**Okinawa Institute of Science and Technology
Graduate University**

Thesis submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

High-Throughput Screening of Cell-Free Riboswitches for Chemical Communication Between Microdroplets

by

Takeshi Ricardo Tabuchi Yagui

Under the supervision of

Yohei Yokobayashi, PhD.

September 2022
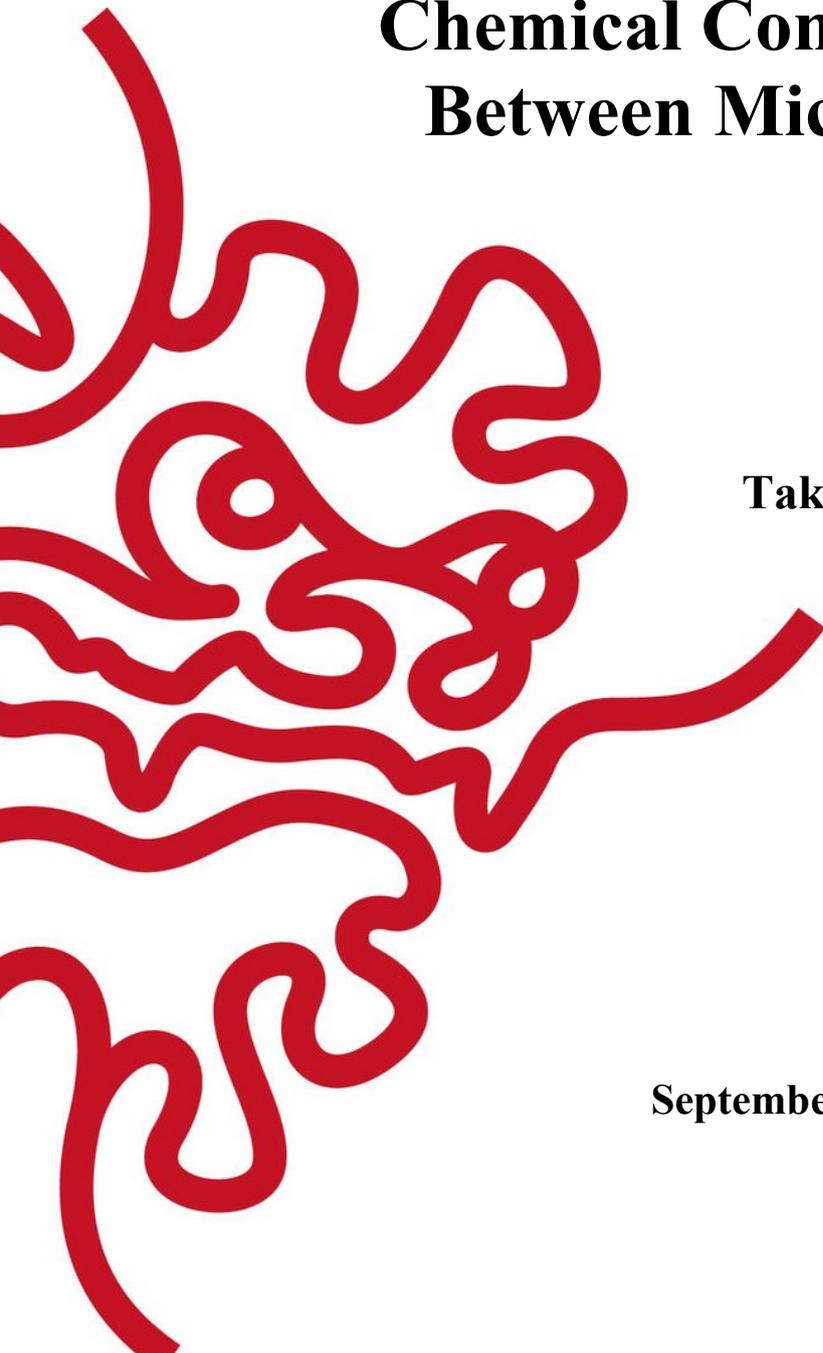
DECLARATION OF ORIGINAL AND SOLE AUTHORSHIP

I, Takeshi Ricardo Tabuchi Yagui, hereby declare that this thesis entitled "High-Throughput Screening of Cell-Free Riboswitches for Chemical Communication Between Microdroplets" and the data presented in it are original and my own unaided work.

I confirm that:

- No part of this work has previously been submitted for a degree at this or any other university.

- All direct or indirect sources used have been clearly acknowledged as references. Quotations from the work of others have been clearly indicated and attributed to them.

- Whenever others have contributed to parts of this work, such contribution has been acknowledged and distinguished from my own work.

- None of this work has been previously published elsewhere, with the exception of the following:

  o Tabuchi, T. and Y. Yokobayashi (2021). "Cell-free riboswitches." *RSC Chem. Biol.* **2**(5): 1430-1440. DOI: 10.1039/d1cb00138h

  o Tabuchi, T. and Y. Yokobayashi (2022). "High-throughput screening of cell-free riboswitches by fluorescence-activated droplet sorting." *Nucleic Acids Res.* **50**(6): 3535-3550. DOI: 10.1093/nar/gkac152

September 1st, 2022

Takeshi Ricardo Tabuchi Yagui

ABSTRACT

Riboswitches have recently attracted the attention of synthetic biologists as an alternative to transcription factors for genetic regulation because of their engineerability, relative simplicity, and potential for responding to a wide array of chemical signals. However, biological constraints such as cell permeability, metabolic stability, and toxicity of their chemical ligands have prevented the development of some of those devices using conventional approaches with living cells. Cell-free systems are generally not subject to such constraints and offer a unique platform for building biochemical and genetic systems that display complex functions without using living cells. Efforts to engineer regulatory components directly in cell-free systems thus far have been based on low-throughput experimental approaches, limiting the availability of basic components for building genetically programmed cell-free systems. Here, I report a high-throughput screening method for engineering riboswitches directly in a cell-free system. Fluorescence-activated droplet-sorting (FADS) of randomized riboswitch libraries in a cell-free protein synthesis (CFPS) system rapidly identified cell-free riboswitches that respond to histamine and ciprofloxacin, compounds that are normally not compatible with conventional bacterial screening methods. Finally, the riboswitches obtained through this method were used to demonstrate chemical communication between microdroplets in a self-contained system.

ACKNOWLEDGEMENTS

Thanks to the OIST Student Support Section, OIST Graduate School, and OIST Resource Center for all the incommensurable support. They took care of me and helped me survive in this beautiful foreign land called Okinawa, Japan.

Thanks to Mo, Otis, Kim, and an extensive list of friends from inside and outside OIST. Friends that are in the *et al.* of this list, please be sure that even if I haven't mention you by name, you still have a special place in my heart and these lines are for you. To all of you, thank you for your company and friendship, which helped me to overcome every roadblock I have faced in this journey, and in general, made my life more fun and enjoyable. I am eternally thankful for you being a part of my life.

Thanks to OIST Football Club, OIST Kickboxing Club, and my MTG playgroup for allowing me to unwind and de-stress from work while having fun doing sports and hobbies. Thank you for letting me kick footballs, kick and punch people in protective gear, and smash imaginary life totals with expensive cardboard pieces.

Thanks to my family and friends far away, back home in Peru, for their unconditional love and support that gave me the strength to keep moving forward. We have been apart for too many years, and I have greatly missed you.

Thanks to my loyal and beautiful Kumiko, my red Honda PCX 125cc motorbike. She treated me well and carried me around on many exciting adventures on this island, as well as very mundane errands, like grocery shopping and bank visits. I'm sorry for the many scratches I let happen to you. Farewell my beloved motorbike, I will miss you.

And special thanks to Annie McGovern, the love of my life, who has stood by my side for all these years (figuratively, not physically, unfortunately). She inspires me to do better and encourages me to keep going. Her love gave me a very special goal to look forward to in the future. Also, thanks for looking into the grammar and spelling *horrors* in my manuscripts.

LIST OF ABBREVIATIONS

Abbreviation	Meaning
aa	Amino acid residues
Abs_{470nm}	Absorbance at 470 nm
AHL	N-Acyl homoserine lactone
AMS	Alternative metastable structures
AUG	Start codon
BEAMing	Beads, emulsion, amplification, and magnetics
BHQ-2	Black Hole Quencher-2
Bio	Biotin
bp	Base pairs
CFPS	Cell-free protein synthesis
CFX	Ciprofloxacin
COP	Cyclo olefin polymer
CS	Competing stem
DAPI	4′,6-diamidino-2-phenylindole
DEP	Dielectrophoresis
DF	Dilution factor
DFHBI	3,5-difluoro-4-hydroxybenzylidene imidazolidinone
DIB	Droplet interphase bilayer
EC_{50}	Half maximal effective concentration
EDTA	Ethylenediaminetetraacetic acid
eGFP	Enhanced green fluorescent protein
em.	Emission wavelength
ePCR	Polymerase chain reaction in emulsion
ex.	Excitation wavelength
FACS	Fluorescence-activated cell sorting
FADS	Fluorescence-activated droplet sorting
FAM	Fluorescein amidite
FL-2	Fluorescence channel 2 (ex. 488 nm, em. 543±22 nm); GFP
FL-5	Fluorescence channel 5 (ex. 561 nm, em. 676±37 nm); ROX
FlAsH	Fluorescein arsenical helix/hairpin binder; Lumio Green
FSC	Forward scatter
GFP	Green fluorescent protein
GFP11	Green fluorescent protein β-sheet 11
GFP1–10	Green fluorescent protein β-barrel (β-sheets 1–10)
HA	Histamine
HDC	Histidine decarboxylase
HF	High fidelity
HFE-7500	Hydroxyfluoroether 7500; Novec 7500

Abbreviation	Meaning
His	L-histidine
HTS	High-throughput sequencing
IDT	Integrated DNA Technologies, inc. (IA, USA)
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IS	Inhibitory stem
ISL	Inhibitory stem-loop
iSp18	18-atom hexa-ethyleneglycol spacer
K_d	Dissociation constant
LB	Luria-Bertani medium
Lig.	Ligand
mRNA	Messenger RNA
MWCO	Molecular weight cut-off
NEB	New England Biolabs, inc. (MA, USA)
Ni-NTA	Nickel-nitrilotriacetic acid
nt	Nucleotides
NTP	Nucleoside triphosphate (ATP, GTP, CTP, TTP)
OD_{600}	Optical density at 600 nm
OMB	2'-O-methyl-RNA molecular beacon
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PGI	Post Genome Institute Co., Ltd. (Tokyo, Japan)
PLP	Pyridoxal phosphate
P_{T7}	T7 promoter
PURE	Protein synthesis using recombinant elements
RBS	Ribosome binding site
ReAsH	Resorufin arsenical helix/hairpin binder; Lumio Red
RNAP	RNA polymerase
ROX	Rhodamine Red-X or 6-carboxy-X-rhodamine
S30	*Escherichia coli* S30 cell extract
SD	Shine-Dalgarno sequence
SDS	Sodium dodecyl sulfate
SELEX	Systematic evolution of ligands by exponential enrichment
sfGFP	Superfolder green fluorescent protein
SSC	Side scatter
TE	Tris-HCl EDTA buffer
TEK	Tris-HCl EDTA potassium buffer
TENa	Tris-HCl EDTA sodium buffer
TexRd	Texas Red or sulforhodamine 101
TF	Transcription factors
TK	Tris-HCl potassium buffer
TPP	Thiamine pyrophosphate
tRNA	Transfer RNA
UTR	Untranslated region
WGE	Wheat germ extract

GLOSSARY

Term	Definition
BEAMing (Beads, Emulsion, Amplification, and Magnetics)	Highly sensitive method developed by Vogelstein group (Diehl et al., 2006) which combines digital PCR in emulsion, magnetic beads, and flow cytometry. It can be used to identify and quantify mutations present in DNA samples. In this work, BEAMing was employed to produce beads coated with monoclonal DNA templates encoding unique riboswitch variants.
CFPS (Cell-Free Protein Synthesis)	*In vitro* protein expression of either *in vitro* transcribed mRNA or mRNA isolated from cells, using a cell-free system that replicates the translation process. These cell-free systems usually also include *in vitro* transcription machinery to synthesize the mRNA directly from DNA templates.
Cell-free system	*In vitro* tool or platform that allows the study of complex biological processes without the use of intact cells, thus reducing the complexity and interactions typically present when working with whole living cells.
Dissociation constant (K_d)	In the context of this work, K_d is an equilibrium constant that measures the propensity of the aptamer–ligand complex to reversibly separate (dissociate). Generally, the smaller the value of K_d, the higher the affinity between the aptamer and the ligand.
Enrichment trend	Parameter employed in this work to estimate the tendency of the riboswitch variants to become enriched or depleted during the sorting cycles, which was used to rank and identify potential functional riboswitches. It was calculated by the slope formula of the linear regression of the abundance of each variant through the sorting cycles (see **H.4. Sequencing data processing**, page **XIII**).
Half maximal effective concentration (EC_{50})	In the context of this work, EC_{50} measures the concentration of a metabolite (ligand) which induces a riboswitch response halfway between the baseline and the maximum gene expression level achieved under specific experimental conditions.
Mock sorting	Process employed to validate the sorting method presented in this work. "Mock" DNA constructs were used instead of real riboswitch candidates and were designed to simulate the gene expression output of a riboswitch in an ON and OFF state. These "mock" constructs were mixed in specific ratios and subjected to the same gene expression and sorting process as the real riboswitches, after which the enrichment of the constructs was verified (see **3.5. Validation of the sorting method through mock sorting**, page **39**).
PURE system (Protein synthesis Using Recombinant Elements)	A type of cell-free system originally developed by Shimizu et al. (Shimizu et al., 2001, Shimizu et al., 2005) and produced by individually synthesizing and purifying each component necessary for protein expression. It is commercially available under different brands and formulations: PURESYSTEM (PGI), PURE*frex* (GeneFrontier), PURExpress (NEB), etc.
SELEX (Systematic Evolution of Ligands by EXponential enrichment)	Also called *in vitro* selection or *in vitro* evolution. It is a combinatorial chemistry technique used for obtaining oligonucleotides (DNA or RNA), commonly referred to as aptamers, that specifically bind to a target ligand. Large, randomized oligonucleotide libraries are subjected to multiple iterative cycles of amplification, mutagenesis, binding, and selection.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1:
INTRODUCTION

All living organisms must possess internal sensory and regulatory mechanisms that allow them to coordinate their functions and respond to different stimuli. For instance, the ability to precisely regulate gene expression in response to different environmental conditions is critical for the adaptation and survival of any living cell (Machtel et al., 2016). Along with the well-known protein gene switches (i.e., transcription factors), bacteria commonly employ a much simpler regulatory mechanism embedded in some of their mRNAs called riboswitches. These riboswitches enable cells to respond to specific metabolites and control expression of appropriate genes. Riboswitches have recently attracted the attention of researchers because they have a number of favorable characteristics and offer some advantages over their protein counterparts. For instance, diverse RNA aptamers sequences obtained through *in vitro* selection (Ellington and Szostak, 1990, Stoltenburg et al., 2007, Codrea et al., 2010, Szeto et al., 2013, Zhuo et al., 2017, Ogawa and Itoh, 2020) can be potentially adapted into riboswitches, allowing cells to be engineered to respond to a plethora of molecules, both natural and synthetic.

Because functional riboswitches require active transcription and translation, the process of engineering and optimization of synthetic riboswitches has generally involved the use of living cells. However, such conventional approaches occasionally fail due to the intrinsic biological constraints of the *in vivo* systems. For example, optimization of riboswitches in bacteria is not always possible because many desirable ligands may be toxic to the cells, impermeable through the cell wall or membrane, unstable inside living cells, rapidly metabolized, or their optimal binding conditions do not match the intracellular environment. Therefore, although numerous RNA aptamers have been obtained by *in vitro* selection, many of them cannot be adapted to riboswitches that function *in vivo*.

Bottom-up synthetic biology is emerging as a new frontier in synthetic biology, aiming to build biochemical or genetic systems that display complex functions without using living cells (cell-free systems) (Rollin et al., 2013, Lu, 2017, Laohakunakorn et al., 2020). Such cell-free systems exhibit fewer experimental and biological constraints compared to living cells allowing the implementation of functions or properties that would otherwise be difficult or impossible using living cells, for example, using or producing toxic secondary metabolites, or building hybrid systems with nonbiological components. Cell-free protein synthesis (CFPS) systems based on cell extracts or reconstituted protein translation machinery are key platforms in the emerging cell-free systems, as they can recreate the biological processes of transcription and translation, and thus gene expression. Consequently, riboswitches that function in CFPS systems can substantially extend the capabilities of cell-free systems to detect and respond to chemical signals, including those not compatible with living cells.

Thus far, most cell-free riboswitches that are not derived from natural or synthetic bacterial counterparts have been engineered by trial-and-error or rational design followed by extensive

screening of riboswitch variants. This approach, while sometimes effective, is not very efficient as the throughput is limited and labor intensive. The lack of efficient design strategies has, at least partially, limited the development of cell-free synthetic riboswitches compared to those that function in living cells. Since most successful design strategies for synthetic riboswitches in living cells involve some form of high-throughput screening or selection, a similar strategy for cell-free riboswitches should greatly facilitate the design and engineering process. It is worth noting that the maximum library size that can be screened using *in vivo* systems is limited by the transformation (or transfection) efficiencies, making the screening capacity several orders of magnitude smaller than the capacity of producing such libraries by chemical or biochemical synthesis of DNA or RNA. CFPS systems offer a potential platform for high-throughput screening of riboswitches on par with other *in vitro* selection methods (Mishler and Gallivan, 2014) such as SELEX used for the screening of huge RNA libraries for aptamer discovery. The idea of screening riboswitches directly in cell-free systems has been suggested (Mishler and Gallivan, 2014), but no such effort has been reported to date.

To address this challenge, I developed a novel high-throughput screening strategy to engineer riboswitches directly in a CFPS system. By combining single-template DNA amplification on magnetic microbeads, cell-free gene expression in a CFPS system encapsulated in water-in-oil microdroplets, and microfluidic fluorescence-activated droplet sorting (FADS), it was possible to rapidly enrich functional cell-free riboswitches from over 4000 riboswitch variants. Subsequently, cell-free riboswitches obtained through this method were employed to build simple chemical communication circuits between cell-sized droplets, demonstrating that such riboswitches can be readily available for use in cell-free application. I believe that this strategy substantially expands the possibilities for increasing the variety of cell-free riboswitches and their performance, which in turn will contribute to the efforts of building cell-free systems and artificial cells with more complex functions.

The following chapters of this thesis are organized as follows. Chapter 2 provides an overview on the relevant topics including natural riboswitches and their mechanisms, synthetic riboswitches in living cells and their design strategies, a brief background on CFPS and cell-free systems, and a comprehensive review on cell-free riboswitch research to date. In Chapter 3, I present a novel high-throughput screening method for cell-free riboswitches, and the validation process of the method through mock sorting experiments, proving its capacity for sorting riboswitches in both ON and OFF states. In addition, I describe some of the issues addressed during its development, as well as the current limitations and areas for improvement. Chapter 4 contains successful examples of the screening method used to obtain riboswitches that respond to histamine and ciprofloxacin, compounds that are normally not compatible with conventional bacterial screening methods. In Chapter 5, I demonstrate the utility of these newly developed cell-free riboswitches as genetic devices for driving chemical communication between microdroplets. Finally, brief concluding remarks are provided in Chapter 6.

CHAPTER 2:
RIBOSWITCHES IN LIVING CELLS AND CELL-FREE SYSTEMS

In this chapter, an extensive literature review on the state of the art of synthetic riboswitches is presented with particular emphasis on those employed in cell-free systems.

Parts of this chapter have been published as a review article (Tabuchi and Yokobayashi, 2021): "Cell-free riboswitches." *RSC Chem. Biol.* **2**(5): 1430-1440. DOI: 10.1039/d1cb00138h

2.1. NATURAL RIBOSWITCHES

Riboswitches are among the oldest and simplest gene regulatory mechanisms in nature (Vitreschak et al., 2004). They are composed of RNA that can intervene at different stages of the gene expression process such as transcription, translation, mRNA processing, and RNA splicing. Some researchers have even hypothesized that riboswitches may be remnants of ancient metabolic regulatory mechanisms in the primordial life before the emergence of protein-based regulatory factors (Nahvi et al., 2002, Breaker, 2012).

Riboswitches have been found mainly across the bacteria domain (with a few instances in archaea and eukaryotes) (Barrick and Breaker, 2007, Kazanov et al., 2007, McCown et al., 2017, Pavlova et al., 2019) and represent a common widespread system for sensing intracellular metabolites (Nahvi et al., 2002, Winkler and Breaker, 2005). In numerous microorganisms, various metabolic pathways are regulated by riboswitches, including biosynthesis of fundamental metabolites such as coenzymes, nucleobases, ions, and amino acids (Nahvi et al., 2002, Vitreschak et al., 2004, Winkler and Breaker, 2005, Kazanov et al., 2007, McCown et al., 2017, Lotz and Suess, 2018, Pavlova et al., 2019). The riboswitches allow the cells to maintain appropriate levels of these metabolites by modulating the expression of key proteins involved in their biosynthesis or catabolism pathways. Interestingly, the thiamine pyrophosphate (TPP) riboswitch class is the only natural metabolite-binding riboswitch class found outside of bacteria; TPP riboswitches have been found in euryarchaeota, fungi, and plantae species, usually associated with alternative RNA splicing (Kubodera et al., 2003, Sudarsan et al., 2003, Barrick and Breaker, 2007, Cheah et al., 2007, Li and Breaker, 2013).

Canonical bacterial riboswitches are *cis*-acting genetic elements located in untranslated regions (UTR) of some messenger RNAs (mRNA) that specifically binds to certain metabolites or small molecules (also called effector molecules or ligands) and modulate the expression of the gene(s) encoded in the same mRNA (Nahvi et al., 2002, Sudarsan et al., 2003, Nudler and Mironov, 2004, Vitreschak et al., 2004, Tucker and Breaker, 2005, Winkler and Breaker, 2005, Garst et al., 2011, Lotz and Suess, 2018, Pavlova et al., 2019). Riboswitches can adopt alternative mutually exclusive conformations, switching between activating (ON-state) and repressing (OFF-state) conformations in response to the presence or absence of the effector

molecule (Vitreschak et al., 2004, Mandal and Breaker, 2004, Winkler and Breaker, 2005, Batey, 2006, Garst et al., 2011, Porter et al., 2017). The structure adopted by the riboswitch subsequently dictates the outcome of the expression of the associated gene(s).
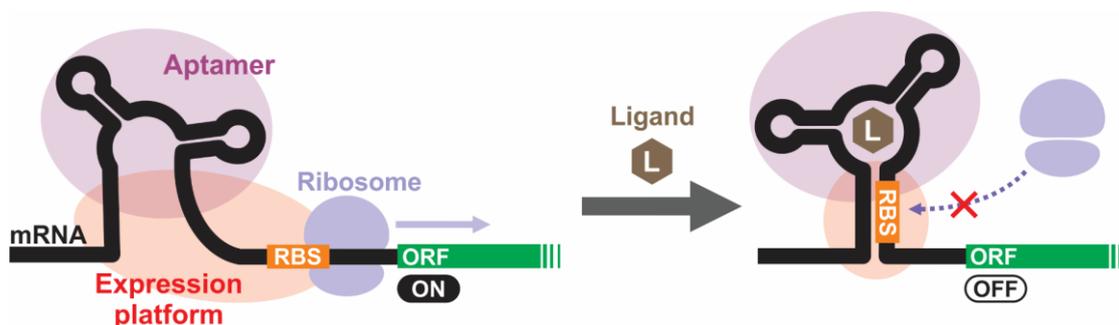
Figure 1. Schematic of a canonical translationally regulated riboswitch composed of an aptamer domain and an expression platform, located at the 5' end of the mRNA controlling the translation of a downstream open reading frame (ORF).

Conceptually, the structure of a canonical riboswitch consists of two elements (**Figure 1**): (i) an **aptamer domain**, which directly and specifically binds to a small molecule metabolite (ligand) and can discriminate closely related analogs (Breaker, 2008); and (ii) an **expression platform**, which mediates a local structural change (e.g., alternative base-pairing, secondary structure changes, self-cleaving, etc.) of the mRNA upon aptamer–ligand interaction, leading to a change in the gene expression level of a downstream open reading frame (ORF). Occasionally, the expression platform also involves part of the aptamer domain.

2.1.1. MECHANISMS OF GENE REGULATION

There are two major mechanisms by which natural riboswitches modulate gene expression (**Figure 2**): (i) by modulating formation of a transcription terminator structure upstream of the start codon; or (ii) by modulating the accessibility of the ribosome binding site (RBS) to the ribosomes thus controlling the translation initiation. In nature it is not uncommon to find riboswitches, such as the TPP-responsive *thiC* riboswitch in *Escherichia coli* (Chauvier et al., 2017), that use a combination of both transcriptional and translational regulation mechanisms. Barrick and Breaker (Barrick and Breaker, 2007) grouped natural riboswitches into 5 categories based on how these two mechanisms act on the gene regulation: (i) transcription attenuation (formation of a termination hairpin is controlled by the aptamer formation); (ii) dual transcription and translation attenuation (the terminator hairpin also acts as the RBS sequestering structure); (iii) translation attenuation (the RBS sequestering structure is separated from the aptamer); (iv) direct translation attenuation (the aptamer structure itself sequesters the RBS); and (v) antisense regulation (the riboswitch is located downstream the ORF). The antisense regulation mechanism, however, is based only on bioinformatic analysis with no experimental evidence.

Mapping riboswitch mechanism predictions onto phylogenetic trees revealed that most Gram-negative bacteria, including proteobacteria such *E. coli*, seem to prefer translationally regulated riboswitches, whereas most Gram-positive bacteria favor transcriptionally regulated ones (Winkler and Breaker, 2003, Nudler and Mironov, 2004, Soukup and Soukup, 2004, Barrick and Breaker, 2007, Pavlova et al., 2019). It also seems that riboswitches are overall more common in Firmicutes (low C+G Gram-positive bacteria) than any other group (Barrick and Breaker, 2007). Coincidentally, genes of the Gram-positive bacteria are more frequently organized in larger polycistronic operons which are probably regulated more efficiently by

transcriptionally regulated riboswitches (Winkler and Breaker, 2003, Nudler and Mironov, 2004).
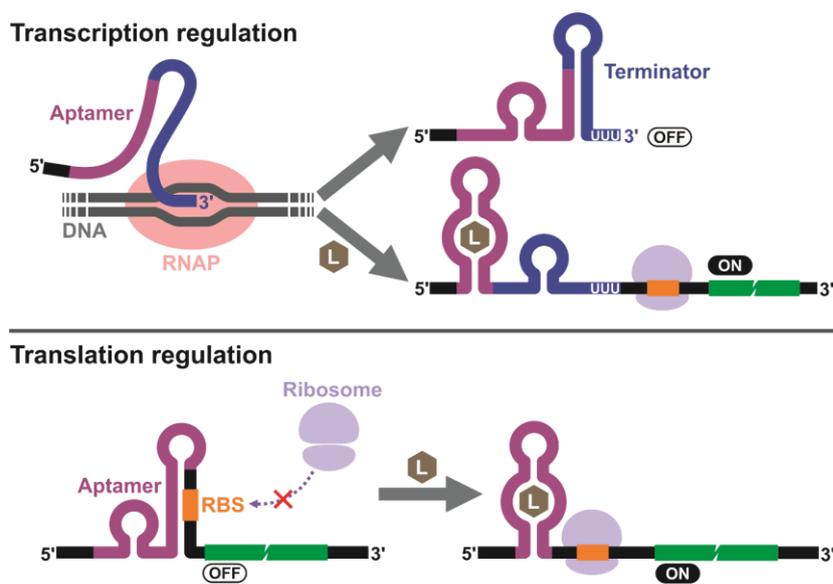
Figure 2. Two main mechanisms for riboswitch gene regulation. During transcription, riboswitch can modulate the formation of a transcription termination hairpin which releases the RNA polymerase (RNAP) and truncate the transcript. During translation initiation, riboswitches can modulate the accessibility of the ribosome binding site (RBS) blocking the ribosome from loading onto the mRNA or scanning the RBS and/or start codon. L: ligand, RNAP: RNA polymerase.

A third but less common mechanism is associated with self-cleaving ribozymes (RNA sequences capable of cleaving phosphodiester bonds). One example is the GlcN6P-responsive *glmS* riboswitch of *Bacillus subtilis* (Winkler et al., 2004, Collins et al., 2007) located in the 5' UTR of the mRNA encoding glutamine-fructose-6-phosphate aminotransferase (*glmS* gene). Glucosamine-6-phosphate (GlcN6P) acts as a cofactor to induce self-cleavage of the mRNA at the ribozyme active site. RNase J1 recognizes the resulting 5'-OH terminus and degrades the cleaved mRNA, resulting in reduced expression of the *glmS* gene (Collins et al., 2007). In contrast to other synthetic ligand-dependent ribozymes (also called aptazymes), in which the ligand acts as an allosteric effector, the binding of GlcN6P does not induce a conformational change in the ribozyme active site, but instead participates in the catalytic reaction as an essential cofactor (Klein and Ferre-D'Amare, 2006). Another example is the c-di-GMP-II allosteric riboswitch associated with a group I self-splicing ribozyme to control the expression of the open reading frame *CD3246* (putative virulence gene) of *Clostridium difficile* (Chen et al., 2011). In the presence of cyclic dimeric guanosine monophosphate (c-di-GMP), the ribozyme generates a spliced transcript with a newly accessible RBS for the downstream open reading frame. In the absence of c-di-GMP, however, the riboswitch mediates an alternative GTP attack that predominantly generates truncated RNAs that are not translated (Chen et al., 2011).

Another type of naturally occurring riboswitches are the T-boxes. This kind of riboswitches, however, are very different from the ones mentioned above as their effector ligand is not a small molecule metabolite but a transfer RNA (tRNA), and the mechanism of action involve Watson−Crick base pairing interactions. In short, the T-box leader sequences binds directly to specific tRNAs; when such tRNAs are aminoacylated, they act as an anti-terminator and controls bacterial transcription (Green et al., 2010, Raina and Ibba, 2014).

2.1.2. MECHANISMS OF RIBOSWITCH ACTIVATION

The activity of the riboswitches relies on the formation of the proper secondary structures within the mRNA in relation to certain concentrations of a specific ligand. Discrepancies

between the ligand concentrations required for activating riboswitches *in vivo* and the concentrations required for aptamer–ligand binding *in vitro* (Gilbert et al., 2006, Forster et al., 2012), together with studies on transcriptionally regulated riboswitches (Wickiser et al., 2005a, Wickiser et al., 2005b, Lang et al., 2007, Lemay et al., 2011, Frieda and Block, 2012) and riboswitch folding simulations (Lutz et al., 2014), have suggested the importance of co-transcriptional folding of the aptamer structure for the correct function of some riboswitches.
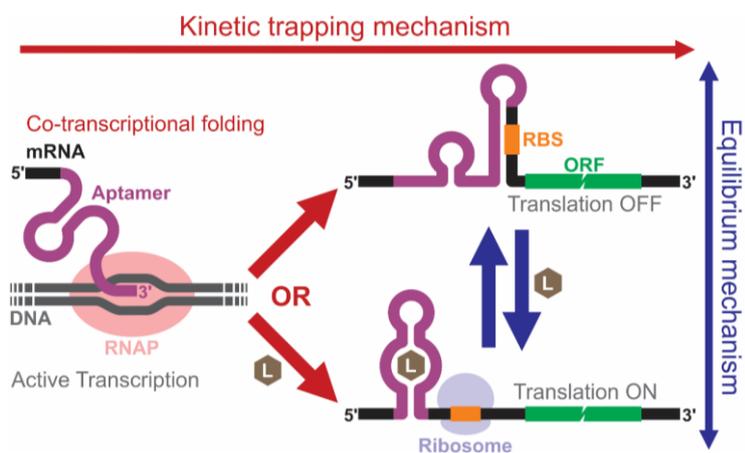
Figure 3. Mechanistic models of riboswitches. In the kinetic trapping mechanism, presence of the ligand during transcription is necessary for the aptamer/riboswitch structure to form. In the thermodynamic equilibrium mechanism, the active and inactive riboswitch structures are in dynamic equilibrium; addition of the ligand can shift the equilibrium and affect protein expression even after transcription. L: ligand, RBS: ribosome binding site, RNAP: RNA polymerase. Adapted from (Tabuchi and Yokobayashi, 2021) with permission of the authors.

Two main mechanisms govern the activation of riboswitches (Coppins et al., 2007, Lemay et al., 2011, Quarta et al., 2012, Mishler and Gallivan, 2014, Jones and Ferre-D'Amare, 2017, Chushak et al., 2021) (**Figure 3**): (i) a **co-transcriptional kinetic trapping mechanism**, in which active transcription is required for riboswitch activation because the local concentration of the ligand at the moment of the initial folding of the aptamer during transcription determines the final structure adopted, as the alteration of the stability of some transition structures can force ("trap") the mRNA into an alternative conformation; and (ii) a **post-transcriptional thermodynamic equilibrium mechanism**, which does not require active transcription since the different RNA conformations are in a thermodynamic equilibrium, and changes in the ligand concentration shift that equilibrium to a particular conformation. A series of parameters such as the transcription rate of the RNA polymerase, the stability of the RNA structures formed, the kinetics of the aptamer–ligand interaction, etc. determine which mechanism dominates the activation of a particular riboswitch under certain reaction conditions.

2.2.   SYNTHETIC RIBOSWITCHES

Soon after the discovery of natural riboswitches, researchers realized their practical potential. The possibility of chemical regulation of gene expression without direct involvement of proteins (i.e., transcription factors) motivated many researchers to develop synthetic riboswitches (Koizumi et al., 1999, Hermann and Patel, 2000, Nahvi et al., 2002). Synthetic riboswitches, like their natural counterparts, regulate gene expression in response to small molecules via RNA aptamers; however, synthetic riboswitches can employ RNA aptamers acquired by *in vitro* selection (SELEX) allowing them to respond to natural and non-natural molecules, thus greatly expanding the repertoire of chemical inputs.

The modularity, plasticity, versatility, and engineerability of synthetic riboswitches, together with a relatively simple architecture and regulation mechanisms, make them very attractive tools for synthetic biology and other applications (Groher and Suess, 2014, Machtel et al., 2016,

Rossmanith and Narberhaus, 2017, Nshogozabahizi et al., 2019). Riboswitches also offer distinct advantages over protein transcription factors (TF) as genetic devices for chemical regulation of gene expression:

- There is no need to express or add large TF or any other regulatory protein, which translates to a smaller genetic footprint and less metabolic burden on the host.
- Riboswitches (specifically their aptamer domain) can be engineered to recognize various molecules, both natural and synthetic. In addition, compared to TF, aptamer sequences are relatively easy to obtain through techniques such as SELEX.
- Due to their relatively simple architecture and regulatory mechanisms, riboswitches are easier to engineer than TF. In addition, riboswitches (RNA) are made up of only 4 different ribonucleotides, compared to >20 amino acid residues in the TF.
- Gene regulation is more direct and has faster response times since synthesis of additional regulatory proteins is not necessary.

Synthetic riboswitches can be engineered by mutating natural riboswitches to obtain new functionality, or by modifying existing aptamer sequences and adapting them to a different expression platform (Martini and Mansy, 2014, Etzel and Morl, 2017). While many synthetic riboswitches have been designed mimicking natural riboswitch mechanisms, synthetic riboswitches based on alternative mechanisms of action not found in nature have also been engineered (Tucker and Breaker, 2005, Wieland and Hartig, 2008, Suess and Weigand, 2008, Wieland et al., 2009, Nomura et al., 2013, Scott et al., 2013, Ketzer et al., 2014, Takahashi and Yokobayashi, 2019, Mustafina et al., 2020, Ender et al., 2021, Tickner and Farzan, 2021). One example of synthetic riboswitches based on non-natural mechanisms are the ones that employ aptazymes. Such aptazymes combine an RNA aptamer and a self-cleaving ribozyme so that the catalytic activity of the ribozyme is regulated by the aptamer–ligand interaction (Muller et al., 2016). This allosteric regulation allows the riboswitch to control gene expression by (i) self-cleaving and freeing a sequestered RBS (ON-state) (Tucker and Breaker, 2005, Wieland and Hartig, 2008, Wieland et al., 2009) or by (ii) cleaving the mRNA on critical regions and inducing the degradation of the transcript, effectively interrupting or inhibiting translation (OFF-state) (Nomura et al., 2013, Scott et al., 2013, Ketzer et al., 2014). In bacteria, there are no natural analogs for this type of riboswitch. It should be noted that the *glmS* riboswitch discussed before (see **2.1.1. Mechanisms of gene regulation**, page **4**) is not allosterically regulated by the ligand, and that the c-di-GMP-II riboswitch in *C. difficile* does not include the ribozyme domain but it is associated to it.

While synthetic aptamers and aptazymes can be obtained and engineered independently from the rest of the riboswitch through *in vitro* evolution techniques (Tuerk and Gold, 1990, Stoltenburg et al., 2007, Szeto et al., 2013, Blind and Blank, 2015), or screening and selection methods *in vitro* and *in vivo* (Link et al., 2007, Chen et al., 2009, Wittmann and Suess, 2011, Goler et al., 2014, Kobori et al., 2015, Kobori and Yokobayashi, 2016, Nomura et al., 2017, Endoh et al., 2019, Yokobayashi, 2019a), creating new riboswitches is not as trivial as arbitrarily fusing a synthetic aptamer (or an aptazyme) to an expression platform (Wachsmuth et al., 2013). Each domain needs to be engineered in the context of the other. For example, to make an aptamer and expression platform work together, it is necessary to either (i) design an overlapping region between both domains that allows them to switch conformations (Garst et al., 2011, Barsacchi et al., 2016) or (ii) have no overlap between domains, but instead include a communication module (a functional sequence between the aptamer and the expression platform) to relay structural changes from one domain to the other (Ceres et al., 2013a, Ceres et al., 2013b). Some other important parameters to consider are the relative stability of the

alternative RNA structures formed, the kinetics of the aptamer–ligand interaction, the kinetics of the co-transcriptional folding of the mRNA, etc.

2.3.   STRATEGIES FOR ENGINEERING SYNTHETIC RIBOSWITCHES

In the following section, recent progress in synthetic riboswitch engineering is reviewed according to design and screening strategies.

2.3.1.   *IN VIVO* SCREENING

In vivo screening of riboswitches starts with a library of riboswitch mutants that control the expression of a reporter gene (e.g., GFP, luciferase, β-galactosidase, etc.), which are constructed on a plasmid. Typically, the screening process is performed in two steps. First, single colonies are picked from agar plates based on the expression of the reporter gene, and then they are sub-cultured in microtiter plates, usually in 96-well or 384-well format, in the presence and absence of the aptamer ligand. Subsequently, reporter gene expression of the individual mutants are measured to identify those that respond to the chemical input.

The Gallivan group used this approach to screen theophylline-responsive ON riboswitches controlling β-galactosidase expression in *E. coli* (Desai and Gallivan, 2004, Lynch et al., 2007). They identified candidate riboswitches by plating the transformed cells onto selective media containing X-gal. In order to minimize the background levels of gene expression (leakage), the whitest colonies were picked from the initial screening plates that did not contain theophylline. Candidates were sub-cultured in 96-well microtiter plates and β-galactosidase activity was measured for each individual.

The Hartig group employed a similar strategy to screen aptazyme-based riboswitches in *E. coli*. The riboswitch libraries were designed by inserting the theophylline aptamer into the *Schistosoma mansoni* hammerhead ribozyme connected by a randomized "communication module" or "connector sequence" (Wieland and Hartig, 2008, Rehm and Hartig, 2014). They also inserted the theophylline and/or thiamine pyrophosphate (TPP) aptamers into a twister ribozyme (Felletti et al., 2016). The riboswitch sequences were engineered to sequester the RBS controlling the expression of eGFP. After the *in vivo* screening in *E. coli*, they identified clones that displayed strong induction of eGFP in the presence of the ligand. They also studied two-input riboswitches by introducing two aptamer domains regulating the same ribozyme (Felletti et al., 2016).

Kirchner et al. envisioned a method for high-throughput screening of riboswitch ligands *in vivo* that can potentially also be used for studying and screening OFF-riboswitches (Kirchner et al., 2017). Their system included a reverse reporter gene setup consisting of a set of two gene locus, one containing the guanine-responsive *xpt* riboswitch (OFF switch) from *Bacillus anthracis* controlling the expression of BlaI repressor protein of *B. licheniformis*, and the other containing a BlaI-repressible promoter and the reporter gene (luciferase or β-galactosidase). Their screening was performed using *B. subtilis* cultured in a 384 well microtiter plates which allowed the researchers to identify several hit compounds from a library of 1280 molecules.

The Kelley-Loughnane group developed a new riboswitch-based dual-color sensor (Harbaugh et al., 2017) and employed it to screen synthetic riboswitches in *E. coli* (Harbaugh et al., 2018). Their dual-color reporter system consisted of two parts: a riboswitch controlling the FimE recombinase from the fimbriae phase variation system of *E. coli*, and the corresponding invertible DNA segment (*fimS*) containing a constitutive promoter located in between the two

fluorescence protein genes (*gfpa1* and *mkate2*). Depending on the presence of the ligand (TNT) and the activity of the riboswitch, the expression of FimE recombinase causes the *fimS* segment, and the promoter within, to become inverted, switching the reporter gene expression from GFPa1 (green) to mKate2 (red). The screening and selection of candidates was performed on microtiter plates based on the green/red fluorescence ratio.

