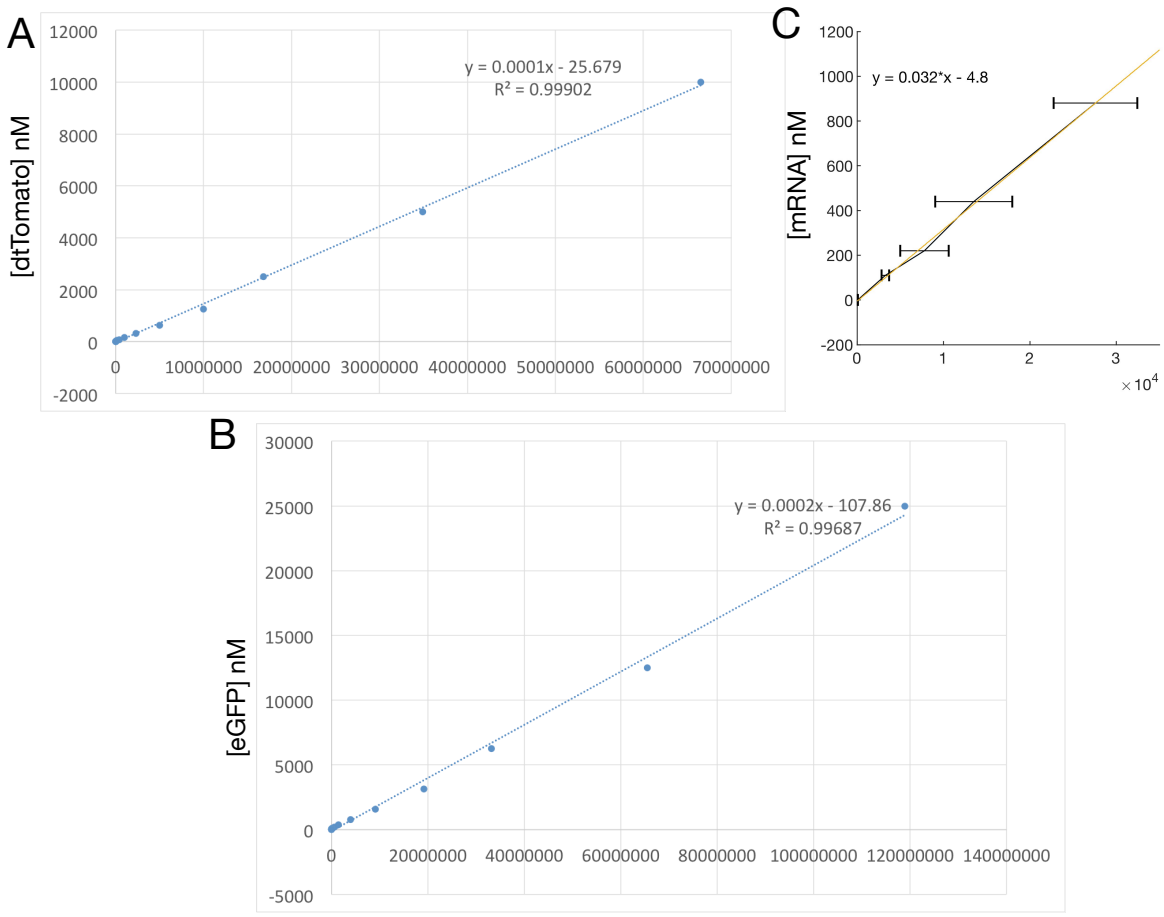


Fig. 6.1 Calibration curves for TX-TL experiments. A) Calibration curve for dtTomato. B) Calibration curve for sfGFP. C) Calibration curve for sfGFP-MGA4x mRNA.

In vitro transcription with T7 polymerase

The DNA template used for *in vitro* transcription was first purified with sodium acetate/ethanol precipitation. HiScribe T7 Quick High Yield RNA Synthesis Kit (NEB) was used to generate the RNA required for the MGA degradation assays and for the pre-expressed RNA activators experiments. The reaction containing the elements listed in table 6.1 were assembled at room temperature. The mix was then incubated for at least 2 hours at 37 °C. The sample was finally treated with DNaseI and purified with a RNA-clean and concentrator kit (ZymoResearch).

Protein purification

GamS: *E. coli* strain BL21 DE3 was transformed with a plasmid encoding for GamS with a polyhistidine tag under the control of an inducible promoter (see Table 6.17). One single

Table 6.1 T7 transcription assembly mix.

Nuclease free water	x µl
NTPs buffer mix	10 µl (10 mM each NTP final)
DNA template	x µl (500 ng to 2 µg)
T7 RNA Polymerase	2 µl
Total reaction volume	20 µl

Table 6.2 Buffer composition for GamS purification.

Buffer	Composition	Final concentration
GamS Lysis	Tris-HCl pH 8.0	50 mM
	EDTA	1 mM
	NaCl	500 mM
	Imidazole	5 mM
GamS Wash	Tris-HCl pH 8.0	50 mM
	EDTA	1 mM
	NaCl	500 mM
	Imidazole	25 mM
GamS Dialysis	Tris-HCl pH 7.5	50 mM
	EDTA	1 mM
	NaCl	100 mM
	Imidazole	250 mM
	DMSO	2%

colony was picked up and grown as a starter culture in 4ml of LB Medium with the addition of 4 µl (100 mg/ml) ampicillin over night. 1ml of the starter culture was transferred into a flask with 30 ml of LB Medium and grown to an OD_{600} of 0.5 at 37 °C. The culture was induced with 0.025 % [w/v] of arabinose and incubated at 30 °C for 3 h 30 min. The grown cultures were split equally into 50 ml falcon tubes and centrifuged at 4000 g for 20 min at 4 °C. The resulting supernatant was discarded and the pellets were flash-frozen in liquid nitrogen and stored at -80 °C. The pellets were weighed and 10x the weight of the pellet was added in lysis buffer (see Table 6.2). The suspended cells were sonicated at 20 % power for 1 minute, four times every 5 minutes. The lysate was cleared by centrifugation at 4000 g and 4 °C for 20 minutes. Ni-NTA agarose (Qiagen) was added into a 15 ml Falcon tube centrifuged at 4000 g for one minute and the supernatant was discarded. 15 ml of the cleared lysate were added to the Ni-NTA slurry agarose and incubated after vortexing. The mix was applied to a drip column and washed twice with 15 column volumes of native wash buffer. The protein was eluted in 2 column volumes of elution buffer in 2 ml fractions. A SDS-PAGE was used to analyse the cleared lysate, the wash fractions and the elution fractions.

Table 6.3 Buffer composition for SDS-PAGE electrophoresis.

Buffer	Composition	Final concentration
TAE 50x	EDTA disodium salt	50 mM
	Tris base	2 M
	Glacial / acetic acid	1 M
Laemmli buffer 2x	Tris-HCl pH 6.8	100 mM
	SDS	2%
	Glycerol	20%
	β-mercaptoethanol	25 mM

SDS-PAGE for protein electrophoresis

Samples were incubated at 95 °C for 5 minutes in 2x Laemmli buffer (sample and buffer 1:1, see Table 6.3). Samples were loaded onto the BioRadTM© mini-protean precast gel. The electrophoresis was run at 140 V for one hour.

6.1.3 Methods for Molecular Cloning

Plasmid assembly

The construction of the plasmid library is based on plasmids previously published: pJBL2801 was a gift from Julius Lucks (Addgene plasmid 71207) and was used for cloning the STARs, triggers, sgRNAs and UTRs variants; pJBL2807 was a gift from Julius Lucks (Addgene plasmid 71203) and was used for cloning the sense, toehold and the nine combined gates. Golden gate and Gibson assembly were used for cloning the inserts with length longer than 100 bps. Inserts were obtained by PCR, and backbones either by PCR or restriction enzymes. Q5 site directed mutagenesis kit from NEB was used for cloning the inserts inferiors at 100 bps. Some of the constructs, such as the sRNA operon, were synthesised from Genscript, and inserted into pUC57-mini vector. Oligos used for the various assemblies or mutagenesis were ordered from Sigma Aldrich. Sanger sequencing verified all constructed plasmids.

The plasmids or linear fragments used in the 4th chapter are based in plasmids used and published by Shin and Noireaux [106], in particular pBEST-p70a-deGFP-t500. pBEST-p70a-UTR-deGFP-6xHis-T500 was a gift from Lital Alfonta (Addgene plasmid 92224).

Plasmids were prepared using a NucleoBond Extra Midi Plus (Macherey-Nagel) and followed by isopropanol precipitation and eluted with 10 mM Tris-HCl, pH=8.5, and quantified using a NanoDrop 1000. Those having an insufficient purity were subject to a Phenol-

Table 6.4 PCR mixture.

Reagent	Final concentration
Forward Primer	2 µM
Reverse Primer	2 µM
dNTPs	200 µM
DNA template	5 to 10 ng
Q5/phusion Polymerase	1 U
Q5/phusion Buffer 5x	1 x
Nuclease free water	to 50 µl

Table 6.5 Thermocyclar program for Q5/phusion PCR. Steps highlighted in cyan are cycled for 30 to 35 times.

Step	Duration(s)	Temperature (°C)
Initial denaturation	120	98
Denaturation	15	98
Annealing	15 to 30	50-70
Elongation	30 s/kb	72
Final elongation	180	72
Cool-down	∞	4

chloroform extraction followed by a novel step of isopropanol precipitation. The purity quality of the plasmid was critical to achieve reproducible results in the CFEs.

Plamid names, resistances, and part composition are listed in Table 6.17.

PCR of DNA fragments

Two different high-fidelity polymerases, Q5 and Phusion (NEB) were used to amplify insert or templates sequences with two overlapping primers from a vector to create a double stranded DNA that can be used for cloning. For colony PCR, Taq polymerase (NEB) was used. Temperature for primer hybridization and extension elongation times were adjusted to target length and primer-vector hybridization. temperatures. Mixture details and protocol are given in table 6.4 and 6.5.

Preparation of *E.coli* competent cells

All transformations with *E.coli* strain DH5α are performed following the protocol of zymo research Mix Go kit. Cells were grown to mid-log phase (OD_{600} = 0.4 - 0.6) and placed

on ice for 10 min . The cells were harvested by centrifugation at 2500 g for 10 min at 4 °C. Supernatant was removed and cells were washed twice with 5 ml ice-cold wash buffer. The final step is the resuspension of the pellet in 5 ml cold competent buffer. The cells are aliquoted and stored in a -80 °C refrigerator. The transformation includes thawing on ice, addition of 1-5 µL circular DNA, an hour of incubation at 37 °C under constant shacking and last the spreading onto pre-heated agar plates with the corresponding resistant marker. In this thesis, Ampicillin, Chloramphenicol and Kanamycin were used as selective pressure.

Flow cytometry analysis

The plasmid containing the regulatory element with the reporter and the plasmid containing the activating RNA were transferred into *E. coli* TOP10 cells by the calcium chloride method. The cells were grown on LB agar plates containing the antibiotics ampicillin and chloramphenicol over night at 37 °C. A single colony was used to inoculate 4 ml of LB containing the needed antibiotics for selection. The cells were grown over night (37 °C, 150 rpm). For sfGFP analysis the cells were diluted 1:10 with PBS into a 96-well plate (flat bottom, Greiner). Fluorescence was analysed by using the CytoFlex S cytometer from Beckmann. For sfGFP excitation a 488 nm laser was used and emission was detected by a 510/20 nm filter. For each sample 25 000 events were measured. Each experiment was repeated three times.

Nucleic acid quantification

To measure the DNA or RNA concentration, a Nanodrop ND-1000 was used by loading 1.5 µM of nucleic acid on the pedestal. The absorbance was read at 260 nM and the concentration was determined using the Beer-Lambert's law.

Gel electrophoresis

For DNA extraction and analysis, agarose gel was used. After casting the gels with agarose concentrations ranging from 1% to 3%, samples were mixed with DNA loading dye (5:1) and run alongside a DNA ladder of adequate size. Gels were run in 1x TAE buffer containing 1 µM of SYBR safe stain, at 6 V/cm for 20-60 min. DNA bands were visualised with 254 nm light for analytical purposes or extracted with 360 nm light.

Fluorescence Activated Cell Sorting for selection of the S6T3-sfGFP mutants

E. coli Top10 were co-transformed with the mutated library of S6T3-Nx10-sfGFP and with the corresponding activators plasmid (STAR 6 and trigger 3). Several plates were streaked together and grow in 50 mL culture of LB overnight. 1 mL is taken out, centrifuged 5 min at 3000 rpm at 4 °C. Cells are washed with 1 mL 1x PBS and centrifuged again in the same conditions. Cells are then diluted (1:10) and resuspended in 1 mL PBS and filtrated through a 40 µM cell strainer. After sorting, cells were plated to with 100 colonies per plate. Flowjo was used for gating the data and plotting the fluorescence level per cell count.

6.1.4 Microfluidic Chips

Master mold fabrication

All layouts were designed with AutoCAD (AutoDesk). Corresponding foil plots were ordered for each layer of the final design. For structures below 30 µM resolution, chrome masks were ordered. The steps of the fabrication and exposure of multiple layers are described in Table 6.7. At the end of the exposure, the wafer is coated with a highly hydrophobic chemical. The wafer is placed inside a vacuum chamber, together with 100 µL of 1H,1H,2H,2H-Perfluorodecyltrichlorosilane sealed and vacuum is applied. Through evaporation, a thin layer is applied to the silane oxid wafer. This coating increases the robustness of the wafer and can therefore be re-used for casting PDMS chips multiple times. The hydrophobic coating is repeated every 20 batches of chips fabrication.

PDMS chips replica

A two compound polymer, polydimethylsiloxane (PDMS) is mixed in weighting trays and placed into a desiccating chamber for degassing. For the control layer, 40 g liquid PDMS and 4 g of the cross-linking agent are mixed. The flow layer has a ratio of 20:1. After 2 h of degassing, the mixed PDMS is poured on the master mold, either in a petri-dish for the control layer or directly done in the spin coater for the flow layer. The thickness of the flow layer is adjusted by rotation speed and duration. The typical program used for the flow layer is described in Table 6.6.

Both wafers were degassed again for 1 h. A preheated oven was then used to cure the molded wafers at 75 °C.The flow layer was taken out after 15 min to have partially cured PDMS, offering better bonding property for the later baking step. The control layer was

Table 6.6 Spin-coater program for the flow layer. The acceleration step high-lighted in cyan is the critical step where the thickness of the PDMS layer can be adjusted between 5 and 20 µm.

Step	Duration (s)	Speed (rpm)
Initialisation	30	400
Acceleration	60	1600
Deceleration	30	400-0

removed from the oven after 30 min and cooled down at room temperature. Next, the cured PDMS of the control layer were cut into rectangular chip size with the help of the border marks. The control layer chips were then aligned onto the flow layer according to the alignment marks (a big and a small cross that fall together). The process of alignment is critical for the proper working of the final microfluidic chips and requires to be undertaken under a stereo microscope. The layers' assembly was baked overnight in the oven at 75 °C in order to seal the bond between them. Finally, the chips were peeled off the flow wafer and a 0.5 mm punch marker were used to make the holes for every flow and control channel connection. The resulting PDMS chips were bonded to cover slips by activating the surfaces with a plasma oven treatment of 10 secs.

6.1.5 Mathematical Modeling

This section describes the Bayesian inference framework developed by Maleen Hanst for the modelling of the RNA-based gates described in the 2nd chapter.

Parallel tempering:

In Bayesian parameter inference, we are concerned with finding the posterior distribution of the parameters θ included in our model, given the experimental data Y;

$$p(\theta \,|\, Y) \propto p(Y \,|\, \theta)\; p(\theta).$$

To identify this distribution, MCMC parallel tempering [30] was used. This method improves the exploration of the whole parameter space with faster convergence. In the following, the main formulas of the algorithm used in this work are stated.

Table 6.7 Soft-lithography protocols for the silicon wafers.

Layer 1 : Herringbone / Array structures	
Step	10 µm with SU8-2005
1.	Wash wafer with acetone, then isopropanol
2.	Coating with TI Prime "Haftvermittler"
3.	Spin-coating : 500 rpm 10 s at 500 rpm; 30 s at 1500 rpm
4.	Softbake: 5 min at 50 °C
5.	Exposure: 120 mJ/cm^2 at 365 nm, hard contact mode
6.	Resting for 10 min
7.	Post exposure bake (PEB) 3 min to 50 °C, 3 min to 65 °C, 3 min to 95 °C
8.	Developing for 40 s
9.	Rinse with isopropanol, then water
10.	Hard bake: 5 min at 130 °C
Layer 2 : Flow/control channels	
Step	20 µm with SU8-25
1.	Wash wafer with acetone, then isopropanol
2.	Coating with TI Prime "Haftvermittler"
3.	Spin-coating : 500 rpm 10 s at 500 rpm; 30 s at 4000 rpm
4.	Softbake: 5 min at 50 °C, 5 min to 65 °C, 5 min to 95 °C
5.	Exposure: 300 mJ/cm^2 at 365 nm, hard contact mode
6.	Resting for 10 min
7.	PEB 6 min at 95 °C
8.	Developing for 3:30 min
9.	Rinse with isopropanol, then water
10.	Hard bake: 2 min at 130 °C and slow cool down
Layer 3 : Round channels	
Step	10 µm with AZ4562
1.	Wash wafer with acetone, then isopropanol
2.	Coating with TI Prime "Haftvermittler"
3.	Spin-coating : 500 rpm 10 s at 500 rpm; 30 s at 1700 rpm
4.	Softbake: 10 min at 100 °C
5.	Exposure: 700 mJ/cm^2 at 405 nm, hard contact mode
6.	Resting for 5 min
7.	Developing for 3:30 min
8.	Rinse with isopropanol, then water
9.	Hard bake: 2 min at 130 °C and slow cool down

Our log-likelihood is determined by the common assumption of independent normally distributed measurement deviations and we denote it similar to Moore et al. [79] by

$$\ln p(Y \mid \theta) = \\ -\sum_{\substack{c\in\mathcal{C},\\ t\in\mathcal{T}}} \sum_{i=1,2,3} \left\{ \frac{1}{2}\ln(2\pi(\sigma_{c,t}+\sigma_b)^2) + \frac{(Y_{c,t,i}-X_{c,t}(\theta))^2}{2(\sigma_{c,t}+\sigma_b)^2} \right\},$$

where $\mathcal{C}$ contains the experimental conditions (in our case the different input concentrations of the activator plasmids), and $\mathcal{T}$ contains the different time points of the measurements (in our case, measurements were taken every two minutes). The second sum stands for the three replicates measured for each concentration. The simulated model data of the reporter fluorescence intensity are indicated by $X_{c,t}(\theta)$ due to their dependence on the activator input concentration, the evaluation time, and the parameter θ. The experimental data as well are indicated by the activator concentration, the time point of the measurement and the enumeration of the replicate. From the three replicates for each concentration and each time point, the empirical standard deviation $\sigma_{c,t}$ was computed. Further, an additional "base noise" σ_b was set to capture the experimental irregularities up to 25 minutes.

As a prior distribution, we choose the uniform distribution on the interval of $\pm$ five orders of magnitude around parameter literature values (e.g. taken from Hu et al. [49]).

Samples are drawn from the tempered posteriors, i.e. from the posterior distributions of different temperatures given as follows

$$p_i(\theta \mid Y) \propto p(Y \mid \theta)^{\beta_i}\pi(\theta),$$

where β_i is the ith inverse temperature (thermodynamic parameter). The inverse temperatures were chosen according to the following – *geometric* – temperature schedule: Denoting the predefined number of different temperatures by N_t, e.g. 50, one further chooses a minimal inverse temperature, here 10^{-6}, and identifies $q \in (0,1)$ such that $q^{N_t-1} = 10^{-6}$. Therewith, the set of inverse temperatures is given by

$$\{\beta_i := q^{i-1} \mid i = 1,\dots N_t\},$$

where $\beta_1 = 1$ corresponds to the "original temperature", e.g. to the true posterior distribution.

The acceptance probability for the ith chain to switch to a proposed sample θ' is determined by the Metropolis criterion and given by

$$a(\theta' \,|\, \theta) = \min\left(1, \frac{p(Y \,|\, \theta')^{\beta_i}\ \pi(\theta')}{p(Y \,|\, \theta)^{\beta_i}\ \pi(\theta)}\right).$$

The proposal distribution for a new parameter set is given by the normal distribution with mean equal to the current parameter values and predefined proper variances that are scaled with some predefined constant greater than 1 which scales exponentially with the enumeration of the temperatures, i.e., the highest temperature possesses the largest variance.

Finally, in each sampling step, there is a predefined number of possible swaps of current samples belonging to the differently tempered chains: picking randomly one of the chains, its corresponding current sample can possibly be changed with the current sample from the chain with next higher temperature according to the following swap acceptance probability [50, 118]

$$a_s(\theta_i, \theta_{i+1}) = \min\left(1, \left[\frac{p(Y \,|\, \theta_{i+1})}{p(Y \,|\, \theta_i)}\right]^{(\beta_i - \beta_{i+1})}\right).$$

Having characterised the parameter posterior distribution, the computation of the posterior predictive distribution is given for a new data point y' at time t' regarding to an input concentration c' via

$$\begin{aligned} p(y' \,|\, Y) &= \int p(y' \,|\, \theta, Y)\, p(\theta \,|\, Y)\, \mathrm{d}\theta \\ &= \int p(y' \,|\, x) \left(\int p(x \,|\, \theta)\, p(\theta \,|\, Y)\, \mathrm{d}\theta\right) \mathrm{d}x, \end{aligned}$$

where in the case of a deterministic model, we have $p(x \,|\, \theta) = \delta(x - X_{c',t'}(\theta))$.

Table 6.8 List of chemicals and reagents for cell-extract preparation.

Chemicals/Reagents for cell-extract	Manufacturer
2x YT media	Sigma Aldrich, USA
3-PGA	Sigma Aldrich, USA
HEPES	Carl Roth, Karlsruhe
tRNA	Roche Applied Sciences, Penzberg
CoA	Sigma Aldrich, USA
cAMP	Sigma Aldrich, USA
Folinic Acid	Sigma Aldrich, USA
Magnesium L-glutamate ($C_{10}H_{16}MgN_2O_8$)	Sigma Aldrich, USA
NAD	Carl Roth, Karlsruhe
Potassium L-glutamate ($C_5H_8KNO_4$)	Sigma Aldrich, USA
PEG-8000	Carl Roth, Karlsruhe
RTS Amino Acid Sampler	Biotech Rabbit, USA
Spermidine	Roth, Karlsruhe

Table 6.9 List of chemicals and reagents.