The Suess group employed *Saccharomyces cerevisiae* to develop a two-stage strategy for screening neomycin-responsive OFF-riboswitches (Weigand et al., 2008). First, an aptamer pool was subjected to several cycles of *in vitro* selection to enrich aptamers that binds to neomycin B. The enriched library was then subcloned into an expression vector upstream a reporter gene (*gfp*) and transformed into yeast. Finally, the screening *in vivo* was performed first on agar plates and then in microtiter plates, in a similar way as described in the first paragraph. Goler et al. employed a similar two-stage screening strategy but using theophylline-sensitive hammerhead aptazyme-based ON-riboswitches in *E. coli* (Goler et al., 2014). Likewise, aptazyme candidates were enriched first via several rounds of a PAGE-based *in vitro* selection. Then, the selected candidates were subcloned upstream a red fluorescent protein (RFP) reporter gene and transformed into *E. coli* for the *in vivo* screening process.

In vivo screening by reporter gene assay of riboswitch mutants tends to be very accurate because each mutant is assayed individually. Although the process can be partially automated by using robots for colony picking and subculturing, the throughput is usually limited by the time-consuming and labor-intensive nature of the assay. With recent development of microfluidic devices and flow cytometry, specifically fluorescence-activated cell sorting (FACS) and droplet sorting (FADS), high-throughput screening of individual cells can now be achieved (Marcoux et al., 2011, Kaminski et al., 2016, Abatemarco et al., 2017, Lu et al., 2017, Terekhov et al., 2017). FACS-based strategies have been successfully applied with *E. coli* (Fowler et al., 2008, Lynch and Gallivan, 2009) and yeast cells (Liang et al., 2012) to screen for functional riboswitches in a high-throughput manner.

Recently, cDNA/RNA sequencing has also been employed for high-throughput screening of active aptazymes-based riboswitches in mammalian cells (Strobel et al., 2020, Xiang et al., 2019). This approach requires extensive computational analysis and is susceptible to bias introduced during every step of the sample preparation and sequencing process, therefore can only provide qualitative estimations on the activity of potential candidates that need to be validated individually. However, no other high-throughput screening method has been developed for riboswitches that function in mammalian cells.

2.3.2. *In vivo* genetic selection

For *in vivo* selection of riboswitches, a library of riboswitch mutants controlling the expression of a selection marker is constructed. The selection marker can be a gene that confers a function such as antibiotic resistance or other selectable characteristics (motility, auxotrophy, etc.). The pool of mutants is then subjected to a series of positive and negative selection rounds in the presence or absence of the chemical input. This results in the enrichment of riboswitch mutants that respond to the aptamer ligand. After several rounds of selection, individual clones are evaluated and sequenced to confirm the riboswitch function.

Genetic selection allows a larger number of riboswitch mutants (typically limited by transformation efficiency) to be examined compared to *in vivo* screening. Because selection rounds can be performed in a single culture, it is less costly and labor intensive than screening

individual mutants. However, it is difficult to adjust the selection stringency, and no information about the unselected variants is obtained.

The Yokobayashi group developed a strategy for riboswitch selection in *E. coli* (Muranaka et al., 2009b) by linking the TPP-responsive *thiM* riboswitch to the *tetA-GFPuv* fusion gene while randomizing the spacer sequence between the aptamer and the Shine-Dalgarno (SD) sequence. The *tetA* gene acted as both a positive and negative selection marker, as it confers resistance to tetracycline (ON selection) but makes the bacteria more susceptible to toxic ions such as Ni^{2+} (OFF selection). After a few rounds of dual-selection, colonies were isolated and GFPuv fluorescence was used to measure the gene expression level. The enrichment efficiency of the single dual-selection round was >7000 and they were able to obtain riboswitches sequences with an ON/OFF ratio up to 58.

Another interesting approach was reported by Topp and Gallivan (Topp and Gallivan, 2008) who used *cheZ* gene as the selection marker. Because protein phosphatase CheZ is critical for chemotaxis in *E. coli*, they used bacterial motility to select cells for ON or OFF outputs of the riboswitches. The researchers fused a theophylline riboswitch library to the *cheZ* gene, and the *E. coli* mutants were seeded at the center of an agar plate with or without theophylline. The selection was accomplished by recovering the cells near the center (OFF selection) or from the outskirts of the motility halo (ON selection). After isolating promising candidates, they found riboswitches with ON/OFF ratios greater than 15.

2.3.3. RATIONAL DESIGN OF RIBOSWITCHES

Apart from the dynamic nature of riboswitches that involves multiple secondary structures, their relative thermodynamic stabilities, and more subtle factors such as transcription speed and pausing, aptamer–ligand binding kinetics, and co-transcriptional folding have been implicated in riboswitch function. This increases the difficulty to accurately estimate *a priori* some riboswitch parameters, and thus rational design of synthetic riboswitches remains challenging. Some notable efforts toward rational design of riboswitches are reviewed here.

The Mörl group proposed a rational riboswitch design approach (Wachsmuth et al., 2013, Domin et al., 2017) as an alternative to selection or screening of riboswitch mutants. Their design algorithm starts with a known aptamer sequence and a secondary structure, and then designs the rest of the transcriptionally regulated riboswitch (spacer and terminator). A large set of candidates are then evaluated against the aptamer by analyzing the folding paths and secondary structure stability *in silico* using the RNAfold tool from ViennaRNA package (Zuker and Stiegler, 1981, McCaskill, 1990, Hofacker and Stadler, 2006, Gruber et al., 2008, Lorenz et al., 2011). A series of filters are applied based on biological, structural, and physicochemical constraints, and candidates are discarded if determined unviable or nonfunctional. After several folding simulations, some parameters (such as folding energy, minimum free energy, and z-score) are calculated to estimate the stability required by the ligand-bound aptamer to prevent the formation of the terminator.

Similarly, the Salis group built a biophysical model for rational design of translationally regulated riboswitches (Espah Borujeni et al., 2016). Their model considers the folding trajectory of the mRNA while being transcribed, and its RNA–RNA, RNA–ribosome, and RNA–ligand interactions. The model also accounts for the concentrations and molar volumes of the mRNA and the ligand to estimate the translation initiation rate and the riboswitch activation ratio which ultimately controls the expression level. The activation ratios are calculated from the difference of the ribosome's binding free energy, in the presence and

absence of the ligand, of the different conformations of the mRNA. These calculations also include energy calculations and penalties for secondary structures of the mRNA, RBS optimality, tRNA$_{Met}$ binding to the start codon, etc.

The Batey group showed that riboswitches can be engineered modularly following a set of design considerations, such as non-overlapping domains, alternative secondary structures, RNA sequence, and ligand binding requirements. They developed a series of synthetic riboswitches by combining various aptamer domains with different transcription termination expression platforms (Ceres et al., 2013a, Ceres et al., 2013b). In this case, their analysis and design strategy was not based on *in silico* simulations, but instead on secondary structure and kinetic information of the aptamers and expression platforms that were derived from previous reports and *in vitro* assays.

It is worth noting that rational design is not limited to transcriptionally regulated riboswitches, but it can be applied to the design of other types of riboswitches since the activity of all riboswitches follows an energy function of the mRNA folding and its interaction with the different components of the reaction (ligand, ribosomes, etc.). For example, Zhong et al. used an energy-like function of the sequence of the communication module to rationally design aptazyme-based riboswitches in mammalian cells by predicting the dynamic range of the aptazyme activity when the ligand was present (Zhong et al., 2016).

2.4. CELL-FREE SYSTEMS AND RIBOSWITCH RESEARCH

Cell-free systems includes all kinds of *in vitro* tools that allow the study of complex biological processes without the use of intact cells (Swartz, 2006), thus reducing the complexity and interactions typically present when working with whole living cells. Interestingly, cell-free systems have been around for more than a century, since Eduard Buchner demonstrated that sugars can be fermented into alcohol using nonliving yeast cell extracts, for which he was awarded the Nobel Prize in Chemistry in 1907. In the past 10 to 20 years, recent advances in terms of efficiency, versatility, and production cost of cell-free technologies have resulted in a rapidly increasing interest—especially from the synthetic biology and biotechnology communities—, as well as a growing number of applications such as in artificial cells, metabolic engineering, prototyping, biomanufacturing, biosensors, and lab-on-a-chip technologies.

Cell-free systems are promising tools for synthetic biology as they offer a unique platform for building biochemical and genetic systems that display complex functions without using living cells (Rollin et al., 2013, Lu, 2017, Laohakunakorn et al., 2020). Constructing dynamically regulated cell-free gene expression systems, for example, allows deeper understanding of biological systems as well as provide tools for more practical applications such as biosensors and biomanufacturing (Brookwell et al., 2021). However, current options for chemical regulation of gene expression in cell-free systems are limited. Very often, canonical transcription factors (TF), such as LacI–IPTG or LuxR–AHL systems, have been used as a generic inducible platform to activate cell-free protein synthesis (Sun et al., 2014, Schwarz-Schilling et al., 2016, Dupin and Simmel, 2019, Jeong et al., 2019). However, there are a number of limitations in using TF-based gene switches in cell-free systems. For example, the switch performance is dependent on the levels of the transcription factor and the DNA template, and expression of a transcription factor *in situ* can affect the response time and may impose additional metabolic burden to the system. It has been shown that the translational machinery of the cell-free system can become a bottleneck, generating time delays and reducing the output with each additional protein synthesis step in a genetic cascade (Noireaux et al., 2003).

Furthermore, despite recent advances, engineering TF to respond to new ligands remains challenging (Gutierrez-Gonzalez et al., 2019, Wan et al., 2019, Li et al., 2020, Ding et al., 2021). In that context, riboswitches have attracted increasing attention as an alternative strategy to protein factors as gene regulators and sensory devices, because they offer some advantages that may be useful and beneficial in certain situations, such as fewer off-target effects, faster response time, and reduced metabolic burden.

Because riboswitch function can only be properly assessed during active transcription and translation, engineering, screening, and optimization of riboswitches have historically involved the use of living cells (*in vivo* systems). Thus far, most synthetic riboswitches have been engineered to work in bacteria, yeast, mammalian, and other types of living cells (Groher and Suess, 2014, Etzel and Morl, 2017, Hallberg et al., 2017, Nshogozabahizi et al., 2019, Ge and Marchisio, 2021); however, traditional approaches occasionally fail due to the intrinsic biological constraints of the *in vivo* systems. For instance, optimization of synthetic riboswitches in bacteria is not always possible because many desirable ligands may be toxic to the cells, impermeable through the cell wall or membrane, rapidly metabolized or unstable inside living cells. Cell-free systems offers an alternative platform for engineering such riboswitches in a simplified environment with fewer biological constraints, while having more control over the experimental conditions and parameters such as composition of the system, pH, and temperature (Martini and Mansy, 2014, Khambhati et al., 2019).

When working with riboswitches, the biological process of interest is gene expression and protein synthesis; therefore, the cell-free system to be employed must be capable of replicating transcription and translation outside living cells (Martini and Mansy, 2011). Several different systems have already been developed under the general term cell-free protein synthesis (CFPS) systems, also called *in vitro* transcription-translation (IVTT) systems. CFPS systems can be categorized into two main types (Lu, 2017, Rollin et al., 2013):

- **Cell extract-based systems**, obtained by lysing and clarifying living cell cultures. They contain all components from inside the cell required for transcription, translation, protein folding, and energy regeneration (Perez et al., 2016), usually supplemented with additional essential metabolites such as NTPs, tRNAs, and amino acids. Although cell extracts from a variety of organisms exist—some of them being commercially available—, only three cell extract types have been successfully employed with riboswitches: *Escherichia coli* cell extracts (termed "S30 extracts"), wheat germ extract (WGE), and rabbit reticulocyte lysate (RRL).

- **Reconstituted systems**, made by individually synthesizing and purifying each component necessary for protein expression, and mixing them at well-defined concentrations. The most commonly used system is based on the PURE system (stands for Protein synthesis Using Recombinant Elements) developed by Shimizu et al. (Shimizu et al., 2001, Shimizu et al., 2005) which is commercially available under different brands and formulations: PURESYSTEM (PGI), PURE*frex* (GeneFrontier), PURExpress (NEB), etc. It consists of *E. coli*'s components for translation, aminoacylation, and energy regeneration, coupled to a viral T7 RNA polymerase (sometimes replaced with *E. coli*'s own RNA polymerase or other viral polymerases) for transcription.

Synthetic riboswitches have been engineered and studied in both prokaryotic and eukaryotic CFPS systems (**Table 1**, **Table 2**). Prokaryotic CFPS systems include conventional *E. coli* S30

extracts (Pratt, 1984, Nevin and Pratt, 1991, Swartz et al., 2004, Kwon and Jewett, 2015, Krinsky et al., 2016), and the reconstituted PURE systems (Shimizu et al., 2001, Shimizu et al., 2005). S30 extracts are relatively inexpensive and simple to produce, can be scaled up, and exhibit high protein yields (Zawada et al., 2011, Perez et al., 2016, Krinsky et al., 2016). They also contain numerous nonessential cellular components and may reflect the intracellular environment more accurately. On the other hand, PURE systems mainly consist of purified components from recombinant *E. coli*, as a result, they are much more expensive and laborious to produce, and are less scalable. However, they offer a more flexible, well-defined, and reproducible platform that allows more freedom for controlling biochemical parameters. They also contain low amounts of ribonucleases which can significantly influence riboswitch performance (Chushak et al., 2021).

Although a number of eukaryotic CFPS systems are currently available, nearly all of the eukaryotic cell-free riboswitches have been studied using wheat germ extract (WGE) (**Table 2**). In dry state, wheat germ embryos naturally contain all the components required for translation, ready to start protein synthesis as germination begins (Morita et al., 2003, Endo and Sawasaki, 2006). The Endo group developed a stable WGE system with high translation efficiency after removing endogenous translation inhibitors that limited the life span of the conventional WGEs (Madin et al., 2000). The current WGE systems offer the highest translation efficiency among eukaryotic CFPS systems (Harbers, 2014) and can produce high quality proteins in folded state (Endo and Sawasaki, 2006).

Cell extracts and reconstituted CFPS systems generally behave in similar ways, but they could show some differences in terms of transcription and translation kinetics of riboswitch-controlled genes. A recent study from Chushak et al. used a theophylline-responsive riboswitch to explore those processes in the context of an *E. coli* S30 cell extract and a PURE system (PURExpress, NEB) (Chushak et al., 2021). They learned, for example, that mRNA degraded over time in the cell extract, whereas mRNA levels in the PURE system increased continuously until reaching a plateau; interestingly, the protein (sfGFP) production was actually higher in the cell extract than in the PURE system. They also discovered that in the cell extract, the riboswitch exhibited a sharper response to the increasing ligand concentration compared to the PURE system, which they hypothesized to be caused by the presence of a variety of components in the cell extract that enhanced the mRNA folding and/or the ligand binding. Finally, they also showed that both systems are prone to become saturated, for example, an increase in DNA template above a certain concentration would result in lower protein yield likely due to the saturation of the translation machinery, and the transcription machinery competing for resources at the expense of the protein synthesis (Chushak et al., 2021).

2.4.1. PROKARYOTIC CELL-FREE RIBOSWITCHES

Thus far, prokaryotic CFPS systems (cell extracts and PURE systems alike) have been the preferred type of system for studying cell-free riboswitches. The earliest attempt to use riboswitches in a cell-free system was reported by Ogawa and Maeda in 2007 (Ogawa and Maeda, 2007), when they used the PURESYSTEM classic to develop aptazyme-based riboswitches, intended to be applied as label-free detector-free biosensors. A few years later, Yokobayashi group (Muranaka et al., 2009a) adapted TPP-responsive riboswitches, initially engineered in *E. coli* using an *in vivo* dual genetic selection, and tested them in a cell-free *E. coli* S30 extract. Afterward, in the span of about a decade, articles related to prokaryotic cell-free riboswitches have been published at an average rate of about one publication per year (**Table 1**).

Table 1. List of riboswitches in prokaryotic cell-free systems.
Modified and updated from (Tabuchi and Yokobayashi, 2021) with permission of the authors.

Mechanism	Riboswitch	Ligand	ORF	ON/OFF ratio (Ligand conc.)	CFPS*	Year	Ref.
Aptazyme-driven RBS release	duplex5-Theo	Theophylline	Luc LacZ	ON ~9× (2 mM) ON ~30× (2 mM)	PURE[a]	2007	1
	duplex5-cGMP	cGMP	Luc	ON ~10× (10 mM)	PURE[a]	2007	1
	DelC-M3	TPP	GFP	ON ~48× (300 μM)	PURE[c]	2012	2
	Theo/HHR	Theophylline	GFP	ON ~15× (2 mM)	PURE[c]	2012	2
Nonsense suppression	AST4m	Theophylline	Luc	ON 11.6× (1 mM)	PURE[b]	2008	3, 4
RBS sequestration	pLac-thiM#2	TPP	GFPuv mCherry Luc	ON 6× (1 mM) ON 14× (1 mM) ON 20× (1 mM)	CE[f]	2009	5
	pLac-tenA#59	TPP	Luc	ON 6× (1 mM)	CE[f]	2009	5
	add	Adenine	LacZ	ON (N/D)	CE[f]	2011	6
	pT7-theo	Theophylline	YPet	ON 6× (500 μM)	PURE[d]+AC[i,j]	2011	7, 8
	pTac-theo	Theophylline	YPet	ON 8× (500 μM)	CE[f]+AC[i,j]	2011	7
	JF001A	Theophylline	αHL	ON ~10× (1.5 mM)	PURE[d]+AC[j]	2014	9
	pTac-C	Theophylline	Luc	ON ~65× (2 mM)	CE[f]	2014	10
	TMR-10	TMR	Luc	ON ~16.5× (30 μM)	CE[f]	2016	11
	Dopa-5	Dopamine	Luc	ON 2× (1 mM)	CE[f]	2016	11
	T4-2	Thyroxine	Luc	ON 2.4× (150 μM)	CE[f]	2016	11
	Theo-αHL	Theophylline	αHL	ON (N/D)	CE[g]+AC[j]	2017	12
	H2	Histamine	mCherry αHL PLC	ON 30.7× (5 mM) ON (N/D) ON (N/D)	PURE[e]+AC[j]	2019	13
	Theo-GFP-MG	Theophylline	sfGFP	ON ~12× (10 mM)	CE[h], PURE[d]	2021	14
	HA-C1g-19	Histamine	eGFP	ON 20.6× (5 mM)	PURE[e]+AC[i]	2022	15
	HA-OFF4-a9-13	Histamine	eGFP	OFF 8.5× (2.5 mM)	PURE[e]	2022	15
	CFX-a1-sr5-19	Ciprofloxacin	eGFP	ON 9.3× (100 μM)	PURE[e]	2022	15
Transcription termination	pTac-ade/ydhL	Adenine	YPet	ON 1.7× (1 μM)	CE[f]+AC[i,j]	2011	7
	FRR/CrcB	Fluoride (F⁻)	sfGFP C23DO	ON ~20× (3.5 mM) ON (N/D)	CE[g]	2020	16

*CFPS (cell-free protein synthesis) systems. PURE: PURE systems, (a) PURESYSTEM classic II (Post-Genome Institute), (b) PURESYSTEM custom (Post-Genome Institute) without RF1, (c) modified Shimizu's PURE system with T7 RNAP replaced by T3 RNAP, (d) PURExpress (New England Biolabs), (e) PURE*frex* 1.0 (Gene Frontier). CE: cell extracts, (f) commercial S30 cell extract from *E. coli* B strain SL119 (Promega), (g) in-house prepared *E. coli* cell extract from Rosetta2 (DE3), (h) in-house prepared *E. coli* cell extract from BL21 Star (DE3). AC: artificial cells, (i) water-in-oil emulsions, (j) lipid vesicles/liposomes. cGMP: cyclic guanosine monophosphate, TMR: tetramethylrhodamine, TPP: thiamine pyrophosphate, αHL: α-hemolysin, C23DO: catechol-2,3-dioxygenase, GFP: green fluorescent protein (e-: enhanced, sf-: super-folder, -uv: ultraviolet), LacZ: β-galactosidase, Luc: firefly luciferase, ORF: open reading frame, PLC: phospholipase C, RNAP: RNA polymerase, YPet: yellow fluorescent protein for energy transfer.

References: 1 (Ogawa and Maeda, 2007); 2 (Kobori et al., 2012); 3 (Ogawa and Maeda, 2008b); 4 (Ogawa, 2021); 5 (Muranaka et al., 2009a); 6 (Lemay et al., 2011); 7 (Martini and Mansy, 2011); 8 (Martini and Mansy, 2014); 9 (Lentini et al., 2014); 10 (Mishler and Gallivan, 2014); 11 (Espah Borujeni et al., 2016); 12 (Adamala et al., 2017); 13 (Dwidar et al., 2019); 14 (Chushak et al., 2021); 15 (Tabuchi and Yokobayashi, 2022); 16 (Thavarajah et al., 2020).

2.4.1.1. APTAZYME-BASED RIBOSWITCHES

Aptazyme-based riboswitches employ interesting regulation mechanisms not found in nature. This mechanism involves the irreversible self-cleaving of a synthetic aptazyme (aptamer + ribozyme) upon ligand–aptamer interaction. The hammerhead ribozyme (HHR), initially used by the Breaker group (Soukup and Breaker, 1999, Soukup et al., 2000) and later modified by

the Hartig group (Wieland and Hartig, 2008, Wieland et al., 2009), has been engineered in several occasions as aptazyme-based riboswitches in cell-free systems.

Ogawa and Maeda (Ogawa and Maeda, 2007) adapted a theophylline-activated HHR in a PURE system so that the ribosome binding site (RBS) of the mRNA was sequestered by a complementary sequence upstream of the ribozyme (anti-RBS). Irreversible self-cleavage of the ribozyme in the presence of theophylline releases the RBS, allowing protein translation to be activated (**Figure 4A**). They also constructed a riboswitch that responds to cyclic guanosine monophosphate (cGMP) based on the same strategy (Ogawa and Maeda, 2007). The same theophylline riboswitch and another riboswitch based on a TPP-activated aptazyme (Wieland et al., 2009, Wieland and Hartig, 2008) (**Figure 4B**) were further analyzed by Kobori et al. in a modified PURE system for the purpose of optimizing riboswitch performance based on kinetic modeling (Kobori et al., 2012). It should also be noted that synthetic riboswitches developed based on the same design strategy have been shown to function in *E. coli* by Ogawa and Maeda (Ogawa and Maeda, 2008a), and the Hartig group (Wieland and Hartig, 2008, Wieland et al., 2009).

In another study, Ogawa and Maeda tethered the theophylline-responsive aptazyme to the 5' end of a nonsense-suppressor tRNA (sup-tRNA) to regulate translation read-through of a gene that contains an amber stop codon in a customized PURE system (Ogawa and Maeda, 2008b). In this strategy, the aptazyme activity mimics the canonical 5' terminus cleavage mediated by RNase P *in vivo* which is absent in the PURE system (**Figure 4C**).

2.4.1.2. TRANSLATIONALLY REGULATED RIBOSWITCHES

The majority of prokaryotic cell-free riboswitches operate by regulating the translation efficiency of the associated mRNA in response to aptamer–ligand binding. While the aptazyme-based riboswitches achieve translational regulation by irreversible self-cleavage that physically separates the RBS from the anti-RBS sequence, translationally regulated riboswitches modulate the local structure near the RBS through a structural change triggered by aptamer–ligand interaction (**Figure 4D**). A series of cell-free riboswitches have been designed inspired by these natural prokaryotic riboswitches and comprise the majority of cell-free prokaryotic riboswitches reported to date (**Table 1**) (Muranaka et al., 2009a, Martini and Mansy, 2011, Lentini et al., 2014, Mishler and Gallivan, 2014, Espah Borujeni et al., 2016, Adamala et al., 2017, Dwidar et al., 2019, Drachuk et al., 2020, Chushak et al., 2021).

One of the first reports belongs to Muranaka et al., who used an *E. coli* S30 extract system to characterize several TPP-responsive synthetic riboswitches that were originally engineered in *E. coli*, and observed comparable riboswitch performance in the cell-free system (Muranaka et al., 2009a). Similarly, a good number of cell-free riboswitches of this type were originally engineered in *E. coli* and simply adapted to or further optimized in cell-free systems (Martini and Mansy, 2011, Mishler and Gallivan, 2014) while just a few cell-free riboswitches have been directly engineered in prokaryotic cell-free systems (Espah Borujeni et al., 2016, Dwidar et al., 2019).

2.4.1.3. TRANSCRIPTIONALLY REGULATED RIBOSWITCHES

Another major class of natural riboswitches regulate premature transcription termination upstream of the start codon by modulating the transcription terminator structure upon aptamer–ligand binding (**Figure 4E**). In spite of the high abundance of transcriptionally regulated

riboswitches in nature (Barrick and Breaker, 2007), there are only two reported cell-free riboswitches of this type, both based on naturally occurring sequences.

Martini and Mansy studied the adenine-responsive riboswitch associated with the *ydhL* gene of *Bacillus subtillis* using an *E. coli* S30 extract, and observed modest activation (1.7-fold) of gene expression in the presence of adenine (Martini and Mansy, 2011).

More recently, Thavarajah et al. used the fluoride-responsive riboswitch that activates expression of the efflux pump CrcB of *Bacillus cereus* in lyophilized *E. coli* cell-free extract in their effort to develop a biosensor for fluoride detection in water (Thavarajah et al., 2020). Both systems used *E. coli* RNA polymerase rather than T7 RNA polymerase for cell-free transcription which is reasonable considering the low termination efficiency of T7 RNA polymerase at canonical bacterial transcription terminators (Jeng et al., 1990, Jeng et al., 1992).

The scarcity of engineered riboswitches that are transcriptionally regulated also extends to *in vivo* studies in *E. coli* and other bacteria, with only a handful of such riboswitches reported (Ceres et al., 2013a, Ceres et al., 2013b, Wachsmuth et al., 2013, Wachsmuth et al., 2015, Domin et al., 2017, Günzel et al., 2020). Transcriptionally regulated synthetic riboswitches are notoriously difficult to engineer, in despite of numerous efforts to systematically generate them (Ceres et al., 2013a, Ceres et al., 2013b); as a result, there are much fewer synthetic transcriptionally regulated riboswitches reported in literature compared to synthetic translationally regulated ones, and the ones reported usually show medium-to-low activation ratios and significant leakage levels. Some of the reasons why this may happen could be the complexity of the terminator–antiterminator RNA structures, the importance of the co-transcriptional kinetic component during the riboswitch folding and activation, the dependency on the catalytic rate of the RNA polymerase used, and that additional accessory protein factors are often needed for efficient regulation (Wickiser et al., 2005a, Wickiser et al., 2005b, Coppins et al., 2007, Lang et al., 2007, Lemay et al., 2011, Frieda and Block, 2012, Quarta et al., 2012, Ceres et al., 2013a, Ceres et al., 2013b). It is worth noting that Gram-negative bacteria seem to prefer translationally regulated riboswitches over transcriptional ones (Barrick and Breaker, 2007); however, the model organism on which, both, cell-free systems and *in vivo* studies are usually based on is *E. coli*, a Gram-negative bacterium.

2.4.1.4. DNA/RNA-RESPONSIVE CELL-FREE SWITCHES

It is worth noting the existence of related synthetic RNA switches/sensors in cell-free systems that specifically respond to nucleic acids (DNA or RNA) instead of small molecule metabolites. These RNA switches modulate transcription or translation through mechanisms similar to those of the natural bacterial riboswitches except that they respond to specific DNA or RNA sequences through Watson–Crick base pairing. Earlier examples come from Aoyama and coworkers who employed a molecular beacon-like structure to regulate cell-free translation in response to short oligo DNA or RNA (Sando et al., 2005, Narita et al., 2006). More recent and sophisticated sensors called toehold switches (Green et al., 2014) also regulate cell-free translation in response to specific RNA sequences (**Figure 4F**). Likewise, Small Transcription Activating RNAs (STARs) (Chappell et al., 2015) exploit RNA–RNA hybridization to regulate premature transcription termination in cell-free systems (**Figure 4G**). These RNA/DNA-responsive cell-free switches have been used to prototype genetic circuits (Pardee et al., 2014, Takahashi et al., 2015, Lehr et al., 2019, Jeong et al., 2019) and to develop biosensors for detecting RNAs from viruses (Pardee et al., 2016a, Ma et al., 2018a), bacteria (Takahashi et al., 2018), and a variety of RNA markers (Takahashi et al., 2018, Chau and Lee, 2020, Chau et al., 2020).
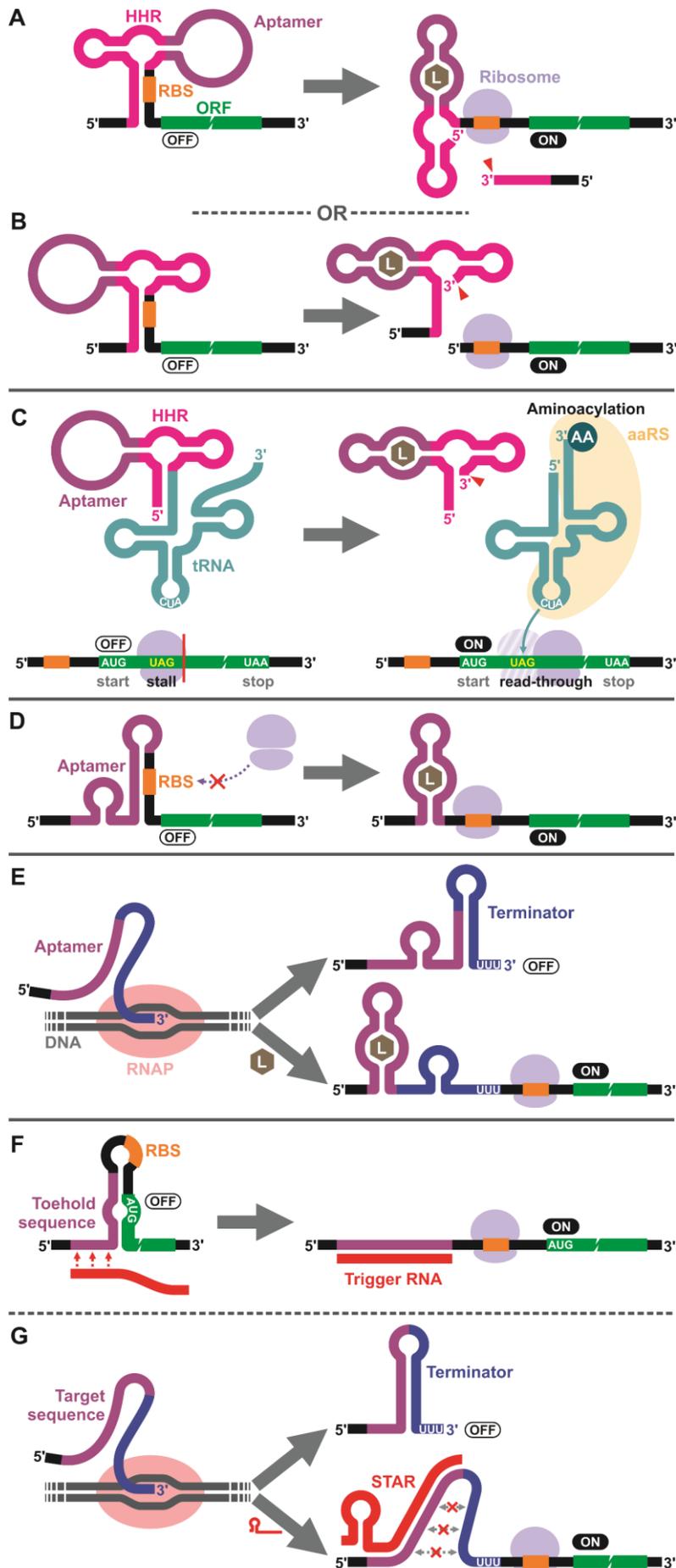
Figure 4. Riboswitch mechanisms in prokaryotic cell-free systems. (**A**, **B**) Aptazyme-driven RBS release: aptamer–ligand binding activates self-cleavage of the ribozyme resulting in physical separation of the anti-RBS sequence from the mRNA. (**C**) Aptazyme-driven sup-tRNA release: self-cleavage of the ribozyme upon aptamer–ligand binding releases the sup-tRNA which undergoes aminoacylation and allows translation elongation past the amber stop codon (UAG). (**D**) RBS sequestration: aptamer–ligand binding induces a structural change near the RBS modulating the accessibility of the sequence to ribosomes. (**E**) Regulation of transcription termination: co-transcriptional folding of the riboswitch in the presence of the ligand prevents formation of the transcription terminator hairpin structure, allowing the full-length mRNA to be transcribed and the protein to be translated. (**F**, **G**) DNA/RNA-responsive switches: they employ similar regulatory mechanisms as riboswitches but respond to specific DNA or RNA sequences through Watson–Crick base pairing. Toehold switches regulate translation initiation (**F**), while small transcription activating RNAs (STARs) regulate transcription termination (**G**). AA: amino acid, aaRS: aminoacyl tRNA synthetase, HHR: hammerhead ribozyme, L: ligand, RBS: ribosome binding site, RNAP: RNA polymerase. Reproduced from (Tabuchi and Yokobayashi, 2021) with permission of the authors.

In addition, T-boxes are a naturally occurring type of riboswitches whose mechanism of action also involves Watson−Crick base pairing interactions with the effector molecule (tRNA) (Green et al., 2010, Raina and Ibba, 2014). There are no examples of T-boxes directly employed as riboswitches in cell-free systems; however, there are some examples of tRNA used to control cell-free gene expression by different mechanisms. Ishida et al. created a system that used a flexizyme (a ribozyme capable of aminoacylate tRNA substrates) associated with a *Bacillus subtilis glyQS* T-box riboswitch to aminoacylate a $tRNA^{Gly}_{GCC}$ only in the presence of N-biotinyl-L-phenylalanine cyanomethyl ester (Bio-^{L}Phe-CME) (Ishida et al., 2020). They introduced their constructs into the FIT (Flexible *In vitro* Translation) system developed by Goto et al. (Goto et al., 2011), and controlled the translation of a DNA template with a GGC start codon.

2.4.2. EUKARYOTIC CELL-FREE RIBOSWITCHES

Designing synthetic riboswitches for eukaryotic CFPS systems can be more challenging than prokaryotic ones. For starters, apart from the TPP riboswitches that modulate pre-mRNA splicing in fungi and plants (Kubodera et al., 2003, Sudarsan et al., 2003, Cheah et al., 2007, Wachter, 2010), there are no other examples of natural eukaryotic riboswitches reported to date (Wachter, 2010) that can be used as a model to engineer synthetic riboswitches.

Furthermore, transcription and translation processes in eukaryotes are significantly more complex than in prokaryotes. For example, canonical translation in eukaryotes generally requires assembly of an initiation complex, comprised of a variety of eukaryotic initiation factors (eIFs) in association with the 5'-cap and the 3' poly(A) tail of the mRNA, before ribosome recruitment. Also, eukaryotic mRNAs do not contain well-defined signal sequence for translation initiation as is the case with prokaryotes (Shine-Dalgarno sequence). Consequently, synthetic riboswitches that function in eukaryotic systems, both in living cells and CFPS systems, have historically been designed to operate via simpler non-natural regulatory mechanisms, such as aptazymes (Mustafina et al., 2020, Shanidze et al., 2020, Yokobayashi, 2019b). Alternatively, riboswitches have been designed based on long viral sequences such as the internal ribosomal entry site (IRES) or the cap-independent translation elements (CITE) that bypass the canonical translation initiation mechanism (**Table 2**).

2.4.2.1. RIBOSOME BLOCKING

Interestingly, earlier reports on eukaryotic cell-free riboswitches predated the discovery of bacterial riboswitches. Inspired by the observation that a stable RNA structure at the 5' terminus of an eukaryotic mRNA can lower gene expression, as it may interfere with ribosome loading or mRNA scanning (Vega Laso et al., 1993, Babendure et al., 2006), Werstuck and Green inserted an aptamer that binds tobramycin or Hoechst dye 33258 in the 5' UTR of an mRNA. Addition of the aptamer's ligand suppressed translation in wheat germ extract (WGE) presumably due to the stabilization of the aptamer structure upon ligand binding (Werstuck and Green, 1998). Similarly, Harvey et al. controlled the synthesis of chloramphenicol acetyltransferase (CAT) in WGE and rabbit reticulocyte lysate (RRL) by inserting multiple copies of theophylline or biotin aptamers in the 5' UTR (**Figure 5A**) (Harvey et al., 2002).

More recently, Ogawa et al. designed a riboswitch that induces mRNA stabilization at the 5' UTR and blocks ribosome loading in WGE (**Figure 5B**) (Ogawa et al., 2018). It only involves a split aptamer at a 5' terminal stem loop, followed by a translation enhancer (E01). When the aptamer binds to the ligand, it induces the formation of a hairpin structure at the 5' end of the mRNA, blocking the ribosome loading. As previously reported in eukaryotes *in vivo*, highly

stable secondary structures at the 5' mRNA leader sequence can inhibit translation due to structural hindrance (Vega Laso et al., 1993, Babendure et al., 2006). Ogawa and collaborators found that the translation inhibition in their riboswitches depends more on the structure's stability than on its size, and its function is affected by its position within the 5' UTR. It was suggested that the preferred position of the inhibitory structure may also depend on whether the translation initiation efficiency is more affected by ribosome loading, as in higher eukaryotes (plants and mammals), or by ribosome scanning, as in lower eukaryotes (yeast) (Vega Laso et al., 1993, Babendure et al., 2006, Ogawa et al., 2018). It is worth noting that similar riboswitch designs have been shown to function in yeast (Grate and Wilson, 2001, Hanson et al., 2003, Suess et al., 2003, Weigand et al., 2008) and in mammalian cells (Werstuck and Green, 1998).

Table 2. List of riboswitches in eukaryotic cell-free systems.
Modified from (Tabuchi and Yokobayashi, 2021) with permission of the authors.

Mechanism	Riboswitch	Ligand	ORF	ON/OFF ratio (Ligand conc.)	CFPS*	Year	Ref.
Ribosome blocking	tob3-RSETA	Tobramycin	RSETA	OFF ~7× (60 µM)	WGEa	1998	1
	H2-RSETA	H33258	RSETA	OFF ~12× (80 µM)	WGEa	1998	1
	(Th)$_3$	Theophylline	CAT	OFF ~8× (1 mM)	WGEa	2002	2
	B$_3$	Biotin	CAT	OFF ~9× (1mM)	WGEa, RRL	2002	2
	th+0-U7	Theophylline	YPet NLuc	OFF 5.9× (1 mM) OFF 7.2× (1 mM)	WGEb	2018	3
	TMR+0-U7	TMR	NLuc	OFF 5.1× (500 µM)	WGEb	2018	3
Aptazyme-regulated translation	mRNA6-Theo	Theophylline	Luc	ON ~50× (500 µM)	WGEb	2009	4
	mRNA6-cGMP	cGMP	Luc	ON ~10× (750 µM)	WGEb	2009	4
IRES-modulated translation	theo5	Theophylline	Luc eGFP	ON 9.6× (1mM) ON 5.1× (1 mM)	WGEb	2011	5, 6
	FMN4	FMN	Luc	ON 7.5× (300 µM)	WGEb	2011	5
	tc7	Tetracycline	Luc eGFP	ON 29× (300 µM) ON 35× (300 µM)	WGEb	2011	5
	sr4	SRB	Luc	ON 4.2× (300 µM)	WGEb	2011	5
	theoN5	Theophylline	NLuc	OFF 5.8× (1 mM)	WGEb	2012	7
	tc-N5	Tetracycline	NLuc	OFF 4.9× (300 µM)	WGEb	2012	7
	FMN-N5	FMN	NLuc	OFF 4.7× (300 µM)	WGEb	2012	7
	theoA3-rS	Theophylline	Luc	OFF 14× (1 mM)	WGEb	2017	8
	nDNA-547-N4	nDNA	NLuc	ON 21× (300 µM)	WGEb	2020	9
Ribosomal shunting	theoS1	Theophylline	Luc	ON 14.4× (1 mM)	WGEb	2013	10, 11
	tmrS1	TMR	Luc	ON 5.4× (333 µM)	WGEb	2013	12
Nonsense suppression	theo(th1)-MS(4)	Theophylline	Luc	ON 7.8× (1 mM)	WGEb	2015	13, 14
	tc(th1)-MS(4)	Tetracycline	Luc	ON 81× (100 µM)	WGEb	2015	13, 14
	FMN(th1)-MS(4)	FMN	Luc	ON ~4× (30 µM)	WGEb	2015	13, 14
3' CITE-modulated translation	5SL-BYm2-theo	Theophylline	YPet NLuc	ON 7.7× (1 mM) ON 7.3× (1 mM)	WGEb	2017	15
	5SL-BYm2-TMR	TMR	NLuc	ON 5.8× (100 µM)	WGEb	2017	15

*CFPS (cell-free protein synthesis) systems. WGE: wheat germ extracts, (a) Promega, (b) WEPRO1240 (CellFree Sciences). RRL: rabbit reticulocyte lysate system (Promega). cGMP: cyclic guanosine monophosphate, FMN: flavin mononucleotide, H33258: Hoechst dye 33258, nDNA: nano-sized ssDNA or pentadeoxyribonucleotide, TMR: tetramethylrhodamine, SRB: sulphorhodamine B, CAT: chloramphenicol acetyltransferase, eGFP: enhanced green fluorescent protein, Luc: firefly luciferase, NLuc: NanoLuc, ORF: open reading frame, RSETA: ORF of the undigested cloning site from pRSET-A plasmid (Invitrogen), YPet: yellow fluorescent protein for energy transfer.