Chemicals/Reagents for molecular biology	Manufacturer
Agar	Carl Roth, Karlsruhe
Agarose	Carl Roth, Karlsruhe
Ampicilin	Carl Roth, Karlsruhe
Calcium chloride	Carl Roth, Karlsruhe
Deoxynucleoside triphosphate (dNTPs)	Thermofischer Scientific, USA
Dimethyl sulfoxide	Carl Roth, Karlsruhe
Dulbecco's Phosphate buffered Saline (PBS)	Life Technologies, USA
Dithiothreitole (DTT)	Roth, Karlsruhe
Ethanol, denatured	VWR, Darmstadt
Ethanol, absolute	VWR, Darmstadt
FlowClean Cleaning Agent	Beckman Coulter
Glycerol	Carl Roth, Karlsruhe
Hydrochloric acid (HCl)	Carl Roth, Karlsruhe
Isopropanol, p.a.	VWR, Darmstadt
Kanamycin	Carl Roth, Karlsruhe
Acrylamide (Rotiphorese Gel 40, 19:1)	Sigma Aldrich, USA
Potassium chloride (KCl)	Carl Roth, Karlsruhe
Sodium chloride (NaCl)	Carl Roth, Karlsruhe
Sodium dodecyl sulfate (SDS), pellets	Roth, Karlsruhe
SYBR Green	Thermofischer Scientific, USA
TEMED	Roth, Karlsruhe
Tris	Carl Roth, Karlsruhe
Tryptone	Carl Roth, Karlsruhe
Urea	Carl Roth, Karlsruhe
Yeast extract	Carl Roth, Karlsruhe
Chemicals/Reagents for microfluidics	**Manufacturer**
PDMS	Corning, USA
SU8 2005	Micro Resist Technologies, Berlin
SU8 developer	Micro Resist Technologies, Berlin
SU8 2025	Micro Resist Technologies, Berlin
AZ4562	Micro Resist Technologies, Berlin
Methyltrichlorosilane	Sigma Aldrich, USA
SU8 2025	Corning, USA
Cover glas 50 x 24 mm	Carl Roth, Karlsruhe
Metal plugs 23G	Darwin Microfluidics, Paris
Tygon Tubing 0.51 mm ID	Darwin Microfluidics, Paris

Table 6.10 List of utilised softwares.

Software	Purpose
Matlab 2018a	Data analysis, model generation
ImageJ 1.52n	Image analysis, data visualization
NIS-Elements 4,00	Microscope control and imaging software
OmmniGraffle 6.6.2	Creation of graphics
Nupack 3.2	RNA secondary structure computation
Python 3.1	Scripting routines
FlowJo v10	Flow cytometry data analysis
AutoCAD 2015	Microfluidic design

Table 6.11 List of utilised instruments.

Instrument	Manufacturer
Centrifuges	CellStar,
CytoFlex S	Beckman-Coulter, Krefeld
Heating block	30
Incubator (1)	Heidolph, Schwabach
Incubator (2)	Heraeus Christ, Osterode
Infinite© M200 plate reader	Tecan Trading AG, Switzerland
CLARIOstar plus	BMG Labtech, Offenburg
PHERAstar	BMG, Labtech, Offenburg
Milli-Q water purification system	EMD Millipore, France
NanoDrop© ND-1000	Peqlab, Erlangen
SimplyAmp© PCR	LifeTechnologies, Carlsbad
Inverted microscope Eclipse Ti	Nikon, Tokyo
Gel iX20 Imager	INTAS, Göttingen
Plasma oven	Diener electronics, Ebhausen
Spin Coater	Laurell Technologies, Hong Kong
Solenoid Valves	The Lee Company, Westbrook
I-Dot One (Nanoliter dispenser)	Dispendix
Cell sorter sh800	Sony, Tokyo
Balance	Acculab, USA

Table 6.12 List of utilised kits.

Kit	Manufacturer
QIAprep© Spin Miniprep Kit	QIAGEN, Hilden
NucleoBond© Xtra Spin Midiprep Kit	Macherey-Nagel, Düren
NucleoSpin© Gel-PCR Clean-up© Xtra Spin Midiprep Kit	Macherey-Nagel, Düren
RNA Clean Concentrator-25 Kit	Zymo Research, USA

6.2 Material

6.2.1 List of Chemicals Used

Table 6.13 List of utilised enzymes and ladders.

Enzymes and purified proteins	Manufacturer
RNase inhibitor	MOLOX, Hilden
Q5 DNA-polymerase	New England Biolabs, USA
Phusion DNA-polymerase	New England Biolabs, USA
Kld enzyme mix	New England Biolabs, USA
T4 DNA ligase	New England Biolabs, USA
SpydCas9	Abm Good, USA
Taq DNA-ligase	New England Biolabs, USA
T5 exonuclease	New England Biolabs, USA
GamS	Arbor Biosciences, USA
myTXTL-96	Arbor Biosciences, USA
Dyes and purified fluorophores	**Manufacturer**
Enhanced Green Fluorescent Protein	BioCat GmbH, Heildeberg
dTomato	BioCat GmbH, Heildeberg
DFHBI-1T	Lucerna Technologies, USA
Malachite Green solution	Sigma Aldrich, USA
TO3-biotin	University of Strasbourg
TMR-DN	University of Heildeberg
Restriction endunucleases	**Manufacturer**
NheI-HF [20 U/µl]	New England Biolabs, USA
EcoRI-HF [20 U/µl]	New England Biolabs, USA
SpeI-HF [20 U/µl]	New England Biolabs, USA
XbaI-HF [20 U/µl]	New England Biolabs, USA
Ladders	**Manufacturer**
Low-range GeneRuler	ThermoFisher Scientific, USA
1kb GeneRuler	Thermofisher Scientific, USA
RiboRuler Low Range (RNA)	ThermoFisher Scientific, USA

Table 6.14 Buffer composition for diverse routine protocols.

Buffer	Composition	Final concentration
2x YT + P	2x YT	31 g/L
	Potassium phosphate dibasic	7 g/L
	Potassium phosphate monobasic	3 g/L
LB medium	Tryptone	1% (w/v)
	Yeast extract	0.5% (w/v)
	NaCl	1% (w/v)
	Agar (for plates)	2% (w/v)
Ampicilin Stock Solution	Ampicillin in 70% (v/v) EtOH	100 mg/ml
Chloramphenicol Stock Solution	Chloramphenicol in filtered water	25 mg/mL
SOC medium	KCl	2.5 mM
	Yeast extract	0.5% (w/v)
	NaCl	1% (w/v)
	$MgCl_2$	10 mM
	$MgCSO_4$	10 mM
	Glucose	20 mM
	Tryptone	0.2% (w/v)
ISO buffer 5x	Tris-HCl pH 7.5	0.5 M
	$MgCl_2$	50 mM
	dNTP	1 mM
	DTT	50 mM
	NAD	5 mM

Table 6.16 List of utilised consumables.

Consumable	Manufacturer
Dispenser tips	Starlab, Hamburg
Nitrile gloves	Starlab, Hamburg
Petri dishes	Starlab, Hamburg
Pipette tips	Starlab, Hamburg
Reaction tubes (1, 1.5, 2 ml)	Greiner, Nürtingen
384 well plates, low volume, black	Greiner, Nürtingen
384 well plates, white, clear bottom	Greiner, Nürtingen
1536 well plates, clear bottom	Greiner, Nürtingen
Cellstar© cell culture plates, 12-well/ 24-well	Greiner Bio-One, Austria
Breathseal, gas permeable, sterile	Greiner Bio-One, Austria
Multiwell plate sealers	Greiner Bio-One, Austria
PCR tubes	Greiner, Nürtingen

Table 6.15 List of utilised *E. coli* strains with their corresponding genotype.

Bacterial Strains	Genotype
E. coli Top10	F^- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80*lacZ*ΔM15 Δ*lacX*74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G
E. coli Dh5α	F^- *ompT gal dcm lon hsdSB*(rB^- mB^-) Δ(DE3 [*lacI lac*UV5-T7p07 *ind1 sam7 nin5*]) [*mal*B^+]K-12(ΔS)
E. coli BL21 DE3	F^- *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20* γ*80dlacZ*ΔM15 Δ(*lac*ZYA-*argF*)U169, *hsd*R17(rK^- mK^+), δ-
E. coli BL21 Rosetta 2	F^- *ompT hsdSB*(rB^- mB^-) *gal dcm* (DE3) pRARE2 (Cam^R)

6.3 Plasmid Constructs

Table 6.17 Plasmids used for expression of synthetic circuits.

Plasmid Name	Plasmid architecture	Source
pFXL01	J23119 – RBS – sfGFP – T500 – amp – ColE1 origin	this thesis
pNoireaux01	P70a – UTR1– RBS – deGFP – T500 – amp – ColE1 origin	Noireaux
pNoireaux02	P70a – UTR1– RBS – S28 – T500 – amp – ColE1 origin	Noireaux
pFXL02	J23119 – RBS – sfGFP-MGA4x – T500 – amp – ColE1 origin	this thesis
pFXL03	J23119 – RBS – sfGFP-Mango – T500 – amp – ColE1 origin	this thesis
pNoireaux03	P70a – UTR1– RBS – dTomato – T500 – amp – ColE1 origin	Noireaux
pFXL04	P70a – UTR1– RBS – dTomato-iSpinach – T500 – amp – ColE1 origin	This thesis
pFXL05	P70a – UTR1– RBS – TPP-OFF-deGFP – T500 – amp – ColE1 origin	This thesis
pFXL06	P70a – UTR1– RBS – Theo-ON-deGFP – T500 – amp – ColE1 origin	This thesis
pFXL07	P70a – UTR1– RBS – Theo-ON-deGFP – T500 – amp – ColE1 origin	This thesis
pJBL2801	J23119 – Target AD1.S5 – RBS – sfGFP – TrrnB – CmR – p15A origin	Chappell et al. [13]
pJBL4970	J23119 – Target5 – RBS – sfGFP – TrrnB – CmR – p15A origin	Chappell et al. [16]
pJBL5816	J23119 – Target6 – RBS – sfGFP – TrrnB – CmR – p15A origin	Chappell et al. [16]
pJBL5808	J23119 – Target8 – RBS – sfGFP – TrrnB – CmR – p15A origin	Chappell et al. [16]
pJBL2807	J23119 – STAR AD1.A5 – T500 – AmpR – ColE1 origin	Chappell et al. [13]
pJBL4971	J23119 – STAR5 – T500 – AmpR – ColE1 origin	Chappell et al. [16]
pJBL5817	J23119 – STAR6 – T500 – AmpR – ColE1 origin	Chappell et al. [16]
pJBL5809	J23119 – STAR8 – T500 – AmpR – ColE1 origin	Chappell et al. [16]
pJBL-wt	J23119 — RBS – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-Toehold1	J23119 – Toehold1 – RBS – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-Toehold2	J23119 – Toehold2 – RBS – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-Toehold3	J23119 – Toehold3 – RBS – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-trigger1	J23119 – trigger1 – T500 – AmpR – ColE1 origin	this thesis
pJBL-trigger2	J23119 – trigger2 – T500 – AmpR – ColE1 origin	this thesis
pJBL-trigger3	J23119 – trigger3 – T500 – AmpR – ColE1 origin	this thesis
pJBL-Target AD1.S5-Toehold2	J23119 – Target AD1.S5 – Toehold2 – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-Target5-Toehold1	J23119 – Target5 – Toehold1 – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-Target5-Toehold2	J23119 – Target5 – Toehold2 – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-Target5-Toehold3	J23119 – Target5 – Toehold3 – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-Target6-Toehold1	J23119 – Target6 – Toehold1 – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-Target6-Toehold2	J23119 – Target6 – Toehold2 – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-Target6-Toehold3	J23119 – Target6 – Toehold3 – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-Target8-Toehold1	J23119 – Target8 – Toehold1 – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-Target8-Toehold2	J23119 – Target8 – Toehold2 – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-Target8-Toehold3	J23119 – Target8 – Toehold3 – sfGFP – TrrnB – CmR – p15A origin	this thesis
pJBL-STAR AD1.A5-trigger2	t500 – trigger2 – J23119 – spacer – J23119 – STAR AD1.A5 – T500 – AmpR – ColE1 origin	this thesis
pJBL-STAR5-trigger1	t500 – trigger1 – J23119 – spacer – J23119 – STAR5 – T500 – AmpR – ColE1 origin	this thesis
pJBL-STAR5-trigger2	t500 – trigger2 – J23119 – spacer – J23119 – STAR5 – T500 – AmpR – ColE1 origin	this thesis
pJBL-STAR5-trigger3	t500 – trigger3 – J23119 – spacer – J23119 – STAR5 – T500 – AmpR – ColE1 origin	this thesis
pJBL-STAR6-trigger1	t500 – trigger1 – J23119 – spacer – J23119 – STAR6 – T500 – AmpR – ColE1 origin	this thesis

Continued on next page

Plasmid Name	Plasmid architecture	Source
pJBL-STAR6-trigger2	t500 – trigger2 – J23119 – spacer – J23119 – STAR6 – T500 – AmpR – ColE1 origin	this thesis
pJBL-STAR6-trigger3	t500 – trigger3 – J23119 – spacer – J23119 – STAR6 – T500 – AmpR – ColE1 origin	this thesis
pJBL-STAR8-trigger1	t500 – trigger1 – J23119 – spacer – J23119 – STAR8 – T500 – AmpR – ColE1 origin	this thesis
pJBL-STAR8-trigger2	t500 – trigger2 – J23119 – spacer – J23119 – STAR8 – T500 – AmpR – ColE1 origin	this thesis
pJBL-STAR8-trigger3	t500 – trigger3 – J23119 – spacer – J23119 – STAR8 – T500 – AmpR – ColE1 origin	this thesis
pJBL-wt(red)	J23119 — RBS – dTomato – TrrnB – CmR – p15A origin	this thesis
pJBL-Target6-Toehold3	J23119 – Target6 – Toehold3 – dTomato – TrrnB – CmR – p15A origin	this thesis
pFXL-S6T3MGA	J23119 – Target6 – Toehold3 – sfGFP-MGA4x – TrrnB – CmR – p15A origin	this thesis
pFXL-S8T1MGA	J23119 – Target8 – Toehold1 – sfGFP-MGA4x – TrrnB – CmR – p15A origin	this thesis
pFXL-S5T2MGA	J23119 – Target6 – Toehold3 – sfGFP-MGA4x – TrrnB – CmR – p15A origin	this thesis
pFXL-P5-deCas9	J23108 – RBS – deCas9 – T500 – AmpR – ColE1 origin	this thesis
pFXL-P5-Csy4	J23108 – RBS – Csy4 – T500 – AmpR – ColE1 origin	this thesis
pFXL-deCas9-Csy4	J23108 – RBS – Csy4 – T500 – J23108 – RBS – deCas9 – T500 – AmpR – ColE1 origin	this thesis
pJBL-sgRNA1	J23119 – sgRNA1 – T500 – AmpR – ColE1 origin	this thesis
pJBL-sgRNA2	J23119 – sgRNA2 – T500 – AmpR – ColE1 origin	this thesis
pJBL-sgRNA3	J23119 – sgRNA3 – T500 – AmpR – ColE1 origin	this thesis
pJBL-sgRNA4	J23119 – sgRNA4 – T500 – AmpR – ColE1 origin	this thesis
pJBL-sg-trig-star	J23119 – STAR6 – Cys4 - Trigger3 – Cys4 – sgRNA1 – T500 – AmpR – ColE1 origin	this thesis
pFXL-ion	J23119 – ion– RBS – sfGFP-MGA4x – T500 – amp – ColE1 origin	this thesis
pFXL-ionC-	J23119 – ion– Csy4 – RBS – sfGFP-MGA4x – T500 – amp – ColE1 origin	this thesis
pFXL-rpiv	J23119 – rpiv– RBS – sfGFP-MGA4x – T500 – amp – ColE1 origin	this thesis
pFXL-rpivC-	J23119 – rpiv– Csy4 – RBS – sfGFP-MGA4x – T500 – amp – ColE1 origin	this thesis
pFXL-t7	J23119 – t7– RBS – sfGFP-MGA4x – T500 – amp – ColE1 origin	this thesis
pFXL-t7C-	J23119 – t7– Csy4 – RBS – sfGFP-MGA4x – T500 – amp – ColE1 origin	this thesis
pFXL-Ag	J23119 – Ag– RBS – sfGFP-MGA4x – T500 – amp – ColE1 origin	this thesis
pFXL-AgC-	J23119 – Ag– Csy4 – RBS – sfGFP-MGA4x – T500 – amp – ColE1 origin	this thesis
pFXL-Irp	J23119 – Irp– RBS – sfGFP-MGA4x – T500 – amp – ColE1 origin	this thesis
pFXL-IrpC-	J23119 – Irp– Csy4 – RBS – sfGFP-MGA4x – T500 – amp – ColE1 origin	this thesis
pFXL02-Sd	J23119 – Sd– RBS – sfGFP-MGA4x – T500 – amp – ColE1 origin	this thesis
pFXL-SdC-	J23119 – Sd– Csy4 – RBS – sfGFP-MGA4x – T500 – amp – ColE1 origin	this thesis

Table 6.18 Plasmids used for protein production.

Plasmid Name	Plasmid architecture	Source

Continued on next page

Plasmid Name	Plasmid architecture	Source
pBADmod1	AraC – Pbad – RBS – GamS – T500 – amp – ColE1 origin	Addgene #45833
pmal-cs5x-cys4	LacI– ptac – RBS – MBP – Cys5 – rrnB – amp – ColE1 origin	this thesis

Construct examples

Table 6.19 Example Plasmids.

Name	Sequence [5'->3']
pJBL-STAR5-trigger1 t500, trigger, J23119, STAR, ColE1, AmpR	GAATTCaaaaaaaagcccgcctttcggcgggctttgAGATCCctatcttatcttatctatctcgttta tccctgctttactgactattgcacagaatagtcagtcccactagtattatacctaggactgagctagctgtcaa AGATCTTTAACGGGGT-CATCACGGCTCATCATGCGCCAAACAAATGTGTGCAATACACGCTCGGATGACTGCA TGATGACCGCACTGACTGGGGACAGCAGATCCACCTAAGCCTGTGAGAGAAGCAGACACC-CGACAGATCAAGGCAGTTAAATTAAAGATCTttgacagctagctcagtcctaggtat aatactagttgaactgtatacattcccccgcaggataggaattgaagatgaaacgatgagacttgggacgaGGATCT caaagcccgccgaaaggcgggctttttttGGATCCTTACTCGAGTCTAGACTGCAGGCTTCCT CGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTC ACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAggccgcgttgctggcg tttttccacaggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaa gataccaggcgtttcccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttct cccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgt gtgcacgaacccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttat cgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaa ctacggctacactagaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttga tccggcaaacaaaccaccgctggtagcggtggttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaag atctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgaGATTATCA AAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCATCT AAAGTATATATGAGTAAACTTGGTCTGACAGTTAccaatgcttaatcagtgaggcacctatctcagc gatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggcccc agtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgca gaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatag
pJBL-STAR5-trigger1 t500, trigger, J23119, STAR, ColE1, AmpR	tttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaa cgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagt tggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgac tggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataatacc gcgccacatagcagaactttaaaagtgctcatCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCA AGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTG ATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGC AAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATA CATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCC GAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAA AATAGGCGTATCACGAGGCAGAATTTCAGATAAAAAAAATCCTTAGCTTTCGCTAA GGATGATTTCTG

Continued on next page

Name	Sequence [5'->3']
pJBL-Target5-Toehold1 J23119, Target, toehold, sfGFP, TrrnB, CmR, p15A	GAATTCTAAAGATCTttgacagctagctcagtcctaggtataatactagttcgtcccaagtctcatcgtttcatct tcaattcctatcctgcggggaatgtatacagttcatgtatatattccccgctttttttttGGATCTgggtcttatcttatc tatctcgtttatccctgcatacagaaacagaggagatatgcaatgataaacgagaacctggcggcagcgcaaaag atgagcaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaatgggcacaaatttttctgtc cgtggagagggtgaaggtgatgctacaaacggaaaactcacccttaaatttatttgcactactggaaaactacctgttccgtgg ccaacacttgtcactactctgacctatggtgttcaatgcttttcccgttatccggatcacatgaaacggcatgacttttttcaag agtgccatgcccgaaggttatgtacaggaacgcactatatctttcaaagatgacgggacctacaagacgcgtgctgaagtcaag tttgaaggtgatacccttgttaatcgtatcgagttaaagggtattgattttaaagaagatggaaacattcttggacacaaactc gagtacaactttaactcacacaatgtatacatcacggcagacaaacaaaagaatggaatcaaagctaacttcaaaattcgccac aacgttgaagatggttccgttcaactagcagaccattatcaacaaaatactccaattggcgatggccctgtccttttaccagac aaccattacctgtcgacacaatctgtcctttcgaaagatcccaacgaaaagcgtgaccacatggtccttcttgagtttgtaact gctgctgggattacacatggcatggatgagctctacaaaTAAGCGGCCGCGGATCTgaagcttgggcccgaaca aaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagtttaaacggtctccagcttggc tgttttggcggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggtctgataaaacagaatttgcct ggcggcagtagcgcggtggtcccacctgaccccatgccgaactcagaagtgaaacgccgtagcgccgatggtagtgtggggtct ccccatgcgagagtagggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttg tttgtcggtgaactGGATCCTTACTCGAGTCTAGACTGCAGTTGATCGggcacgtaagaggttc caactttcaccataatgaaataagatcactaccgggcgtatttttgagttatcgagattttcaggagctaaggaagctaaaat ggagaaaaaaatcactggatataccaccgttgatatatcccaatggcatcgtaaagaacattttgaggcatttcagtcagttgc tcaatgtacctataaccagaccgttcagctggatattacggccttttaaagaccgtaaagaaaaataagcacaagttttatcc ggcctttattcacattcttgcccgcctgatgaatgctcatccggaatttcgtatggcaatgaaagacggtgagctggtgatatg ggatagtgttcaccttgttacaccgttttccatgagcaaactgaaacgttttcatcgctctggagtgaataccacgacgattt ccggcagtttctacacatatattcgcaagatgtggcgtgttacggtgaaaacctggcctatttcccctaaagggtttattgagaa tatgtttttcgtctcagccaatccctgggtgagtttcaccagttttgatttaaacgtggccaatatggacaacttcttcgcccc cgttttcaccatgggcaaatattatacgcaaggcgacaaggtgctgatgccgctggcgattcaggttcatcatgccgtttgtga tggcttccatgtcggcagaatgcttaatgaattacaacagtactgcgatgagtggcagggcggggcgtaatttgatatcgagct cgcttggactcctgttgatagatccagtaatgacctcagaactccatctggatttgttcagaacgctcggttgccgccgggcgt tttttattGGTGAGAATCCAAGCCTCCGATCAACGTCTCATTTTCGCCAAAAGTTG
pJBL-Target5-Toehold1 J23119, Target, toehold, sfGFP, TrrnB, CmR, p15A	GCCCAGGGCTTCCCGGTATCAACAGGGACACCAGGATTTATTTATTCTGCGAAGT GATCTTCCGTCACAGGTATTTATTCGGCGCAAAGTGCGTCGGGTGATGCTGCCAA CTTACTGATTTAGTGTATGATGGTGTTTTTGAGGTGCTCCAGTGGCTTCTGTTTC TATCAGCTGTCCCTCCTGTTCAGCTACTGACGGGGTGGTGCGTAACGGCAAAAGC ACCGCCGGACATCAgcgctagcggagtgtatactggcttactatgttggcactgatgagggtgtcagtgaagtgcttc atgtggcaggagaaaaaaggctgcaccggtgcgtcagcagaatatgtgatacaggatatattccgcttcctcgctcactgactc gctacgctcggtcgttcgactgcggcgagcggaaatggcttacgaacggggcggagatttcctggaagatgccaggaagatact taacagggaagtgagagggccgcggcaaagccgtttttccataggctccgcccccctgacaagcatcacgaaatctgacgctca aatcagtggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggcggctccctcgtgcgctctcctgttcct gcctttcggtttaccggtgtcattccgctgttatggccgcgtttgtctcattccacgcctgacactcagttccgggtaggcagt tcgctccaagctggactgtatgcacgaaccccccgttcagtccgaccgctgcgccttatccggtaactatcgtcttgagtccaa cccggaaagacatgcaaaagcaccactggcagcagccactggtaattgatttagaggagttagtcttgaagtcatgcgccggtt aaggctaaactgaaaggacaagttttggtgactgcgctcctccaagccagttacctcggttcaaagagttggtagctcagagaa ccttcgaaaaaccgccctgcaaggcggtttttttcgttttcagagcaagagattacgcgcagaccaaaacgatctcaagaagatc atcttattaatcagataaaatatttCTAGATTTCAGTGCAATTTATCTCTTCAAATGTAGCACC TGAAGTCAGCCCCATACGATATAAGTTGTAATTCTCATGTTTGACAGCTTATCAT CGATAAGCTTCCGATGGCGCGCCGAGAGGCTTTACACTTTATGCTTCCGGCT

Table 6.20 Part sequences.