References: 1 (Werstuck and Green, 1998); 2 (Harvey et al., 2002); 3 (Ogawa et al., 2018); 4 (Ogawa, 2009); 5 (Ogawa, 2011); 6 (Ogawa, 2014); 7 (Ogawa, 2012); 8 (Ogawa et al., 2017a); 9 (Ogawa and Itoh, 2020); 10 (Ogawa, 2013); 11 (Ogawa, 2015); 12 (Ogawa, 2013); 13 (Ogawa and Tabuchi, 2015); 14 (Ogawa, 2021); 15 (Ogawa et al., 2017b).
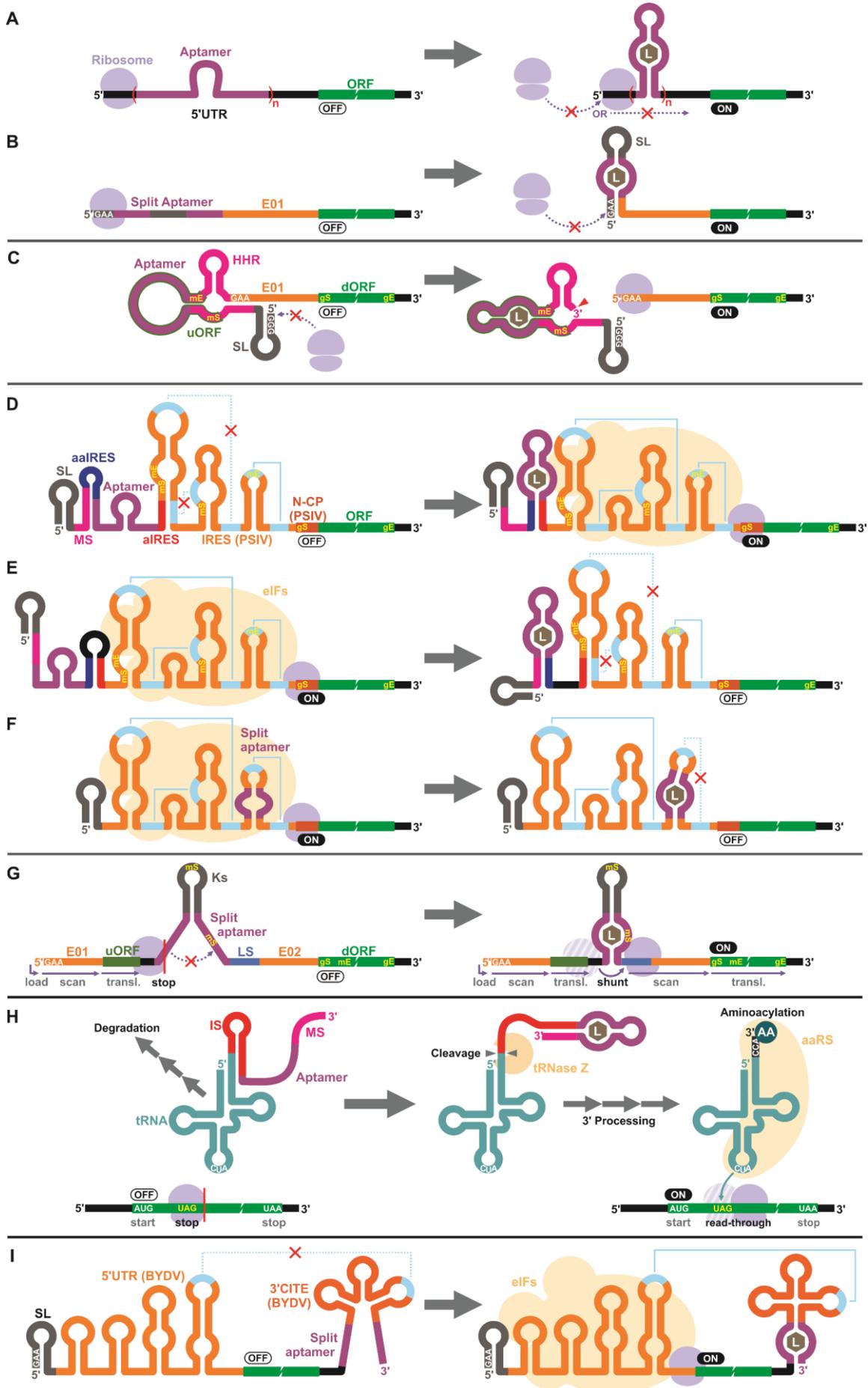
Figure 5. (Previous page) Riboswitch mechanisms in eukaryotic cell-free wheat germ extract (WGE). (**A**, **B**) Ribosome blocking: Stable structure induced by aptamer–ligand binding at the 5' UTR can block ribosome loading or mRNA scanning. (**C**) Aptazyme-regulated translation: activation of the aptazyme cleaves the mRNA and exposes the 5' terminus which can be recognized by the WGE's translational machinery. (**D–F**) IRES-modulated translation: IRES-mediated translation can be controlled by different combinations of sequence elements that include modulator sequence (MS), (split) aptamer, anti-IRES (aIRES), and anti-aIRES (aaIRES). (**G**) Ribosomal shunting: aptamer–ligand binding brings the upstream ORF (uORF) and the downstream ORF (dORF) to close proximity, allowing the ribosome to shunt over to reinitiate translation from the dORF. (**H**) Aptamer-regulated sup-tRNA processing: aptamer–ligand binding activates the tRNase Z-mediated 3' processing of the sup-tRNA–aptamer that would otherwise be rapidly degraded. The mature aminoacylated tRNA suppresses the amber stop codon (UAG) in the coding sequence. (**I**) 3' CITE-modulated translation: binding of the ligand to the aptamer promotes correct folding of the 3' CITE structure to restore the interaction between the 5' UTR and the 3' UTR. AA: amino acid, aaRS: aminoacyl tRNA synthetase, BYDV: barley yellow dwarf virus, E01/E02: translational enhancers, eIFs: eukaryotic translation initiation factors, gE: real gene stop codon, gS: real gene start codon, HHR: hammerhead ribozyme, IRES: internal ribosome entry site, IS: inhibitory sequence, Ks: artificial Kozak stem-loop, L: ligand, LS: landing site, mE: mimic stop codon, mS: mimic start codon, N CP: N-terminus of capsid protein gene, PSIV: *Plautia stali* intestine virus, SL: stem-loop. Reproduced from (Tabuchi and Yokobayashi, 2021) with permission of the authors.

2.4.2.2. APTAZYME-BASED RIBOSWITCHES

Aptazymes are frequently used to regulate gene expression in living eukaryotic cells (Yokobayashi, 2019b, Shanidze et al., 2020, Mustafina et al., 2020). Ogawa, who reported the first prokaryotic cell-free riboswitches based on an aptazyme, also designed the first aptazyme-based cell-free riboswitches in a eukaryotic CFPS system. Incidentally, Ogawa and collaborators have designed the majority of synthetic riboswitches that function in eukaryotic CFPS systems, specifically, wheat germ extract (WGE) (**Table 2**). To engineer their aptazyme-based riboswitches, Ogawa exploited a property of the WGE described by Endo and coworkers (Sawasaki et al., 2002, Kamura et al., 2005): WGE can efficiently initiate translation from an uncapped mRNA, and its efficiency depended on the first three bases at the 5' terminus of the mRNA. In his design, Ogawa strategically inserted a theophylline-activated aptazyme directly upstream of a GAA trinucleotide followed by an enhancer sequence (E01), in such way that the 5' terminal GAA becomes exposed upon ribozyme cleavage. In order to minimize the background noise, he introduced into the aptazyme sequence a less effective 5'-terminal sequence (5'-GGG) and a short upstream open reading frame (uORF), mimicking a gene, embedded within the aptazyme sequence, so any spurious transcription of the non-activated switches would result in a low-efficiency translation of the decoy gene instead of the actual reporter. Without cleavage, the lack of 5' GAA, the presence of the relatively stable aptazyme structure at the 5' terminus, and the existence of the mimic uORF, repress reporter gene expression (**Figure 5C**) (Ogawa, 2009).

2.4.2.3. INTERNAL RIBOSOMAL ENTRY SITES

Next, Ogawa and coworkers shifted their attention to internal ribosomal entry sites (IRES). An IRES is a viral RNA element that recruits the translation initiation complex in the absence of the 5' cap structure. Ogawa identified the minimum functional sequence of the IRES from *Plautia stali* intestine virus (PSIV), and rationally fused an RNA aptamer to disrupt or induce critical IRES structural elements in the presence of the ligand. For example, Ogawa engineered an ON-switch that activates gene expression in the presence of theophylline following a rational design strategy. First, an 8-nt anti-IRES (aIRES) sequence was introduced upstream of the IRES to disrupt a critical pseudoknot structure, successfully repressing gene expression (OFF). Then, an anti-anti-IRES (aaIRES) and the theophylline aptamer was added further upstream to restore gene expression (ON) by sequestering the aIRES. Finally, a modulator

sequence (MS) was appended to the 5' end of the mRNA interfering with the aaIRES and the aptamer to suppress gene expression (OFF) (**Figure 5D**). Ogawa systematically optimized the stability of the MS–aaIRES–aptamer interaction to obtain riboswitches that shift the IRES to its ON structure in the presence of theophylline and other ligands (Ogawa, 2011). Ogawa later designed OFF-switches by rearranging the riboswitch elements (**Figure 5E**) (Ogawa, 2012) demonstrating the flexibility of this design strategy. More recently, Ogawa et al. reported a simpler OFF-switch design with just a split aptamer embedded within the IRES structure (**Figure 5F**) (Ogawa et al., 2017a).

2.4.2.4. OTHER MECHANISMS

Ogawa has developed additional cell-free riboswitches in WGE based on other regulatory mechanisms. For example, the ribosome shunting mechanism found in some viruses (Fütterer et al., 1993, Schmidt-Puchta et al., 1997, Dominguez et al., 1998, Ryabova and Hohn, 2000, Pooggin et al., 2006) enables the ribosomes to scan the mRNA without unwinding some secondary structures, effectively bypassing (or "shunting over") these regions of the mRNA. Ogawa inserted two halves of a split aptamer separated by an intervening sequence between a short upstream ORF (uORF) and the reporter gene downstream (dORF). Aptamer–ligand binding brings the two regions to close proximity and induces ribosome shunting, resulting in efficient gene expression of the reporter gene (**Figure 5G**) (Ogawa, 2013). Once again, Ogawa introduced mimic start (m1, m2) and stop (mE) codons to "trap" the ribosome and decrease the translation efficiency in the absence of ligand, and in the event of the ribosome managing to keep scanning the mRNA without shunting.

In another design, Ogawa and Tabuchi engineered riboswitches that control read-through of premature translation termination codons (nonsense amber mutations) using sup-tRNAs in WGE (Ogawa and Tabuchi, 2015) through a mechanism reminiscent of their aptazyme-based riboswitch that operates in a prokaryotic cell-free system (**Figure 4C**, see **2.4.1.1. Aptazyme-based riboswitches**, page **14**) (Ogawa and Maeda, 2008b). In their eukaryotic riboswitch, the aptamer was fused to the 3' end of the sup-tRNA via an inhibitory sequence (IS) that interferes with the natural 3' processing of the tRNA (cleavage by tRNase Z and addition of CCA by tRNA nucleotidyltransferases). Aptamer–ligand interaction exposes the 3' terminus of the tRNA allowing its rapid maturation followed by suppression of the amber codon (**Figure 5H**) (Ogawa and Tabuchi, 2015). The same design was later adapted by Ogawa et al. to develop sup-tRNA switches triggered by complementary DNA oligonucleotides in WGE (Ogawa et al., 2016).

Later, Ogawa and colleagues manipulated a viral translation initiation mechanism mediated by the 3' cap-independent translation element (3' CITE) from barley yellow dwarf virus (BYDV). The viral RNA element located in the 3' UTR forms a kissing loop interaction with another element located in the 5' UTR, mimicking the canonical 5'cap–3'poly(A) interaction in eukaryotic cells. The 3'CITE–5'UTR pair interacts with the translation initiation factors (eIFs) and recruit the ribosomes to the mRNA. Ogawa and colleagues were able to control translation of a reporter gene by inserting a split aptamer within the 3' CITE to enhance the kissing loop interaction upon aptamer–ligand binding (**Figure 5I**) (Ogawa et al., 2017b).

2.5. APPLICATIONS OF CELL-FREE RIBOSWITCHES

With the rapid progress in cell-free system research, some researchers have explored various applications of cell-free riboswitches which are reviewed below.

2.5.1. CELL-FREE BIOSENSORS

Naturally, researchers thought of employing riboswitches as biosensors because of their capacity for detecting specific metabolites. Natural and laboratory-evolved aptamers have been already reported for a plethora of compounds, and in principle, it is possible to generate aptamers capable of recognizing a variety of analytes, both natural and synthetic. Although aptamers obtained by *in vitro* evolution can potentially work in cellular systems, very few have been used to engineer riboswitches *in vivo*, most likely due to biological factors such as unfavorable intracellular environment, low cell permeability, low stability inside cells, or toxicity of the ligand. Those same riboswitches if applied in cell-free systems could circumvent some of these constraints.

One of the major concerns for applications on-field, is the stability and robustness of the cell-free systems after long-term storage. Fortunately, it has been demonstrated that cell-free systems can be freeze-dried (lyophilized) in a tube (Smith et al., 2014, Pardee et al., 2016b, Gregorio et al., 2020, Thavarajah et al., 2020) and on filter paper (Pardee et al., 2014, Pardee et al., 2016a, Ma et al., 2018a, Takahashi et al., 2018), and then rehydrated on-demand without significant loss of activity if stored properly. Cell-free riboswitches possess several potential advantages as a versatile platform for biosensing. For example, they can be deployed within the cell-free system as DNA templates, which is very stable, especially if the system is lyophilized. Freeze-drying and rehydrating has little to no impact on the final activity of the riboswitch nor the reporter gene, as the mRNA and the reporter proteins are freshly produced once the cell-free system is activated. This has significantly improved the shelf-life and applicability of riboswitches in cell-free systems as biosensors, although there are still some remaining challenges such as susceptibility to contaminants and inhibitors.

Cell-free riboswitches can potentially trigger expression of any gene(s), which broadens the options for the detection method. Sensitivity of the biosensor can be further improved by taking advantage of the signal amplification inherent to transcription/translation (multiple mRNA transcribed per DNA template, then multiple proteins translated per mRNA molecule) and provided by some reporter genes such as enzymes (single enzyme can turn over multiple substrate molecules). Since the readiness and easiness for detecting the output signal is an important factor to consider when performing tests on-site, colorimetric enzymatic reporters are usually better as reporter genes than fluorescent or luminescent proteins (generally preferred for in-lab experiments) because (i) the output is visible to the naked eye and no extra equipment is necessary, (ii) they can amplify weak signals, and (iii) they can respond faster. Cell-free systems also allow tuning of biochemical parameters such as concentration of cofactors, substrates, and DNA template (enzyme), in order to reduce signal leakage and improve the sensitivity and dynamic range of the biosensor (Thavarajah et al., 2020).

Although, the potential use of the cell-free riboswitches as biosensors have been suggested multiple times, and some of the performance indicators, such as detection limit, dose–response, and signal-to-noise ratio, have been analyzed for some riboswitches (Ogawa and Maeda, 2007, Ogawa and Maeda, 2008b, Muranaka et al., 2009a, Ogawa, 2011, Ogawa and Tabuchi, 2015, Espah Borujeni et al., 2016, Drachuk et al., 2020), most of these have never been put to test by measuring more practical and relevant samples. A recent notable exception was reported by Thavarajah et al. who engineered a cell-free fluoride riboswitch as a biosensor for field applications (Thavarajah et al., 2020). They implemented the riboswitch in a lyophilized *E. coli* extract format previously optimized for toehold switches (Pardee et al., 2014, Pardee et al., 2016a, Ma et al., 2018a, Takahashi et al., 2018), and used catechol 2,3-dioxygenase (C23DO) as the reporter gene which provides a visual colorimetric output suitable for field use. The

practical potential as a biosensor was demonstrated by on-site detection of fluoride levels as low as 50 μM (~1 ppm) in real-world water samples (Thavarajah et al., 2020).

2.5.2. ARTIFICIAL CELLS

Artificial cells are emerging as a new frontier in synthetic biology (Xu et al., 2016, Buddingh and van Hest, 2017, Damiano and Stano, 2020). Advances in microfluidic technologies have allowed the encapsulation of cell-free systems into micrometer-size compartments (Griffiths and Tawfik, 2006, Kelly et al., 2007), essentially producing miniature bioreactors in which biochemical processes are carried out *in vitro* at cell-like scales. Emergent properties could also arise from this encapsulation, since each compartment is now an individual entity that can interact with each other.

Although artificial cells are not alive, they are able to mimic some features of natural living cells (Lentini et al., 2014, Karzbrun et al., 2014, Lentini et al., 2016, Guindani et al., 2022). These artificial cells can contain biological molecules and structures (Sunami et al., 2010, Spencer et al., 2013, Guindani et al., 2022), like lipid bilayers, membrane proteins, nucleic acids, enzymes, and even small vesicles as subcompartments (Beneyton et al., 2018); but they can also include engineered materials, like magnetic beads (Diehl et al., 2006, Tabuchi and Yokobayashi, 2022) or synthetic polymers (Lai et al., 2020, Zhou et al., 2018, Jiang et al., 2022), that are not naturally found in living cells but mimic some of their properties or carry out some specific functions. Artificial cells can ultimately act as simplified models of living cells, which could lead to a better understanding of life, as well as offer potential applications in medicine and biotechnology. Xu et al. pointed out several motivations and implications (Xu et al., 2016):

- Artificial cells may provide insight into the origin of life.
- Can be used as a tool for investigating cellular processes in a simplified environment with fewer interference.
- Represents a connection between non-living and living systems.
- May replace engineered organisms for biomanufacturing of biomolecules (bioreactors).
- Can be employed in biomedical applications such as drug delivery or medical imaging.
- Can perform new functions that cannot be naturally found in living cells.

Despite recent significant advances (Ivanov et al., 2021, Guindani et al., 2022), researchers are still far from recreating a truly autonomous artificial cell. To close the gap between artificial and biological cells, researchers still need to address many challenges, including self-replication, nutrient uptake, cell-to-cell communication, and interactions with the environment (Xu et al., 2016). One way for the artificial cells to interact with the environment is to control gene expression in response to chemical signals in the medium, for example, artificial cells encapsulating a CFPS system along with DNA encoding a gene controlled by a genetic switch that responds to a chemical signal. Synthetic riboswitches that can function in cell-free platforms have become attractive tools for interfacing artificial cells with diverse chemical signals. Alongside with canonical transcription factor-based switches well established in bacterial systems (e.g., LacI–IPTG, AraC–arabinose, LuxR–AHL), some synthetic riboswitches have been introduced in different types of artificial cells such as water-in-oil emulsions (Martini and Mansy, 2011, Martini and Mansy, 2014), and lipid vesicles (Martini and Mansy, 2011, Martini and Mansy, 2014, Lentini et al., 2014, Adamala et al., 2017, Dwidar et al., 2019). In a recent report, a riboswitch-regulated DNA was trapped inside silk fibroin-based microcapsules to demonstrate protein expression in the presence of theophylline (Drachuk et al., 2020).
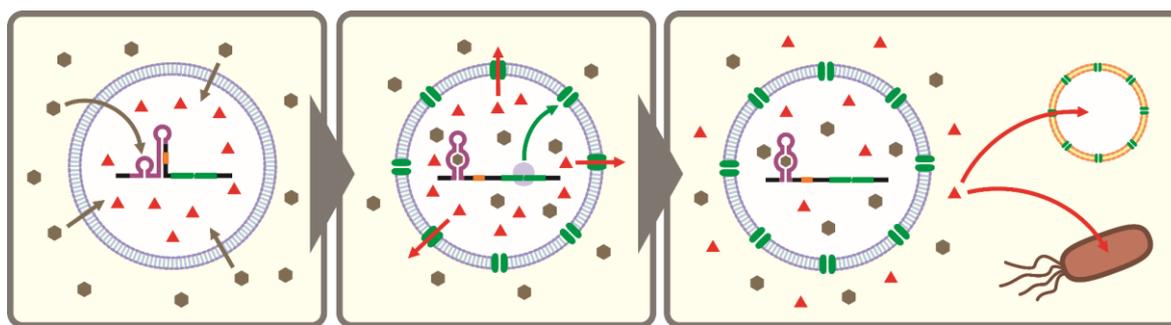
Figure 6. Schematic illustration of artificial cells (e.g., liposomes encapsulating CFPS system) equipped with a cell-free riboswitch. A membrane-permeable ligand (brown hexagons) diffuses into the artificial cells and activates the riboswitch to express α hemolysin (green ovals). The nanopores localize on the artificial cell membrane allowing non-membrane-permeable small molecules (red triangles) entrapped in the artificial cells to be released. The released compound can serve as a chemical signal to trigger response in other artificial cells or living cells. Reproduced from (Tabuchi and Yokobayashi, 2021) with permission of the authors.

On the other hand, intercellular chemical communication has been studied extensively using engineered living cells (Hennig et al., 2015), but remains poorly explored in the context of artificial cells (Lentini et al., 2016). Researchers have started to explore communication between artificial cells and use it to prototype and test genetic circuits and cascades. These networks were often assembled with natural protein components, such as receptor proteins, enzymes, and transcription factors (Schwarz-Schilling et al., 2016, Dupin and Simmel, 2019, Buddingh et al., 2020, Smith et al., 2021, Jiang et al., 2022), or with non-riboswitch RNA switches (Takahashi et al., 2015, Lehr et al., 2019, Smith et al., 2021); with a just few examples of riboswitches playing the role of relay points for signal transduction in such networks (Lentini et al., 2014, Adamala et al., 2017). For the latter, theophylline riboswitches have always been used to mediate chemical signaling between artificial cells, or between artificial cells and bacteria. Lentini et al. encapsulated IPTG inside liposomes along with DNA encoding an α-hemolysin gene controlled by a theophylline-activated riboswitch (Lentini et al., 2014). Addition of theophylline (membrane permeable) triggers expression of α-hemolysin which forms nanometer sized pores on the membrane of the artificial cells. The pores allow the trapped IPTG (membrane impermeable) to diffuse out of the artificial cells and to induce GFP expression in co-cultured *E. coli* cells. Similarly, Adamala et al. constructed riboswitch-controlled artificial cells that release doxycycline to achieve communication between different artificial cells (**Figure 6**) (Adamala et al., 2017).

More recently, Dwidar et al. applied Systematic Evolution of Ligands by EXponential enrichment (SELEX) (Ellington and Szostak, 1990, Tuerk and Gold, 1990) to discover a novel aptamer that recognizes histamine, and they designed and optimized histamine-responsive cell-free riboswitches in the PURE system (Dwidar et al., 2019). They thoroughly characterized the riboswitch in liposome-based artificial cells using fluorescence microscopy and flow cytometry to confirm the robust and dynamic response to histamine. They constructed liposome artificial cells that, depending on the DNA template they contained, can trigger different responses when histamine is in the surrounding media; for example, expressing a reporter protein (mCherry), releasing cargo molecules after expression of α-hemolysin, or triggering self-destruction after expression of phospholipase C or A1 which hydrolyze the liposome membranes.

In the future, chemically responsive artificial cells using riboswitches may enable more practical applications, for example, as vesicles-based drug delivery systems for imaging, diagnosis, and/or treatment of diseases that exhibit specific biochemical fingerprints that can be detected by riboswitches.

2.5.3. MECHANISTIC UNDERSTANDING OF RIBOSWITCHES

Before the advent of cell-free systems, riboswitches were studied with either *in vivo* systems, or *in vitro* biophysical methods; however, very often important details were overlooked or misinterpreted due to lack of clearer information. For example, *in vitro* biophysical methods lack active transcription and translation crucial for the proper function of riboswitches. On the other hand, living organisms exhibit various biological constraints, and their own metabolism can mask or interfere with riboswitch activity. CFPS systems serve as a good proxy for mimicking the riboswitch functions observed *in vivo*, but in a more controlled and simplified environment, providing a valuable platform for prototyping, analyzing, and characterizing riboswitches (Chushak et al., 2021), which can lead to fundamental insights into the regulatory mechanisms and the underlying design principles of natural and synthetic riboswitches. As already mentioned before, the advantage of working with cell-free riboswitches is the presence of fewer experimental constraints compared to working in living cells, in addition to circumventing other biological constraints such as ligand toxicity, ligand permeability, and metabolic burden. Researchers can precisely control various parameters (such as ligand concentration, or the timing of ligand addition), or decouple the gene expression process to study the behavior of the riboswitch at each step.

An elegant example by Lemay et al. illustrates the advantage of using a cell-free system to study riboswitch mechanisms (Lemay et al., 2011). By decoupling the transcription and translation processes in *E. coli* S30 extract, they showed that the natural *add* riboswitch from *Vibrio vulnificus* is governed by an equilibrium mechanism and can modulate translation initiation by reversibly binding to adenine post-transcriptionally. A similar approach was taken by Mishler and Gallivan to study their theophylline-activated riboswitches that they originally engineered in bacteria (Mishler and Gallivan, 2014). Previous experiments *in vivo* let researchers to mistakenly suggest that this kind of riboswitch was also controlled by a thermodynamic equilibrium mechanism (Lynch et al., 2007); however, results from more detailed experiments in CFPS systems pointed towards the presence of an important co-transcriptional kinetic component governing the activation of this riboswitch (Mishler and Gallivan, 2014). By adding theophylline before or after they stopped transcription with rifampicin, they unambiguously showed that theophylline must be available during transcription for their riboswitches to function, demonstrating that the kinetic trapping mechanism plays a more dominant role than the thermodynamic equilibrium mechanism for the activation of their translationally regulated riboswitch. This new observation also explained the discrepancies between the high affinity of the theophylline aptamer measured *in vitro*, and the relatively high concentrations of theophylline required for activating the riboswitches *in vivo*, previously attributed to low membrane permeability or active efflux of theophylline (Desai and Gallivan, 2004, Bayer and Smolke, 2005). The same phenomena have been observed in other riboswitches *in vivo* and *in vitro* and could most likely be attributed to a similar kinetic control on the folding of competing RNA structures (Haller et al., 2011, Peselis and Serganov, 2014, Dwidar et al., 2019).

CFPS systems have also allowed the creation or validation of several mathematical or computational models (Kobori et al., 2012, Mishler and Gallivan, 2014, Espah Borujeni et al., 2016, Chushak et al., 2021) aimed to gain a better understanding of the riboswitch mechanisms and to determine the influence of different parameter on their activity (concentrations, aptamer structure, ligand affinity, folding free energy, macromolecular crowding, transcription rate, RNA and protein dynamics, etc.). Rational design strategies enormously benefit from this as they strongly rely on such models and algorithms for predicting the performance of the

riboswitches designed *in silico*. Independent reports from Kobori et al. (2012) and Chushak et al. (2021) illustrate how kinetic modeling of riboswitch mechanisms based on experimental data of cell-free riboswitches can provide not only insights into the riboswitch mechanisms, but also hints for optimizing riboswitch performance (Kobori et al., 2012, Chushak et al., 2021). Espah Borujeni et al. built a more detailed model considering additional kinetic and biophysical parameters such as co-transcriptional folding and free energy of ribosome binding (Espah Borujeni et al., 2016). Importantly, they used the model to automate riboswitch design using arbitrary aptamers, and they experimentally validated the designed riboswitches in either *E. coli* cells or cell-free S30 extracts. While the performance of some of the riboswitches obtained was not particularly superior, the ability of such a model-driven design to generate functional riboswitches *de novo* is an important achievement. There is hope that new improved models will eventually be able to accurately predict sequence-structure-function relationships of potential designs, improving the efficiency of the designing process, and facilitating engineering of new and better riboswitches.

While many riboswitches have been shown to function similarly *in vivo* and in cell-free systems, it is very important to remember that some parameters obtained from one system may not be directly applicable to other systems (Kobori et al., 2012) if, for example, the objective is to eventually transfer the cell-free riboswitches to an *in vivo* system, or vice versa. For instance, Espah Borujeni et al. tested a counter-intuitive phenomenon predicted by their physics-base model: higher mRNA levels of the riboswitch actually decrease their activation ratio inside cells, but not in CFPS systems (Espah Borujeni et al., 2016). This phenomenon was attributed to the intracellular macromolecular crowding inside living cells, which limits the free space available for the riboswitch-ligand interactions, whereas this is not an issue in a relatively dilute cell-free reaction.

As more research results on cell-free riboswitches are accrued, we are starting to learn some general lessons on their properties and design approaches. For example, minimum free energies of different RNA conformations estimated using RNA folding algorithms and tools such as mfold (Zuker and Stiegler, 1981, Zuker, 2003) and ViennaRNA packages (Gruber et al., 2008, Lorenz et al., 2011) may be a good predictor for the behavior of a riboswitch if it operates under an equilibrium mechanism. However, this may not be enough if the riboswitch exhibits a significant kinetic barrier between its conformations; and if those barriers are not taken into account, the riboswitch will likely fail (Mishler and Gallivan, 2014). Moreover, the equilibrium dissociation constant (K_d) of the aptamer—which was conventionally used for ranking suitable aptamers for riboswitch design under the assumptions of an equilibrium mechanism—, may not be the right parameter to consider. It was suggested that the ligand association kinetics of the aptamer may be a better predictor of the riboswitch performance (Mishler and Gallivan, 2014) because a faster binding of the ligand to the aptamer seems to be more relevant for a kinetically controlled riboswitch than a high affinity (that may be just a result of a slow dissociation rate). Most likely a combination of these parameters—which includes, among others, the aptamer's ligand association kinetics useful in the kinetic-trapping models (Mishler and Gallivan, 2014), and the ligand binding free energy from the thermodynamic-equilibrium models (Espah Borujeni et al., 2016)—is important to ultimately explain why some RNA aptamers are more likely to produce effective riboswitches than others, and why some RNA aptamers behave so differently in different contexts while other do not.

2.6. CONCLUSIONS AND FUTURE DIRECTIONS

As reviewed in this chapter, while a steady number of cell-free riboswitches have been reported over the last two decades, cell-free riboswitches have attracted relatively little attention

compared to cellular riboswitches. This trend may change in the near future with the growing interest in cell-free synthetic biology and its applications. With fewer experimental constraints compared to living cells, cell-free systems can allow more precise and extensive control over the genetic circuits and metabolic pathways built by the researchers. Consequently, cell-free systems offer a unique opportunity for elucidating riboswitch mechanisms in a controlled environment that can lead to fundamental insights on how they work and what their underlying design principles are. Potentially, this knowledge can even provide insights into the early stages of the origin of life (RNA world theory), and the evolution of sensory mechanisms and regulatory elements of gene expression.

It should be acknowledged, however, that cell-free riboswitches still need more efforts to prove that they are valuable tools for cell-free synthetic biology. For example, OFF-switches are conspicuously missing in the prokaryotic cell-free systems. Additionally, the performance and robustness of the current cell-free riboswitches must be improved, while some cell-free riboswitches show high ON/OFF ratios, many others display more modest responses. Nonetheless, the variety of ligands that cell-free riboswitches have been engineered to detect are quite broad compared to cell-based riboswitches. Theophylline has been, by far, the most popular trigger molecule, very likely because its best aptamer have been well-studied and well-minimized, it barely interferes with the *in vitro* transcription-translation, and the aptamer have a relatively high affinity for its ligand (Ogawa and Maeda, 2007, Ogawa, 2014). In addition to theophylline riboswitches, cell-free riboswitches have also been designed to respond to biotin (Harvey et al., 2002), tetramethylrhodamine (Ogawa, 2013, Espah Borujeni et al., 2016, Ogawa et al., 2018), flavin mononucleotide (Ogawa, 2011, Ogawa, 2012, Ogawa and Tabuchi, 2015), cGMP (Ogawa and Maeda, 2007, Ogawa, 2009), TPP (Muranaka et al., 2009a, Kobori et al., 2012), fluoride (Thavarajah et al., 2020), histamine (Dwidar et al., 2019), dopamine (Espah Borujeni et al., 2016), thyroxine (Espah Borujeni et al., 2016), sulphorhodamine B (Ogawa, 2011), Hoechst dye 33258 (Werstuck and Green, 1998), tobramycin (Werstuck and Green, 1998), tetracycline (Ogawa, 2011, Ogawa and Tabuchi, 2015), and pentadeoxyribonucleotides (Ogawa and Itoh, 2020) (**Table 1**, **Table 2**). With the exception of histamine, however, these cell-free riboswitches have been constructed using known aptamers from natural riboswitches or discovered through SELEX intended for other applications. Additional aptamer–ligand combinations adapted to cell-free riboswitches will lead to better understanding of what makes an aptamer–ligand pair more suitable (or not) for cell-free riboswitches.

Finally, cell-free systems offer a very good platform for studying and engineering complex genetic systems from bottom-up, such as artificial cells that can respond and adapt to a dynamic environment. Such systems, much like their cell-based counterparts, need gene switches to sense and respond to a variety of chemical signals. Thus far, chemical gene switches in artificial cells have been limited to known transcription factors, quorum sensing systems, and the theophylline riboswitch. Evidently, there is a need to expand the current genetic toolkit of gene switches for artificial cells. Availability of riboswitches that can interface cell-free systems with diverse chemical signals should accelerate development of innovative cell-free systems with complex functions and practical applications.

CHAPTER 3:
NOVEL HIGH-THROUGHPUT METHOD FOR SCREENING CELL-FREE RIBOSWITCHES

Parts of this chapter have been published as an original research article (Tabuchi and Yokobayashi, 2022): "High-throughput screening of cell-free riboswitches by fluorescence-activated droplet sorting" *Nucleic Acids Res.* **50**(6): 3535-3550. DOI: 10.1093/nar/gkac152

3.1. INTRODUCTION

A number of cell-free riboswitches have been reported in the literature, most of them for the purpose of understanding or demonstrating how aptamer–small molecule interactions affect gene expression (Tabuchi and Yokobayashi, 2021). Despite these examples, design of cell-free riboswitches is not straightforward. Most of the cell-free riboswitches reported to date were either (i) based on natural riboswitches, (ii) originally developed in *E. coli* using high-throughput screening or selection, or (iii) designed by trial-and-error. Although several strategies for high-throughput screening and selection of riboswitches *in vivo* have been successfully applied to engineer bacterial riboswitches (Desai and Gallivan, 2004, Lynch et al., 2007, Topp and Gallivan, 2008, Wieland and Hartig, 2008, Muranaka et al., 2009b, Rehm and Hartig, 2014, Felletti et al., 2016, Kirchner et al., 2017, Harbaugh et al., 2018), they often do not function as well in cell-free systems (Martini and Mansy, 2011, Martini and Mansy, 2014, Adamala et al., 2017). Furthermore, as noted before, many aptamer–ligand combinations are not amenable to bacterial screening or selection. Cell-free riboswitches may also be more difficult and costly to engineer and optimize directly in CFPS systems due to the lack of high-throughput cell-free assays and screening methods, limiting the variety and the performance of cell-free riboswitches that, otherwise, could extend the capabilities of cell-free systems. I addressed this challenge by developing a microfluidic droplet sorting strategy to rapidly enrich functional riboswitches from over 4000 variants directly in a cell-free system.

It is worth mentioning that proof-of-principle sorting of enzymes expressed in monodisperse droplets containing a CFPS system has been demonstrated multiple times (Fallah-Araghi et al., 2012, Holland-Moritz et al., 2020), but directed evolution from a randomized protein library in CFPS systems was only recently achieved by Holstein et al. who evolved serine protease savinase mutants with improved activity (Holstein et al., 2021). Alternatively, Zhang et al. developed femtoliter microfluidics arrays in which enzymes are expressed in the PURE system from single DNA templates (Zhang et al., 2019), with which they achieved 20-fold improvement of alkaline phosphatase activity via directed evolution. Other previous attempts of directed evolution of proteins using microfluidics involved conventional *in vitro* evolution and *in vivo* expression in bacteria and other organisms, while microfluidics fluorescence-activated droplet sorting (FADS) were only employed at the last step to perform a high throughput fluorogenic enzymatic assay after lysing individual cells expressing the protein variants inside the droplets (Obexer et al., 2016, Obexer et al., 2017, Fu et al., 2021).

Additionally, diverse microfluidics approaches and *in vitro* compartmentalization (IVC) techniques have also been employed for high-throughput selection and analysis of aptamers (Karns et al., 2013, Liu et al., 2021), ribozymes (Ryckelynck et al., 2015), and RNA-based biosensors (Autour et al., 2019). Conversely, microfluidics in combination with aptamer or riboswitches have been used for detecting and engineering metabolite-producing bacterial strains (Abatemarco et al., 2017, Jang et al., 2016).

To our knowledge, however, there have been no attempts to evolve dynamic biomolecular genetic devices such as riboswitches directly in CFPS systems. Current efforts to build sophisticated cell-free systems, such as artificial cells and biosensors, mostly rely on parts and devices borrowed from or engineered in living cells. The method described in this chapter should pave the way for directed evolution and engineering of cell-free genetic devices directly in CFPS systems optimized for cell-free applications. This strategy should also be useful for ribozymes, protein-based switches, and other types of genetic devices.

Last but not least, many published microfluidics methods require special setup and custom designed chips. Such methods are difficult or even impossible to replicate without the specific devices, custom chips, or extensive expertise in microfluidics. The technique presented here have been developed exclusively using commercially available microfluidics instruments and chips. With such general-purpose microfluidics platforms becoming increasingly accessible, this and similar techniques should be easier to implement and more accessible to laboratories without specialized equipment and expertise. This new method is also simple and flexible enough to be adapted to custom designed chips if available or desired, expanding the potential of the technique.

3.2. OVERVIEW OF THE CELL-FREE RIBOSWITCH SORTING STRATEGY

The strategy for screening a large number of cell-free riboswitch variants is illustrated in **Figure 7**. First, DNA templates encoding riboswitch variants are clonally amplified on magnetic microbeads following the previously reported strategy called BEAMing (Beads, Emulsion, Amplification, and Magnetics) (Diehl et al., 2006) (**Figure 7D**). Then, the magnetic beads are individually encapsulated in water-in-oil emulsion droplets containing the CFPS system (PURE*frex*) with or without the ligand (**Figure 7F**). Individual droplets (containing only one riboswitch variant) express a fluorescence reporter gene depending on the riboswitch state (**Figure 7G, H**). The droplets are sorted to recover the riboswitch-encoding templates that show the desired output (ON or OFF) (**Figure 7I, J**). These ON and OFF sorting cycles are repeated to enrich functional riboswitches from the original library. Samples from the sorted libraries after each cycle are barcoded and sequenced through high-throughput sequencing (HTS) (**Figure 7K**). The sequencing data is analyzed to obtain an "enrichment trend" for each riboswitch variant, which serves as an estimator of potential functional riboswitches. The top candidates are then individually analyzed to assess their activity in cell-free conditions.