Name	Sequence [5'->3']
Promoter J23119	TTGACAGCTAGCTCAGTCCTAGGTATAATACTAGT
Promoter J23107	tttacggctagctcagccctaggtattatgctagc
Promoter J23108	ctgacagctagctcagtcctaggtataatgctagc

Continued on next page

Name	Sequence [5'->3']
Promoter J23115	tttatagctagctcagcccttggtacaatgctagc
Terminator t500	CAAAGCCCGCCGAAAGGCGGGCTTTTTTTT
Terminator TrrnB	GAAGCTTGGGCCCGAACAAAAACTCATCTCAGAAGAGGATCTGAATAGCGC CGTCGACCATCATCATCATCATCATTGAGTTTAAACGGTCTCCAGCTTGGC TGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAA CCAGAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGG TCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGATG GTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATA AAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTG TCGGTGAACT
sfGFP	ATGAGCAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGA ATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCCGTGGAGAGGGTG AAGGTGATGCTACAAACGGAAAACTCACCCTTAAATTTATTTGCACTACT GGAAAACTACCTGTTCCGTGGCCAACACTTGTCACTACTCTGACCTATGG TGTTCAATGCTTTTCCCGTTATCCGGATCACATGAAACGGCATGACTTTT TCAAGAGTGCCATGCCCGAAGGTTATGTACAGGAACGCACTATATCTTTC AAAGATGACGGGACCTACAAGACGCGTGCTGAAGTCAAGTTTGAAGGTGA TACCCTTGTTAATGTATCGAGTTAAAGGGTATTGATTTTAAAGAAGATGG AAACATTCTTGGACACAAACTCGAGTACAACTTTAACTCACACAATGTAT ACATCACGGCAGACAAACAAAAGAATGGAATCAAAGCTAACTTCAAAATT CGCCACAACGTTGAAGATGGTTCCGTTCAACTAGCAGACCATTATCAACA AAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACC TGTCGACACAATCTGTCCTTTCGAAAGATCCCAACGAAAAGCGTGACCAC ATGGTCCTTCTTGAGTTTGTAACTGCTGCTGGGATTACACATGGCATGGA TGAGCTCTACAAA
MGA4x	TCTAGATGGTGTTTTGGTTTGGTCAACGGATCCCGACTGGCGAGAGCCAG GTAACGAATGGATCCGTTGACACCCAACAAAAAAAAACACGGATCCCGAC TGGCGAGAGCCAGGTAACGAATGGATCCGTGTAAAAAAAAACCCAAAAAG CCGGGATCCCGACTGGCGAGAGCCAGGTAACGAATGGATCCCGGCAACCA ACCAACCAACCAAAAAGGATCCCGACTGGCGAGAGCCAGGTAACGAATGG ATCCTTTTTCTAGCGCGGATCCGAATTCGAGCTCCGTCGACAAGCTT
spacer	AGATCTTTAACGGGGTCATCACGGCTCATCATGCGCCAAACAAATGTGTG CAATACACGCTCGGATGACTGCATGATGACCGCACTGACTGGGGACAGCA GATCCACCTAAGCCTGTGAGAGAAGCAGACACCCGACAGATCAAGGCAGT TAAATTAAAGATCT
Target AD1.S5	AGTTTTTACAGTGAATTGTTTTAATTAGTTGTATAAATGTTGGAGCAGCG GGGAATGTATACAGTTCATGTATATATTCCCCGCTTTTTTTTT
Target5	TCGTCCCAAGTCTCATCGTTTCATCTTCAATTCCTATCCTGCGGGGAATG TATACAGTTCATGTATATATTCCCCGCTTTTTTTTT
Target6	CCAGTCATCAAGTCAGTCCAGTCAAAGTTTCCGTCGTTCAGCGGGGAATGT ATACAGTTCATGTATATATTCCCCGCTTTTTTTTT
Target8	CCATCCTCAATCTCTACCTACTCTCACTTACTCTTATCCTGCGGGGAATGT ATACAGTTCATGTATATATTCCCCGCTTTTTTTTT
Toehold1	GGGTCTTATCTTATCTATCTCGTTTATCCCTGCATACAGAAACAGAGGAGA TATGCAATGATAAACGAGAACCTGGCGGCAGCGCAAAAG
Toehold2	GGGAGTTTGATTACATTGTCGTTTAGTTTAGTGATACATAAACAGAGGAGA TATCACATGACTAAACGAAACCTGGCGGCAGCGCAAAAG
Toehold3	GGGATCTATTACTACTTACCATTGTCTTGCTCTATACAGAAACAGAGGAGA TATAGAATGAGACAATGGAACCTGGCGGCAGCGCAAAAG
STAR AD1.A5	TGAACTGTATACATTCCCCGCTGCTCCAACATTTATACAACTAATTAAAAC AATTCACTGTAAAAACT
STAR5	TGAACTGTATACATTCCCCGCAGGATAGGAATTGAAGATGAAACGATGAGA CTTGGGACGA
STAR6	TGAACTGTATACATTCCCCGCTGAACGACGGAAACTTTGACTGGACTGACT TGATGACTGG
STAR8	TGAACTGTATACATTCCCCGCAGGATAAGAGTAAGTGAGAGTAGGTAGAGA TTGAGGATGG

Continued on next page

Name	Sequence [5'->3']
trigger1	GGGACTGACTATTCTGTGCAATAGTCAGTAAAGCAGGGATAAACGAGATAG ATAAGATAAGATAG
trigger2	GGGACAGATCCACTGAGGCGTGGATCTGTGAACACTAAACTAAACGACAAT GTAATCAAACTAAC
trigger3	GGGTGATGGGACATTCCGATGTCCCATCAATAAGAGCAAGACAATGGTAAG TAGTAATAGATAAG
SgRNA1	tcctaggtataatactagttgtagtgacaagtgttggccaGT TTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTT GAAAAAGTGGCACCGAGTCGGTGCTTTTTTTggatctcaaagcccgccgaa
SgRNA2	gtagtgacaagtgttggccaGTTTTAGAGCTAGAAATAGCAA GTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT
SgRNA3	caaactcaagaaggaccatgGTTTTAGAGCTAGAAATAGCAAGTTAAAATA AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTT
SgRNA4	GGGACAGATCCACTGAGGCGTGGATCTGTGAACACTAAACTAAACGACAAT GTAATCAAACTAAC
Csy4 hairpin	GTTCACTGCCGTATAGGCAGCTAAGAAA
Ion UTR	TTAAACTAAGAGAGAGCTCT
RplY+SD UTR	TTAAACTAAGGAGGAGCTCT
T7 UTR	gctagcaataattttgtttaacttta
ybaB-SD UTR	TTAAACTAACACACAGCTCT
ybaB-AG UTR	CGTGATTCACACAGAAACCT
Lrp UTR	AATACAGAGAGACAATAATT
TPP riboswitch	TTCACAGCCAATCTCGCAGAACAATTCAATATGTATTCGTTTAACCACTAG GGGTGTCCTTCATAAGGGCTGAGATAAAAGTGTGACTTTTAGACCCTCATA ACTTGAACAGGTTCAGACCTGCGTAGGGAAGTGGAGCGGTATTTGTGTTAT TTTACTATGCCAATTCCAAACCACTTTTCCTTGCGGGAAAGTGGTTTTTTT ATTTTTC

6.3.1 Plasmid Maps

Fig. 6.2 Examples of plasmid architectures based on pJBL2807 (A) and on pJBL2801 backbones (B).

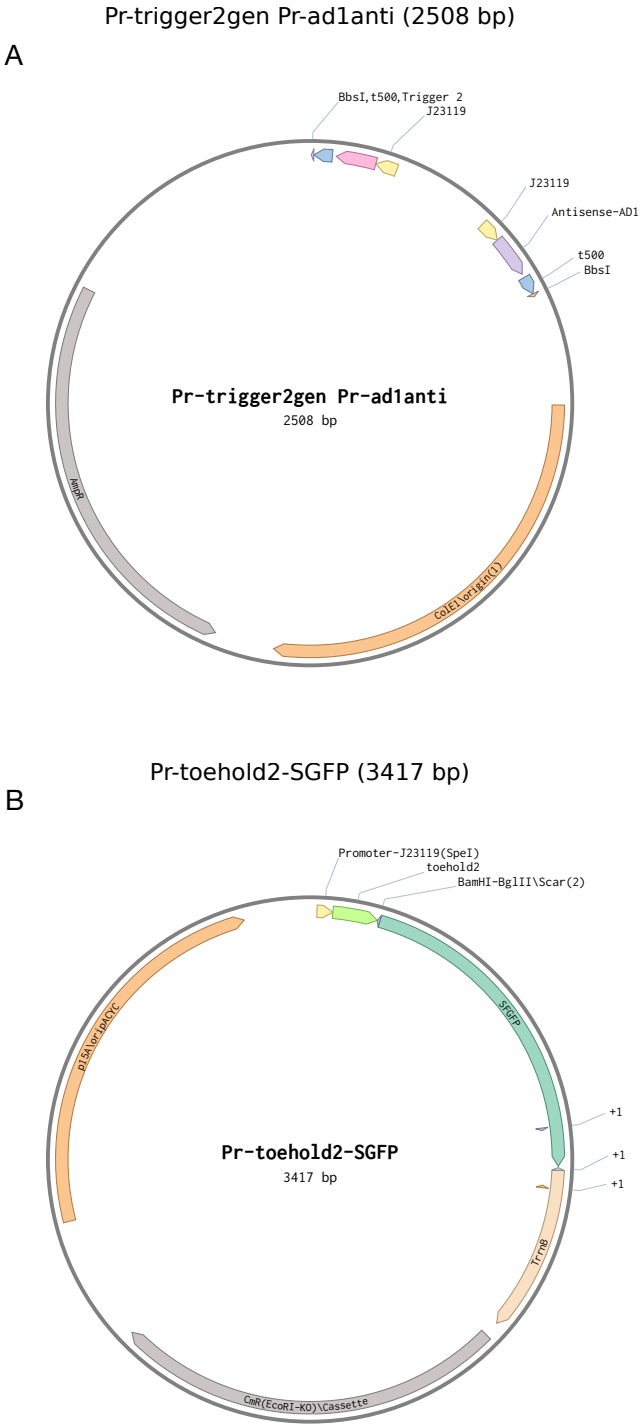
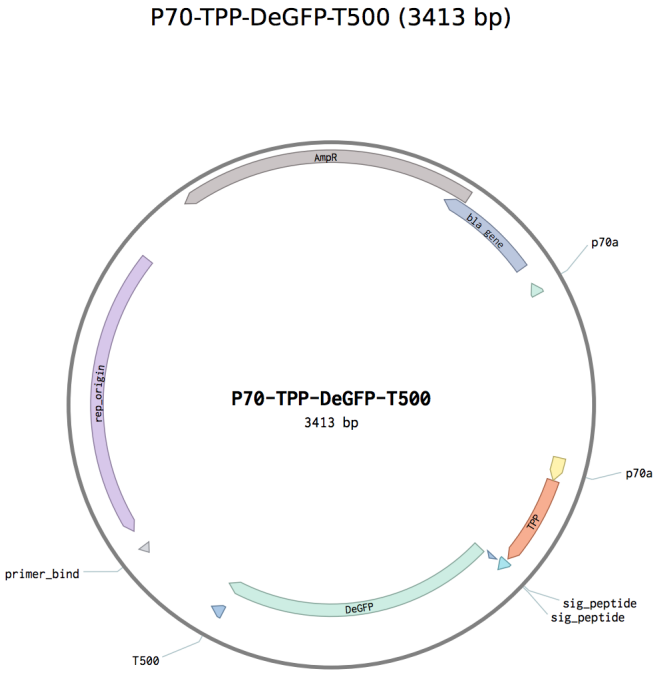


Fig. 6.3 Example of plasmid architecture based on pBEST.



References

[1] Ainla, A., Gözen, I., Orwar, O., and Jesorka, A. (2009). A microfluidic diluter based on pulse width flow modulation. *Analytical Chemistry*, 81(13):5549–5556.

[2] Al-Attar, S., Westra, E. R., van der Oost, J., and Brouns, S. J. J. (2011). Clustered regularly interspaced short palindromic repeats (crisprs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biological chemistry*, 392(4):277—289.

[3] Alam, K. K., Jung, J. K., Verosloff, M. S., Clauer, P. R., Lee, J. W., Capdevila, D. A., Pastén, P. A., Giedroc, D. P., Collins, J. J., and Lucks, J. B. (2019). Rapid, low-cost detection of water contaminants using regulated in vitro transcription. *bioRxiv*.

[4] Angenent-Mari, N. M., Garruss, A. S., Soenksen, L. R., Church, G., and Collins, J. J. (2019). Deep learning for rna synthetic biology. *bioRxiv*.

[5] Au, A. K., Lai, H., Utela, B. R., and Folch, A. (2011). Microvalves and micropumps for BioMEMS. *Micromachines*, 2:179–220.

[6] Autour, A., Westhof, E., and Ryckelynck, M. (2016). ISpinach: A fluorogenic RNA aptamer optimized for in vitro applications. *Nucleic Acids Research*, 44(6):2491–2500.

[7] Borkowski, O., Bricio, C., Murgiano, M., Rothschild-Mancinelli, B., Stan, G.-B., and Ellis, T. (2018). Cell-free prediction of protein expression costs for growing cells. *Nature communications*, 9(1):DOI: 10.1038/s41467–018–03970–x.

[8] Cardinale, S., Joachimiak, M. P., and Arkin, A. P. (2013). Effects of genetic variation on the e. coli host-circuit interface. *Cell Reports*, 4(2):231–237.

[9] Carlson, P. D., Glasscock, C. J., and Lucks, J. B. (2018). De novo Design of Translational RNA Repressors. *bioRxiv*.

[10] Caschera, F. and Noireaux, V. (2014). Synthesis of 2.3Â mg/ml of protein with an all escherichia coli cell-free transcriptionâtranslation system. *Biochimie*, 99:162 – 168.

[11] Chappell, J., Jensen, K., and Freemont, P. S. (2013a). Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic acids research*, 41(5):3471–81.

[12] Chappell, J., Jensen, K., and Freemont, P. S. (2013b). Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic Acids Research*, 41(5):3471–3481.

[13] Chappell, J., Takahashi, M. K., and Lucks, J. B. (2015a). Creating small transcription activating RNAs. *Nature Chemical Biology*, 11(3):214–220.

[14] Chappell, J., Takahashi, M. K., Meyer, S., Loughrey, D., Watters, K. E., and Lucks, J. (2013c). The centrality of RNA for engineering gene expression. *Biotechnology Journal*, 8(12):1379–1395.

[15] Chappell, J., Watters, K. E., Takahashi, M. K., and Lucks, J. B. (2015b). A renaissance in RNA synthetic biology: New mechanisms, applications and tools for the future. *Current Opinion in Chemical Biology*, 28:47–56.

[16] Chappell, J., Westbrook, A., Verosloff, M., and Lucks, J. B. (2017). Computational design of small transcription activating RNAs for versatile and dynamic gene regulation. *Nature Communications*, 8(1):DOI: 10.1038/s41467–017–01082–6.

[17] Chen, Y. J., Liu, P., Nielsen, A. A., Brophy, J. A., Clancy, K., Peterson, T., and Voigt, C. A. (2013). Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nature Methods*, 10(7):659–664.

[18] Cheng, A. A. and Lu, T. K. (2012). Synthetic Biology: An Emerging Engineering Discipline. *Annual Review of Biomedical Engineering*, (14):155–178.

[19] Cheng, Y.-Y., Hirning, A. J., Josić, K., and Bennett, M. R. (2017). The Timing of Transcriptional Regulation in Synthetic Gene Circuits. *ACS Synthetic Biology*, 6(11):1996–2002.

[20] Ching-Yuet To, A., Ho-Ting Chu, D., RuoningWang, A., Cheuk-Yau Li, F., Wai-On Chiu, A., Gao, D. Y., Choi, C. H. J., Kong, S. K., Chan, T. F., Chan, K. M., and Yip, K. Y. (2018). A comprehensive web tool for toehold switch design. *Bioinformatics*, 34(16):2862–2864.

[21] Chizzolini, F., Forlin, M., Yeh Martín, N., Berloffa, G., Cecchi, D., and Mansy, S. S. (2017). Cell-Free Translation Is More Variable than Transcription. *ACS Synthetic Biology*, 6(4):638–647.

[22] Cole, S. D., Beabout, K., Turner, K. B., Smith, Z. K., Funk, V. L., Harbaugh, S. V., Liem, A. T., Roth, P. A., Geier, B. A., Emanuel, P. A., Walper, S. A., Chávez, J. L., and Lux, M. W. (2019). Quantification of Interlaboratory Cell-Free Protein Synthesis Variability. *ACS Synthetic Biology*, pages 2080–2091.

[23] Collins, J. (2012). Synthetic Biology: Bits and pieces come to life. *Nature*, 483:S8–S10.

[24] Davies, J. A. (2018). *Synthetic biology : a very short introduction*. Oxford University Press, Oxford, UK.

[25] Davies, J. A. (2019). Real-world synthetic biology: Is it founded on an engineering approach, and should it be? *Life*, 9(6).

[26] Dillingham, M. S. and Kowalczykowski, S. C. (2008). RecBCD Enzyme and the Repair of Double-Stranded DNA Breaks. *Microbiology and Molecular Biology Reviews*, 72(4):642–671.

[27] Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H. W., Listgarten, J., and Root, D. E. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology*, 34(2):184–191.

[28] Dubuc, E., Pieters, P. A., van der Linden, A. J., van Hest, J. C., Huck, W. T., and de Greef, T. F. (2019). Cell-free microcompartmentalised transcription–translation for the prototyping of synthetic communication networks. *Current Opinion in Biotechnology*, 58:72–80.

[29] Dvir, S., Velten, L., Sharon, E., Zeevi, D., Carey, L. B., Weinberger, A., and Segal, E. (2013). Deciphering the rules by which 5'-UTR sequences affect protein expression in yeast. *Proceedings of the National Academy of Sciences*, 110(30):E2792–E2801.

[30] Earl, D. J. and Deem, M. W. (2005). Parallel tempering: Theory, applications, and new perspectives. *Physical Chemistry Chemical Physics*, 7(23):3910–3916.

[31] Endo, Y. and Sawasaki, T. (2006). Cell-free expression systems for eukaryotic protein production. *Current Opinion in Biotechnology*, (4):373–380.

[32] Evfratov, S. A., Osterman, I. A., Komarova, E. S., Pogorelskaya, A. M., Rubtsova, M. P., Zatsepin, T. S., Semashko, T. A., Kostryukova, E. S., Mironov, A. A., Burnaev, E., Krymova, E., Gelfand, M. S., Govorun, V. M., Bogdanov, A. A., Sergiev, P. V., and Dontsova, O. A. (2017). Application of sorting and next generation sequencing to study 5 -UTR influence on translation efficiency in Escherichia coli. *Nucleic acids research*, 45(6):3487–3502.

[33] Ezure, T., Suzuki, T., and Ando, E. (2014). A cell-free protein synthesis system from insect cells. *Methods in Molecular Biology*, 1118:285–296.

[34] Failmezger, J., Rauter, M., Nitschel, R., Kraml, M., and Siemann-Herzberg, M. (2017). Cell-free protein synthesis from non-growing, stressed Escherichia coli. *Scientific Reports*, 7(1):16524.

[35] Fallmann, J., Will, S., Engelhardt, J., Grüning, B., Backofen, R., and Stadler, P. F. (2017). Recent advances in RNA folding. *Journal of Biotechnology*, 261(July):97–104.

[36] Fredens, J., Wang, K., de la Torre, D., Funke, L. F., Robertson, W. E., Christova, Y., Chia, T., Schmied, W. H., Dunkelmann, D. L., Beránek, V., Uttamapinant, C., Llamazares, A. G., Elliott, T. S., and Chin, J. W. (2019). Total synthesis of Escherichia coli with a recoded genome. *Nature*, 569:514–518.

[37] Fu, P. (2006). A perspective of synthetic biology: Assembling building blocks for novel functions.

[38] Ganser, L. R., Kelly, M. L., Herschlag, D., and Al-Hashimi, H. M. (2019). The roles of structural dynamics in the cellular functions of RNAs. *Nature Reviews Molecular Cell Biology*, 20:474–489.

[39] Garamella, J., Marshall, R., Rustad, M., and Noireaux, V. (2016). The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology. *ACS Synthetic Biology*, 5(4):344–355.

[40] Garcia-Cordero, J. L. and Maerkl, S. J. (2016). Mechanically Induced Trapping of Molecular Interactions and Its Applications. *Journal of Laboratory Automation*, 21(3):356–367.

[41] Garcia-Martin, J. A., Clote, P., and Dotu, I. (2013). RNAiFold: a web server for RNA inverse folding and molecular design. *Nucleic acids research*, 41(Web Server issue):W465–W470.

[42] Gines, G., Menezes, R., Nara, K., Kirstetter, A.-S., Taly, V., and Rondelez, Y. (2019). Isothermal digital detection of microRNA using background-free molecular circuit. *bioRxiv*, page 701276.

[43] Green, A. A., Kim, J., Ma, D., Silver, P. A., Collins, J. J., and Yin, P. (2017). Complex cellular logic computation using ribocomputing devices. *Nature*, 548(7665):117–121.

[44] Green, A. A., Silver, P. A., Collins, J. J., and Yin, P. (2014). Toehold switches: De-novo-designed regulators of gene expression. *Cell*, 159(4):925–939.

[45] Hodgman, C. E. and Jewett, M. C. (2012). Cell-free synthetic biology: Thinking outside the cell. *Metabolic Engineering*, 14(3):261–269.

[46] Hooshangi, S., Thiberge, S., and Weiss, R. (2005). Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proceedings of the National Academy of Sciences*, 102(10):3581–3586.

[47] Hori, Y., Kantak, C., Murray, R. M., and Abate, A. R. (2017). Cell-free extract based optimization of biomolecular circuits with droplet microfluidics. *Lab on a Chip*, 17(18):3037–3042.