Each DNA template encoding a riboswitch variant is flanked by "tag 1" and "tag 2" sequences that serve as primer annealing sites for polymerase chain reaction (PCR) (**Figure 7A**). The T7 promoter is located downstream of the "tag 1" and drives the transcription of the mRNA by the T7 RNA polymerase. The transcript encoded downstream of the T7 promoter contains the riboswitch variant in the 5' UTR followed by a ribosome binding site (RBS), a short open reading frame (ORF) encoding *gfp11*, and finally the "tag 2" after the stop codon (**Figure 7A**). The droplets contain a molecular beacon (OMB) that anneals to the "tag 2" region (3' UTR) of the mRNA activating the fluorescence of 6-carboxy-X-rhodamine (ROX) which serves as an indicator of active transcription. Consequently, only ROX+ droplets are sorted in order to avoid

recovering empty droplets lacking DNA template. The droplets also contain GFP1–10 protein fragment that noncovalently associates with GFP11 peptide encoded in the ORF, reconstituting the GFP fluorescence which reports the riboswitch output (Cabantous et al., 2005).
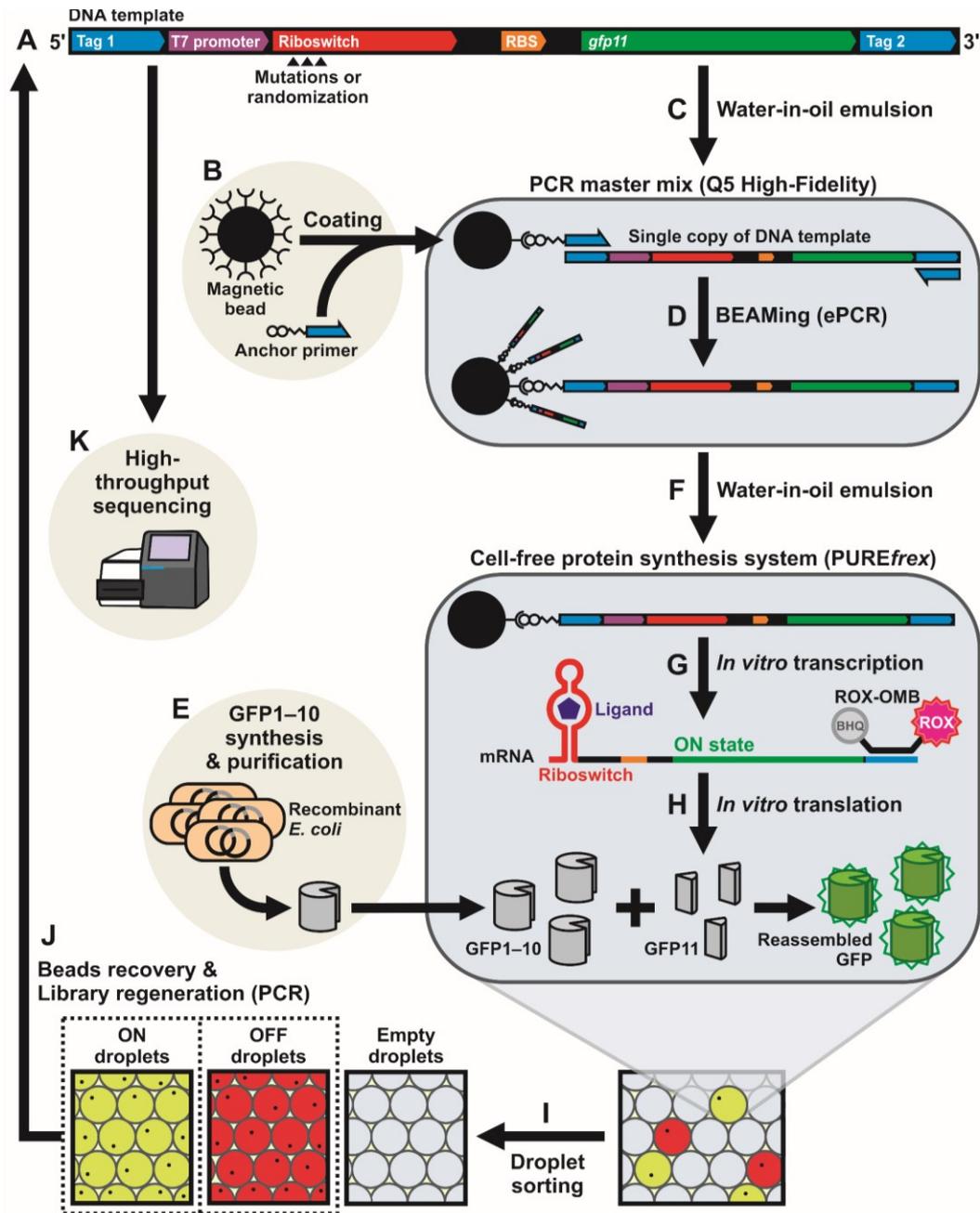
Figure 7. Overview of the high-throughput screening strategy for cell-free riboswitches. (**A**) DNA template encoding the riboswitch variants. (**B**) Streptavidin-coated magnetic microbeads are used to immobilize a biotinylated primer. (**C**, **D**) A single DNA template and a magnetic bead are encapsulated in a water-in-oil emulsion for clonal amplification of the DNA template on the magnetic bead (BEAMing). ePCR: emulsion PCR. (**E**) Recombinant GFP1–10 fragment is produced in *E. coli.* (**F**, **G**, **H**) The DNA template-immobilized magnetic beads are re-encapsulated in water-in-oil emulsions containing the CFPS system (PURE*frex* 1.0) and GFP1–10 fragment. Droplets that produce mRNAs yield ROX fluorescence. Droplets that contain riboswitch variants at ON state yield green fluorescence due to assembly of the split GFP fragments. ROX-OMB: ROX-labeled molecular beacon. (**I**) The droplets with ROX fluorescence are sorted for ON or OFF states of the riboswitch variants. (**J**) The magnetic beads are recovered from the sorted droplets and amplified by PCR for the next round of sorting. (**K**) The enriched riboswitch candidate sequences are analyzed by high-throughput sequencing. Reproduced from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

3.3. MATERIALS AND METHODS

For a detailed protocol of the methods see **Appendix 4: Protocols** (page **VI**).

3.3.1. STRUCTURE OF THE CONSTRUCTS

Each GFP11 construct consists of a 5' UTR sequence to which the anchor primer P0L and forward primer P3L anneal ("tag 1"), T7 promoter, riboswitch sequence, optimized ribosome binding site (RBS), *gfp11* gene, and a 3' UTR sequence ("tag 2") which serves as the binding site for the reverse primer P4L and the ROX-OMB molecular beacon. The eGFP constructs are similar to the GFP11 constructs except that the "tag 1" was shortened and the *gfp11* gene after the first 2 codons was replaced with the complete *egfp* gene. The first 2 codons of the *gfp11* gene were kept because they were usually part of the expression platform of the riboswitch.

3.3.2. DNA TEMPLATE CONSTRUCTION

All the DNA templates were amplified via polymerase chain reaction (PCR). The PCR was performed using Q5 High-Fidelity 2X Master Mix (NEB) with 0.5 µM of each primer unless noted otherwise. The DNA templates for the GFP11 constructs were assembled by PCR from 3 or 4 overlapping oligonucleotides. The randomized libraries were created as mentioned before using chemically synthetized degenerate oligonucleotides containing 3–6 Ns in the desired positions. For the eGFP reporter constructs, the constant region containing the eGFP coding sequence was first cloned into pMD20-T vector. Then, each DNA template was produced by PCR using the eGFP-containing plasmid as a template, a common reverse primer (P-eGFP-R), and a forward primer containing the T7 promoter followed by the respective riboswitch variant sequence. All PCR products were purified by silica columns (DNA Clean & Concentrator Kit-5, Zymo Research) or, when necessary, by agarose gel electrophoresis (Zymoclean Gel DNA Recovery Kit, Zymo Research).

3.3.3. GFP1–10 SYNTHESIS AND PURIFICATION

The *gfp1–10* gene fragment was amplified from pk-thiC#19-gfp(1–10) plasmid (Muranaka and Yokobayashi, 2010) by PCR and was subcloned into pTrcHis vector. The plasmid was transformed into *Escherichia coli* BL21 (DE3). Protein expression was induced by addition of IPTG (0.5 mM) after the bacterial culture reached an OD_{600}~0.6 in LB medium supplemented with 100 µg/ml ampicillin. The cells were further incubated for 3 h at 37°C before they were harvested and lysed. The His-tagged GFP1–10 protein was purified using Ni-NTA magnetic agarose beads (QIAGEN) according to the manufacturer's protocol. The protein was desalted using Amicon Ultra-0.5 ml (MWCO 3 kDa) centrifugal filters (Merck Millipore) and stored at -80°C in storage buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 20 mM KCl).

3.3.4. *IN VITRO* COMPARTMENTALIZATION (DROPLET GENERATION)

The water-in-oil emulsion droplets were generated by On-chip Droplet Generator (On-chip Biotechnologies) in 2D chip-800DG chips (material: COP) at 4°C (**Figure 8**). 5% (w/w) Pico-Surf 1 (Sphere Fluidics) in HFE-7500 3M Novec Engineered Fluid (Fluorochem) was used as the oil phase, while the aqueous phase was either a PCR mix (Q5 High-Fidelity 2X Master Mix, NEB) or a cell-free protein synthesis (CFPS) reaction mix (PURE*frex* 1.0, Gene Frontier). The sample and oil pressures were set to 58 kPa and 79 kPa, respectively. The throughput was approximately 2.2×10^5 droplets/min. The droplets produced were relatively monodisperse and their size was approximately 26 µm in diameter, roughly estimated through light microscopy.
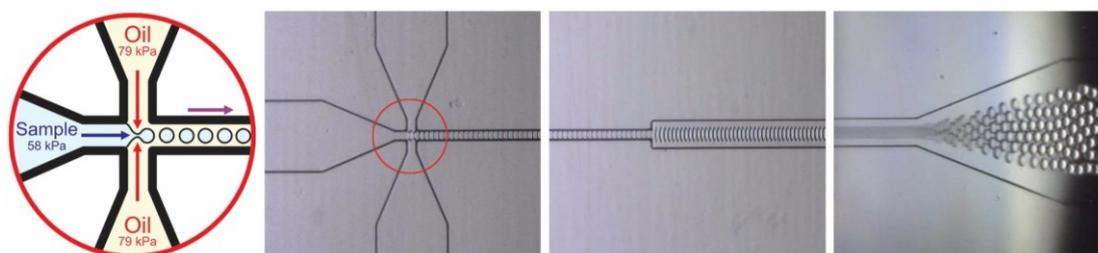
Figure 8. Production of water-in-oil microdroplets using a 2D chip-800DG microfluidics chip and a Droplet Generator (On-chip Biotechnologies).

3.3.5. SINGLE-TEMPLATE AMPLIFICATION ON MAGNETIC BEADS (BEAMING)

The template DNAs encoding the riboswitch variants were clonally amplified by emulsion PCR (ePCR) on magnetic beads according to the BEAMing protocol (Diehl et al., 2006) with few modifications. Streptavidin-conjugated magnetic beads (Dynabeads MyOne Streptavidin C1, 1 μm diameter, Invitrogen) were coated with the anchor primer P0L (see **Appendix 1: List of primers**, page **II**) as described previously (Diehl et al., 2006). The anchor primer was modified at the 5' end with a dual-biotin tag connected via an 18-atom hexa-ethyleneglycol spacer (iSp18) synthesized by IDT. Water-in-oil emulsion droplets were generated as described before using an aqueous phase containing the DNA template (50 fM), P0L-coated magnetic beads (~5×10^5/μl), PCR mix, and the primers (50 nM P3L and 4 μM P4L labeled with Texas Red (TexRd), see **Appendix 1: List of primers**, page **II**). The droplets were subjected to 35 cycles of PCR (two-step: 98°C for 10 s followed by 72°C for 10 s). To ensure that the majority of the beads are monoclonal, the DNA concentration was adjusted so as to yield ~10–30% positive beads (TexRd+), taking Poisson distribution into account (Collins et al., 2015, Lu et al., 2017). To ensure the amplification on the beads happens in every droplet with a DNA template, the beads were added in excess as so every droplet contains multiple beads. Upon completion of the PCR, the emulsion was broken using an anti-static gun (MILTY Zerostat 3), then the magnetic beads were collected using a home-made neodymium magnet stand and washed once with the breaking buffer (15 mM Tris-HC pH 7.5, 0.5 mM EDTA, 50 mM NaCl, 25 mM KCl, 0.5% Triton X-100, 0.5% SDS). The beads were further washed five times with TK buffer (20 mM Tris-HCl pH 8.4, 50 mM KCl) and suspended in TEK buffer (15 mM Tris-HCl pH 7.5, 25 mM KCl, 0.5 mM EDTA) for storage at 4°C.

3.3.6. CELL-FREE PROTEIN SYNTHESIS IN DROPLETS

Water-in-oil emulsion droplets were generated as described before using an aqueous phase that contains magnetic beads coated with the DNA encoding riboswitch-*gfp11*, PURE*frex* 1.0 reaction mix, GFP1–10 (50 μg/ml, approximately 26 μM), ROX-labeled 2'-O-methyl-RNA molecular beacon (ROX-OMB) (0.5 μM), and where applicable, histamine or ciprofloxacin at an appropriate concentration. To ensure that the majority of the droplets statistically contain only one riboswitch variant (one bead) or are empty, following a Poisson distribution, the beads concentration of the was empirically adjusted to yield ~10–30% transcriptionally active droplets (ROX+). The emulsion was incubated at 37°C for 4 h to allow CFPS reaction to proceed.

3.3.7. DROPLET SORTING

The droplets containing the CFPS reaction mix were sorted by fluorescence-activated droplet sorting (FADS) using a commercially available On-chip Sort instrument (On-chip Biotechnologies) and Chip-Z1001 microfluidics chips (material: COP). The sheath fluid used

was 0.1% (w/w) Pico-Surf 1 in HFE-7500, and low-density mineral oil (Sigma) was added to the collection reservoir to trap the sorted droplets and facilitate their recovery. The droplets were sorted according to ROX (FL-5, ex. 561 nm, em. 676±37 nm) and GFP (FL-2, ex. 488 nm, em. 543±22 nm) fluorescence (**Figure 9**). ROX fluorescence was used to detect the presence of transcriptional activity. GFP fluorescence was used to detect the riboswitch output.
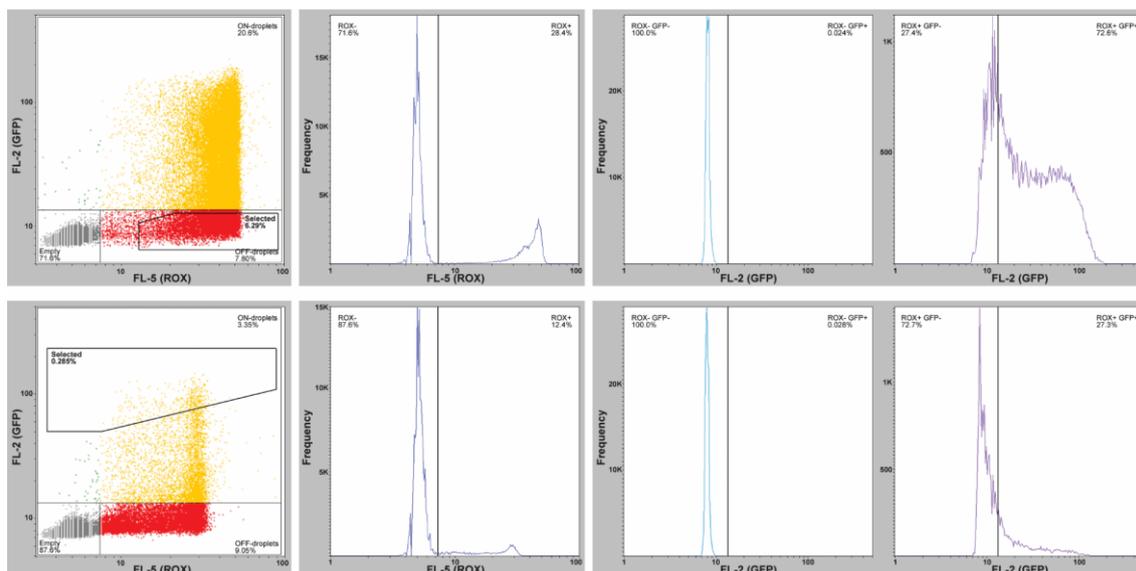
Figure 9. Examples of population thresholds and selection gates used during the OFF-droplet (top row) and ON-droplet sorting (bottom row). The example plots presented here correspond to the cycle 2 and 3 of the histamine ON-switch selection process. The gates labeled "Selected" (thicker border) correspond to the chosen selection gates. In the 2D plots, the vertical axis (log scale) corresponds to GFP fluorescence (FL-2 channel), and the horizontal axis (log scale) corresponds to ROX fluorescence (FL-5 channel). In the histograms, the vertical axis (linear scale) corresponds to the frequency, and the horizontal axis (log scale) corresponds to the respective fluorescence channel (FL-2 or FL-5 channels). Reproduced from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

The thresholds were established using the output of the empty droplets (GFP- ROX-) as the baseline. The empty droplets can be easily identified as they represent more than 70% of the droplets and always cluster at the bottom-left corner of the FL-2 vs. FL-5 plot. The gates for sorting were set according to the targeted riboswitch output (ON or OFF) and adjusted for the population distribution of each library. The gates were arbitrarily set to select the ROX+ droplets with the highest or the lowest GFP fluorescence, representing no more than 10% of the total droplet population (typically between 0.1% and 3%). The sorting gates were adjusted to increase the stringency of selection every cycle.

3.3.8. BEADS RECOVERY AND LIBRARY REGENERATION

The sorted droplets were carefully recovered from the collection reservoir and transferred to a new tube. TK buffer (30 µl) and blank magnetic beads (0.5 µl) were added to the tube to improve the recovery yield. The emulsion was broken using an anti-static gun (MILTY Zerostat 3), and the magnetic beads were collected using a home-made magnetic stand. The beads were washed once with the breaking buffer and three times with TK buffer. The washed beads were resuspended in 12 µl of the PCR reaction mix (Q5 High-Fidelity PCR Master Mix containing primers P3L and P4L). Thirty cycles of PCR (two-step: 98°C for 10 s followed by 72°C for 10 s) was performed to regenerate the riboswitch library. The PCR product was purified using a silica column or agarose gel electrophoresis. The purified DNA was used as template for the BEAMing of the next cycle.

3.3.9. MOCK SORTING

Droplets simulating the ON and the OFF outputs of the riboswitch were generated using DNA templates with a "strong" (P_{T7}-RBS-GFP11) or a "weak" RBS (P_{T7}-wRBS-GFP11), respectively. Additionally, the DNA template for the ON-droplets contains an EcoRI restriction site while that of the OFF-droplets contains a SacI site. The DNA templates were mixed in approximately 5:95 or 95:5 ratios and were used to produce mock libraries by BEAMing and CFPS reactions as described before. The droplets were sorted, and the enrichment efficiencies were estimated by flow cytometry, fluorescence microscopy, and restriction digestion of the DNA recovered from the sorted droplets. The sorting efficiency was estimated from the droplet statistics by dividing the number of sorting events of the desired droplets by the total number of the droplets detected by the sorter within the selected sorting gate. The purity of the sorting was estimated from the proportion of droplets exhibiting the desired fluorescence output within the droplet population recovered after the sorting.

3.3.10. RESTRICTION DIGESTION ANALYSIS

Because the ON and OFF templates used in mock sorting tended to form a significant amount of heteroduplex products after standard PCR, the following steps were included to reduce such products before restriction digestion analysis. The templates (0.5 nM) were re-amplified with 10 cycles of PCR (98°C for 10 s, 72°C for 10 s) with primers P3L and P4L. The product was then diluted 10-fold with fresh PCR reaction mix of the same composition except with 2.5× DNA polymerase and was incubated at 98°C for 2 min followed by 72°C for 2 min. The PCR products were purified with silica columns and 20 ng of each of the PCR products was digested with EcoRI-HF (New England Biolabs) and SacI-HF (New England Biolabs) according to the manufacturer's instruction. Native polyacrylamide gel electrophoresis (PAGE, 8%) was performed to separate the digested samples. The gels were stained with SYBR Gold (Invitrogen) and photographed by LuminoGraph WSE-6100Z imager (ATTO). The ratios of the ON and the OFF templates before and after sorting were estimated from the band intensities analyzed by ImageJ 1.52p.

3.4. ISSUES ADDRESSED DURING THE DEVELOPMENT PROCESS

During the development of this methods there were several issues that needed to be solved. Here I present the most important ones and how they were addressed.

3.4.1. SIGNAL AMPLIFICATION

Because a single copy of DNA template is not enough to produce reliable protein expression and output signal within a reasonable time frame, it was decided to clonally amplify the templates on magnetic beads through a previously described protocol named BEAMing (Diehl et al., 2006). Similar to droplet digital PCR (ddPCR), single template molecules are encapsulated and amplified inside microdroplets (emulsion PCR), with the difference that the amplification occurs on the surface of co-encapsulated magnetic beads instead of in free solution. This way, millions of magnetic beads individually coated with multiple copies of a unique template from a pool of thousands of variants can be obtained. Reportedly, the beads produced with BEAMing accurately reflect the diversity of the template population and can be used to determine the frequency of a specific mutations within a DNA population (Diehl et al., 2005, Diehl et al., 2006, Holdhoff et al., 2011, Lauring and Park, 2011, Chen et al., 2013, Diehl and Smergeliene, 2013, Denis et al., 2017). Because BEAMing (PCR) and CFPS conditions are not compatible, the sorting protocol requires two separate instances of emulsification; using

magnetic beads is helpful for recovering the templates after each step and for transferring them to the next reaction, while preserving the monoclonality of the amplicons, and thus the individuality of the variants.

3.4.2. INDIVIDUALIZATION OF VARIANTS

For this screening method to work, each variant needs to be separated from each other so they can be amplified and assessed individually. The template DNA molecules at very low concentrations can be considered as discrete particles; hence, when partitioned into a set number of uniform compartments (i.e., monodisperse microdroplets), they are stochastically distributed following a Poisson distribution of randomly dispersed objects (Collins et al., 2015, Lu et al., 2017). Uniform dimensions of such compartments is important for achieving a Poisson distribution with DNA molecules (Zhang and Noji, 2017, Zhang et al., 2019).
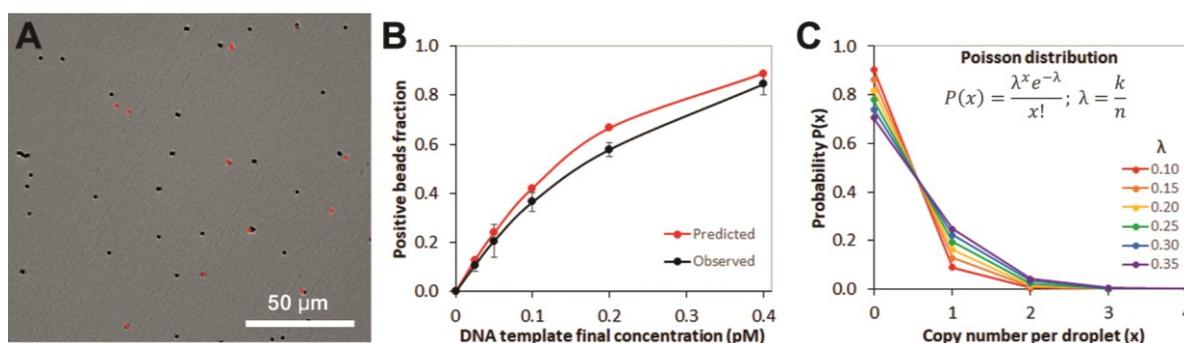
Figure 10. Empirical determination of the ideal DNA template concentration range for clonal amplification using BEAMing. (**A**) Example micrograph (merged: Texas Red + bright field) of magnetic beads after BEAMing with 50 fM of DNA template (P_{T7}-RBS-GFP11). Because a reverse primer labeled with Texas Red (TexRd) was used, beads with successfully amplified template on the surface exhibit red fluorescence. (**B**) Amplification-positive beads fraction (TexRd+) in relation to the template concentration used. The ideal DNA template concentration range was estimated to be around 50 fM final concentration ($\lambda \sim 0.274$). Error bars represent the standard deviation of several micrographs from two independent replicates. (**C**) Poisson probability distribution (P(x)) simulation for ~10–30% positive droplets fraction (i.e., droplets containing one or more DNA molecules). x: number of DNA molecules per droplet, λ: Poisson parameter Lambda, k: total number of events (DNA molecules), n: number of partitions (droplets).

To statistically ensure that the majority of the droplets during the amplification with BEAMing contain only one riboswitch variant (one DNA template molecule) or are empty, the concentration of the DNA template has to be adjusted to yield ~10–30% amplification-positive droplets (**Figure 10**). The ideal template concentration was confirmed by using a reverse primer labeled with Texas Red (TexRd) for the BEAMing, and then counting the fraction of beads that exhibit red fluorescence (TexRd+). For a droplet size of ~26 µm diameter (~9 pl), the DNA template must be diluted down to femtomolar concentrations (~50 fM). The slight difference between "Observed" and "Predicted" (**Figure 10B**) is probably caused by handling errors during the dilution of the samples, or due to minor inaccuracies in the initial DNA concentration measurements. For the CFPS reactions, the same principle was applied, so the ideal beads concentration should be around ~2.5×10^4 beads/µl; however, because the actual concentration of DNA-coated beads is difficult to estimate, the concentration was empirically adjusted to yield ~10–30% transcriptionally active droplets (ROX+).

3.4.3. CONSTRUCT LENGTH AND REPORTER GENE

BEAMing works well with template lengths in the range of 50 to 150 bp, but the amplification efficiency drops sharply when the length exceeds 200 bp, probably caused by steric effects and

molecular crowding on surface of the beads (Diehl et al., 2006, Mercier and Slater, 2005). This excludes the possibility of using canonical reporter genes such as GFP, mCherry, or any enzyme. This problem was addressed by using the split GFP system (Cabantous et al., 2013, Cabantous et al., 2005). It has been shown that GFP protein can be split into two fragments; a short β-strand located in the C-terminus (GFP11 peptide) and the remaining β-barrel structure (GFP1–10 protein). While the individual fragments are non-fluorescent on their own, the two fragments can self-assemble to restore the GFP fluorescence. The final DNA constructs now fall within an acceptable size range as the leader sequences (promoter + riboswitch + ribosome binding site) are normally withing 150 bp, and the GFP11 sequence is just a 17 aa long peptide (54 bp, including the stop codon). The large fragment of the reporter gene (GFP1–10 protein, ~230 aa) can simply be produced separately and added into the reaction mix.

3.4.4. OIL AND EMULSION STABILITY

There are numerous formulations of oils and surfactants/emulsifiers reported for emulsion PCR (ePCR) and other droplets applications (Nakano et al., 2003, Diehl et al., 2006, Marcoux et al., 2011, Schutze et al., 2011, Pandit et al., 2015, Kaminski et al., 2016, Witt et al., 2017, Zubaite et al., 2017, Chai, 2019, Iacumin et al., 2020, Siu et al., 2021). A hydrofluoroether (HFE)-based oil (i.e., Novec 7500) and a fluorinated surfactant (i.e., Pico-Surf 1) were selected for producing the droplets because of their low viscosity (which helps during droplet generation and droplet sorting), compatibility with the material of the microfluidics chips (i.e., cyclo olefin polymer, COP), good thermal conductivity (ideal for fast temperature changes during PCR), good gas solubility (which allow gas-exchange during the reactions), being inert and biocompatible (so it does not interfere with the PCR or CFPS reactions). HFE oils are also relatively safe to use as they are non-flammable, non-ozone-depleting, and exhibit low toxicity.
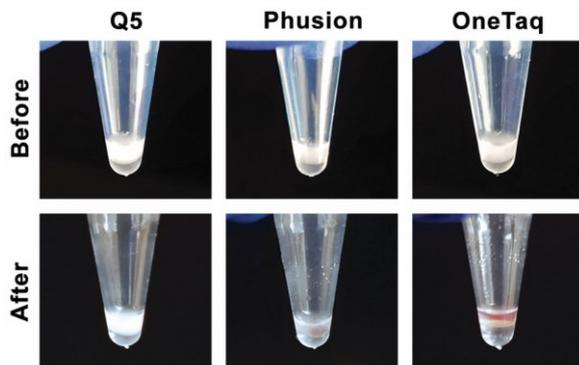
Figure 11. Emulsion PCR (ePCR) reactions before and after thermal cycling under standard BEAMing conditions (see **3.3.5. Single-template amplification on magnetic beads (BEAMing)**, page 33) using Q5 High-Fidelity 2X master mix (NEB), Phusion High-Fidelity 2X Master Mix with HF Buffer (NEB), and OneTaq 2X Master Mix with Standard Buffer (NEB). Q5 was the only master mix that produced a stable emulsion in HFE-7500/Pico-Surf 1.

Although the original protocol employs a different polymerase (Diehl et al., 2006), Q5 High-Fidelity 2X master mix (NEB) was chosen for the BEAMing reaction because of its high processivity and fidelity, and also because it produced an emulsion that was stable through all the PCR cycles under the experimental conditions stablished in the current protocol without additional additives. It has been demonstrated that the emulsion stability is particularly important for reducing the formation of non-specific amplification by-products (Witt et al., 2017) caused by the extremely low concentrations of template during single-molecule amplification reaction in the ePCR (Nakano et al., 2003). Other PCR master mixes and DNA polymerases, such as Phusion High-Fidelity (NEB) and OneTaq (NEB), were also tested but they all failed to produce stable emulsions in HFE-7500/Pico-Surf 1 (**Figure 11**). Although the composition of Q5 buffer is not disclosed, it is very likely that it contains Tween-20 and/or bovine serum albumin (BSA) which are commonly used additives to enhance PCR reaction, which are also known for increasing the stability of the emulsions during PCR in fluorinated oils (Williams et al., 2006, Loveday et al., 2021). Fortunately, there were no problem during

the CFPS reaction as the emulsion of PURE*frex* 1.0 and PURE*frex* 2.0 were relatively stable at 37°C even after incubation for several hours undisturbed, although coalescence of droplets can still occur with a higher incidence the longer the incubation proceeds.

3.4.5. EMULSION BREAKING

Initially, breaking of the emulsion was performed using Pico-Break 1 (Sphere Fluidics) as suggested by the oil-surfactant manufacturer; however, it was readily noticed that sometimes PCR or CFPS did not work properly even after washing the beads. It has been demonstrated that chemical demulsifiers such as 1H,1H,2H,2H-perfluorooctanol (PFO), present in Pico-Break 1, can actually interfere with or inhibit downstream reactions such as PCR (Karbaschi et al., 2017). Therefore, an antistatic-gun was employed instead to gently disrupt and break the emulsion aided by low concentrations of soft surfactants such as triton X-100 and SDS.

3.4.6. DETECTION OF RIBOSWITCHES IN OFF-STATE

This method relies on iterative cycles of droplet sorting, alternating between presence and absence of the ligand, and selecting droplets according to the expected output. Therefore, enrichment of the desired switches depends on the sorting efficiency of the droplets in the desired state (ON or OFF). One of the technical challenges in screening genetic devices in droplet-based methods is the efficiency of OFF-sorting. While selecting for ON-droplets is straightforward as the activity of the switch is directly measured through the expression of the reporter gene (GFP+), isolating rare OFF-droplets (GFP-) is complicated by the large proportion (>70%) of empty droplets without DNA template which are also non-fluorescent (GFP-). Therefore, sorting for OFF droplets based on the gene expression of the reporter (GFP-) would result in recovery of mostly empty droplets. To address this problem, a 2'-O-methyl-RNA molecular beacon (OMB) labeled with 6-carboxy-X-rhodamine (ROX) and Black Hole Quencher-2 (BHQ-2) was designed to anneal to the "tag 2" region (3' UTR) of the mRNA. When the ROX-OMB anneals to the mRNA, its fluorescence is activated and serves as an indicator of active mRNA transcription. This makes it possible to differentiate three droplet populations: ON-droplets (ROX+ GFP+), OFF-droplets (ROX+ GFP-), and empty droplets (ROX- GFP-) (**Figure 12**).
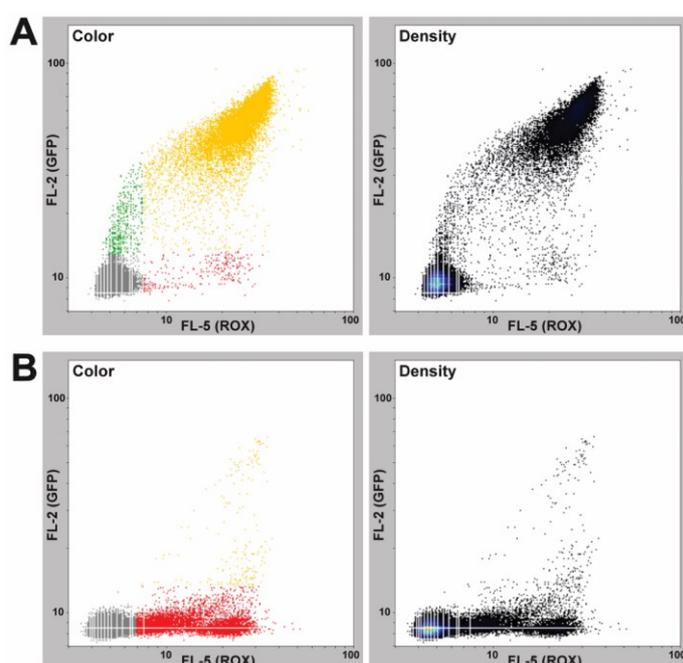
Figure 12. Examples of the dual-color fluorescence density plots of the droplet sorting experiments. The vertical axis (log scale) corresponds to GFP fluorescence (FL-2 channel), and the horizontal axis (log scale) corresponds to ROX fluorescence (FL-5 channel). Three different populations can be observed: empty droplets (grey: ROX-/GFP-, bottom left), OFF-droplets (red: ROX+/GFP-, bottom right), and ON-droplets (yellow/green: ROX+/GFP+, top right). (**A**) A population with high abundance of ON-droplets from the mock sorting experiment. (**B**) A population with high abundance of OFF-droplets from the mock sorting experiment. Reproduced from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

3.5.   VALIDATION OF THE SORTING METHOD THROUGH MOCK SORTING

To validate the sorting strategy, a mock sorting experiment was designed and performed using two DNA templates that simulate ON and OFF riboswitch outputs. The ON template contains a canonical RBS sequence that robustly translates the encoded gene ("strong" RBS). The OFF template contains a weakened RBS sequence with low translation efficiency ("weak" RBS). Additionally, the DNA template for the ON-droplets contains an EcoRI restriction site while that of the OFF-droplets contains a SacI site (**Figure 13A**).
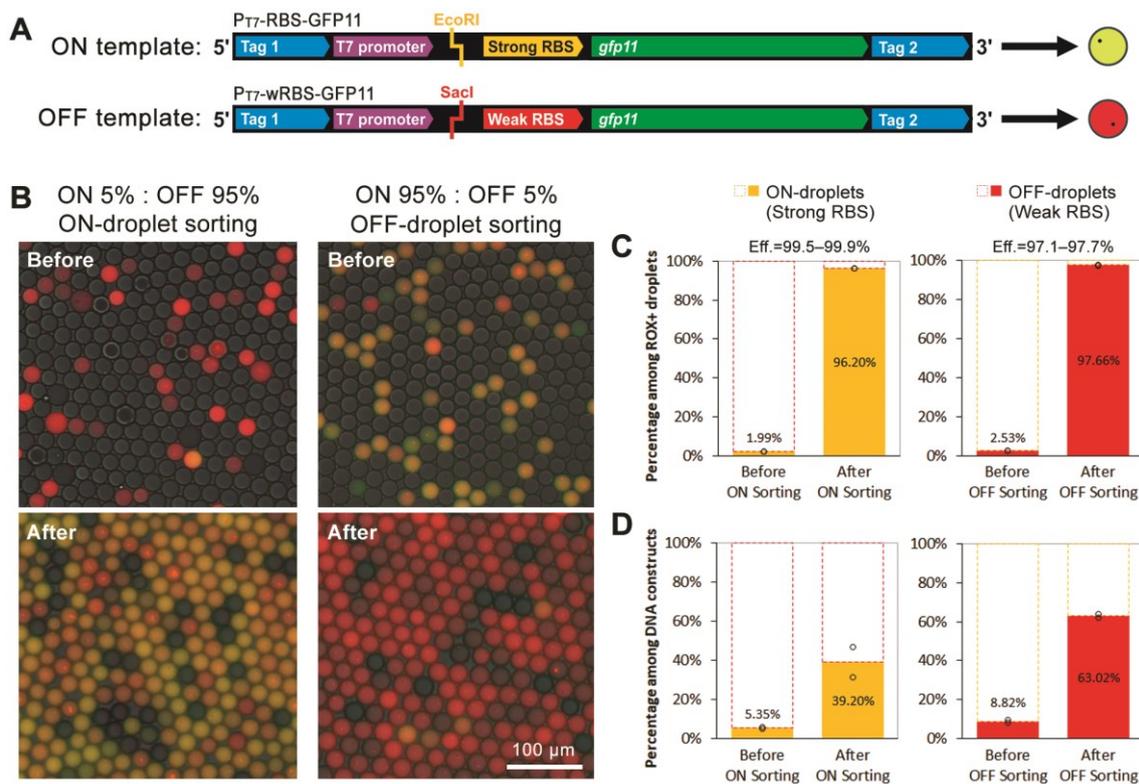
Figure 13. Mock sorting of ON- and OFF-droplets. (**A**) The ON template (P_{T7}-RBS-GFP11) contains a "strong" RBS that regulates translation of the GFP11 peptide. It also contains an EcoRI restriction site upstream of the RBS. The OFF template (P_{T7}-wRBS-GFP11) contains a "weak" RBS to simulate riboswitch variants in an OFF state. It also contains a restriction site recognized by SacI upstream of the RBS. (**B**) Fluorescence micrographs (merged: GFP + ROX + bright field) before and after droplet sorting for ON- (left) or OFF-droplets (right). (**C**) Abundance of ON- or OFF-droplets before and after sorting as estimated by fluorescence microscopy and flow cytometry. (**D**) Abundance of ON or OFF templates before sorting and after sorting and subsequent PCR, as estimated by restriction digestion and gel electrophoresis (**Figure 14**). Two independent mock sorting experiments were performed. The open circles in **C** and **D** represent the two measurements, and the bars represent the average values. Eff.: sorting efficiency or recovery rate of the desired droplets. Reproduced from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

The ON and OFF templates were mixed in approximately 5:95 and 95:5 ratios to prepare mock libraries. The DNA mixtures were clonally amplified on magnetic beads by BEAMing, and then were used to make droplets containing single beads. The droplets were sorted for the minor member of the mock library. The sorted droplets were then imaged by fluorescence microscopy. The template DNAs were also recovered from the sorted droplets by PCR, and the relative abundance of the two species (ON and OFF) was estimated by different methods.

The sorting efficiency of the instrument was around 99.5–99.9% and 97.1–97.7%, with an apparent purity of about 79.9–82.5% and 86.5–89.3%, for ON- and OFF-droplet sorting,

respectively. Fluorescence images of the droplets before and after sorting showed that ON- and OFF-droplets were enriched in average by 48- and 38-fold, respectively (**Figure 13B, C**). It is worth noting that more than 70% of the initial droplets were empty to ensure that the majority of the droplets contained one magnetic bead (or none) based on the Poisson distribution. The sorting efficiently eliminated these empty droplets because a molecular beacon was included to detect droplets that exhibit active transcription.

Using the magnetic beads isolated from the sorted beads, the template DNA was recovered and amplified by PCR. Enrichment of the ON and OFF DNA templates was estimated by restriction digestion analysis of the PCR products (**Figure 14**). Enrichment of 7.3- and 7.1-fold of the ON and OFF templates, respectively, was observed (**Figure 13D**). The lower enrichment of the DNA templates compared to the droplets is likely due to some droplets containing multiple magnetic beads, as the beads tend to aggregate in spite of all the efforts to disperse them (**Figure 15**). Therefore, some OFF templates may be co-sorted along with ON templates within the same droplet during the ON-sorting process. Conversely, some OFF templates may not be recovered due to coexisting ON templates within the same droplet during the OFF-sorting step. Although the presence of heteroduplexes formed after PCR was minimized (see **3.3.10. Restriction digestion analysis**, page **35**), their unavoidable persistence may have affected the restriction enzyme assay results. Nevertheless, it was confirmed that DNA templates in a CFPS system can be enriched for either ON or OFF gene expression using this droplet sorting method.
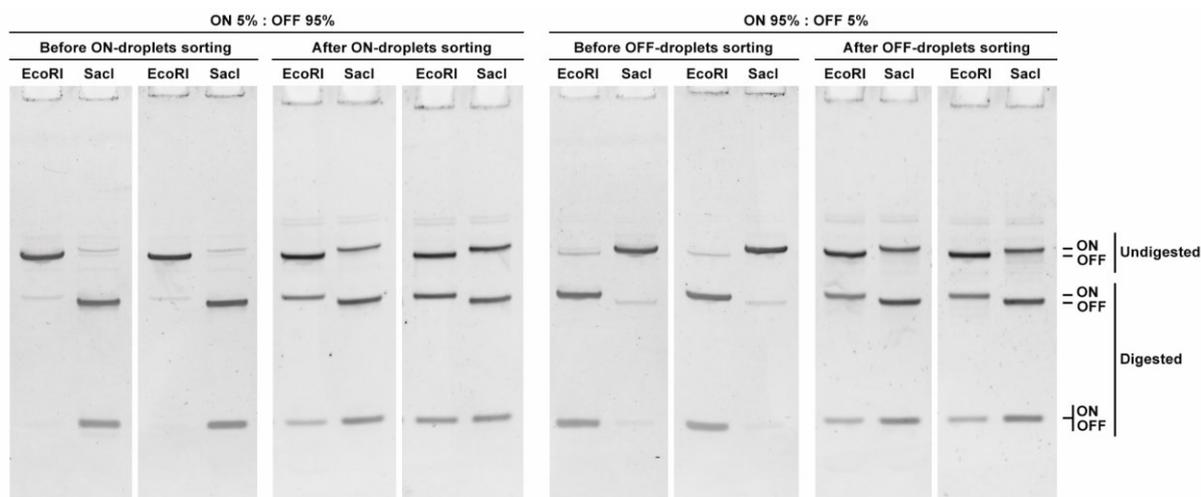
Figure 14. Native PAGE (8%) of the DNA templates recovered from mock sorting after restriction digestion. The gels were stained by SYBR Gold. The ON templates (P_{T7}-RBS-GFP11) are digested by EcoRI, and the OFF templates (P_{T7}-wRBS-GFP11) are digested by SacI. Reproduced from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

3.6. LIMITATIONS, AREAS FOR IMPROVEMENT, AND POTENTIAL MODIFICATIONS TO THE METHOD

The technique presented in this work is a proof-of-concept technique and there is a lot of room for improvement. Fortunately, some of drawbacks and limitations it currently faces can be addressed with further development and optimization.