[48] Hu, C. Y. and Takahashi, M. K. (2018). Engineering a Functional Small RNA Negative Autoregulation Network with Model-Guided Design. *ACS Synthetic Biology*, 7(6):1507–1518.

[49] Hu, C. Y., Varner, J. D., and Lucks, J. B. (2015). Generating Effective Models and Parameters for RNA Genetic Circuits. *ACS Synthetic Biology*, 4(8):914–926.

[50] Hukushima, K. and Nemoto, K. (1996). Exchange Monte Carlo Method and Application to Spin Glass Simulations. *Journal of the Physical Society of Japan*, 65(6):1604–1608.

[51] Hunsicker, A., Steber, M., Mayer, G., Meitert, J., Klotzsche, M., Blind, M., Hillen, W., Berens, C., and Suess, B. (2009). An RNA Aptamer that Induces Transcription. *Chemistry and Biology*, 16(2):173–180.

[52] Hutchison, C. A., Chuang, R. Y., Noskov, V. N., Assad-Garcia, N., Deerinck, T. J., Ellisman, M. H., Gill, J., Kannan, K., Karas, B. J., Ma, L., Pelletier, J. F., Qi, Z. Q., Richter, R. A., Strychalski, E. A., Sun, L., Suzuki, Y., Tsvetanova, B., Wise, K. S., Smith, H. O., Glass, J. I., Merryman, C., Gibson, D. G., and Venter, J. C. (2016). Design and synthesis of a minimal bacterial genome. *Science*, 351(6280).

[53] Isaacs, F. J., Dwyer, D. J., Ding, C., Pervouchine, D. D., Cantor, C. R., and Collins, J. J. (2004). Engineered riboregulators enable post-transcriptional control of gene expression. *Nature biotechnology*, 22(7):841–847.

[54] Kafri, M., Metzl-Raz, E., Jona, G., and Barkai, N. (2016). The Cost of Protein Production. *Cell Reports*, 14(1):22–31.

[55] Karig, D. K., Jung, S. Y., Srijanto, B., Collier, C. P., and Simpson, M. L. (2013). Probing cell-free gene expression noise in femtoliter volumes. *ACS Synthetic Biology*, 2(9):497–505.

[56] Khalil, A. S. and Collins, J. J. (2010). Synthetic biology: Applications come of age. *Nature Reviews Genetics*, 11:367–379.

[57] Kim, D. M. and Swartz, J. R. (1999). Prolonging cell-free protein synthesis with a novel ATP regeneration system. *Biotechnology and Bioengineering*, 66(3):180–188.

[58] Kim, J., Quijano, J. F., Yeung, E., and Murray, R. M. (2014). Synthetic logic circuits using RNA aptamer against T7 RNA polymerase. *bioRxiv*.

[59] Kim, M. Y., Hur, J., and Jeong, S. (2009). Emerging roles of RNA and RNA-binding protein network in cancer cells. *BMB Reports*, 42(3):125–130.

[60] Kitada, T., DiAndreth, B., Teague, B., and Weiss, R. (2018). Programming gene and engineered-cell therapies with synthetic biology. *Science*, 359(651).

[61] Kovtun, O., Mureev, S., Jung, W. R., Kubala, M. H., Johnston, W., and Alexandrov, K. (2011). Leishmania cell-free protein expression system. *Methods*, 55(1):58–64.

[62] Kushwaha, M., Rostain, W., Prakash, S., Duncan, J. N., and Jaramillo, A. (2016). Using RNA as Molecular Code for Programming Cellular Function. *ACS Synthetic Biology*, 5(8):795–809.

[63] Kwon, Y.-C. and Jewett, M. C. (2015). High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Scientific Reports*, 5:8663–8670.

[64] Lavickova, B. and Maerkl, S. J. (2019). A simple, robust, and low-cost method to produce the pure cell-free system. *ACS Synthetic Biology*, 8(2):455–462. PMID: 30632751.

[65] Lee, Y. J., Hoynes-O'Connor, A., Leong, M. C., and Moon, T. S. (2016). Programmable control of bacterial gene expression with the combined CRISPR and antisense RNA system. *Nucleic Acids Research*, 44(5):2462–2473.

[66] Lee, Y. J., Kim, S. J., and Moon, T. S. (2018). Multilevel Regulation of Bacterial Gene Expression with the Combined STAR and Antisense RNA System. *ACS Synthetic Biology*, 7(3):853–865.

[67] Lehr, F.-X., Hanst, M., Vogel, M., Kremer, J., Göringer, H. U., Suess, B., and Koeppl, H. (2019). Cell-Free Prototyping of AND-Logic Gates Based on Heterogeneous RNA Activators. *ACS Synthetic Biology*, 8(9):2163–2173.

[68] Levskaya, A., Chevalier, A. A., Tabor, J. J., Simpson, Z. B., Lavery, L. A., Levy, M., Davidson, E. A., Scouras, A., Ellington, A. D., Marcotte, E. M., and Voigt, C. A. (2005). Engineering Escherichia coli to see light. *Nature*, 438(7067):441–442.

[69] Li, G. W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell*, 157(3):624–635.

[70] Lian, Q., Cao, H., and Wang, F. (2014). The Cost-Efficiency Realization in the Escherichia coli-Based Cell-Free Protein Synthesis Systems. *Applied Biochemistry and Biotechnology*, (174):2351–2367.

[71] Mardanlou, V., Tran, C. H., and Franco, E. (2014). Design of a molecular bistable system with RNA-mediated regulation. *Proceedings of the IEEE Conference on Decision and Control*, 2015(February):4605–4610.

[72] Marshall, R., Maxwell, C. S., Collins, S. P., He, Y., Beisel, C. L., Noireaux, V., Marshall, R., Maxwell, C. S., Collins, S. P., Jacobsen, T., Luo, M. L., Begemann, M. B., Gray, B. N., January, E., Singer, A., He, Y., Beisel, C. L., and Noireaux, V. (2018). Rapid and Scalable Characterization of CRISPR Technologies Using an E . coli Cell-Free Transcription-Translation System. *Molecular Cell*, 69(1):146–157.e3.

[73] Matsubayashi, H. and Ueda, T. (2014). Cell-free protein synthesis system. *Kobunshi*, 63:379–381.

[74] Matsuoka, Y., Ghosh, S., and Kitano, H. (2009). Consistent design schematics for biological systems: Standardization of representation in biological engineering. *Journal of the Royal Society Interface*, 6(4).

[75] Maurizi, M. R. (1992). Proteases and protein degradation in Escherichia coli. *Experientia*, 48(2):178–201.

[76] Maxwell, C. S., Jacobsen, T., Marshall, R., Noireaux, V., and Beisel, C. L. (2018). A detailed cell-free transcription-translation-based assay to decipher crispr protospacer-adjacent motifs. *Methods*, 143:48 – 57. Methods and advances in RNA characterization and design.

[77] Meyer, S., Chappell, J., Sankar, S., Chew, R., and Lucks, J. B. (2016). Improving fold activation of small transcription activating RNAs (STARs) with rational RNA engineering strategies. *Biotechnology and Bioengineering*, 113(1):216–225.

[78] Mikami, S., Masutani, M., Sonenberg, N., Yokoyama, S., and Imataka, H. (2006). An efficient mammalian cell-free translation system supplemented with translation factors. *Protein Expression and Purification*, 46(2):348–357.

[79] Moore, S. J., MacDonald, J. T., Wienecke, S., Ishwarbhai, A., Tsipa, A., Aw, R., Kylilis, N., Bell, D. J., McClymont, D. W., Jensen, K., Polizzi, K. M., Biedendieck, R., and Freemont, P. S. (2018). Rapid acquisition and model-based analysis of cell-free transcription–translation reactions from nonmodel bacteria. *Proceedings of the National Academy of Sciences*, 115(19):E4340—-E4349.

[80] Müller, I. E., Rubens, J. R., Jun, T., Graham, D., Xavier, R., and Lu, T. K. (2019). Gene networks that compensate for crosstalk with crosstalk. *Nature Communications*, 10(1):4028.

[81] Murakami, S., Matsumoto, R., and Kanamori, T. (2019). Constructive approach for synthesis of a functional IgG using a reconstituted cell-free protein synthesis system. *Scientific Reports*, 9(679).

[82] Niederholtmeyer, H., Stepanova, V., and Maerkl, S. J. (2013). Implementation of cell-free biological networks at steady state. *Proceedings of the National Academy of Sciences*, 110(40):15985–15990.

[83] Niederholtmeyer, H., Sun, Z. Z., Hori, Y., Yeung, E., Verpoorte, A., Murray, R. M., and Maerkl, S. J. (2015). Rapid cell-free forward engineering of novel genetic ring oscillators. *eLife*, 4:DOI: 10.7554/eLife.09771.

[84] Nirenberg, M. (2004). Historical review: Deciphering the genetic code - A personal account. 29:46–54.

[85] Noireaux, V., Bar-Ziv, R., and Libchaber, A. (2003). Principles of cell-free genetic circuit assembly. *Proceedings of the National Academy of Sciences*, 100(22):12672–12677.

[86] Noireaux, V. and Libchaber, A. (2004). A vesicle bioreactor as a step toward an artificial cell assembly. *Proceedings of the National Academy of Sciences of the United States of America*, 101(51):17669–17674.

[87] Norred, S. E., Caveney, P. M., Chauhan, G., Collier, L. K., Collier, C. P., Abel, S. M., and Simpson, M. L. (2018). Macromolecular crowding induces spatial correlations that control gene expression bursting patterns. *ACS Synthetic Biology*, 7(5):1251–1258. PMID: 29687993.

[88] Orth, J. H., Schorch, B., Boundy, S., Ffrench-Constant, R., Kubick, S., and Aktories, K. (2011). Cell-free synthesis and characterization of a novel cytotoxic pierisin-like protein from the cabbage butterfly Pieris rapae. *Toxicon*, 57:199–207.

[89] Paddon, C. J. and Keasling, J. D. (2014). Semi-synthetic artemisinin: A model for the use of synthetic biology in pharmaceutical development. 12(5):355–367.

[90] Pardee, K., Green, A. A., Ferrante, T., Cameron, D. E., Daleykeyser, A., Yin, P., and Collins, J. J. (2014). Paper-based synthetic gene networks. *Cell*, 159(4):940–954.

[91] Pardee, K., Green, A. A., Takahashi, M. K., Braff, D., Lambert, G., Lee, J. W., Ferrante, T., Ma, D., Donghia, N., Fan, M., Daringer, N. M., Bosch, I., Dudley, D. M., O'Connor, D. H., Gehrke, L., and Collins, J. J. (2016). Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell*, 165(5):1255–1266.

[92] Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006). Engineering and characterization of a superfolder green fluorescent protein. *Nature Biotechnology*, 24(1):79–88.

[93] Pingoud, A., Wilson, G. G., and Wende, W. (2014). Type II restriction endonucleases - A historical perspective and more. 42(12):7489–7527.

[94] Pouvreau, B., Vanhercke, T., and Singh, S. (2018). From plant metabolic engineering to plant synthetic biology: The evolution of the design/build/test/learn cycle. *Plant Science*, 273:3–12.

[95] Prangemeier, T., **Lehr, F-X**., Schoeman, R., and Koeppl, H. (2020). Microfluidic Platforms for the Dynamic Characterisation of Synthetic Circuitry. *Current Opinion in Biotechnology*.

[96] Qi, L., Haurwitz, R. E., Shao, W., Doudna, J. A., and Arkin, A. P. (2012). RNA processing enables predictable programming of gene expression. *Nature Biotechnology*, 30(10):1002–1006.

[97] Rath, D., Amlinger, L., Rath, A., and Lundgren, M. (2015). The crispr-cas immune system: Biology, mechanisms and applications. *Biochimie*, 117:119 – 128. Special Issue: Regulatory RNAs.