3.6.1. THROUGHPUT

The droplet sorter employed (On-chip sort instrument) performs the sorting via pulses of air pressure to hydrodynamically redirect the flow of the sheath fluid, instead of dielectrophoresis (DEP) most commonly employed in more sophisticated microfluidic devices (Fiedler et al.,

1998, Hu et al., 2005, Shields et al., 2015, Lee et al., 2016, Schutz et al., 2019, Zhang et al., 2021). In DEP-based devices, a nonuniform electric field induces the polarization of the droplets and forces their movement towards a desired direction (Pamme, 2007); however, if two or more droplet are close together, the application of an electric field can cause the coalescence of the droplets (Fallah-Araghi et al., 2012, Karbaschi et al., 2017, Zhou et al., 2020) (see **3.4.5. Emulsion breaking**, page **38**). Using pulses of air pressure is relatively gentler to the droplets but require mechanical actuators (pumps and valves) to timely dispense the air pressure, which compromises the accuracy and speed of the sorting. In consequence, the throughput of this method is currently limited by that of the droplet sorting machine, which is approximately 6.3×10^5 droplets/h under the current settings—a few orders of magnitude lower of what have been reported using more sophisticated devices (Fiedler et al., 1998, Hu et al., 2005, Lee et al., 2016). Due to the requirement of maintaining ~10–30% magnetic beads encapsulation rate to ensure single variant per droplet, the effective throughput decreases to $\sim 6 \times 10^4$–1.9×10^5 droplets/h. Nevertheless, 1–6 h of sorting can achieve sufficient coverage (10–40×) of a library with 6 randomized nucleotides (4096 variants). While this has been sufficient for the riboswitch libraries tested, higher throughput may be desirable for more complex systems and larger libraries, for example, Boolean logic gate riboswitches (Sharma et al., 2008). The throughput of the On-chip droplet generator ($\sim 1.3 \times 10^7$ droplets/h), on the other hand, is currently high enough for our needs, but it also may not be sufficient for very large libraries.

Until better instruments become more widely available, large libraries can still be processed with the current settings through several hours and multiple sessions, albeit very time-consuming. Pre-selection rounds could be performed to reduce the size or complexity of the library. Alternatively, the number of variants per droplets could be increased and then rely on probabilities and statistical analysis to estimate the presence of functional riboswitches. Of course, further improvements in throughput and sorting efficiency may be possible by adapting custom chips and specialized instruments and techniques such as pico-injection (Abate et al., 2010, Ryckelynck et al., 2015), droplet fusion/coalescence (Fidalgo et al., 2007, Mazutis et al., 2009, Xu et al., 2011, Fallah-Araghi et al., 2012, Ryckelynck et al., 2015, Autour et al., 2019, Zhou et al., 2020), dielectrophoresis-based droplet concentration (Han et al., 2017), femtoliter droplet arrays (Zhang et al., 2019), and tandem or parallel multichannel sorting (Caen et al., 2018, Vyawahare et al., 2021) that are yet to be commoditized.

3.6.2. SORTING EFFICIENCY

This technique's effectiveness depends on the sorting efficiency of the instrument and on the ability of generating droplets containing monoclonal templates. For the former, the instrument employed have to compromise sorting speed (and throughput) for better accuracy, and vice versa, so finding the right balance is important. Furthermore, the size of the droplet is also an important factor, as the smaller the size the more individual positive droplets can be analyzed per volume of reaction; however, it affects the sorting efficiency due to the geometry of the chip's channels and the non-uniform flow of the droplets. Undoubtedly, using a more capable and accurate microfluidic device or technique, would greatly improve the sorting efficiency.

Regarding the droplet population, an ideal Poisson distribution is very hard to achieve consistently. First, it is difficult to accurately dilute the templates or beads when working with such low concentrations as the ones required for obtaining droplets with a single template or a single bead (see **3.4.2. Individualization of variants**, page **36**). Moreover, magnetic beads tend to stick together in spite of every effort to disperse them (**Figure 15**); using chemicals such as detergents may affect the CFPS reaction or the droplet stability, and physical methods such as sonication may damage the DNA templates. And interesting approach that could

minimize the problem of selecting droplet containing multiple beads is by analyzing and sorting only the beads, instead of the droplets. For that, it would be necessary to capture and amplify the DNA template, the transcribed mRNAs, and the synthetized proteins on the surface of the same bead, avoiding cross-contamination with other beads, in a similar way as done with the RNA-Capturing Microsphere Particles (R-CAMP) system (Endoh et al., 2019) or the bead display system (Nord et al., 2003, Wang et al., 2012, Diamante et al., 2013, Huang et al., 2013, Paul et al., 2013, Wang et al., 2014, Mankowska et al., 2016, Gordon et al., 2019, Lindenburg and Hollfelder, 2021).
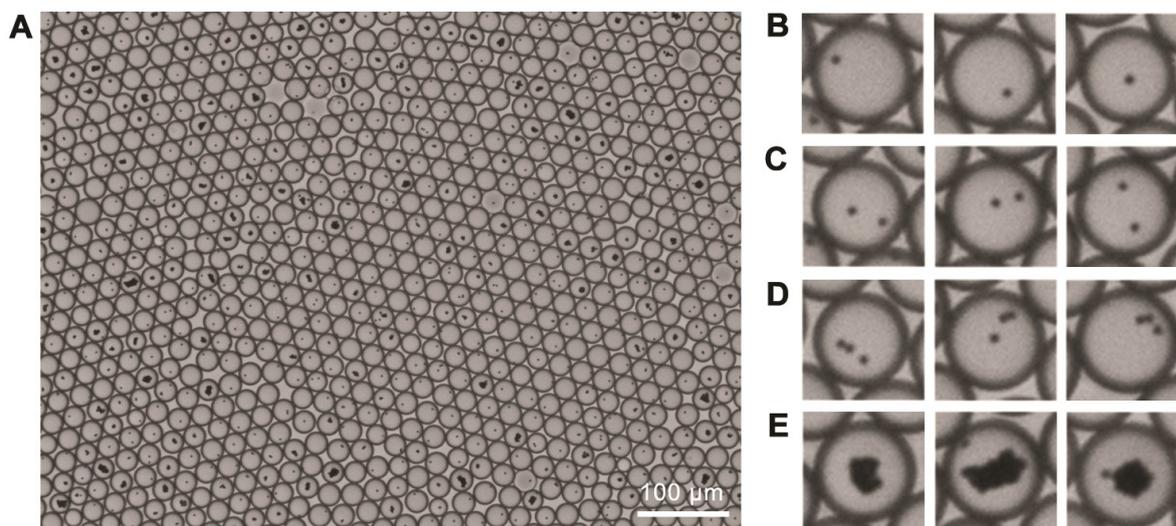
Figure 15. An example of aggregation of the magnetic beads inside droplets. (**A**) Droplets recovered after ON-selection (mock sorting experiment) viewed under a light microscope (bright field). Magnification of the droplets containing (**B**) a single bead, (**C**) two beads, (**D**) multiple isolated beads, and (**E**) aggregated beads. Adapted from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

Additionally, randomized libraries produced or amplified by conventional PCR methods will inherently contain hybrid heteroduplexes (Thompson et al., 2002, Michu et al., 2010), even after gel extraction or conditioning processes are performed in an attempt to reduce them (**Figure 16**). Such heteroduplexes carry at least two different variants in a single DNA molecule and ultimately decrease the overall selection efficiency as unwanted templates will be co-selected with the desired ones. Finally, stochastic effects due to the micro-compartmentalization can lead to differences in the amount of enzymes inside the droplets (Nakano et al., 2003), or differences in the amount of amplified template on the beads, creating artifacts during the BEAMing, CFPS, and droplet sorting.
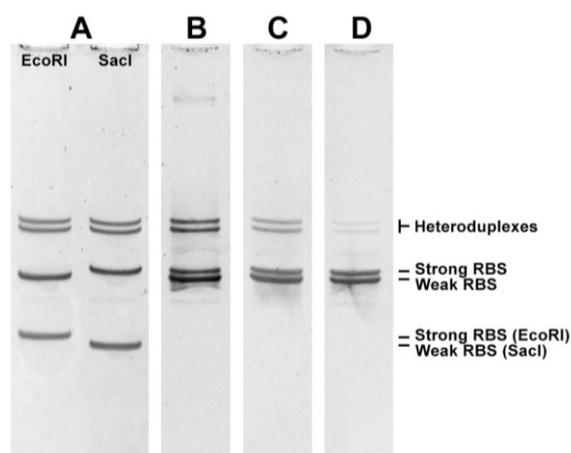
Figure 16. Native PAGE (8%) of the amplification products of a 1:1 mixture of "strong" (P_{T7}-RBS-GFP11) and "weak" RBS (P_{T7}-wRBS-GFP11) constructs, showing the reduction of heteroduplexes formed during PCR amplification of 0.5 nM DNA template mix. The gels were stained by SYBR Gold. (**A**) Standard 30-cycles PCR, purified by agarose gel extraction, and digested by EcoRI or SacI restriction enzymes. (**B**) Standard 30-cycles PCR, purified by silica column. (**C**) 10-cycles PCR. (**D**) 10-cycles PCR followed by 10-fold dilution and a single elongation step (98°C for 2 min, 72°C for 2 min) with fresh PCR reaction mix and 2.5× DNA polymerase, as described in **3.3.10. Restriction digestion analysis** (page **35**).

3.6.3. DNA TEMPLATE SIZE

The DNA template amplified on beads through BEAMing is limited to sizes under 200 bp, as the amplification efficiency drops sharply when it exceeds that length, probably due to steric constrains on the bead surface (Diehl et al., 2006, Mercier and Slater, 2005). Similar techniques for single-template amplification on beads could be employed to produce longer DNA templates, such as the Ion OneTouch instrument (Thermo Fisher Scientific) able to amplify fragments up to 400 bp on non-magnetic beads intended for ion torrent sequencing (Ion PGM and Ion S5 systems, Thermo Fisher Scientific), which has been successfully adapted by Endoh et al. to produce RNA-Capturing Microsphere Particles (R-CAMP) for fluorogenic aptamer selection and optimization (Endoh et al., 2019).

In addition, it has been known that interactions at the droplet interface tend to be dominant during emulsion PCR (ePCR), in particular the adsorption of the polymerase at the interface which leads to inefficient amplification (Pandit et al., 2015). This has been somewhat alleviated by using different kinds of surfactants; however, DNA amplification with conventional ePCR remains challenging, as long amplicons (>500 bp) cannot be amplified efficiently in droplets and the integrity of the droplets can be affected by the high temperatures required for PCR (Zubaite et al., 2017). As an alternative, it has been suggested the use of isothermal amplification (ITA) for producing long DNA templates from single molecules in emulsion. As demonstrated by Zubaite et al., single DNA molecules can be condensed into DNA-magnesium-pyrophosphate microparticles through isothermal Multiple Displacement Amplification (MDA), and then be used as templates for CFPS reactions (Zubaite et al., 2017).

With less restrictive size limitation for the DNA template, fluorogenic aptamers could be included to label the transcribed mRNA (Paige et al., 2011, Pothoulakis et al., 2014, Warner et al., 2014, Ouellet, 2016), or different reporter genes could be employed, such as bigger split fragments from other protein-fragment complementation systems, or even whole enzymes.

3.6.4. SIMULTANEOUS TEMPLATE AMPLIFICATION AND CFPS REACTION

This technique currently requires two separate steps of emulsification (droplet generation), one for the DNA amplification on beads, and one for the CFPS reaction and droplet sorting. The emulsion breaking, washing, and re-emulsification steps between the emulsion PCR (BEAMing) and CFPS entails the loss of sample and increases the chance of producing droplets containing 2 or more different DNA templates, because of the tendency of the magnetic beads to stick together or onto the surfaces. If template amplification and CFPS could be combined in a single reaction, the protocol could be simplified, and the cross-contamination could be reduced. Unfortunately, PCR and CFPS reaction are incompatible due to the chemical composition of the buffers and the different temperatures of the reactions.

Isothermal amplification (ITA) of templates is a plausible option as it usually proceed at similar temperatures as the CFPS reaction. Diverse ITA methods, such as Multiple Displacement Amplification (MDA) (Zubaite et al., 2017), Rolling Circle Amplification (RCA) (Hadi et al., 2020), and Nucleic Acid Sequence-Based Amplification (NASBA) (Pardee et al., 2016a, Takahashi et al., 2018, Ma et al., 2018a) have been employed for amplifying templates prior to CFPS reactions. However, in all cases, the ITA and CFPS reactions occurred in separate instances. It appears that the reaction conditions of the different ITA techniques available are not very compatible with the CFPS systems. Most recently, Sakatani et al. managed to perform *in vitro* evolution of phi29 polymerase using RCA coupled to a CFPS system (Sakatani et al., 2019); however, the total reaction time of each RCA-CFPS reaction was four times longer than

a standard CFPS reaction (up to 16 h). In addition, the RCA reaction was the selection marker and thus occurred after the CFPS. Nevertheless, they found several phi29 polymerase mutants that exhibit higher rolling-circle activity that the wild type when expressed in the CFPS system. Those highly active mutants could potentially enable the future development of more efficient ITA-coupled CFPS reactions.

Alternatively, by using a much more sophisticated microfluidics system as the one employed by the Griffiths group (Fallah-Araghi et al., 2012, Ryckelynck et al., 2015), it is possible to separately perform a standard single-template PCR reaction in emulsion, and then fuse each individual droplet with a fresh droplet containing CFPS reaction components without breaking the original emulsion, preserving this way the monoclonality of the templates.

3.6.5. CELL-FREE PROTEIN SYNTHESIS SYSTEM

The CFPS in droplets is currently performed as a batch reaction, with a limited amount of substrates, enzymes, and ribosomes. It has been shown that PURE systems, as well as cell extracts, are prone to become saturated, probably due to competition of the different metabolic processes for the limited resources, which unevenly depletes the substrates and creates a disproportion between DNA template, mRNA, and protein produced at different time points (Chushak et al., 2021). Polyacrylamide hydrogel-based artificial cells capable of long-term CFPS reaction in a continuous feed system developed by the Zheng group (Lai et al., 2020, Zhou et al., 2018) could be an alternative to the current water-in-oil droplets system.

Naturally, since all CFPS systems available are constructed from biological sources, they are sensitive to the same inhibitors that directly affect transcription and translation processes, such as chloramphenicol (ribosome inhibitor), T7 lysozyme (T7 RNA polymerase inhibitor), or rifampicin (bacterial RNA polymerases inhibitor). Unless orthogonal synthetic components for the CFPS are created, some chemicals will remain incompatible with the current CFPS systems.

It is important to mention that this technique is not restricted to using only reconstituted CFPS systems (i.e., PURE system). In theory, this method should work with any kind of CFPS system, which is particularly appealing because using other CFPS systems with modified parameters or based on cell extracts—especially the ones obtained from uncommon organisms—would allow the engineering of riboswitches and other genetic devices under different metabolic contexts.

3.6.6. TRANSCRIPTION AND TRANSLATION REPORTERS

The reporters used for assessing transcription (ROX-OMB) and translation (split GFP) can become saturated under the current conditions, with the maximum output being limited by the initial concentrations of ROX-OMB oligonucleotide and GFP1–10 protein. Although, the current concentrations are enough for a semi-quantitative assessment of the riboswitch activity prior and during the sorting, it cannot discriminate riboswitches with very high expression levels, as the reporters quickly reach the maximum output. There are multiple alternative solutions that can be implemented if that kind of differentiation is really necessary. For instance, increasing and optimizing the concentrations of those components could help to increase the dynamic range of the signals, keeping in mind that high concentrations of the reporters could interfere with the gene expression process. For instance, adding 5 times more ROX-OMB (2.5 μM instead of 0.5 μM) seems to be enough to prevent saturation when using 20 nM of template (**Figure 17**), at least within 4 hours of incubation, but the eGFP signal seems to be lower. One

big limitation, however, is that ROX-OMB and GFP1–10 protein are unfortunately not readily available in large quantities.

Alternatively, GFP1–10 protein could be replaced with a DNA template containing *gfp1–10* gene, and the ROX-OMB could be replaced with a fluorogenic RNA aptamer (e.g., broccoli, spinach, mango, malachite green aptamer, etc.)—encoded in the same DNA template as the riboswitch—and their corresponding proto-fluorophore ligand (e.g., DFHBI, thiazole orange, dimethyl indole red, malachite green, etc.). However, it must be taken into consideration that this may increase the metabolic burden on the system, and that DNA templates on beads currently has a size limitation (see **3.6.3. DNA template size**, page **43**). Another option would be using small single reporter peptides instead of a split system, for example tetracystein peptide motifs—which are also comparatively smaller than GFP11—, in combination with fluorogenic arsenic-containing compounds like FlAsH and ReAsH (Adams et al., 2002, Martin et al., 2005, Adams and Tsien, 2008), as they have been already used in a technique called PERSIA (PURExpress-ReAsH-Spinach *In vitro* Analysis) for real-time measurement of translation in a CFPS system (Wick et al., 2019).
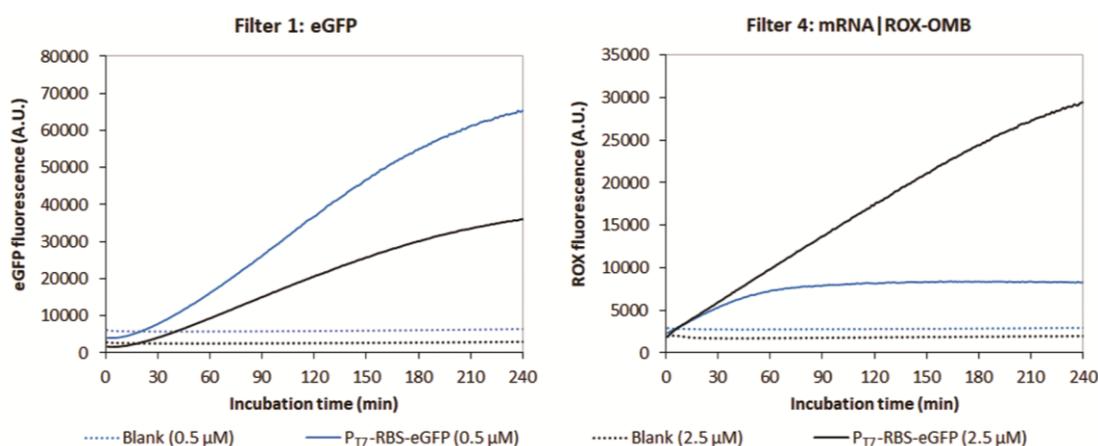
Figure 17. Real-time measurement of eGFP (left) and mRNA (right) synthesis of the no-riboswitch control (P_{T7}-RBS-eGFP) in a PURE*frex* 1.0 reaction with 0.5 µM (blue line) and 2.5 µM (black line) of ROX-OMB. Blank: PURE*frex* reaction without DNA template (dotted line). When 0.5 µM ROX-OMB is used, ROX fluorescence reach saturation within 1 h of incubation. The graphs are based on average values of 2 replicate reactions. The incubation (37°C for 4 h) and real-time fluorescence measurements were performed in a Step One Plus Real-Time PCR System (Applied Biosystems) using filter 1 (GFP/FAM) and filter 4 (ROX).

3.6.7. QUANTITATIVE NATURE OF THE METHOD

Under the current setup, this method can assess the riboswitch libraries in a semi-quantitative manner but requires separate assays to obtain more quantitative information. By optimizing and calibrating the transcription and translation reporters, as mentioned previously (see **3.6.6. Transcription and translation reporters**, page **44**), it would be possible to obtain more quantitative data directly from the flow-cytometry and sorting runs. With such information it would be possible to estimate additional parameters such as the gene expression variability of the sample and the expected dynamic range of the potential riboswitches. This in combination with some statistical analysis, could provide valuable information of the current library and the status of the riboswitch enrichment, which in turn would help to improve the effectiveness of the screening/selection process, and even prevent wasting time processing useless libraries that do not contain functional riboswitches (Zhang et al., 2019).

Additional quantitative information, such as switching activity and dynamic range of each variant, could potentially be inferred from the sequencing data as well. Unfortunately, the current droplet sorting setup cannot provide such information, unless a more sophisticated sorting strategy (e.g., multichannel sorting) and statistical analysis are employed, akin to SORT-seq and FACS-seq approaches (Kinney et al., 2010, Sharon et al., 2012, Kosuri et al., 2013, Noderer et al., 2014, Peterman et al., 2014, Peterman and Levine, 2016, Rohlhill et al., 2017, Koberstein et al., 2021).

3.6.8. MUTAGENIC LIBRARY REGENERATION AND DIRECTED EVOLUTION

The technique can potentially allow the introduction of a diversification step during the library regeneration. It is possible to employ methods such as error-prone PCR (epPCR) or staggered extension process (StEP) for *in vitro* recombination to introduce mutations and increase the variability of the library (Cobb et al., 2013). This way, it is possible to use the sorting cycles for directed evolution of cell-free riboswitches and other genetic devices. The only limitation would be the throughput capacity of the machine currently employed, which may not be enough for completely covering the increasing number of variants created during each diversification step. An obvious solution would be employing a different sorting system with better throughput and sampling capacity.

3.7. CONCLUSIONS

With the new method presented in this work it is possible to enrich riboswitches in both ON and OFF states, albeit some limitations. This proof-of-concept technique represent one of the first attempts to stablish a methodology for engineering riboswitches directly in a cell-free system in a high-throughput manner. Similar cell-free approaches have been employed in the past with enzymes, aptamers, and ribozymes, but not yet with genetic regulators such as riboswitches, which are relatively more difficult to work with due to the dynamic nature of their function, the dichotomy of the expected outputs, and the requirement of active transcription and translation.

This method works without special customized instrumentation as it was purposedly developed using only commercially available microfluidics devices. Although this makes it more accessible to laboratories without specialized equipment and expertise, also restricts the capability of the method to the technical limitations of the instruments and chips available. However, the technique is simple and flexible enough to be adapted for use with more sophisticated custom microfluidics devices already reported in the literature. Such platforms have not been commoditized yet, but are capable of achieving better sorting efficiencies, higher throughputs—a few orders of magnitude higher on some occasions—, and acquiring more quantitative data, which would undoubtedly help improving some of the shortcomings of the method.

CHAPTER 4:
DEVELOPMENT OF SYNTHETIC CELL-FREE RIBOSWITCHES THROUGH DROPLET SORTING

4.1. RIBOSWITCH ENGINEERING PROCESS

After confirming that the method can efficiently enrich ON and OFF droplets, it was put to test by obtaining cell-free riboswitches that respond to histamine (HA) and ciprofloxacin (CFX), two ligands that are difficult to engineer using *in vivo* approaches due to low bioavailability or cytotoxicity. Engineering riboswitches is not straightforward, and it involves multiple steps regardless of the method employed. All the constructs presented in this work have been obtained through the following process (**Figure 18**).
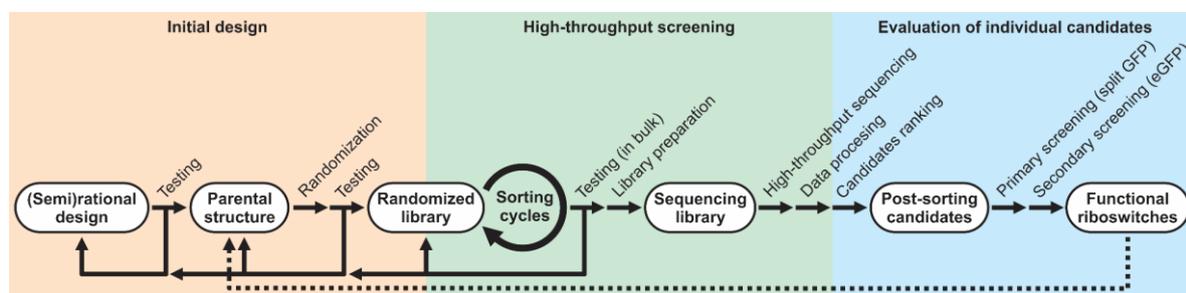
Figure 18. Schematic of the riboswitch engineering process using the droplet sorting method.

4.1.1. INITIAL DESIGN

This is one of the most important steps as the initial design usually determine the success or failure of obtaining good riboswitches. It should be noted that a successful screening with this method—or any other method for that matter—is achievable only if the initial library contains functional riboswitches. For the initial riboswitch design, there are two main approaches: (i) take an existing riboswitch and start introducing modifications; or (ii) de *novo* designs where the riboswitch is created by fusing a known aptamer to a potential expression platform. All three riboswitches presented in this chapter falls into the latter approach. At this stage, using a rational or semi-rational design approach is recommended as it increases the chances for obtaining functional riboswitches structures. The initial designs were synthetized as GFP11 constructs (see **3.3.1. Structure of the constructs**, page **32**) and then tested with a CFPS system in solution (see **4.1.3. Evaluation of individual candidates**, page **48**). The designs were modified gradually until the candidate structure showed a glimpse of riboswitch activity. Once a potential structure was obtained, it is used as the parental sequence for creating the

randomized library (see **3.3.2. DNA template construction**, page **32**) in which the randomized regions are rationally selected with the hope that some of the new variants exhibit better performance than the parental structure, assuming some expected base-pairs interaction and secondary structures.

4.1.2. HIGH-THROUGHPUT SCREENING

The next step is to screen for variants that display the desired output in the presence and absence of the ligand using the newly developed high-throughput sorting method described in the previous chapter. After 4 or 5 sorting cycles, the libraries recovered were tested in bulk to verify the enrichment of functional riboswitches, evidenced by the increase of the overall riboswitch activity of the library. If the libraries did not show any sign of enrichment, either the sorting cycles were repeated with the same library but using different selection parameters; or the riboswitch sequence and/or randomized regions were redesigned, and the process was repeated from the beginning.

After confirming riboswitch enrichment, the libraries were labeled with a custom DNA barcode (6–8 nt) to link each library with its corresponding sorting cycle. The custom DNA barcode was inserted during 6 cycles of PCR (98°C for 10 s, 69°C for 10 s, and 72°C for 15 s) using primers P-HA-Nova-T1 and P-HA-Nova-B1 with the library DNA (25 nM) as template. The PCR product was column-purified and then used to produce the sequencing libraries. The final sequencing libraries were generated by PCR in 20 µl volume containing the barcoded PCR product (0.5 ng) and primers TruSeq-i5 and TruSeq-i7. Twelve cycles of PCR (98°C for 10 s, 69°C for 10 s, 72°C for 15 s) were performed, and the PCR product was purified by agarose gel electrophoresis. The sequencing was performed using Illumina MiSeq (150 bp, single-end) or Illumina NovaSeq (150 bp, paired-end). The sequencing data was processed using a custom script (see **Appendix 4: Protocols**, **H.4. Sequencing data processing**, page **XIII**). Briefly, the raw reads were merged (if paired-end reads were used) and sorted according to the corresponding droplet sorting cycles based on the custom barcode sequence introduced during the sequencing library preparation. The reads were filtered to remove those containing low quality base calls and errors in unintended positions. Finally, the number of reads for each variant was counted and expressed as a percentage of the total number of reads within each cycle (abundance).

The variants were ranked by their "enrichment trend" which was calculated based on the slope of the simple linear regression of the abundance of each variant through the sorting cycles (change in abundance) (see **Appendix 4: Protocols**, **H.4. Sequencing data processing**, page **XIII**). This parameter was used to estimate the tendency and rate of the variants to become enriched (positive trend) or depleted (negative trend) during the sorting cycles. A common pattern was observed in all the libraries presented in this work: as the sorting cycles increased, some variants were no longer detected, while some sequences became enriched, which was an indication that some variants were being preferentially selected with each sorting cycle. Due to the alternating conditions of each subsequent sorting cycle, it is assumed that the variants that placed higher in the ranking (higher "enrichment trend") have a higher probability to be functional riboswitches. The activity of selected candidates can then be individually assessed with regular CFPS reactions in solution.

4.1.3. EVALUATION OF INDIVIDUAL CANDIDATES

Initial library constructs designs and promising riboswitch candidates from the sequencing data were individually synthetized and tested with a CFPS system in solution. First, a primary

screening was performed using the split GFP system (GFP11 + GFP1–10). The riboswitch variants controlling the *gfp11* gene were evaluated in PURE*frex* 1.0 reaction mix (8 μl total reaction volume) containing the DNA template (20 nM), ROX-labeled molecular beacon (0.5 μM), GFP1–10 (50 μg/ml, approximately 26 μM), and where appropriate, the ligand (histamine or ciprofloxacin) at the indicated concentration. The samples were incubated at 37°C for 4 h in 0.2 ml PCR tubes and then transferred to a 384-well microplate (Greiner Bio-One, black, non-binding, flat-bottom). GFP fluorescence (ex. 484 nm, em. 510 nm) and ROX fluorescence (ex. 587 nm, em. 599 nm) were measured using Infinite M1000 Pro (Tecan). Some of the validated riboswitch variants that looked promising were further subjected to a secondary screening step using the full-length reporter gene eGFP (239 amino acids) instead of GFP11, because the primary screening assay does not directly report the expression level of GFP11, as the output is also dependent of the initial concentration of GFP1–10 protein. The riboswitch variants were fused to the *egfp* gene and characterized in CFPS reactions performed as described before except without GFP1–10. All GFP fluorescence values (split GFP and eGFP) were normalized by their ROX fluorescence, and then further normalized by that of a no-riboswitch positive control ("strong" RBS only: P_{T7}-RBS-GFP11 or P_{T7}-RBS-eGFP). If desired, final functional variants can be used as a parental sequence for a new randomized library and another selection process, aimed for further optimizing the riboswitch and finding better derived variants.

Any secondary structure analysis presented here was based on published structures and computational simulations using ViennaRNA (Lorenz et al., 2011) web services. The folding free energies of the riboswitch centroid structures were estimated via RNAfold/ViennaRNA (Zuker and Stiegler, 1981, Lorenz et al., 2011) and used for making comparisons and predictions (Ding et al., 2005). The RNA folding kinetics simulations were performed using the Barriers/Treekin server of ViennaRNA web services (Flamm et al., 2002, Lorenz et al., 2011, Wolfinger et al., 2018) and KineFold web service (Xayaphoummine et al., 2005).

4.2. HISTAMINE ON-SWITCH

Histamine is a hydrophilic vasoactive amine and neurotransmitter involved in local immune response, inflammatory processes, allergic reactions, and regulation of gut functions inside the human body (Panula et al., 2015, Nieto-Alamilla et al., 2016, Wouters et al., 2016). Recently, it has been suggested that it may also play an important role as a signaling molecule mediating bacteria–host interaction (Krell et al., 2021). Our group recently discovered an aptamer that binds to histamine with a K_d ~371 nM (Dwidar et al., 2019). Due to histamine's low bioavailability inside *E. coli*, probably caused by low membrane permeability or by a rapid inactivation by degradative enzymes inside the cell, conventional cell-based riboswitch selection strategies were unsuccessful. Consequently, to engineer histamine riboswitches it was initially resorted to a semi-rational iterative design process which was slow, time consuming, with limited throughput, and not generally applicable to other aptamer sequences (Dwidar et al., 2019). For that reason, as the newly developed cell-free selection method potentially offers a much higher throughput, it was envisioned that it could aid the engineering of those riboswitches unamenable with conventional cell-based methods such as the histamine one.

A new riboswitch library was designed based on the histamine aptamer A1-949 (Dwidar et al., 2019). The aptamer was fused to a linker sequence designed to sequester the region spanning the RBS and the start codon (inhibitory stem-loop, ISL in **Figure 19**) which resulted in repressed protein expression. Four to six nucleotides upstream of the aptamer were randomized (5376 variants) with an expectation that some of the variants would activate translation upon aptamer–ligand binding by competing with the ISL structure for the RBS/AUG sequestering sequence.
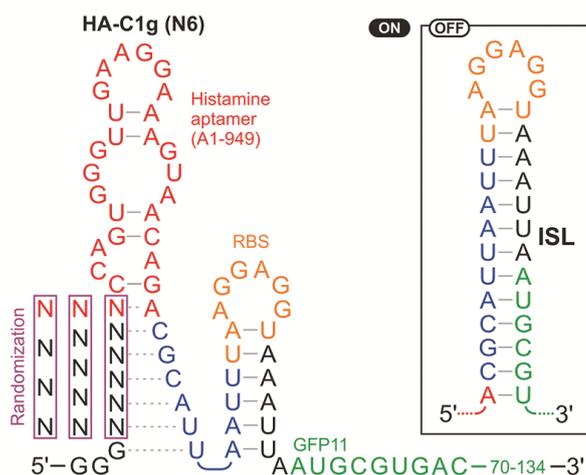
Figure 19. Library design and the hypothetical structure of histamine-responsive riboswitch library HA-C1g (N6). Formation of the inhibitory stem loop (ISL) is expected to repress translation by sequestering the RBS and the start codon. The randomized region was expected to compete with ISL by interacting with the 5' portion of the ISL in the presence of histamine. Adapted from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

Four rounds of sorting was performed: two rounds each of an ON-sorting in the presence of 5 mM histamine and an OFF-sorting without histamine in alternating cycles. Although the initial library when assayed as a mixture showed no response to histamine, the riboswitch library after four rounds of sorting showed 10.8× activation in the presence of histamine, suggesting that the sequential sorting rounds enriched histamine ON-switches (**Figure 20**). The riboswitch populations from all sorting rounds were analyzed by high-throughput sequencing (Illumina). All 5376 variants were detected in the initial library and after the first cycle of sorting, indicating that the library coverage was complete (see **Appendix 3: Summary of the sequencing runs**, page **V**).
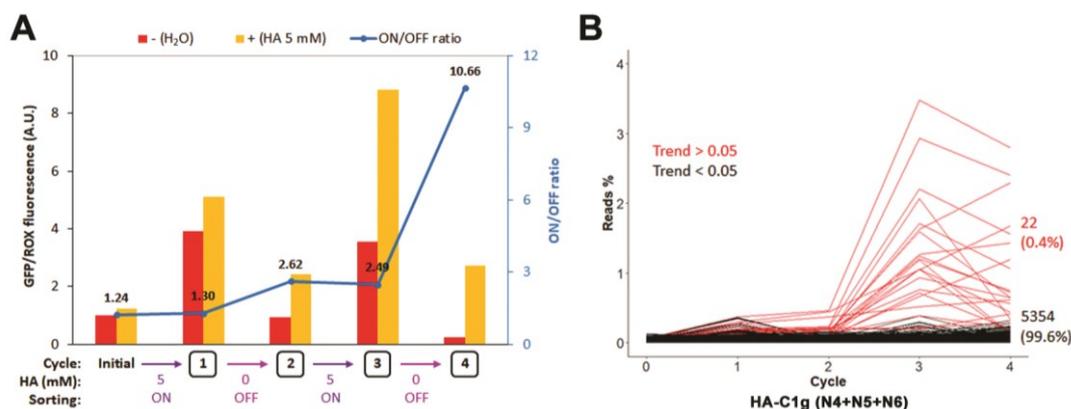
Figure 20. Enrichment of riboswitches after each sorting cycle of histamine ON-riboswitch HA-C1g (N6) library. (**A**) PURE*frex* reaction in solution showed an increase of the bulk ON/OFF ratio of the library after each sorting cycle. HA: histamine. (**B**) Enrichment trend of individual variants through the sorting cycles. Adapted from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

The top 20 sequences, sorted by their "enrichment trend", were individually synthesized and assayed in bulk PURE*frex* reactions using the same GFP11/GFP1–10 reporter system employed in droplet sorting. Five additional variants that ranked lower in the sequencing analysis were chosen as negative controls. All of the top 20 variants clearly activated expression in response to histamine albeit to different degrees. In contrast, none of the five negative controls showed any response to histamine (**Figure 21A**).