[98] Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D. S., Zhang, L. V., Wong, S. L., Franklin, G., Li, S., Albala, J. S., Lim, J., Fraughton, C., Llamosas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R. S., Vandenhaute, J., Zoghbi, H. Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M. E., Hill, D. E., Roth, F. P., and Vidal, M. (2005). Towards a proteome-scale map of the human protein-protein interaction network. *Nature*, 437(7062):1173–1178.

[99] Salis, H. M. (2011). The ribosome binding site calculator. *Methods in Enzymology*, 498:19–42.

[100] Schneider, C., Bronstein, L., Diemer, J., Koeppl, H., and Suess, B. (2017). ROC'n'Ribo: Characterizing a Riboswitching Expression System by Modeling Single-Cell Data. *ACS Synthetic Biology*, 6(7):1211–1224.

[101] Senoussi, A., Lee Tin Wah, J., Shimizu, Y., Robert, J., Jaramillo, A., Findeiss, S., Axmann, I. M., and Estevez-Torres, A. (2018). Quantitative Characterization of Translational Riboregulators Using an in Vitro Transcription-Translation System. *ACS Synthetic Biology*, 7(5):1269–1278.

[102] Shen, S., Rodrigo, G., Prakash, S., Majer, E., Landrain, T. E., Kirov, B., Daròs, J. A., and Jaramillo, A. (2015). Dynamic signal processing by ribozyme-mediated RNA circuits to control gene expression. *Nucleic Acids Research*, 43(10):5158–5170.

[103] Shi, X., Wu, T., Cole, C. M., Devaraj, N. K., and Joseph, S. (2018). Optimization of ClpXP activity and protein synthesis in an E. coli extract-based cell-free expression system. *Scientific Reports*, 8(3488).

[104] Shimizu, Y. and Ueda, T. (2010). PURE technology. *Methods in molecular biology (Clifton, N.J.)*, 607:11–21.

[105] Shin, J. and Noireaux, V. (2010). Study of messenger RNA inactivation and protein degradation in an Escherichia coli cell-free expression system. *Journal of Biological Engineering*, 4(July).

[106] Shin, J. and Noireaux, V. (2012). An E. coli cell-free expression toolbox: Application to synthetic gene circuits and artificial cells. *ACS Synthetic Biology*, 1(1):29–41.

[107] Shojaeian, M., Lehr, F.-X., Göringer, H. U., and Hardt, S. (2019a). On-demand production of femtoliter drops in microchannels and their use as biological reaction compartments. *Analytical Chemistry*, 91(5):3484–3491.

[108] Shojaeian, M., **Lehr, F-X.**, Göringer, H. U., and Hardt, S. (2019b). On-demand production of femtoliter drops in microchannels and their use as biological reaction compartments. *Analytical Chemistry*, 91(5):3484–3491.

[109] Siegal-gaskins, D., Noireaux, V., and Murray, R. M. (2013). Biomolecular resource utilization in elementary cell-free gene circuits. *American Control Conference (ACC)*, pages 1531–1536.

[110] Siegal-Gaskins, D., Tuza, Z. A., Kim, J., Noireaux, V., and Murray, R. M. (2014). Gene circuit performance characterization and resource usage in a cell-free "breadboard". *ACS Synthetic Biology*, 3(6):416–425.

[111] Silverman, A. D., Kelley-Loughnane, N., Lucks, J. B., and Jewett, M. C. (2019). Deconstructing cell-free extract preparation for in vitro activation of transcriptional genetic circuitry. *ACS Synthetic Biology*, 8(2):403–414. PMID: 30596483.

[112] Smolke, C. D. (2009). Building outside of the box: IGEM and the BioBricks Foundation. *Nature Biotechnology*, 27(2009):1099–1102.

[113] Sorek, R., Kunin, V., and Philip, H. (2008). Crispr - a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature reviews. Microbiology*, 6:181–6.

[114] Spirin, A. S., Baranov, V. I., Ryabova, L. A., Ovodov, S. Y., and Alakhov, Y. B. (1988). A continuous cell-free translation system capable of producing polypeptides in high yield. *Science*, 242(4882):1162–1164.

[115] Standards, B. A. (2014). DNA Cloning and Assembly Methods. *Methods in Molecular Biology*, 116:1–24.

[116] Storz, G., Altuvia, S., and Wassarman, K. M. (2005). An abundance of RNA regulators. *Annual Review of Biochemistry*, 74(1):199–217.

[117] Sudarsan, N., Cohen-Chalamish, S., Nakamura, S., Emilsson, G. M., and Breaker, R. R. (2005). Thiamine pyrophosphate riboswitches are targets for the antimicrobial compound pyrithiamine. *Chemistry Biology*, 12(12):1325 – 1335.

[118] Sugita, Y. and Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, 314(1-2):141–151.

[119] Sun, Z. Z., Yeung, E., Hayes, C. A., Noireaux, V., and Murray, R. M. (2014). Linear DNA for rapid prototyping of synthetic biological circuits in an escherichia coli based TX-TL cell-free system. *ACS Synthetic Biology*, 3(6):387–397.

[120] Takahashi, M. K., Chappell, J., Hayes, C. a., Sun, Z. Z., Kim, J., Singhal, V., Spring, K. J., Al-Khabouri, S., Fall, C. P., Noireaux, V., Murray, R. M., and Lucks, J. B. (2014). Rapidly Characterizing the Fast Dynamics of RNA Genetic Circuitry with Cell-Free Transcription-Translation (TX-TL) Systems. *ACS synthetic biology*, 4(5):503–515.

[121] Takahashi, M. K., Hayes, C. A., Chappell, J., Sun, Z. Z., Murray, R. M., Noireaux, V., and Lucks, J. B. (2015). Characterizing and prototyping genetic networks with cell-free transcription-translation reactions. *Methods*, 86:60–72.

[122] Tayar, A. M., Karzbrun, E., Noireaux, V., and Bar-Ziv, R. H. (2015). Propagating gene expression fronts in a one-dimensional coupled system of artificial cells. *Nature Physics*, 11(12):1037–1041.

[123] **Lehr, F-X.**, Hanst, M., Vogel, M., Kremer, J., Göringer, H. U., Suess, B., and Koeppl, H. (2019). Cell-Free Prototyping of AND-Logic Gates Based on Heterogeneous RNA Activators. *ACS Synthetic Biology*, 8(9):2163–2173.

[124] Tinafar, A., Jaenes, K., and Pardee, K. (2019). Synthetic Biology Goes Cell-Free. *BMC Biology*, 17(64).

[125] To, A. C.-Y., Chu, D. H.-T., Wang, A. R., Li, F. C.-Y., Chiu, A. W.-O., Gao, D. Y., Choi, C. H. J., Kong, S.-K., Chan, T.-F., Chan, K.-M., and Yip, K. Y. (2018). A comprehensive web tool for toehold switch design. *Bioinformatics (Oxford, England)*, 34(16):2862–2864.

[126] Wandera, K. G., Collins, S. P., Wimmer, F., Marshall, R., Noireaux, V., and Beisel, C. L. (2020). An enhanced assay to characterize anti-crispr proteins using a cell-free transcription-translation system. *Methods*, 172:42 – 50. Methods for characterizing, applying, and teaching CRISPR-Cas systems.

[127] Weber, W. and Fussenegger, M. (2012). Emerging biomedical applications of synthetic biology. *Nature Reviews Genetics*, 13(2012):21–35.

[128] Westbrook, A. M. and Lucks, J. B. (2017). Achieving large dynamic range control of gene expression with a compact RNA transcription-translation regulator. *Nucleic Acids Research*, 45(9):5614–5624.

[129] Witherell, G. (2000). In Vitro Translation Using HeLa Extract. *Current Protocols in Cell Biology*, 6(1):11.8.1–11.8.10.

[130] Wittmann, A. and Suess, B. (2012). Engineered riboswitches: Expanding researchers' toolbox with synthetic RNA regulators. *FEBS Letters*, 586(15):2076–2083.

[131] Wu, J., Zhou, P., Zhang, X., and Dong, M. (2017). Efficient de novo synthesis of resveratrol by metabolically engineered Escherichia coli. *Journal of Industrial Microbiology and Biotechnology*, 44(7):1083–1095.

[132] Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R., and Benenson, Y. (2011). Multi-Input RNAi-Based Logic Circuit for Identification of Specific Cancer Cells. *Science*, 333(6047):1307–1311.

[133] Yao, R., Liu, D., Jia, X., Zheng, Y., Liu, W., and Xiao, Y. (2018). CRISPR-Cas9/Cas12a biotechnology and application in bacteria. *Synthetic and Systems Biotechnology*, 3(3):135–149.

[134] Yerramilli, V. S. and Kim, K. H. (2018). Labeling RNAs in Live Cells Using Malachite Green Aptamer Scaffolds as Fluorescent Probes. *ACS Synthetic Biology*, 7(3):758–766.

[135] Yim, S. S., Johns, N. I., Park, J., Gomes, A. L., McBee, R. M., Richardson, M., Ronda, C., Chen, S. P., Garenne, D., Noireaux, V., and Wang, H. H. (2019). Multiplex transcriptional characterizations across diverse bacterial species using cell-free systems. *Molecular Systems Biology*, 15(8):e8875.

[136] Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., Dirks, R. M., and Pierce, N. A. (2011). NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry*, 32(1):170–173.

[137] Zhang, F., Carothers, J. M., and Keasling, J. D. (2012). Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nature Biotechnology*, 30(4):354–359.

[138] Zubay, G. (1973). In vitro synthesis of protein in microbial systems. *Annual Review of Genetics*, 7(1):267–287. PMID: 4593305.

Appendix A

Additional Abbreviations

Table A.1 List of units.

Abbreviation	Corresponding unit
V	Volt
J	Joule
L	Liter
M	Molar
g	gram
h	hour
min	minute
s	second
U	Unit
Hz	Herz
Pa	Pascal
rpm	round per minute
RT	Room Temperature
OD_{600}	Optical Density at a wavelength of 600 nM
°C	degree Celsius

Table A.2 List of nucleobases.

Abbreviation	Corresponding unit
A	Adenine
T	Thymine
C	Cytosine
G	Guanine
U	Uracile
N	A, T, C or G

Appendix B

Curriculum Vitae

Name: François-Xavier Lehr
Date of Birth: 01/08/1990
Address: 62 Lichtenbergstr. 64289 Darmstadt, Germany

Professional Experience

Bioinspired Communication Systems, TU Darmstadt, Germany

2015-2020. Graduate student in the laboratory of Prof. Dr. Heinz Koeppl.
Doctoral thesis: Engineering and testing RNA-circuits in cell-free systems.

Education

Master of Sciences, University of Paris-Saclay, France

2013-2014. Master of Systems and Synthetic Biology.
Master Thesis: Mutation analysis: development of a spot detection application along with novel methods for cells segmentation and tracking.

Bachelor of Sciences, University of Grenoble, France

2009-2012. Bachelor of Bioinformatics.

Publications

Shojaeian, M., **Lehr, F-X.**, Göringer, H. U., and Hardt, S. (2019b). On-demand production of femtoliter drops in microchannels and their use as biological reaction compartments. *Analytical Chemistry*, 91(5):3484–3491

Lehr, F-X., Hanst, M., Vogel, M., Kremer, J., Göringer, H. U., Suess, B., and Koeppl, H. (2019). Cell-Free Prototyping of AND-Logic Gates Based on Heterogeneous RNA Activators. *ACS Synthetic Biology*, 8(9):2163–2173

Prangemeier, T., **Lehr, F-X**., Schoeman, R., and Koeppl, H. (2020). Microfluidic Platforms for the Dynamic Characterisation of Synthetic Circuitry. *Current Opinion in Biotechnology*