The riboswitches were then fused to the full-length reporter gene eGFP (239 amino acids) because the previous assay does not directly report the expression level of GFP11. Interestingly, the eGFP expression levels and ON/OFF ratios of the riboswitches varied greatly compared to the GFP11 assay (**Figure 21B, C**). For example, the ON/OFF ratio of the variant 19 (HA-C1g-

19) was only 2.3 in the GFP11 assay but it was over 18 in the eGFP assay, mainly due to the higher ON level in absence of ligand but overall high dynamic range compared to the other variants that was not evident in the GFP11 assay. It can be speculated that the ON levels for some variants in the GFP11 assay were saturated due to the limiting concentration of the GFP1–10 fragment (approximately 26 µM) in the solution. This is also evident from the other switching variants and the no-riboswitch control (P$_{T7}$-RBS-GFP11) exhibiting similarly high fluorescence levels at the ON-state in the GFP11 assay (**Figure 21A**). On the other hand, the OFF levels of the switches in the absence of histamine follow a similar pattern in both GFP11 and eGFP assays (**Figure 21A, C**) because they are not likely to be affected by signal saturation. Moreover, it cannot be excluded the possibility that the full-length mRNA may affect the riboswitch performance. Importantly, none of the negative controls exhibited any response to histamine in the eGFP assay as well.
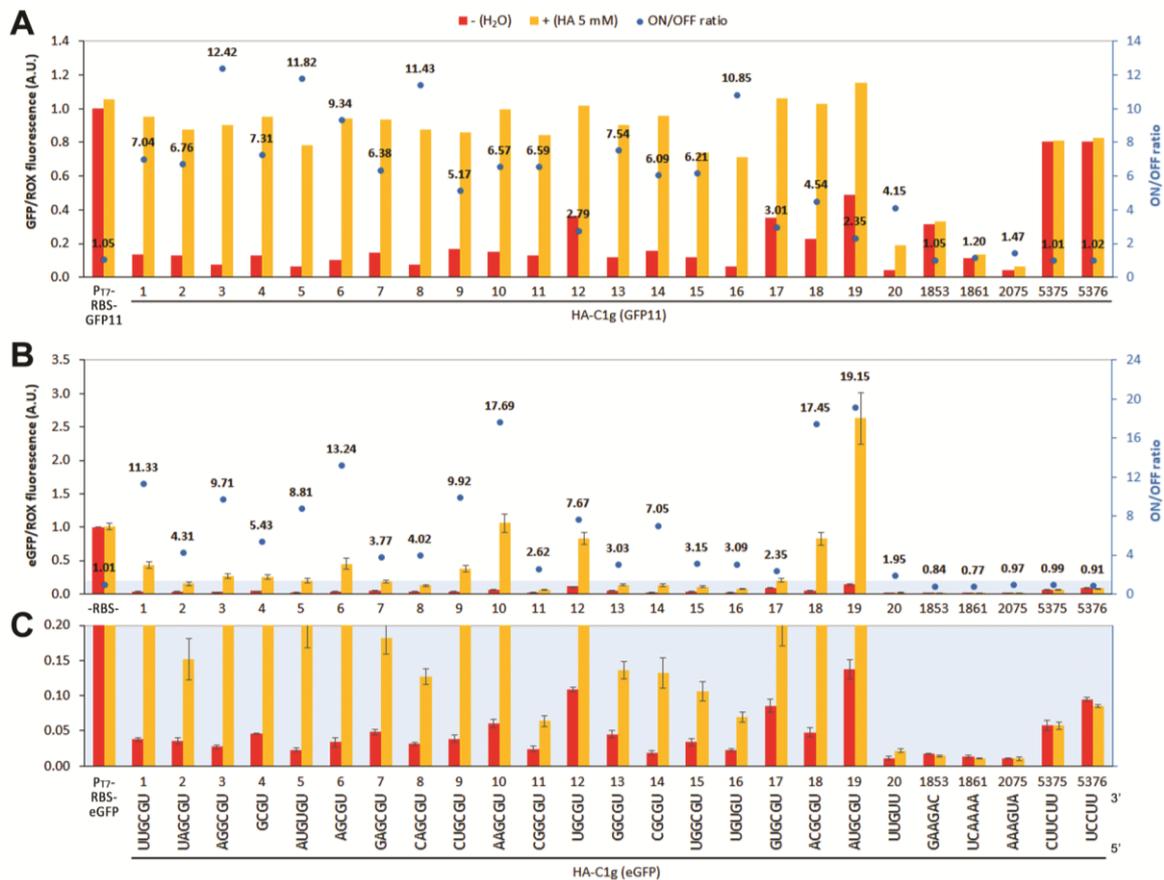
Figure 21. Screening of individual riboswitch variants (histamine ON-switches). (**A**) Primary screening of the riboswitch variants using GFP11/GFP1–10 (split GFP) assay. Performed once for rapid screening. The expression levels were normalized by the no-riboswitch control P$_{T7}$-RBS-GFP11 in the absence of histamine. (**B**) Secondary screening of the riboswitch variants fused to full-length eGFP as a reporter gene. (**C**) Expanded view of the region highlighted in light blue in **B**. The error bars represent standard deviations of three independent assays. The expression levels were normalized by the no-riboswitch control P$_{T7}$-RBS-eGFP in the absence of histamine. The variant number corresponds to the respective ranking based on the "enrichment trend". HA: histamine. Reproduced from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

Inspection of the enriched variants confirms that these riboswitches most likely function by disrupting the ISL structure through the formation of a competing stem (CS) structure at the base of the aptamer (**Figure 22A**). All of the top 19 variants share the consensus sequence GYGU (Y = C or U) at the four nucleotides proximal to the aptamer (**Figure 21C**) that are

likely to form a CS structure which is stabilized upon aptamer–ligand binding (**Figure 22A**). Consistent with this mechanism and thermodynamic predictions, introduction of mismatches within the randomized region that destabilize the CS in HA-C1g-19 resulted in low gene expression with or without histamine (HA-C1g-19/OFF, **Figure 22B, C**). Conversely, stabilization of the CS by increasing its length resulted in a higher expression level in the absence of histamine (HA-C1g-19/ON, **Figure 22B, C**). Finally, destabilization of both the ISL and the CS structures by mutating the 5' end of the ISL resulted in constitutive expression of eGFP as expected (HA-C1g-19/MM, **Figure 22B, C**).
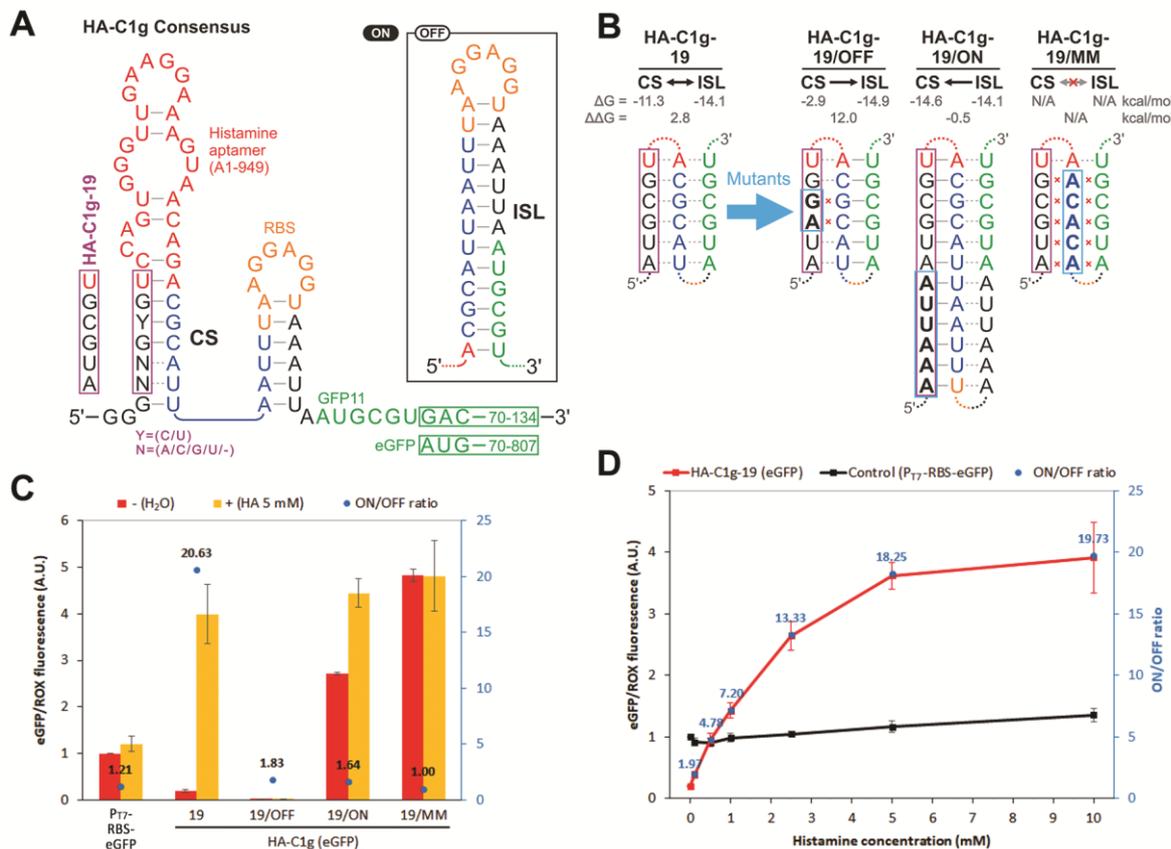
Figure 22. Characterization of the histamine ON-switches. (**A**) The consensus sequence of the enriched variants suggests formation of a competing stem (CS) in the presence of histamine resulting in destabilized inhibitory stem loop (ISL) to activate translation. (**B**) Sequences of HA-C1g-19 and its mutants. HA-C1g-19/OFF contains a double mutation that is designed to destabilize the CS. HA-C1g-19/ON contains an extended and stabilized CS. HA-C1g-19/MM was designed to disrupt both CS and ISL. The calculated free energy values (ΔG and ΔΔG) of the respective structures by ViennaRNA are shown above each sequence. (**C**) Histamine responses of HA-C1g-19 and its mutants fused to eGFP. The expression levels were normalized by the no-riboswitch control P_{T7}-RBS-eGFP in the absence of histamine. (**D**) Dose-dependent expression of the HA-C1g-19 riboswitch (red) and the no-riboswitch control P_{T7}-RBS-eGFP (black). The ON/OFF ratios were corrected for the nonspecific effects of histamine on P_{T7}-RBS-eGFP at each concentration. Error bars represent the standard deviations of three independent assays. HA: histamine. Reproduced from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

Dose dependence analysis of the riboswitch HA-C1g-19 (eGFP) up to 10 mM histamine showed its ON/OFF ratio reaching 19.7 with an EC_{50} of approximately 1.6 mM (**Figure 22D**) after correcting for the moderate nonspecific activation of P_{T7}-RBS-eGFP by histamine (up to ~36% at 10 mM histamine) which does not contain the histamine aptamer. The higher concentration of the effector molecule necessary to activate the riboswitch (EC_{50}) relative to the K_d of the aptamer (371 nM in this case) could be attributed to the presence of a kinetic component during the folding of the riboswitch (Dwidar et al., 2019). Similar phenomena have

been observed and studied on the theophylline riboswitch (Mishler and Gallivan, 2014) and other natural and synthetic riboswitches (Lynch et al., 2007, Lemay et al., 2011, Haller et al., 2011, Peselis and Serganov, 2014), in which switching of the mRNA folding to the ON state is determined by a co-transcriptional trapping mechanism.

Interestingly, HA-C1g-19-eGFP shows a much higher ON level compared to the no-riboswitch control (P_{T7}-RBS-eGFP) even though they share the same RBS sequence. Real-time measurements of eGFP and mRNA levels of HA-C1g-19 confirmed that the stronger signal is due to the higher translation efficiency in the presence of histamine, and that the mRNA level is not significantly affected by histamine (**Figure 23**). The mechanistic basis of this unexpectedly high expression level remains to be investigated.
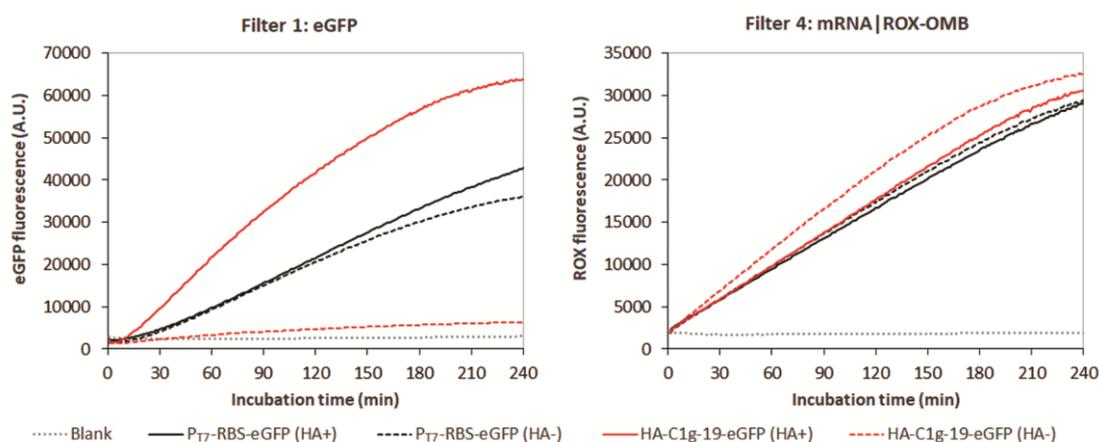
Figure 23. Real-time measurement of eGFP (left) and mRNA (right) synthesis of the no-riboswitch control (P_{T7}-RBS-eGFP, black line) and HA-C1g-19-eGFP (red line) in a PURE*frex* 1.0 reaction with 2.5 µM of ROX-OMB. Blank: PURE*frex* reaction without DNA template (grey dotted line); HA+: 5 mM histamine (solid line); HA-: no histamine (dashed line). A substantial increase (~2-fold) in eGFP signal of HA-C1g-19-eGFP was observed during the course of the reaction compared to the P_{T7}-RBS-eGFP control. The graphs are based on average values of 2 replicate reactions. The incubation (37°C for 4 h) and real-time fluorescence measurements were performed in a Step One Plus Real-Time PCR System (Applied Biosystems) using filter 1 (GFP/FAM) and filter 4 (ROX). Reproduced from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

Table 3. List of the best riboswitches obtained with this technique compared to similar ones reported in the literature.

Ligand	Riboswitch	ORF	ON/OFF ratio (Ligand conc.)	System	RS length (AP+EP)	Reference
Histamine	HA-C1g-10	eGFP	ON 17.7× (5 mM)	PURE	66 nt	This work
	HA-C1g-18	eGFP	ON 17.5× (5 mM)	PURE	66 nt	This work
	HA-C1g-19	eGFP	ON 20.6× (5 mM)	PURE	66 nt	This work
	H2	mCherry	ON 30.7× (5 mM)	PURE	103 nt	(Dwidar et al., 2019)
	HA-OFF4-a9-1	eGFP	OFF 10.8× (2.5 mM)	PURE	67 nt	This work
	HA-OFF4-a9-3	eGFP	OFF 7.8× (2.5 mM)	PURE	67 nt	This work
	HA-OFF4-a9-13	eGFP	OFF 8.5× (2.5 mM)	PURE	67 nt	This work
	HA-OFF4-a9-14	eGFP	OFF 9.3× (2.5 mM)	PURE	67 nt	This work
	HA-OFF4-a9-18	eGFP	OFF 8.1× (2.5 mM)	PURE	67 nt	This work
Ciprofloxacin	CFX-a1-sr5	eGFP	ON 7.8× (100 µM)	PURE	112 nt	This work
	CFX-a1-sr5-2	eGFP	ON 8.3× (100 µM)	PURE	112 nt	This work
	CFX-a1-sr5-19	eGFP	ON 9.3× (100 µM)	PURE	112 nt	This work
	CFX-RS	GFP	OFF 7.5× (1 mM)	Yeast	103 nt	(Groher et al., 2018)
	CFX-RS (GAAA)	Luc	OFF 1.8× (250 µM)	HeLa	91 nt	(Groher et al., 2018)

*PURE: PURE*frex* 1.0, Gene Frontier (cell-free system), Yeast: *Saccharomyces cerevisiae* (*in vivo*), HeLa: HeLa cells (*in vivo*). AP: aptamer, EP: expression platform, GFP: green fluorescent protein (e-: enhanced), Luc: firefly luciferase, ORF: open reading frame, RS: riboswitch.

Compared to the histamine riboswitches reported by Dwidar et al. (Dwidar et al., 2019), the best riboswitch obtained through this method (HA-C1g-19), exhibit a similar sensitivity but a lower ON/OFF ratio (**Table 3**). The length of the new riboswitch, however, is much more compact (37 nt smaller) due to the use of a minimized aptamer sequence and an expression platform with a simpler architecture. It is important to mention that the reporter gene employed by them (mCherry) is different from the one used here (eGFP), and thus different genetic context may influence the activity of the riboswitches.

4.3.   HISTAMINE OFF-SWITCH

To further validate the usefulness of the method, it was employed to engineer OFF-riboswitches that downregulate gene expression in response to histamine. OFF-riboswitches are relatively common in nature where riboswitches are often used to implement negative-feedback regulation of biosynthetic pathways. However, synthetic OFF-riboswitches in living cells are noticeably less common compared to ON-riboswitches. While natural and synthetic OFF-riboswitches in bacteria are expected to function in prokaryotic CFPS systems, there have been no previous reports of OFF-riboswitches implemented in prokaryotic CFPS systems (Tabuchi and Yokobayashi, 2021).
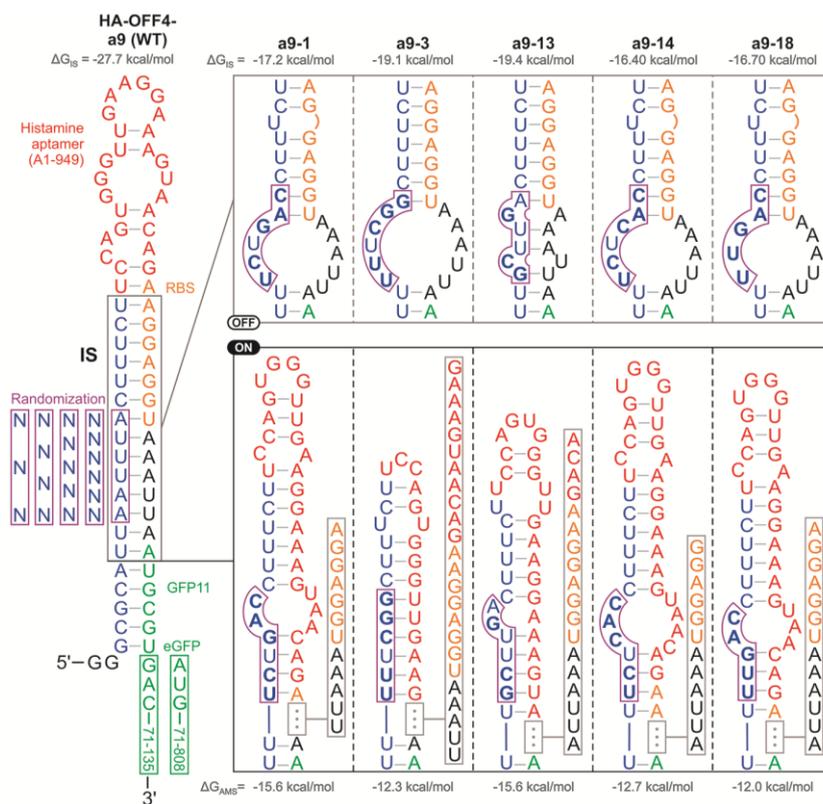
Figure 24. Histamine OFF-switch library design and predicted OFF- and ON-structures of the variants discovered by droplet sorting. The library was based on HA-OFF4-a9 (WT) that is predicted to form an extended inhibitory stem (IS) structure sequestering the RBS and the start codon. The boxed nucleotides were replaced with random sequences. The variants discovered by droplet sorting are suggested to form multiple alternative metastable structures (AMS) that disrupts the RBS-sequestering stem (IS-like structure) in the absence of histamine, the diagram only shows those with the lowest folding energy. The calculated free energy values (ΔGIS and ΔGAMS) are shown above or below each structure. Adapted from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

The histamine ON-riboswitch library was modularly redesigned, aiming to increase the chances of discovering OFF-riboswitches. After a few trials, a new library was created based on the HA-OFF4-a9 construct that contains an extended inhibitory stem (IS) that blocks both RBS and the start codon with the histamine aptamer inserted in the loop region (**Figure 24**). As expected, HA-OFF4-a9 strongly represses gene expression with or without histamine

(WT = HA-OFF4-a9, **Figure 26A**). Three to six bases were randomized upstream of the aptamer near the middle of the IS with an expectation that some of the variants would destabilize the IS structure in the absence of histamine (**Figure 24**). It was also expected that aptamer–ligand binding would force the formation of an IS-like structure repressing translation. This library (up to 5440 variants) was subjected to alternating OFF-sorting in the presence of histamine (5 or 2.5 mM) and ON-sorting in the absence of it, for a total of five rounds (three OFF-sorting and two ON-sorting). Similar to the histamine ON-riboswitch sorting, it was observed that the response to histamine of the bulk library population improved after sorting (**Figure 25A**).
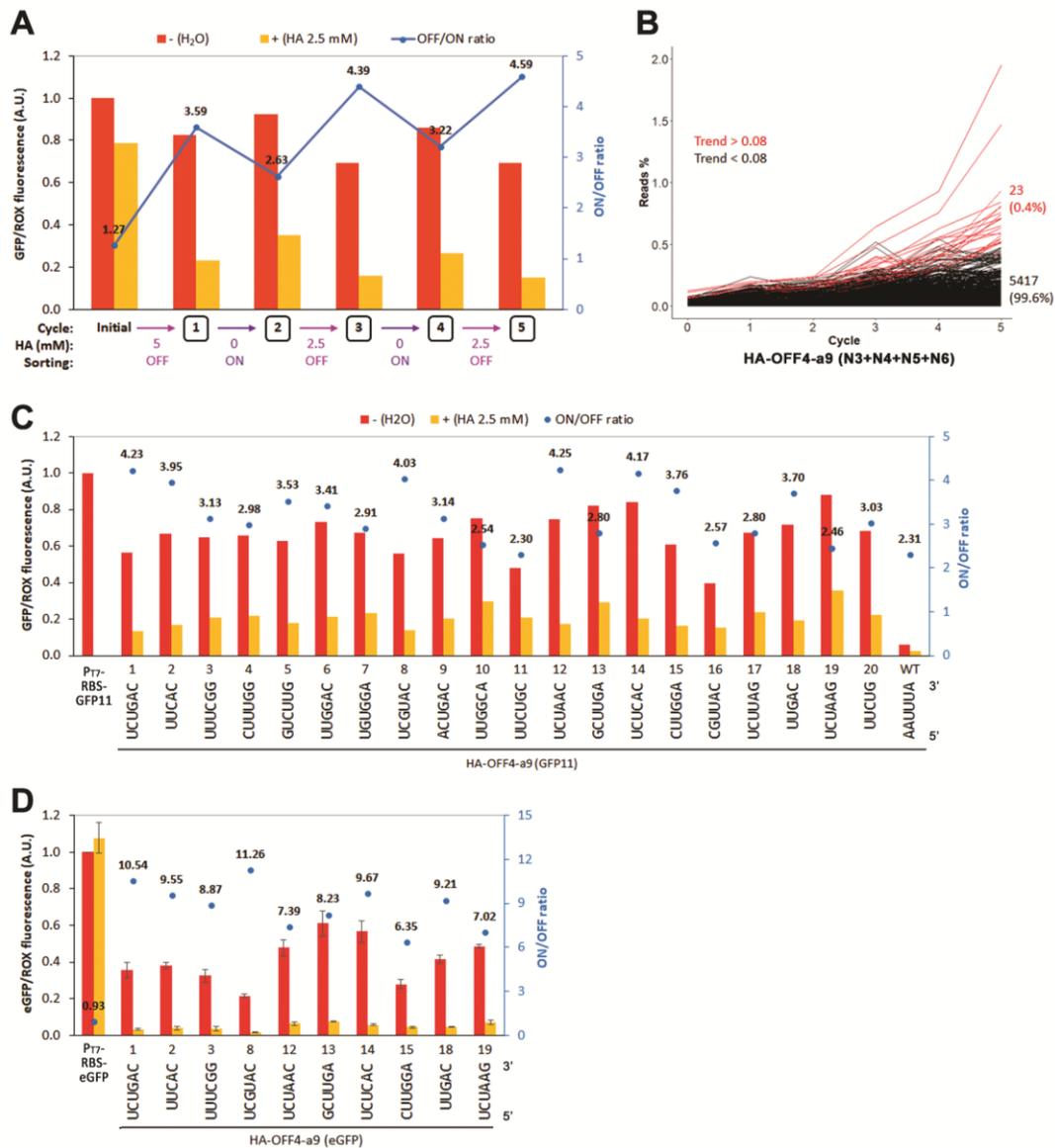
Figure 25. Sorting cycles of histamine OFF-switch HA-OFF4-a9 (N3+N4+N5+N6) library and screening of candidates. (**A**) Response of the bulk riboswitch population in PURE*frex* 1.0 shows slight enrichment of riboswitches. (**B**) Enrichment trend of individual variants through the sorting cycles. (**C**) Primary screening of the riboswitch variants using split GFP assay. The parental construct HA-OFF4-a9 is denoted as WT. Performed once for rapid screening. (**D**) Secondary screening of the riboswitch variants fused to full-length eGFP. The assay was performed in three independent replicates with error bars representing standard deviations. The expression levels were normalized by a no-riboswitch control (P$_{T7}$-RBS-GFP11 or P$_{T7}$-RBS-eGFP) in the absence of histamine. The numbers of the variants correspond to their respective rankings based on the "enrichment trend". HA: histamine. Adapted from (Tabuchi and Yokobayashi, 2022) with permission of the authors.
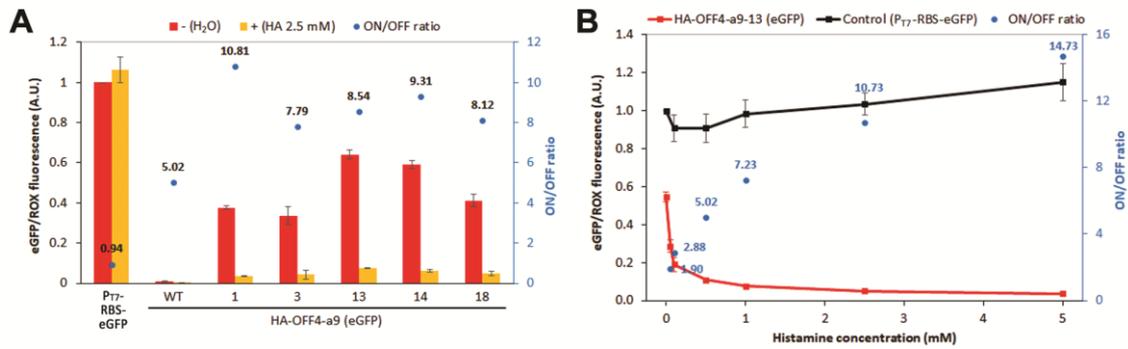
Figure 26. Characterization of histamine OFF-switches. (**A**) Individual characterization of the discovered variants using full-length eGFP as a reporter gene. The expression levels were normalized by the no-riboswitch control P_{T7}-RBS-eGFP in the absence of histamine. WT correspond to the parental sequence HA-OFF4-a9. (**B**) Dose-dependence of HA-OFF4-a9-13 riboswitch (red) and a no-riboswitch control P_{T7}-RBS-eGFP (black). The ON/OFF ratios were corrected for the non-specific effects of histamine on P_{T7}-RBS-eGFP. The error bars represent standard deviations of three independent assays. HA: histamine. Adapted from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

After sequence analysis, the top 20 variants were individually screened by GFP11 assay. All variants were downregulated by histamine, with ON/OFF ratios ranging from 2.3 to 4.5 (**Figure 25C**). Ten variants were further screened by eGFP assay and all of them showed ON/OFF ratios over 6.1 (**Figure 25D**). Detailed analysis of the best five variants confirmed ON/OFF ratios between 7.7 and 10.8, and a much higher overall signal (gene expression) compared to the parental sequence (WT = HA-OFF4-a9, **Figure 26A**). Interestingly, dose-dependent assay of HA-OFF4-a9-13 showed a more sensitive response compared to the ON-riboswitch, with an EC_{50} of approximately 50 µM (**Figure 26B**).
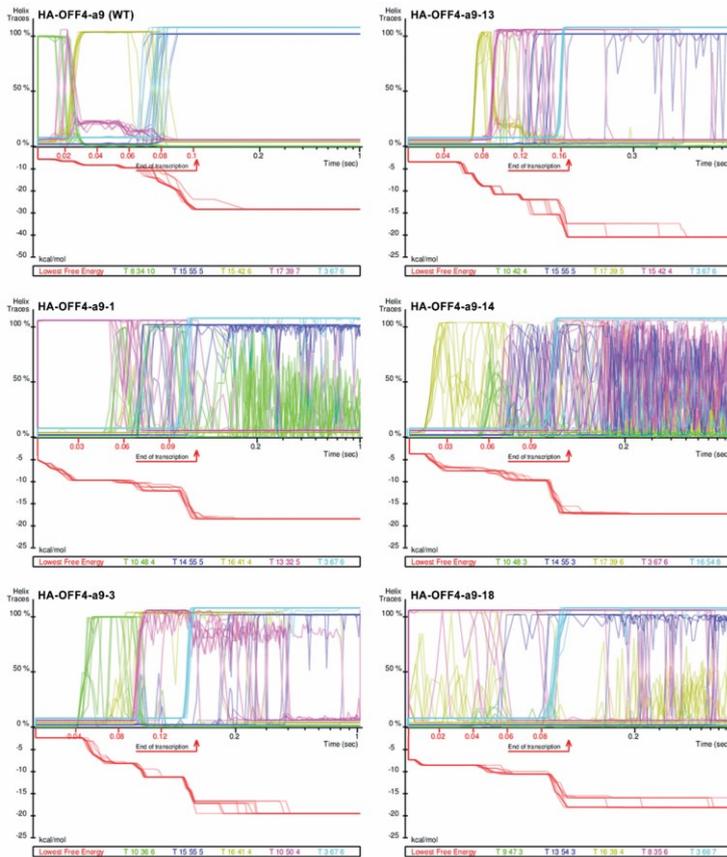
Figure 27. Helix tracing graphs overlapping ten stochastic co-transcriptional folding simulations each. "Helix" refer to a paired strand of RNA (stem). Occasionally, an ON-conformation (green, magenta, or brown) will get trapped and persist for some time after end of transcription before transitioning into an OFF-conformation (blue). In other cases, the RBS-sequestering helix (blue) seems to be unstable and switches multiple times to conformations with exposed RBS (green, magenta, or brown). AUG-sequestering helix (cyan) is always present in all variants regardless the conformation of the rest of the riboswitch. The graphs were generated with KineFold (pseudoknots and entanglement allowed). Red: lowest free energy; blue: IS-like RBS-sequestering helix; cyan: IS-like AUG-sequestering helix; green, magenta, and brown: helices present in alternative conformations with fully or partially exposed RBS, specific for each variant. AUG: start codon; IS: inhibitory stem; RBS: ribosome binding site.

Secondary structure predictions of the OFF-switches by ViennaRNA (Lorenz et al., 2011) suggest that the randomized region destabilizes the IS structure due to multiple mismatches within the stem, and it potentially allows alternative metastable structures (AMS) to form by interacting with the aptamer sequence (**Figure 24**). The AMS are presumably kinetically stabilized in the absence of histamine, as co-transcriptional folding simulations suggest that the mRNA could become temporarily trapped into metastable ON-conformations (**Figure 27**). These AMS may make the RBS more accessible, exposing it long enough for the ribosome to initiate translation. The presence of histamine may stabilize the aptamer structure and force the mRNA to collapse into an OFF-conformation, locking the RBS sequestering stem (IS-like structure) and strongly repressing translation. These AMS, however, are not necessarily thermodynamically favored over the IS-like structures, which could explain the relative low gene expression levels of these riboswitches. In addition, sequestering of the start codon (AUG) alone does not seem to be enough for repressing translation, as a short AUG-sequestering stem seems to be present in the predicted conformations of both ON and OFF states. Further mutational and kinetic analysis are necessary to validate this mechanistic hypothesis.

4.4.   CIPROFLOXACIN ON-SWITCH

To further demonstrate the generality of the droplet sorting strategy, cell-free riboswitches that respond to ciprofloxacin were developed. Ciprofloxacin is a broad-spectrum fluoroquinolone antibiotic and potentially cytotoxic to even mammalian cells (Gurbay et al., 2002, Lawrence et al., 1996, Salimiaghdam et al., 2020); therefore, it is not amenable to the existing bacterial screening or selection methods unless fluoroquinolone-resistant strains are employed. Plasmid-mediated quinolone resistance (PMQR) mechanisms by themselves can only provide low-level resistance to ciprofloxacin (Jacoby et al., 2014), whereas higher-level resistance often requires multiple mutations in the genes of the antibiotic targets (DNA gyrase *gyrA* and DNA topoisomerase IV *parC*) (Hamed et al., 2018, Ma et al., 2018b). Therefore, ciprofloxacin riboswitches represent another example where cell-free riboswitches cannot be easily transferred from bacterial riboswitches.
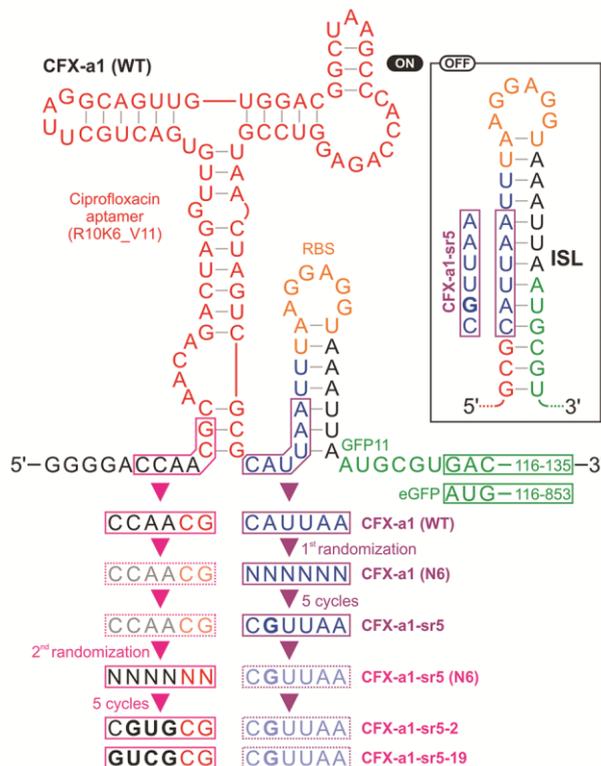
Figure 28. Ciprofloxacin ON-switch libraries design and predicted structure of CFX-a1 (WT) riboswitch. Inhibitory stem loop (ISL) was designed to repress gene expression in the absence of histamine. Nucleotides in the boxed regions were randomized sequentially to discover variants that respond to ciprofloxacin. A summary of the droplet sorting results are shown at the lower section of the figure. In the first round of sorting cycles, six nucleotides within the ISL of CFX-a1 (WT) were randomized to identify CFX-a1-sr5. In the second round of sorting cycles, six nucleotides upstream of CFX-a5-sr5 were randomized to isolate CFX-a5-sr5-2 and CFX-a5-sr5-19. Adapted from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

The Suess group discovered the first ciprofloxacin aptamers by *in vitro* selection (SELEX), and then used cellular screening in yeast to obtain riboswitches with ON/OFF ratios up to 7.5 (Groher et al., 2018). Another ciprofloxacin aptamer R10K6 (K_d=31 nM) was later identified and was used to engineer paper-based biosensors (Jaeger et al., 2019).

A new riboswitch was created through a modular approach by replacing the histamine aptamer in the histamine ON-riboswitch library described before (see **4.2. Histamine ON-switch**, page **49**), with a shorter version of the R10K6 aptamer (R10K6_V11, K_d=36 nM) (Jaeger et al., 2019) (**Figure 28A**). The resulting construct CFX-a1 exhibited attenuated expression in the presence and absence of ciprofloxacin (WT = CFX-a1, **Figure 31A**). Because the CFX-a1 construct exhibited low levels of gene expression and a relatively low ON/OFF ratio, it was speculated that this may be probably caused by the ISL being too strong to be efficiently disrupted by the aptamer formation. Therefore, it was decided to first randomize six nucleotides within the ISL to search for sufficiently destabilized ISL variants that still exhibit ciprofloxacin-responsive gene expression (**Figure 28**). Five cycles of sorting of this library (**Figure 29**) resulted in variant CFX-a1-5 (later renamed CFX-a1-sr5, "sr" standing for "second round") with an improved ON/OFF ratio of 5.9–7.8 (**Figure 29C, Figure 31A**).
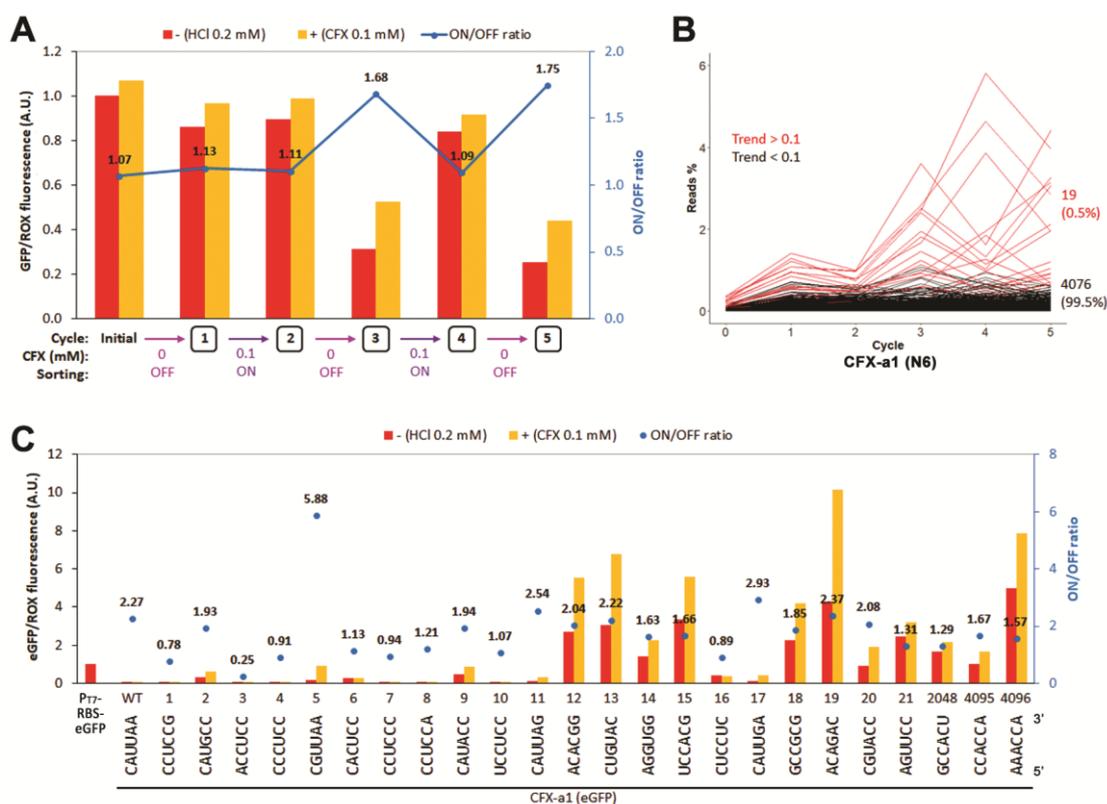
Figure 29. Sorting cycles of ciprofloxacin ON-switch library (N6) based on CFX-a1 (first randomization), and screening of candidates. (**A**) Response of the bulk riboswitch population in PURE*frex* 1.0. (**B**) Enrichment trend of individual variants through the sorting cycles. (**C**) Screening of the riboswitch variants (based on CFX-a1) using eGFP assay. Performed once for rapid screening. The original parental construct (CFX-a1) is denoted as WT. The variant number 5 of this screening was renamed CFX-a1-sr5 and used as the new parental sequence for the next round of sorting cycles and screening (**Figure 30**). The expression levels were normalized by a no-riboswitch control P$_{T7}$-RBS-eGFP in the absence of ciprofloxacin. The numbers of the variants correspond to their respective rankings based on the "enrichment trend". CFX: ciprofloxacin. Adapted from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

From the predicted secondary structure of CFX-a1-sr5 it was noticed that only 3 nucleotides interrupted by a bulge at the distal section of the aptamer stem were competing against the ISL.

Thus, it was thought that CFX-a1-sr5 could be further optimized by randomizing six nucleotides upstream of the aptamer hoping that some of the variants would increase the number of base pairs competing with the ISL structure and improve the activity of the riboswitch. This second round of sorting cycles (**Figure 30**) yielded CFX-a1-sr5-2 and CFX-a1-sr5-19 variants that exhibit a slightly extended predicted aptamer stems and a moderately improved ON/OFF ratios of 8.3 and 9.3, respectively (**Figure 31A**). Dose-dependence analysis of the CFX-a1-sr5-19 riboswitch showed up to 9.8-fold activation with an EC_{50} of approximately 15 μM (**Figure 31B**) after correcting for non-specific gene expression caused by high concentrations of ciprofloxacin (up to ~35% at 0.25 mM ciprofloxacin).
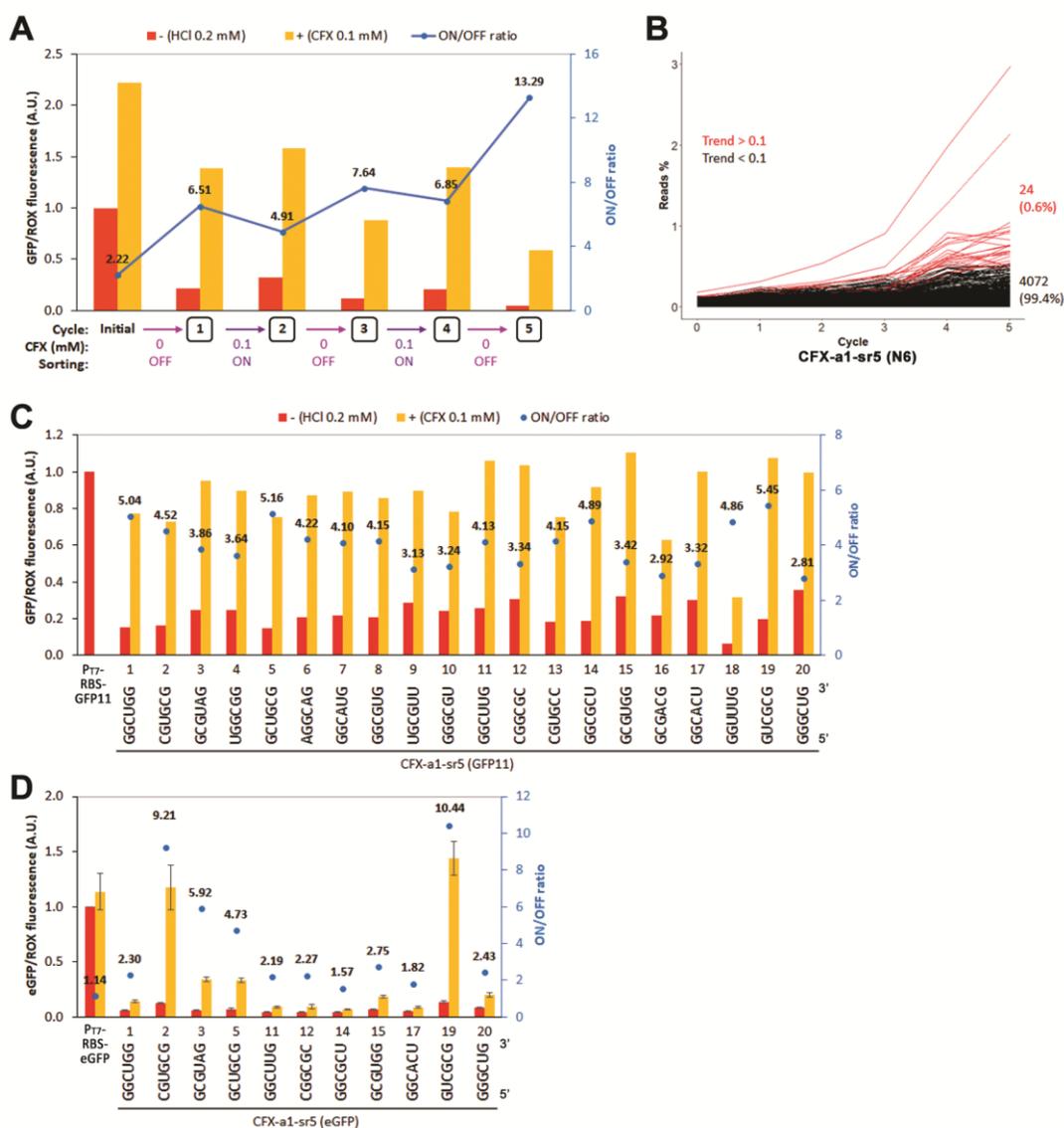
Figure 30. Sorting cycles of ciprofloxacin ON-switch library (N6) based on CFX-a1-sr5 (second randomization), and screening of candidates. (**A**) Response of the bulk riboswitch population in PURE*frex* 1.0. (**B**) Enrichment trend of individual variants through the sorting cycles. (**C**) Primary screening of the riboswitch variants (based on CFX-a1-sr5) using GFP11/GFP1–10 (split GFP) assay. Performed once for rapid screening. (**D**) Secondary screening of the riboswitch variants fused to full-length eGFP as a reporter gene. The assay was performed in three independent replicates for each variant with the error bars representing standard deviations. The expression levels were normalized by a no-riboswitch control (P_T7-RBS-GFP11 or P_T7-RBS-eGFP) in the absence of ciprofloxacin. The numbers of the variants correspond to their respective rankings based on the "enrichment trend". CFX: ciprofloxacin. Adapted from (Tabuchi and Yokobayashi, 2022) with permission of the authors.
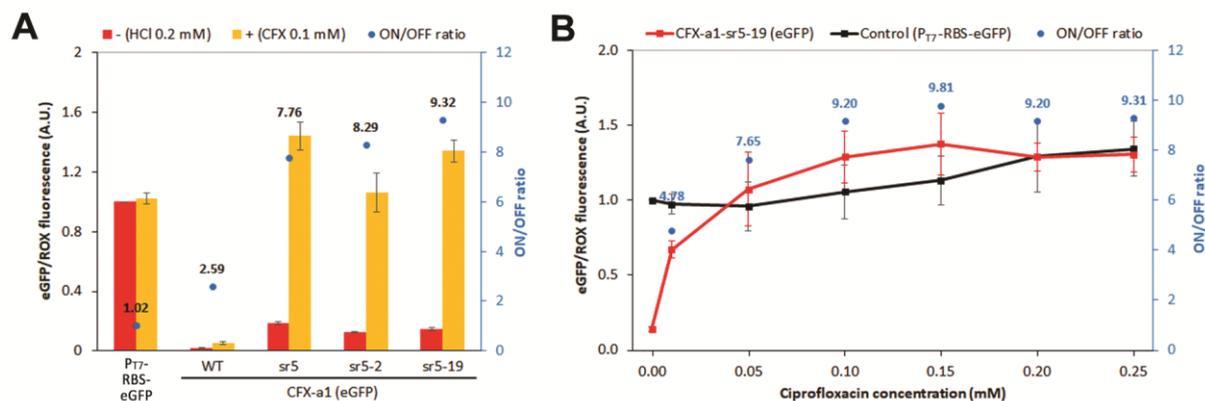
Figure 31. Characterization of ciprofloxacin ON-switches. (**A**) Individual characterization of the discovered variants using full-length eGFP as a reporter gene. The expression levels were normalized by the no-riboswitch control P_{T7}-RBS-eGFP in the absence of ciprofloxacin. (**B**) Dose-dependence of CFX-a1-sr5-19 riboswitch (red) and a no-riboswitch control P_{T7}-RBS-eGFP (black). The ON/OFF ratios were corrected for the non-specific effects of ciprofloxacin on P_{T7}-RBS-eGFP. The error bars represent the standard deviation of three independent assays. CFX: ciprofloxacin. Adapted from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

In comparison to the other only known ciprofloxacin riboswitches, reported by the Suess group (Groher et al., 2018), the best riboswitches obtained with this method exhibit slightly higher ON/OFF ratios and respond to lower ciprofloxacin concentrations (**Table 3**, page **53**). However, it must be noted that the riboswitches from Suess group are vastly different from ours in all aspects. For instance, their riboswitches are OFF-switches and use a different ciprofloxacin aptamer. On top of that, those riboswitches have been developed and tested only *in vivo* using eukaryotic cells (yeast and HeLa cells), which means the regulatory mechanisms may also be very different.

4.5. CONCLUSIONS

Riboswitches responsive to molecules such as histamine and ciprofloxacin are generally difficult to engineer utilizing conventional methods that employ *in vivo* bacterial systems due to biological constraints. The novel method presented in this thesis bypasses this issue by performing the screening directly in a cell-free system and in a high-throughput manner. With only 4 to 5 iterative sorting cycles, the technique allowed the selection of new riboswitches with relatively good ON/OFF ratios (>10) from randomized libraries of more than 4000 variants in less than a week of work each. Additionally, the technique allowed development of not only ON riboswitches but also OFF riboswitches, which have been notably absent in prokaryotic cell-free systems. The results presented in this chapter demonstrate the value of this novel technique as a tool for cell-free riboswitch engineering.

CHAPTER 5:
CHEMICAL COMMUNICATION BETWEEN CELL-SIZED MICRODROPLETS

5.1. INTRODUCTION

Cell-free riboswitches can enable cell-free systems to detect chemical signals from the environment and to respond accordingly by modulating protein expression. By encapsulating such system into water-in-oil droplets or lipid-bilayer membrane vesicles it is possible to create individual entities capable of sensing their environment. If another encapsulated cell-free system can synthesize the chemical signal and release it to the medium, the two entities can communicate as a sender–receiver par, albeit only one-way in this simple example, via a chemical signal. This way it is possible to create artificial cells capable of interacting with each other and with their environment. Previously, liposome-based artificial cells have been used to demonstrate chemical communication using a theophylline-responsive riboswitch (Lentini et al., 2014, Adamala et al., 2017). In these examples, the sender artificial cells contained theophylline-responsive riboswitches controlling the synthesis of α-hemolysin. Once theophylline is added to the medium, it triggers the formation of α-hemolysin pores on the lipid bilayer membrane allowing the release of a different signaling molecule (e.g., IPTG, doxycycline, etc.) trapped inside the artificial cell. In these examples, cell-free riboswitches were simply used as a trigger to release pre-encapsulated chemical signal molecules to the environment, while the ligand for the riboswitch was externally supplied to the system.

In the work presented here, histamine is used as a chemical signal to mediate communication between droplets in a self-contained system, where the signal (histamine) is synthetized within the system itself, either constitutively or triggered by another signaling molecule (ciprofloxacin), which then was sensed by riboswitches, triggering the expression of the reporter gene (GFP).

5.2. ONE-WAY COMMUNICATION BETWEEN MICRODROPLETS

For the one-way communication circuit, the receiver droplets were produced as described in chapter 3 (see **3.3.4. *In vitro* compartmentalization (Droplet generation)**, page **32**) by encapsulating PURE*frex* 2.0 reaction mix (Gene Frontier) containing equimolar concentrations (50 nM) of a DNA template encoding GFP11 under the control of the histamine ON-riboswitch (HA-C1g-19-GFP11), and another DNA template that constitutively expresses the GFP1–10 fragment (P_{T7}-RBS-GFP(1–10)) (**Figure 34A**). It was decided to use the split GFP system

instead of eGFP, because the former exhibits stronger fluorescence signal in the droplets than eGFP, as the split GFP used derives from sfGFP which is significantly brighter than eGFP (Soleja et al., 2018). The sender droplets contain PURE*frex* 2.0, DNA template P_{T7}-*hdc*-6xHis encoding histidine decarboxylase (25 nM), ROX-labeled molecular beacon (1 µM), L-histidine (50 mM), and pyridoxal phosphate (20 µM) (**Figure 34A**). Histidine decarboxylase (HDC) enzyme is synthetized inside the sender droplets and catalyzes the production of histamine from L-histidine which then acts as the signaling molecule. The presence of the cofactor pyridoxal phosphate (PLP) is essential for the enzyme activity (**Figure 32**) and does not intervene in the riboswitch activation (**Figure 33**).

	All	His-	PLP-	DNA-	CFX+	CFX-	Control
HDC template (10 nM)	WT	WT	WT	-	C19	C19	-
PLP (20 µM)	+	+	-	+	+	+	-
L-Histidine (20 mM)	+	-	+	+	+	+	-
Ciprofloxacin (0.1 mM)	-	-	-	-	+	-	-
Histamine (20 mM)	-	-	-	-	-	-	+
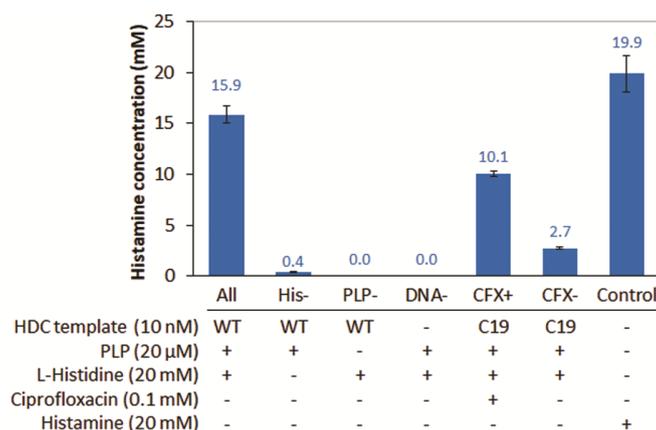
Figure 32. Histidine decarboxylase (HDC) activity in a CFPS system (PURE*frex* 1.0, Gene Frontier). The HDC enzyme and its cofactor, pyridoxal phosphate (PLP), are both essential for converting L-histidine (substrate) into histamine (product). The small amount of histamine produced in the sample without L-histidine (His-) comes from the L-histidine present as part of the amino acids pool of the CFPS reaction. The HDC was directly produced inside the PURE*frex* 1.0 reaction mix (5 µl final volume) containing the DNA template encoding the HDC (10 nM), PLP (20 µM), and L-histidine (20 mM). The HDC was either constitutively expressed using the P_{T7}-*hdc*-6xHis (WT) construct or triggered by ciprofloxacin (0.1 mM) using P_{T7}-CFXsr5.19-*hdc* (C19) construct which encodes an *hdc* gene controlled by the CFX-a1-sr5-19 riboswitch. A histamine solution (20 mM) was used as a positive control. The samples were incubated at 37°C for 4 h, and then the histamine produced during the CFPS reaction was measured with a "Histamine Test" colorimetric kit (Kikkoman Biochemifa) and Infinite M1000 Pro (Tecan) plate reader (Abs_{470nm}). The error bars represent the standard deviation of three independent assays. Adapted from (Tabuchi and Yokobayashi, 2022) with permission of the authors.
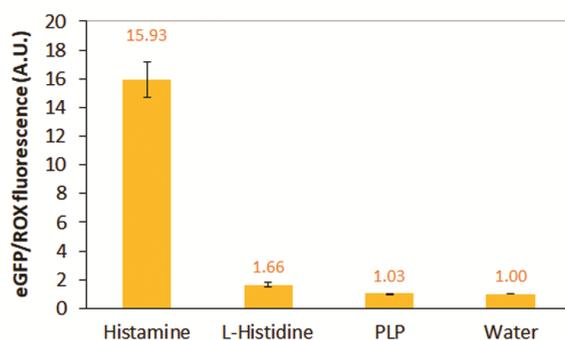
Figure 33. HA-C1g-19 riboswitch is selectively activated by histamine. L-histidine due to its structural similarity to histamine slightly activates the riboswitch but barely above the background level (Water), as the K_d of the aptamer for histamine (~371 nM) is two orders of magnitude smaller than for L-histidine (~23 µM) (Dwidar et al., 2019). Pyridoxal phosphate (PLP) has no effect on the riboswitch activation. The CFPS reaction was performed as previously described (see **4.1.3. Evaluation of individual candidates**, page **48**), with a DNA template (20 nM) encoding eGFP controlled by the HA-C1g-19 riboswitch, and either histamine (5 mM), L-histidine (5 mM), or PLP (20 µM) as the ligand. The expression levels were normalized by that of the untreated sample (Water). The error bars represent standard deviation of three independent assays. Reproduced from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

The *hdc* gene used in this experiment was derived from *Raoultella planticola* ATCC 43176, but the codons were optimized for translation in *Escherichia coli*. All HDC constructs for this and the following experiments were cloned into pTwist Amp vector, while all GFP(1–10) constructs were cloned into pTrpHis vector. The sequence of each cloned construct was verified with Sanger sequencing. Each HDC template was produced by PCR using the aforementioned plasmids as a template, and the P3L and P4L primers. The GFP(1–10) templates were produced in a similar way but using unique forward primers for each construct and a common reverse primer (P-GFP1.10-R). The GFP11 DNA templates were assembled by PCR from 3 or 4 overlapping oligonucleotides, as previously described (see **3.3.2. DNA template construction**, page **32**).

Sender droplets without PLP were used as a negative control, and droplets supplemented with 50 mM histamine instead of L-histidine were used as a positive control. The droplets were mixed in appropriate ratios and incubated at 37°C for 4 h inside the channels (cross-section: 250 μm × 30 μm) of 2D chip-800DG chips (On-chip Biotechnologies). The droplets were observed under an EVOS FL digital inverted fluorescence microscope (Thermo Fisher Scientific). The relative fluorescence intensity of the droplets was estimated from the raw fluorescent micrographs using ImageJ 1.52p. The intensity and the contrast of the fluorescence images presented here were uniformly adjusted for better visualization using the microscope's built-in software and Adobe Photoshop CS6.

The receiver droplets that were surrounded by the sender droplets containing PLP clearly showed higher GFP expression (**Figure 34B, F**) compared to those surrounded by the sender droplets without PLP (**Figure 34C, F**). When the receiver droplets were mixed in excess to the sender droplets, a spatial gradient of GFP expression in the receiver droplets was observed (**Figure 34E**), with the droplets closer to the sender droplets showing stronger GFP fluorescence, in concordance with the diffusion of the histamine from the sender droplets and the dose-dependent behavior of the riboswitch. The receiver droplets adjacent to multiple sender droplets exhibited higher GFP fluorescence than the ones contacting just one sender. The sender droplets containing only histamine confirm that the receiver signal is triggered by histamine, and not PLP or L-histidine (**Figure 34D, F**). This artificial chemical communication mediated by histamine is analogous to bacterial cell-cell communication and quorum sensing systems in which the diffusive signals are produced by cells via an enzymatic reaction and then detected by the receiving cells using a chemical sensor (e.g., transcription factors) (Miller and Bassler, 2001, Horinouchi et al., 2010).

Although some nonpolar and hydrophobic substances inside the droplets may be able to dissolve into the HFE oil used as the medium, hydrophilic substances such as histamine (Lagunoff et al., 1983, Nakao et al., 2011, Nieto-Alamilla et al., 2016)—and ciprofloxacin (Klosinska-Szmurlo et al., 2014) used in the next experiment (see **5.3. Signal transduction with microdroplets**, page **65**)—are unlikely to do so; yet they are still able to diffuse from droplet to droplet as evidenced here. These hydrophilic compounds can probably move across the droplet interphase bilayers (DIB)—a membrane-like bilayer spontaneously formed in between droplets in close contact, mainly composed of surfactants and other amphipathic molecules (Bayley et al., 2008, Taylor and Sarles, 2015, Huang et al., 2022)—without the involvement of membrane channels or pores, via a diffusive translocation mechanism through the DIB (Cramariuc et al., 2012) or other yet undetermined processes.
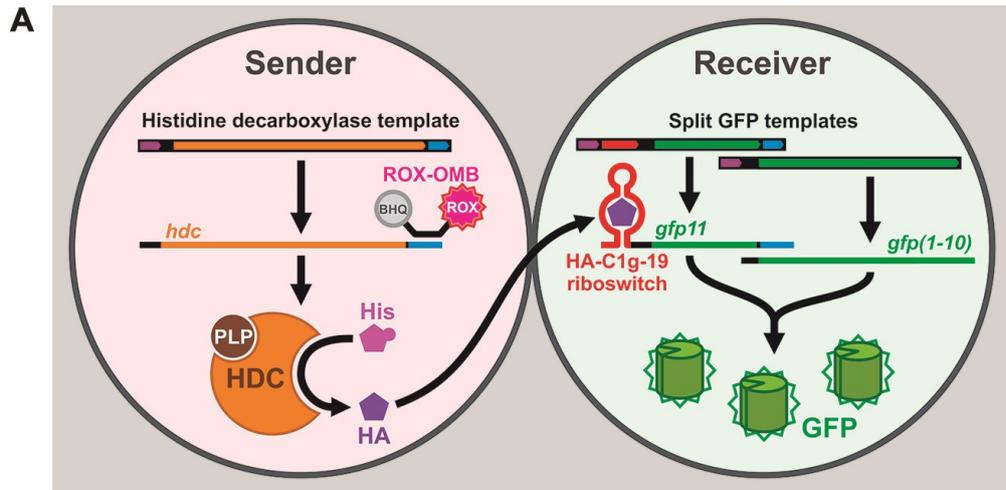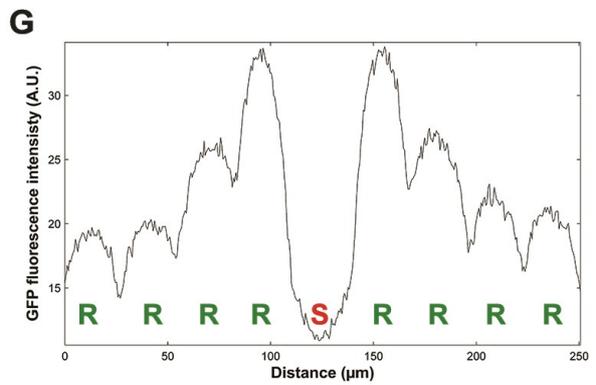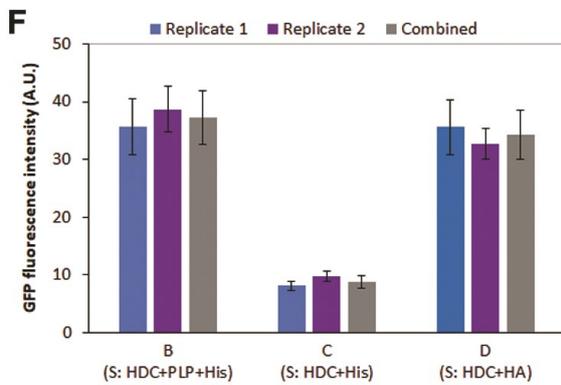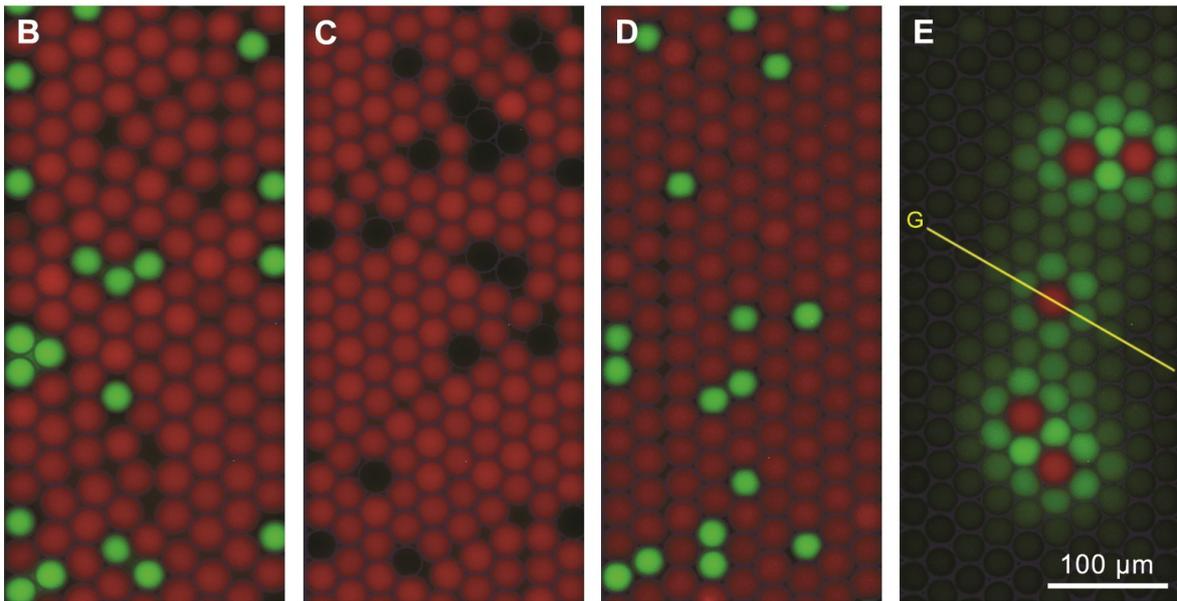
Figure 34. (Previous page) Chemical communication between droplets via histamine. (**A**) Overview of the chemically communicating droplets. The sender droplets (red) contain the DNA template P$_{T7}$-*hdc*-6xHis that encodes histidine decarboxylase (HDC) which converts L-histidine (His) into histamine (HA). Histamine synthesized by the sender droplets diffuses to the receiver droplets (green or non-fluorescent) which contains the histamine ON-riboswitch HA-C1g-19 controlling GFP11 expression, while the GFP1–10 fragment is constitutively expressed. ROX-OMB was included only in the sender droplets to distinguish the two species. (**B–E**) Fluorescence micrographs (merged: GFP + ROX + bright field) of the droplet populations. The droplets were prepared separately and mixed as described, then they were incubated at 37°C for 4 h. The same experiment was performed twice and yielded similar results. Contrast and brightness were enhanced uniformly for visualization. S/R: sender/receiver ratio. (**B**) The sender droplets in excess relative to the receiver droplets. (**C**) Same as **B** but without the essential HDC cofactor pyridoxal phosphate (PLP) in the sender droplets. (**D**) Same as **C** but L-histidine was substituted with histamine. (**E**) The receiver droplets in excess relative to the sender droplets. (**F**) GFP intensity of the receiver droplets calculated from the GFP channel of multiple fluorescence micrographs of conditions **B**, **C**, and **D** using ImageJ 1.52p. The experiment was performed twice yielding similar results. (**G**) GFP intensity profile of the cross-section (yellow line) of the GFP/green channel in panel **E** shows a fluorescence gradient in the receiver droplets (R) surrounding the sender droplet (S). Adapted from (Tabuchi and Yokobayashi, 2022) with permission of the authors.

5.3. SIGNAL TRANSDUCTION WITH MICRODROPLETS

A simple signal transduction cascade (trigger → mediation → response) was recreated using riboswitches and microdroplets. Ciprofloxacin was employed as the first chemical signal (trigger) to induce the production of histamine (mediation), which then would act as a second chemical signal that activates the synthesis of GFP (response). This experiment was performed under similar condition as mentioned before in the previous experiment (see **5.2. One-way communication between microdroplets**, page **61**) with a few modifications in order to adjust the timing of the signaling cascade. The primary sender droplets (trigger) only contain PURE*frex* 2.0 and ciprofloxacin (7 mM). The secondary sender or transducer droplets (mediation) contain the same components as the sender droplets of the previous experiment, with the exception that the DNA template P$_{T7}$-CFXsr5.19-*hdc* (25 nM) now encodes an *hdc* gene controlled by the CFX-a1-sr5-19 ciprofloxacin riboswitch. The receiver droplets (response) also contain the same components as in the previous experiment, but the concentration of the DNA templates for the HA-C1g-19-GFP11 and P$_{T7}$-RBS-GFP(1–10) constructs were adjusted to 20 nM and 10 nM, respectively (**Figure 35A**). The droplets were mixed in appropriate ratios and incubated at 37°C for 5 h inside the channels of 2D chip-800DG chips. Similar as before, the intensity and the contrast of the fluorescence images were adjusted for better visualization using the microscope's built-in software and Adobe Photoshop CS6.

The receiver droplets should only respond to the histamine diffusing from the transducer droplets, which in turn can only be produced in the presence of ciprofloxacin coming from the primary sender droplets. As expected, when observed under a microscope, receiver droplets only showed an increased GFP expression if they were in contact with a transducer droplet immediately adjacent to a primary sender droplet (**Figure 35B, C2**). Receiver droplets surrounding only a primary sender, or a transducer droplet alone did not show any increase in GFP fluorescence (**Figure 35B, C1, C3**). The intensity of the GFP signal of the activated receivers, however, was not as high as in the previous experiment. This is probably due to timing issues regarding the synchronization of the ligand diffusion and the riboswitch activation downstream the signaling cascade. In addition, the histamine yield of the HDC controlled by the CFX-sr5-19 riboswitch is only about 50–60% that of the wild type HDC under the similar conditions (**Figure 32**); therefore, the amount of histamine sensed by the receiver droplets is very likely to be lower than if the wild type HDC is used. Nevertheless, a

signaling cascade entirely controlled and sensed by cell-free riboswitches was successfully achieved with microdroplets.
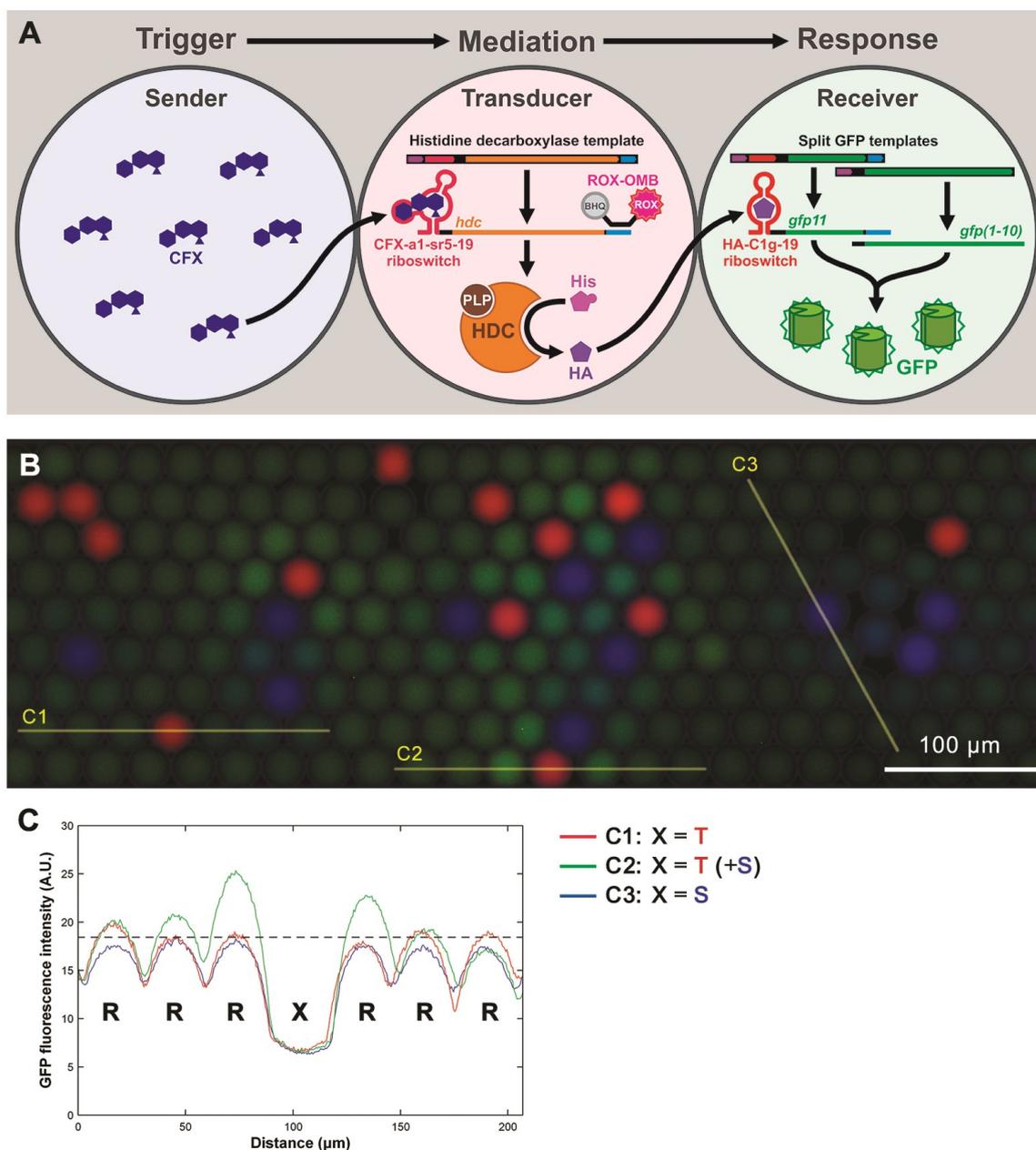
Figure 35. Chemical signal transduction in droplets using riboswitches. (**A**) Overview of the signal transduction cascade in droplets. Trigger: Ciprofloxacin (CFX) working as the initial chemical signal diffuses from the sender droplet (S, blue), and then is detected by a riboswitch in the transducer droplet (T, red). Mediation: Histidine decarboxylase (HDC) is synthetized as a consequence, and the L-histidine (His) inside the droplet is converted into histamine (HA). Response: The histamine, acting as a second chemical signal, then diffuses into a receiver droplet (R, green or non-fluorescent) where it is detected by another riboswitch, triggering the expression of GFP11, while the GFP1–10 fragment is constitutively expressed. (**B**) Fluorescence micrograph (merged: CFX + GFP + ROX + bright field) of the droplet mixture. The droplets were prepared separately and mixed in the appropriate ratios, then they were incubated at 37°C for 5 h. ROX-OMB was included only in the transducer droplets to distinguish them from the other species. Ciprofloxacin can be easily detected using the DAPI/blue channel of the EVOS FL microscope because it is fluorescent (ex. 335 nm, em. 420 nm). Contrast and brightness were enhanced uniformly for visualization. (**C**) GFP intensity profile of the cross-section (yellow lines) of the GFP/green channel in panel **B** shows an increase in fluorescence only in the droplets surrounding a receiver droplet immediately adjacent to a sender droplet (C2).

5.4. OTHER POTENTIAL CIRCUITS USING RIBOSWITCHES

5.4.1. SIGNAL AMPLIFICATION LOOPS

If a histamine-responsive riboswitch is fused to the *hdc* gene, in theory, it is possible to create a positive feedback loop that amplifies a weak initial histamine signal. The HDC enzyme produced as consequence of the first histamine trigger will produce additional histamine that then could trigger the expression of another reporter gene controlled by a less sensitive histamine riboswitch. Such droplet can even become a secondary sender, once the histamine produced inside start diffusing to neighboring droplets. In practice, however, it is very difficult to recreate, and every attempt to produce a controlled signal cascade failed, as evidenced by the random spotted patterns of GFP expression even in the absence of an initial histamine signal (**Figure 36**). First of all, none of the riboswitches are able to completely switch off the expression of HDC; therefore, a single HDC enzyme molecule from a leaky riboswitch could produce enough histamine to trigger a self-amplification loop in an uncontrollable way. Second, due to the chemical similarity, the histamine riboswitch is slightly responsive to L-histidine, the precursor molecule of histamine and substrate for the HDC, especially at high concentrations; consequently, the presence of the precursor molecule within the same droplet can also trigger the riboswitch and start the feedback loop. In order to achieve such circuit, it would be necessary to keep inactive the leaked HDC (e.g., via competitive inhibitors) until the proper activation of the riboswitch occurs, or alternatively, introduce a histamine degradation pathway that can counteract the activity of the leaked HDC.
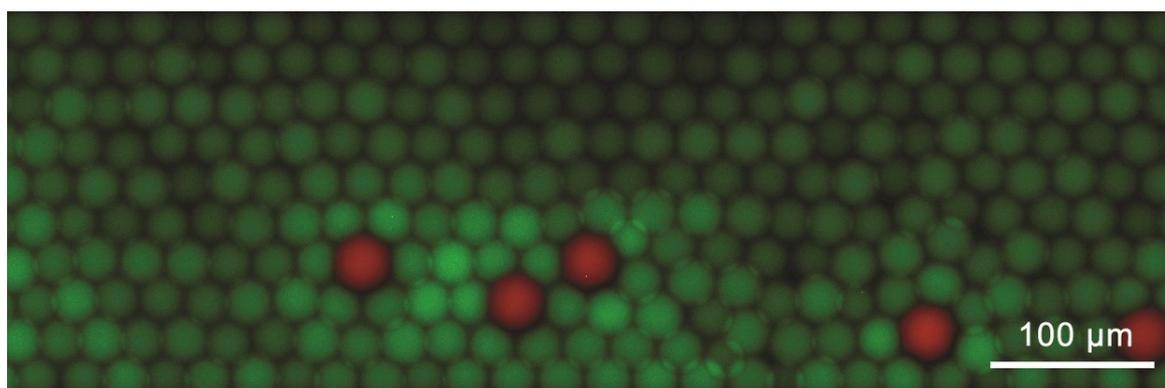
Figure 36. Droplets containing a histamine positive feedback loop intended for signal amplification. Random pattern of GFP expression evidence a stochastic uncontrolled triggering of the histamine riboswitches. Both sender and receiver droplets contain PURE*frex* 2.0, DNA template P_{T7}-HA-C1g-3-*hdc* (20 mM) which encodes *hdc* controlled by a histamine riboswitch HA-C1g-3, L-histidine (25 mM), and pyridoxal phosphate (20 µM). Sender droplets additionally contain ROX-labeled molecular beacon (1 µM), and histamine (50 mM) as the initial trigger signal; while receiver droplets contain the DNA templates for the split GFP system: P_{T7}-RBS-GFP(1–10) (20 mM) and HA-C1g-19-GFP11 (20 nM). The droplets were mixed in the appropriate ratios and incubated at 37°C for 4 h.

5.4.2. BAND-PASS AND BAND-STOP FILTERS

Another interesting application is creating band-pass and band-stop filters by using the split GFP system, where the expression of each fragment was controlled by different riboswitches, so the assembly of the full GFP and the activation of its fluorescence are only obtained within a specific range of the ligand. The OFF-riboswitch was mutated in an attempt to match the sensitivity range and expression level of the ON-switch (**Figure 37**); however, the difference in the sensitivity range of the ON- and OFF- riboswitches was still not ideal, which created a very narrow band window (**Figure 38**).
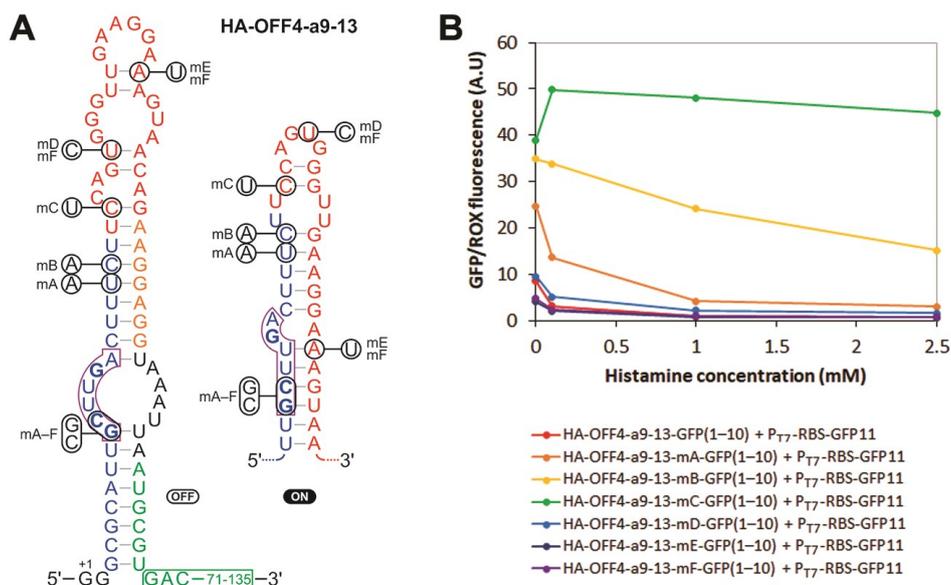
Figure 37. Mutation of HA-OFF4-a9-13 riboswitch. (**A**) Diagram of the mutations. (**B**) Dose-dependence response of the mutated riboswitches controlling the expression of the GFP1–10 fragment (HA-OFF4-a9-13-m#-GFP(1–10)), while GFP11 is constitutively expressed (P_{T7}-RBS-GFP11). The reaction was performed as described in **4.1.3. Evaluation of individual candidates** (page **48**) using 20 nM of each DNA template.
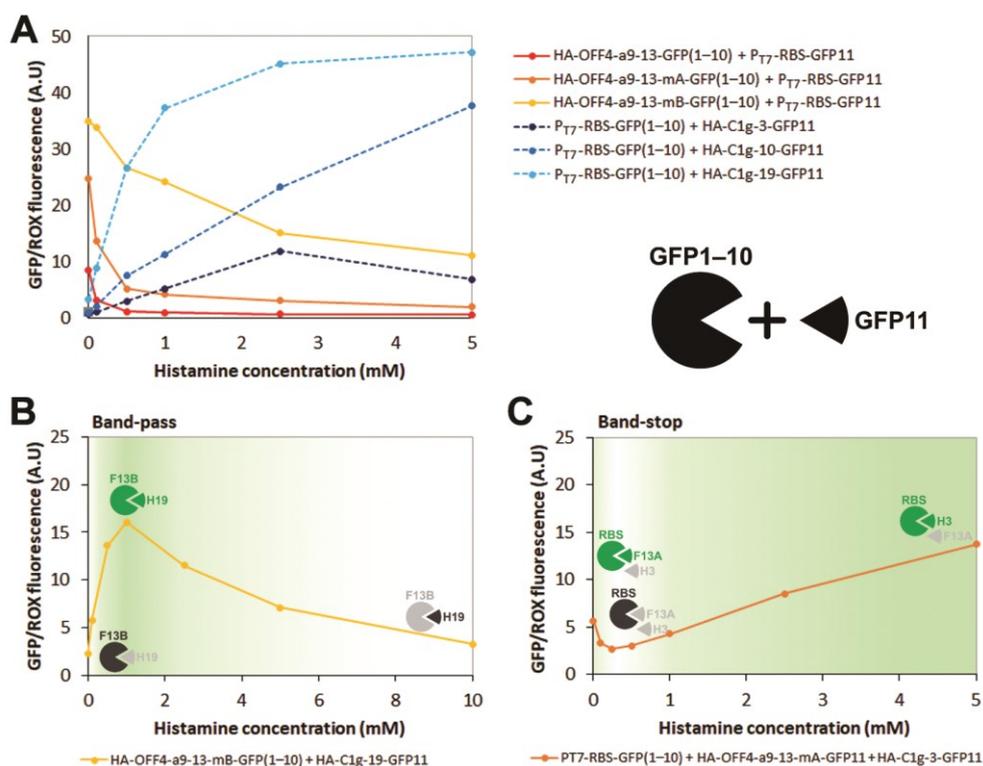
Figure 38. Band-pass and band-stop filters using riboswitches and split GFP. (**A**) Dose-dependent response of histamine OFF-riboswitches controlling the expression of GFP1–10, and histamine ON-riboswitches controlling the expression of GFP11. The other GFP fragment was constitutively expressed. The reaction was performed as described before (see **4.1.3. Evaluation of individual candidates**, page **48**) using 20 nM of each DNA template. (**B**) Band-pass filter in PURE*frex* 1.0 solution created by mixing an OFF-riboswitch controlling GFP1–10 (F13B, HA-OFF4-a9-13-mB-GFP(1–10)) and an ON-riboswitch controlling GFP11 (H19, HA-C1g-19-GFP11). (**C**) Band-stop filter in PURE*frex* 1.0 solution created by mixing a constitutively expressed GFP1–10 (RBS, P_{T7}-RBS-GFP(1–10)) and GFP11 controlled by either an OFF-riboswitch (F13A, HA-OFF4-a9-13-mA-GFP11) or an ON-riboswitch (H3, HA-C1g-3-GFP11). The green and white background represent the concentration ranges in which the GFP signal is allowed (band-pass) or prevented (band-stop).

– 68 –

Preliminary proof-of-concept experiments showed that it is possible to create band-pass and band-stop filters in bulk solution (**Figure 38**); however, they did not function properly when transferred to droplets, as the expected concentric pattern did not appear. These band filters in droplets depend on the formation of a diffusion gradient of the ligand from the sender droplets. Achieving the right concentrations in the gradient within the time frame of the CFPS reaction is crucial, especially if the riboswitch regulation exhibit a kinetical component; however, in practice it is very hard to orchestrate, as both the ligand diffusion and the riboswitch activation are dynamic processes with different timings. It is also possible that the ideal concentration ranges were never achieved within single droplets, maybe due to the droplets being too big and/or the concentration windows of the band filters being too narrow. Finally, there is no protein or mRNA turnover mechanisms in the CFPS system used; therefore, if the riboswitch is activated at some point of the formation of the gradient, the protein produced will remain in the droplet even after the riboswitch is turned off once the gradient changes.

5.4.3. LOGIC GATES

There are already multiple examples of biological logic gate circuits demonstrated to work, both, *in vivo* and *in vitro* using a variety of mechanisms (e.g., genetic cascades, enzymatic reactions, etc.) and components (e.g., transcription factors, enzymes, and/or nucleic acids) (Seelig et al., 2006, Zhou et al., 2009, Qian and Winfree, 2011, Wang et al., 2011, Karig et al., 2012, Iyer et al., 2013, Shis and Bennett, 2013, Sun et al., 2014, Nomura and Yokobayashi, 2015, Katz, 2017, Katz et al., 2017, Sherlock et al., 2018, Bordoy et al., 2019, Jeong et al., 2019, Lehr et al., 2019).
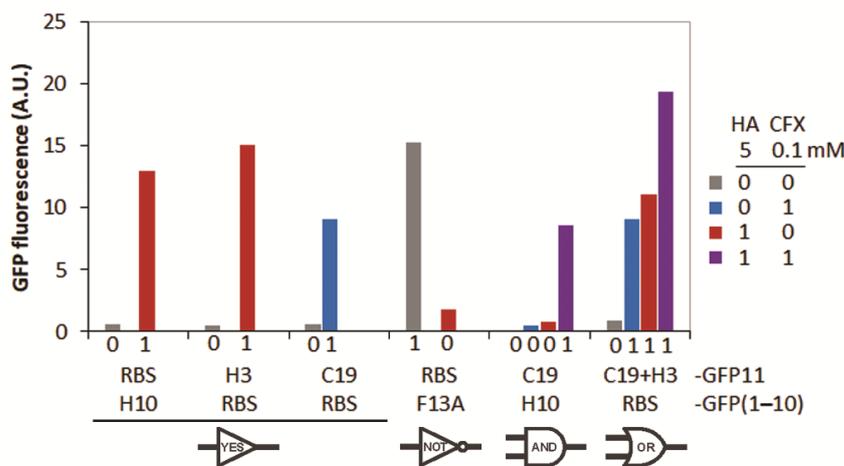
Figure 39. Simple logic gates using riboswitches and split GFP. Histamine (5 mM) and ciprofloxacin (0.1 mM) are used as the inputs for the logic gates, while GFP fluorescence represent the output. RBS (P_{T7}-RBS); H3 (HA-C1g-3); H10 (HA-C1g-10); C19 (CFX-a1-sr5-19); F13A (HA-OFF4-a9-13-mA).

Using a similar split system as with the band-pass/stop described before, simple logic gates that do not involve genetic cascades nor enzymes could be created with riboswitches. For instance, a simple OFF-switch would be the equivalent of a NOT gate; then, different ON-switches controlling each fragment of the split GFP simulate an AND gate, while those same two riboswitches independently controlling the same fragment represent an OR gate (**Figure 39**). NAND and NOR gates could be created in a similar way by replacing the ON-switches with their respective OFF counterparts. In theory, more complex logic circuit could be implemented if different riboswitches and more interacting components are added to the system. Implementation of these kind of logic gates in artificial cells is feasible; however, it would require the fine tuning, in time and space, of the diffusion of each signaling molecule and the activation of the riboswitches.

5.5.   CONCLUSIONS

The riboswitches obtained through the novel screening method were employed to create communication circuits and signaling cascades between artificial cells (more specifically, between water-in-oil microdroplets) in a self-contained system. This droplet-to-droplet communication was entirely controlled and sensed by cell-free riboswitches, which not only demonstrates that the new technique can produce riboswitches readily available for use in cell-free systems, but also shows the potential utility of riboswitches as main sensors and gene regulators in such systems.

CHAPTER 6:
CONCLUSIONS AND FINAL REMARKS

Synthetic cell-free riboswitches suffer fewer biological constraints, such as ligand toxicity and cell permeability, compared to their counterparts functioning in living cells. Ironically, however, cell-free riboswitches may be more difficult and costly to engineer due to the lack of high-throughput assays and screening methods in cell-free systems. Conventional low-throughput assays of individual devices in bulk solutions can also be very expensive, especially when reconstituted CFPS systems (i.e., PURE systems) are used.

To address this problem, a novel high-throughput method for screening riboswitches directly in a cell-free system was developed and put to test. The method allowed the creation of three new types of cell-free riboswitches that respond to two small molecules (histamine and ciprofloxacin) normally not compatible with conventional screening method using living cells. This method is relatively straightforward and does not require special customized instrumentation, as it was developed using only commercially available general-purpose microfluidics devices. This new method should greatly facilitate cell-free riboswitch development which in turn can diversify the chemical interface available to cell-free systems and artificial cells.

For instance, some of the riboswitches obtained through this method were used to demonstrate artificial chemical communication between microdroplets using histamine and ciprofloxacin as signaling molecules, mimicking bacterial cell-to-cell communication, quorum sensing systems, and signal transduction cascades, showing that cell-free riboswitches may enable the development of additional synthetic chemical communication systems for basic and applied research. Furthermore, since RNA aptamers can be selected *in vitro* to recognize diverse molecules, cell-free riboswitches may also enable those systems to detect and respond to a wide variety of chemical signals.

I believe that this work paves the way for cell-free riboswitches to become powerful chemical interface devices for future cell-free systems such as artificial cells and cell-free biomanufacturing platforms.

REFERENCES

Abate, A. R., Hung, T., Mary, P., Agresti, J. J. and Weitz, D. A. (2010) 'High-throughput injection with microfluidics using picoinjectors', *Proc. Natl. Acad. Sci. U.S.A.,* 107(45), 19163-19166.

Abatemarco, J., Sarhan, M. F., Wagner, J. M., Lin, J. L., Liu, L., Hassouneh, W., Yuan, S. F., Alper, H. S. and Abate, A. R. (2017) 'RNA-aptamers-in-droplets (RAPID) high-throughput screening for secretory phenotypes', *Nat. Commun.,* 8(1), 332.

Adamala, K. P., Martin-Alarcon, D. A., Guthrie-Honea, K. R. and Boyden, E. S. (2017) 'Engineering genetic circuit interactions within and between synthetic minimal cells', *Nat. Chem.,* 9(5), 431-439.

Adams, S. R., Campbell, R. E., Gross, L. A., Martin, B. R., Walkup, G. K., Yao, Y., Llopis, J. and Tsien, R. Y. (2002) 'New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications', *J. Am. Chem. Soc.,* 124(21), 6063-6076.

Adams, S. R. and Tsien, R. Y. (2008) 'Preparation of the membrane-permeant biarsenicals FlAsH-EDT2 and ReAsH-EDT2 for fluorescent labeling of tetracysteine-tagged proteins', *Nat. Protoc.,* 3(9), 1527-1534.

Autour, A., Bouhedda, F., Cubi, R. and Ryckelynck, M. (2019) 'Optimization of fluorogenic RNA-based biosensors using droplet-based microfluidic ultrahigh-throughput screening', *Methods,* 161, 46-53.

Babendure, J. R., Babendure, J. L., Ding, J. H. and Tsien, R. Y. (2006) 'Control of mammalian translation by mRNA structure near caps', *RNA,* 12(5), 851-861.

Barrick, J. E. and Breaker, R. R. (2007) 'The distributions, mechanisms, and structures of metabolite-binding riboswitches', *Genome Biol.,* 8(11), R239.

Barsacchi, M., Novoa, E. M., Kellis, M. and Bechini, A. (2016) 'SwiSpot: modeling riboswitches by spotting out switching sequences', *Bioinformatics,* 32(21), 3252-3259.

Batey, R. T. (2006) 'Structures of regulatory elements in mRNAs', *Curr. Opin. Struct. Biol.,* 16(3), 299-306.

Bayer, T. S. and Smolke, C. D. (2005) 'Programmable ligand-controlled riboregulators of eukaryotic gene expression', *Nat. Biotechnol.,* 23(3), 337-343.

Bayley, H., Cronin, B., Heron, A., Holden, M. A., Hwang, W. L., Syeda, R., Thompson, J. and Wallace, M. (2008) 'Droplet interface bilayers', *Mol. Biosyst.,* 4(12), 1191-1208.

Beneyton, T., Krafft, D., Bednarz, C., Kleineberg, C., Woelfer, C., Ivanov, I., Vidakovic-Koch, T., Sundmacher, K. and Baret, J. C. (2018) 'Out-of-equilibrium microcompartments for the bottom-up integration of metabolic functions', *Nat. Commun.,* 9(1), 2391.

Blind, M. and Blank, M. (2015) 'Aptamer Selection Technology and Recent Advances', *Mol. Ther. Nucleic Acids,* 4, e223.

Bordoy, A. E., O'Connor, N. J. and Chatterjee, A. (2019) 'Construction of Two-Input Logic Gates Using Transcriptional Interference', *ACS Synth. Biol.,* 8(10), 2428-2441.

Breaker, R. R. (2008) 'Complex riboswitches', *Science,* 319(5871), 1795-1797.

Breaker, R. R. (2012) 'Riboswitches and the RNA world', *Cold Spring Harb. Perspect. Biol.,* 4(2), a003566.

Brookwell, A., Oza, J. P. and Caschera, F. (2021) 'Biotechnology Applications of Cell-Free Expression Systems', *Life,* 11(12), 1367.

Buddingh, B. C., Elzinga, J. and van Hest, J. C. M. (2020) 'Intercellular communication between artificial cells by allosteric amplification of a molecular signal', *Nat. Commun.,* 11(1), 1652.

Buddingh, B. C. and van Hest, J. C. M. (2017) 'Artificial Cells: Synthetic Compartments with Life-like Functionality and Adaptivity', *Acc. Chem. Res.,* 50(4), 769-777.

Cabantous, S., Nguyen, H. B., Pedelacq, J. D., Koraichi, F., Chaudhary, A., Ganguly, K., Lockard, M. A., Favre, G., Terwilliger, T. C. and Waldo, G. S. (2013) 'A new protein-protein interaction sensor based on tripartite split-GFP association', *Sci. Rep.,* 3, 2854.

Cabantous, S., Terwilliger, T. C. and Waldo, G. S. (2005) 'Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein', *Nat. Biotechnol.,* 23(1), 102-107.

Caen, O., Schutz, S., Jammalamadaka, M. S. S., Vrignon, J., Nizard, P., Schneider, T. M., Baret, J. C. and Taly, V. (2018) 'High-throughput multiplexed fluorescence-activated droplet sorting', *Microsyst. Nanoeng.,* 4, 33.

Ceres, P., Garst, A. D., Marcano-Velazquez, J. G. and Batey, R. T. (2013a) 'Modularity of select riboswitch expression platforms enables facile engineering of novel genetic regulatory devices', *ACS Synth. Biol.,* 2(8), 463-472.

Ceres, P., Trausch, J. J. and Batey, R. T. (2013b) 'Engineering modular 'ON' RNA switches using biological components', *Nucleic Acids Res.,* 41(22), 10449-10461.

Chai, C. (2019) 'Principle of Emulsion PCR and Its Applications in Biotechnology', *J. Anim. Reprod. Biotechnol.,* 34(4), 259-266.

Chappell, J., Takahashi, M. K. and Lucks, J. B. (2015) 'Creating small transcription activating RNAs', *Nat. Chem. Biol.,* 11(3), 214-220.

Chau, T. H. T. and Lee, E. Y. (2020) 'Development of cell-free platform-based toehold switch system for detection of IP-10 mRNA, an indicator for acute kidney allograft rejection diagnosis', *Clin. Chim. Acta,* 510, 619-624.

Chau, T. H. T., Mai, D. H. A., Pham, D. N., Le, H. T. Q. and Lee, E. Y. (2020) 'Developments of Riboswitches and Toehold Switches for Molecular Detection-Biosensing and Molecular Diagnostics', *Int. J. Mol. Sci.,* 21(9), 3192.

Chauvier, A., Picard-Jean, F., Berger-Dancause, J. C., Bastet, L., Naghdi, M. R., Dube, A., Turcotte, P., Perreault, J. and Lafontaine, D. A. (2017) 'Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation', *Nat. Commun.,* 8, 13892.

Cheah, M. T., Wachter, A., Sudarsan, N. and Breaker, R. R. (2007) 'Control of alternative RNA splicing and gene expression by eukaryotic riboswitches', *Nature,* 447(7143), 497-500.

Chen, A. G., Sudarsan, N. and Breaker, R. R. (2011) 'Mechanism for gene control by a natural allosteric group I ribozyme', *RNA,* 17(11), 1967-1972.

Chen, W. W., Balaj, L., Liau, L. M., Samuels, M. L., Kotsopoulos, S. K., Maguire, C. A., Loguidice, L., Soto, H., Garrett, M., Zhu, L. D., Sivaraman, S., Chen, C., Wong, E. T., Carter, B. S., Hochberg, F. H., Breakefield, X. O. and Skog, J. (2013) 'BEAMing and Droplet Digital PCR Analysis of Mutant IDH1 mRNA in Glioma Patient Serum and Cerebrospinal Fluid Extracellular Vesicles', *Mol. Ther. Nucleic Acids,* 2, e109.

Chen, X., Denison, L., Levy, M. and Ellington, A. D. (2009) 'Direct selection for ribozyme cleavage activity in cells', *RNA,* 15(11), 2035-2045.

Chushak, Y., Harbaugh, S., Zimlich, K., Alfred, B., Chavez, J. and Kelley-Loughnane, N. (2021) 'Characterization of synthetic riboswitch in cell-free protein expression systems', *RNA Biol.,* 18(11), 1727-1738.

Cobb, R. E., Chao, R. and Zhao, H. (2013) 'Directed Evolution: Past, Present and Future', *AIChE J.,* 59(5), 1432-1440.

Codrea, V., Hayner, M., Hall, B., Jhaveri, S. and Ellington, A. (2010) 'In vitro selection of RNA aptamers to a small molecule target', *Curr. Protoc. Nucleic Acid Chem.,* Chapter 9, Unit 9.5.1-23.

Collins, D. J., Neild, A., deMello, A., Liu, A. Q. and Ai, Y. (2015) 'The Poisson distribution and beyond: methods for microfluidic droplet production and single cell encapsulation', *Lab Chip.,* 15(17), 3439-3459.

Collins, J. A., Irnov, I., Baker, S. and Winkler, W. C. (2007) 'Mechanism of mRNA destabilization by the glmS ribozyme', *Genes Dev.,* 21(24), 3356-3368.

Coppins, R. L., Hall, K. B. and Groisman, E. A. (2007) 'The intricate world of riboswitches', *Curr. Opin. Microbiol.,* 10(2), 176-181.

Cramariuc, O., Rog, T., Javanainen, M., Monticelli, L., Polishchuk, A. V. and Vattulainen, I. (2012) 'Mechanism for translocation of fluoroquinolones across lipid membranes', *Biochim. Biophys.*

Acta, 1818(11), 2563-2571.

Damiano, L. and Stano, P. (2020) 'On the "Life-Likeness" of Synthetic Cells', *Front. Bioeng. Biotechnol.,* 8, 953.

Denis, J. A., Guillerm, E., Coulet, F., Larsen, A. K. and Lacorte, J. M. (2017) 'The Role of BEAMing and Digital PCR for Multiplexed Analysis in Molecular Oncology in the Era of Next-Generation Sequencing', *Mol. Diagn. Ther.,* 21(6), 587-600.

Desai, S. K. and Gallivan, J. P. (2004) 'Genetic screens and selections for small molecules based on a synthetic riboswitch that activates protein translation', *J. Am. Chem. Soc.,* 126(41), 13247-13254.

Diamante, L., Gatti-Lafranconi, P., Schaerli, Y. and Hollfelder, F. (2013) 'In vitro affinity screening of protein and peptide binders by megavalent bead surface display', *Protein Eng. Des. Sel.,* 26(10), 713-724.

Diehl, F., Li, M., Dressman, D., He, Y., Shen, D., Szabo, S., Diaz, L. A., Jr., Goodman, S. N., David, K. A., Juhl, H., Kinzler, K. W. and Vogelstein, B. (2005) 'Detection and quantification of mutations in the plasma of patients with colorectal tumors', *Proc. Natl. Acad. Sci. U.S.A.,* 102(45), 16368-16373.

Diehl, F., Li, M., He, Y., Kinzler, K. W., Vogelstein, B. and Dressman, D. (2006) 'BEAMing: single-molecule PCR on microparticles in water-in-oil emulsions', *Nat. Methods,* 3(7), 551-559.

Diehl, F. and Smergeliene, E. (2013) 'BEAMing for Cancer', *Genet. Eng. Biotechnol. News,* 33(15), 48-49.

Ding, N., Zhou, S. and Deng, Y. (2021) 'Transcription-Factor-based Biosensor Engineering for Applications in Synthetic Biology', *ACS Synth. Biol.,* 10(5), 911-922.

Ding, Y., Chan, C. Y. and Lawrence, C. E. (2005) 'RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble', *RNA,* 11(8), 1157-1166.

Domin, G., Findeiß, S., Wachsmuth, M., Will, S., Stadler, P. F. and Mörl, M. (2017) 'Applicability of a computational design approach for synthetic riboswitches', *Nucleic Acids Res.,* 45(7), 4108-4119.

Dominguez, D. I., Ryabova, L. A., Pooggin, M. M., Schmidt-Puchta, W., Futterer, J. and Hohn, T. (1998) 'Ribosome shunting in cauliflower mosaic virus. Identification of an essential and sufficient structural element', *J. Biol. Chem.,* 273(6), 3669-3678.

Drachuk, I., Harbaugh, S., Chavez, J. L. and Kelley-Loughnane, N. (2020) 'Improving the Activity of DNA-Encoded Sensing Elements through Confinement in Silk Microcapsules', *ACS Appl. Mater. Interfaces,* 12(43), 48329-48339.

Dupin, A. and Simmel, F. C. (2019) 'Signalling and differentiation in emulsion-based multi-compartmentalized in vitro gene circuits', *Nat. Chem.,* 11(1), 32-39.

Dwidar, M., Seike, Y., Kobori, S., Whitaker, C., Matsuura, T. and Yokobayashi, Y. (2019) 'Programmable Artificial Cells Using Histamine-Responsive Synthetic Riboswitch', *J. Am. Chem. Soc.,* 141(28), 11103-11114.

Ellington, A. D. and Szostak, J. W. (1990) 'In vitro selection of RNA molecules that bind specific ligands', *Nature,* 346(6287), 818-822.

Ender, A., Etzel, M., Hammer, S., Findeiß, S., Stadler, P. and Mörl, M. (2021) 'Ligand-dependent tRNA processing by a rationally designed RNase P riboswitch', *Nucleic Acids Res.,* 49(3), 1784-1800.

Endo, Y. and Sawasaki, T. (2006) 'Cell-free expression systems for eukaryotic protein production', *Curr. Opin. Biotechnol.,* 17(4), 373-380.

Endoh, T., Ohyama, T. and Sugimoto, N. (2019) 'RNA-Capturing Microsphere Particles (R-CAMPs) for Optimization of Functional Aptamers', *Small,* 15(26), e1805062.

Espah Borujeni, A., Mishler, D. M., Wang, J., Huso, W. and Salis, H. M. (2016) 'Automated physics-based design of synthetic riboswitches from diverse RNA aptamers', *Nucleic Acids Res.,* 44(1), 1-13.

Etzel, M. and Morl, M. (2017) 'Synthetic Riboswitches: From Plug and Pray toward Plug and Play', *Biochem.,* 56(9), 1181-1198.

Fallah-Araghi, A., Baret, J. C., Ryckelynck, M. and Griffiths, A. D. (2012) 'A completely in vitro ultrahigh-throughput droplet-based microfluidic screening system for protein engineering and directed evolution', *Lab Chip.,* 12(5), 882-891.

Felletti, M., Stifel, J., Wurmthaler, L. A., Geiger, S. and Hartig, J. S. (2016) 'Twister ribozymes as highly versatile expression platforms for artificial riboswitches', *Nat. Commun.,* 7, 12834.

Fidalgo, L. M., Abell, C. and Huck, W. T. (2007) 'Surface-induced droplet fusion in microfluidic devices', *Lab Chip.,* 7(8), 984-986.

Fiedler, S., Shirley, S. G., Schnelle, T. and Fuhr, G. (1998) 'Dielectrophoretic sorting of particles and cells in a microsystem', *Anal. Chem.,* 70(9), 1909-1915.

Flamm, C., Hofacker, I. L., Stadler, P. F. and Wolfinger, M. T. (2002) 'Barrier Trees of Degenerate Landscapes', *Z. Phys. Chem.,* 216(2).

Forster, U., Weigand, J. E., Trojanowski, P., Suess, B. and Wachtveitl, J. (2012) 'Conformational dynamics of the tetracycline-binding aptamer', *Nucleic Acids Res.,* 40(4), 1807-1817.

Fowler, C. C., Brown, E. D. and Li, Y. (2008) 'A FACS-based approach to engineering artificial riboswitches', *ChemBioChem,* 9(12), 1906-1911.

Frieda, K. L. and Block, S. M. (2012) 'Direct observation of cotranscriptional folding in an adenine riboswitch', *Science,* 338(6105), 397-400.

Fu, X., Zhang, Y., Xu, Q., Sun, X. and Meng, F. (2021) 'Recent Advances on Sorting Methods of High-Throughput Droplet-Based Microfluidics in Enzyme Directed Evolution', *Front. Chem.,* 9, 666867.

Fütterer, J., Kiss-László, Z. and Hohn, T. (1993) 'Nonlinear ribosome migration on cauliflower mosaic virus 35S RNA', *Cell,* 73(4), 789-802.

Garst, A. D., Edwards, A. L. and Batey, R. T. (2011) 'Riboswitches: structures and mechanisms', *Cold Spring Harb. Perspect. Biol.,* 3(6), a003533.

Ge, H. and Marchisio, M. A. (2021) 'Aptamers, Riboswitches, and Ribozymes in S. cerevisiae Synthetic Biology', *Life,* 11(3), 248.

Gilbert, S. D., Stoddard, C. D., Wise, S. J. and Batey, R. T. (2006) 'Thermodynamic and kinetic characterization of ligand binding to the purine riboswitch aptamer domain', *J. Mol. Biol.,* 359(3), 754-768.

Goler, J. A., Carothers, J. M. and Keasling, J. D. (2014) 'Dual-selection for evolution of in vivo functional aptazymes as riboswitch parts', *Methods Mol. Biol.,* 1111, 221-235.

Gordon, C. K. L., Wu, D., Pusuluri, A., Feagin, T. A., Csordas, A. T., Eisenstein, M. S., Hawker, C. J., Niu, J. and Soh, H. T. (2019) 'Click-Particle Display for Base-Modified Aptamer Discovery', *ACS Chem. Biol.,* 14(12), 2652-2662.

Goto, Y., Katoh, T. and Suga, H. (2011) 'Flexizymes for genetic code reprogramming', *Nat. Protoc.,* 6(6), 779-790.

Grate, D. and Wilson, C. (2001) 'Inducible regulation of the S. cerevisiae cell cycle mediated by an RNA aptamer–ligand complex', *Bioorg. Med. Chem.,* 9(10), 2565-2570.

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.

Green, N. J., Grundy, F. J. and Henkin, T. M. (2010) 'The T box mechanism: tRNA as a regulatory molecule', *FEBS Lett.,* 584(2), 318-324.

Gregorio, N. E., Kao, W. Y., Williams, L. C., Hight, C. M., Patel, P., Watts, K. R. and Oza, J. P. (2020) 'Unlocking Applications of Cell-Free Biotechnology through Enhanced Shelf Life and Productivity of E. coli Extracts', *ACS Synth. Biol.,* 9(4), 766-778.

Griffiths, A. D. and Tawfik, D. S. (2006) 'Miniaturising the laboratory in emulsion droplets', *Trends Biotechnol.,* 24(9), 395-402.

Groher, F., Bofill-Bosch, C., Schneider, C., Braun, J., Jager, S., Geissler, K., Hamacher, K. and Suess, B. (2018) 'Riboswitching with ciprofloxacin-development and characterization of a novel RNA regulator', *Nucleic Acids Res.,* 46(4), 2121-2132.

Groher, F. and Suess, B. (2014) 'Synthetic riboswitches - A tool comes of age', *Biochim. Biophys. Acta,* 1839(10), 964-973.

Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R. and Hofacker, I. L. (2008) 'The Vienna RNA websuite', *Nucleic Acids Res.,* 36(Web Server issue), W70-4.

Guindani, C., da Silva, L. C., Cao, S., Ivanov, T. and Landfester, K. (2022) 'Synthetic Cells: From Simple Bio-Inspired Modules to Sophisticated Integrated Systems', *Angew. Chem. Int. Ed.,* 61(16), e202110855.

Günzel, C., Kühnl, F., Arnold, K., Findeiß, S., Weinberg, C. E., Stadler, P. F. and Mörl, M. (2020)

'Beyond Plug and Pray: Context Sensitivity and in silico Design of Artificial Neomycin Riboswitches', *RNA Biol.,* 18(4), 457-467.

Gurbay, A., Garrel, C., Osman, M., Richard, M. J., Favier, A. and Hincal, F. (2002) 'Cytotoxicity in ciprofloxacin-treated human fibroblast cells and protection by vitamin E', *Hum. Exp. Toxicol.,* 21(12), 635-641.

Gutierrez-Gonzalez, M., Latorre, Y., Zuniga, R., Aguillon, J. C., Molina, M. C. and Altamirano, C. (2019) 'Transcription factor engineering in CHO cells for recombinant protein production', *Crit. Rev. Biotechnol.,* 39(5), 665-679.

Hadi, T., Nozzi, N., Melby, J. O., Gao, W., Fuerst, D. E. and Kvam, E. (2020) 'Rolling circle amplification of synthetic DNA accelerates biocatalytic determination of enzyme activity relative to conventional methods', *Sci. Rep.,* 10(1), 10279.

Hallberg, Z. F., Su, Y., Kitto, R. Z. and Hammond, M. C. (2017) 'Engineering and In Vivo Applications of Riboswitches', *Annu. Rev. Biochem.,* 86, 515-539.

Haller, A., Souliere, M. F. and Micura, R. (2011) 'The dynamic nature of RNA as key to understanding riboswitch mechanisms', *Acc. Chem. Res.,* 44(12), 1339-1348.

Hamed, S. M., Elkhatib, W. F., El-Mahallawy, H. A., Helmy, M. M., Ashour, M. S. and Aboshanab, K. M. A. (2018) 'Multiple mechanisms contributing to ciprofloxacin resistance among Gram negative bacteria causing infections to cancer patients', *Sci. Rep.,* 8(1), 12268.

Han, S. I., Soo Kim, H. and Han, A. (2017) 'In-droplet cell concentration using dielectrophoresis', *Biosens. Bioelectron.,* 97, 41-45.

Hanson, S., Berthelot, K., Fink, B., McCarthy, J. E. and Suess, B. (2003) 'Tetracycline-aptamer-mediated translational regulation in yeast', *Mol. Microbiol.,* 49(6), 1627-1637.

Harbaugh, S. V., Goodson, M. S., Dillon, K., Zabarnick, S. and Kelley-Loughnane, N. (2017) 'Riboswitch-Based Reversible Dual Color Sensor', *ACS Synth. Biol.,* 6(5), 766-781.

Harbaugh, S. V., Martin, J. A., Weinstein, J., Ingram, G. and Kelley-Loughnane, N. (2018) 'Screening and selection of artificial riboswitches', *Methods,* 143, 77-89.

Harbers, M. (2014) 'Wheat germ systems for cell-free protein expression', *FEBS Lett.,* 588(17), 2762-2773.

Harvey, I., Garneau, P. and Pelletier, J. (2002) 'Inhibition of translation by RNA-small molecule interactions', *RNA,* 8(4), 452-463.

Hennig, S., Rodel, G. and Ostermann, K. (2015) 'Artificial cell-cell communication as an emerging tool in synthetic biology applications', *J. Biol. Eng.,* 9, 13.

Hermann, T. and Patel, D. J. (2000) 'Adaptive recognition by nucleic acid aptamers', *Science,* 287(5454), 820-825.

Hofacker, I. L. and Stadler, P. F. (2006) 'Memory efficient folding algorithms for circular RNA secondary structures', *Bioinformatics,* 22(10), 1172-1176.

Holdhoff, M., Schmidt, K., Diehl, F., Aggrawal, N., Angenendt, P., Romans, K., Edelstein, D. L., Torbenson, M., Kinzler, K. W., Vogelstein, B., Choti, M. A. and Diaz, L. A., Jr. (2011) 'Detection of tumor DNA at the margins of colorectal cancer liver metastasis', *Clin. Cancer Res.,* 17(11), 3551-3557.

Holland-Moritz, D. A., Wismer, M. K., Mann, B. F., Farasat, I., Devine, P., Guetschow, E. D., Mangion, I., Welch, C. J., Moore, J. C., Sun, S. and Kennedy, R. T. (2020) 'Mass Activated Droplet Sorting (MADS) Enables High-Throughput Screening of Enzymatic Reactions at Nanoliter Scale', *Angew. Chem. Int. Ed.,* 59(11), 4470-4477.

Holstein, J. M., Gylstorff, C. and Hollfelder, F. (2021) 'Cell-free Directed Evolution of a Protease in Microdroplets at Ultrahigh Throughput', *ACS Synth. Biol.,* 10(2), 252-257.

Horinouchi, S., Ueda, K., Nakayama, J. and Ikeda, T. (2010) 'Cell-to-Cell Communications among Microorganisms', in Mander, L. & Liu, H.W. (eds.) *Comprehensive Natural Products II: Chemistry and Biology*. Kidlington: Elsevier, 283-337.

Hu, X., Bessette, P. H., Qian, J., Meinhart, C. D., Daugherty, P. S. and Soh, H. T. (2005) 'Marker-specific sorting of rare cells using dielectrophoresis', *Proc. Natl. Acad. Sci. U.S.A.,* 102(44), 15757-15761.

Huang, L. C., Pan, X., Yang, H., Wan, L. K., Stewart-Jones, G., Dorrell, L. and Ogg, G. (2013) 'Linking genotype to phenotype on beads: high throughput selection of peptides with biological function', *Sci. Rep.,* 3, 3030.

Huang, Y., Fuller, G. G. and Chandran Suja, V. (2022) 'Physicochemical characteristics of droplet interface bilayers', *Adv. Colloid Interface Sci.,* 304, 102666.

Iacumin, L., Cecchini, F., Vendrame, M. and Comi, G. (2020) 'Emulsion PCR (ePCR) as a Tool to Improve the Power of DGGE Analysis for Microbial Population Studies', *Microorganisms,* 8(8), 1099.

Ishida, S., Terasaka, N., Katoh, T. and Suga, H. (2020) 'An aminoacylation ribozyme evolved from a natural tRNA-sensing T-box riboswitch', *Nat. Chem. Biol.,* 16(6), 702-709.

Ivanov, I., Castellanos, S. L., Balasbas, S., 3rd, Otrin, L., Marusic, N., Vidakovic-Koch, T. and Sundmacher, K. (2021) 'Bottom-Up Synthesis of Artificial Cells: Recent Highlights and Future Challenges', *Annu. Rev. Chem. Biomol. Eng.,* 12, 287-308.

Iyer, S., Karig, D. K., Norred, S. E., Simpson, M. L. and Doktycz, M. J. (2013) 'Multi-input regulation and logic with T7 promoters in cells and cell-free systems', *PLoS One,* 8(10), e78442.

Jacoby, G. A., Strahilevitz, J. and Hooper, D. C. (2014) 'Plasmid-mediated quinolone resistance', *Microbiol. Spectr.,* 2(5), 2.5.33.

Jaeger, J., Groher, F., Stamm, J., Spiehl, D., Braun, J., Dorsam, E. and Suess, B. (2019) 'Characterization and Inkjet Printing of an RNA Aptamer for Paper-Based Biosensing of Ciprofloxacin', *Biosensors,* 9(1), 7.

Jang, S., Lee, B., Jeong, H. H., Jin, S. H., Jang, S., Kim, S. G., Jung, G. Y. and Lee, C. S. (2016) 'On-chip analysis, indexing and screening for chemical producing bacteria in a microfluidic static droplet array', *Lab Chip.,* 16(10), 1909-1916.

Jeng, S. T., Gardner, J. F. and Gumport, R. I. (1990) 'Transcription Termination by Bacteriophage T7 Rna-Polymerase at Rho-Independent Terminators', *J. Biol. Chem.,* 265(7), 3823-3830.

Jeng, S. T., Gardner, J. F. and Gumport, R. I. (1992) 'Transcription termination in vitro by bacteriophage T7 RNA polymerase. The role of sequence elements within and surrounding a rho-independent transcription terminator', *J. Biol. Chem.,* 267(27), 19306-19312.

Jeong, D., Klocke, M., Agarwal, S., Kim, J., Choi, S., Franco, E. and Kim, J. (2019) 'Cell-Free Synthetic Biology Platform for Engineering Synthetic Biological Circuits and Systems', *Methods Protoc.,* 2(2), 39.

Jiang, S., Caire da Silva, L., Ivanov, T., Mottola, M. and Landfester, K. (2022) 'Synthetic Silica Nano-Organelles for Regulation of Cascade Reactions in Multi-Compartmentalized Systems', *Angew. Chem. Int. Ed.,* 61(6), e202113784.

Jones, C. P. and Ferre-D'Amare, A. R. (2017) 'Long-Range Interactions in Riboswitch Control of Gene Expression', *Annu. Rev. Biophys.,* 46, 455-481.

Kaminski, T. S., Scheler, O. and Garstecki, P. (2016) 'Droplet microfluidics for microbiology: techniques, applications and challenges', *Lab Chip.,* 16(12), 2168-2187.

Kamura, N., Sawasaki, T., Kasahara, Y., Takai, K. and Endo, Y. (2005) 'Selection of 5′-untranslated sequences that enhance initiation of translation in a cell-free protein synthesis system from wheat embryos', *Bioorg. Med. Chem. Lett.,* 15(24), 5402-5406.

Karbaschi, M., Shahi, P. and Abate, A. R. (2017) 'Rapid, chemical-free breaking of microfluidic emulsions with a hand-held antistatic gun', *Biomicrofluidics,* 11(4), 044107.

Karig, D. K., Iyer, S., Simpson, M. L. and Doktycz, M. J. (2012) 'Expression optimization and synthetic gene networks in cell-free systems', *Nucleic Acids Res.,* 40(8), 3763-3774.

Karns, K., Vogan, J. M., Qin, Q., Hickey, S. F., Wilson, S. C., Hammond, M. C. and Herr, A. E. (2013) 'Microfluidic screening of electrophoretic mobility shifts elucidates riboswitch binding function', *J. Am. Chem. Soc.,* 135(8), 3136-3143.

Karzbrun, E., Tayar, A. M., Noireaux, V. and Bar-Ziv, R. H. (2014) 'Synthetic biology. Programmable on-chip DNA compartments as artificial cells', *Science,* 345(6198), 829-832.

Katz, E. (2017) 'Enzyme-Based Logic Gates and Networks with Output Signals Analyzed by Various Methods', *ChemPhysChem,* 18(13), 1688-1713.

Katz, E., Poghossian, A. and Schoning, M. J. (2017) 'Enzyme-based logic gates and circuits-analytical applications and interfacing with electronics', *Anal. Bioanal. Chem.,* 409(1), 81-94.

Kazanov, M. D., Vitreschak, A. G. and Gelfand, M. S. (2007) 'Abundance and functional diversity of riboswitches in microbial communities', *BMC Genomics,* 8, 347.

Kelly, B. T., Baret, J. C., Taly, V. and Griffiths, A. D. (2007) 'Miniaturizing chemistry and biology in microdroplets', *Chem. Commun.,* 38(18), 1773-1788.

Ketzer, P., Kaufmann, J. K., Engelhardt, S., Bossow, S., von Kalle, C., Hartig, J. S., Ungerechts, G. and Nettelbeck, D. M. (2014) 'Artificial riboswitches for gene expression and replication control of DNA and RNA viruses', *Proc. Natl. Acad. Sci. U.S.A.,* 111(5), E554-62.

Khambhati, K., Bhattacharjee, G., Gohil, N., Braddick, D., Kulkarni, V. and Singh, V. (2019) 'Exploring the Potential of Cell-Free Protein Synthesis for Extending the Abilities of Biological Systems', *Front. Bioeng. Biotechnol.,* 7, 248.

Kinney, J. B., Murugan, A., Callan, C. G., Jr. and Cox, E. C. (2010) 'Using deep sequencing to characterize the biophysical mechanism of a transcriptional regulatory sequence', *Proc. Natl. Acad. Sci. U.S.A.,* 107(20), 9158-9163.

Kirchner, M., Schorpp, K., Hadian, K. and Schneider, S. (2017) 'An in vivo high-throughput screening for riboswitch ligands using a reverse reporter gene system', *Sci. Rep.,* 7(1), 7732.

Klein, D. J. and Ferre-D'Amare, A. R. (2006) 'Structural basis of glmS ribozyme activation by glucosamine-6-phosphate', *Science,* 313(5794), 1752-1756.

Klosinska-Szmurlo, E., Plucinski, F. A., Grudzien, M., Betlejewska-Kielak, K., Biernacka, J. and Mazurek, A. P. (2014) 'Experimental and theoretical studies on the molecular properties of ciprofloxacin, norfloxacin, pefloxacin, sparfloxacin, and gatifloxacin in determining bioavailability', *J. Biol. Phys.,* 40(4), 335-345.

Koberstein, J. N., Stewart, M. L., Mighell, T. L., Smith, C. B. and Cohen, M. S. (2021) 'A Sort-Seq Approach to the Development of Single Fluorescent Protein Biosensors', *ACS Chem. Biol.,* 16(9), 1709-1720.

Kobori, S., Ichihashi, N., Kazuta, Y., Matsuura, T. and Yomo, T. (2012) 'Kinetic analysis of aptazyme-regulated gene expression in a cell-free translation system: modeling of ligand-dependent and -independent expression', *RNA,* 18(8), 1458-1465.

Kobori, S., Nomura, Y., Miu, A. and Yokobayashi, Y. (2015) 'High-throughput assay and engineering of self-cleaving ribozymes by sequencing', *Nucleic Acids Res.,* 43(13), e85.

Kobori, S. and Yokobayashi, Y. (2016) 'High-Throughput Mutational Analysis of a Twister Ribozyme', *Angew. Chem. Int. Ed.,* 55(35), 10354-10357.

Koizumi, M., Soukup, G. A., Kerr, J. N. and Breaker, R. R. (1999) 'Allosteric selection of ribozymes that respond to the second messengers cGMP and cAMP', *Nat. Struct. Biol.,* 6(11), 1062-1071.

Kosuri, S., Goodman, D. B., Cambray, G., Mutalik, V. K., Gao, Y., Arkin, A. P., Endy, D. and Church, G. M. (2013) 'Composability of regulatory sequences controlling transcription and translation in Escherichia coli', *Proc. Natl. Acad. Sci. U.S.A.,* 110(34), 14024-14029.

Krell, T., Gavira, J. A., Velando, F., Fernandez, M., Roca, A., Monteagudo-Cascales, E. and Matilla, M. A. (2021) 'Histamine: A Bacterial Signal Molecule', *Int. J. Mol. Sci.,* 22(12), 6312.

Krinsky, N., Kaduri, M., Shainsky-Roitman, J., Goldfeder, M., Ivanir, E., Benhar, I., Shoham, Y. and Schroeder, A. (2016) 'A Simple and Rapid Method for Preparing a Cell-Free Bacterial Lysate for Protein Synthesis', *PLoS One,* 11(10), e0165137.

Kubodera, T., Watanabe, M., Yoshiuchi, K., Yamashita, N., Nishimura, A., Nakai, S., Gomi, K. and Hanamoto, H. (2003) 'Thiamine-regulated gene expression of Aspergillus oryzae thiA requires splicing of the intron containing a riboswitch-like domain in the 5′-UTR', *FEBS Lett.,* 555(3), 516-520.

Kwon, Y. C. and Jewett, M. C. (2015) 'High-throughput preparation methods of crude extract for robust cell-free protein synthesis', *Sci. Rep.,* 5, 8663.

Lagunoff, D., Martin, T. W. and Read, G. (1983) 'Agents that release histamine from mast cells', *Annu. Rev. Pharmacol. Toxicol.,* 23, 331-351.

Lai, S. N., Zhou, X., Ouyang, X., Zhou, H., Liang, Y., Xia, J. and Zheng, B. (2020) 'Artificial Cells Capable of Long-Lived Protein Synthesis by Using Aptamer Grafted Polymer Hydrogel', *ACS Synth. Biol.,* 9(1), 76-83.

Lang, K., Rieder, R. and Micura, R. (2007) 'Ligand-induced folding of the thiM TPP riboswitch investigated by a structure-based fluorescence spectroscopic approach', *Nucleic Acids Res.,* 35(16), 5370-5378.

Laohakunakorn, N., Grasemann, L., Lavickova, B., Michielin, G., Shahein, A., Swank, Z. and Maerkl, S. J. (2020) 'Bottom-Up Construction of Complex Biomolecular Systems With Cell-Free Synthetic Biology', *Front. Bioeng. Biotechnol.,* 8, 213.

Lauring, J. and Park, B. H. (2011) 'BEAMing sheds light on drug resistance', *Clin. Cancer Res.*, 17(24), 7508-7510.

Lawrence, J. W., Claire, D. C., Weissig, V. and Rowe, T. C. (1996) 'Delayed cytotoxicity and cleavage of mitochondrial DNA in ciprofloxacin-treated mammalian cells', *Mol. Pharmacol.*, 50(5), 1178-1188.

Lee, D., Hwang, B. and Kim, B. (2016) 'The potential of a dielectrophoresis activated cell sorter (DACS) as a next generation cell sorter', *Micro Nano Syst. Lett.*, 4(1), 2.

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

Lemay, J. F., Desnoyers, G., Blouin, S., Heppell, B., Bastet, L., St-Pierre, P., Masse, E. and Lafontaine, D. A. (2011) 'Comparative study between transcriptionally- and translationally-acting adenine riboswitches reveals key differences in riboswitch regulatory mechanisms', *PLoS Genet.*, 7(1), e1001278.

Lentini, R., Santero, S. P., Chizzolini, F., Cecchi, D., Fontana, J., Marchioretto, M., Del Bianco, C., Terrell, J. L., Spencer, A. C., Martini, L., Forlin, M., Assfalg, M., Dalla Serra, M., Bentley, W. E. and Mansy, S. S. (2014) 'Integrating artificial with natural cells to translate chemical messages that direct E. coli behaviour', *Nat. Commun.*, 5, 4012.

Lentini, R., Yeh Martin, N. and Mansy, S. S. (2016) 'Communicating artificial cells', *Curr. Opin. Chem. Biol.*, 34, 53-61.

Li, J. W., Zhang, X. Y., Wu, H. and Bai, Y. P. (2020) 'Transcription Factor Engineering for High-Throughput Strain Evolution and Organic Acid Bioproduction: A Review', *Front. Bioeng. Biotechnol.*, 8, 98.

Li, S. and Breaker, R. R. (2013) 'Eukaryotic TPP riboswitch regulation of alternative splicing involving long-distance base pairing', *Nucleic Acids Res.*, 41(5), 3022-3031.

Liang, J. C., Chang, A. L., Kennedy, A. B. and Smolke, C. D. (2012) 'A high-throughput, quantitative cell-based screen for efficient tailoring of RNA device activity', *Nucleic Acids Res.*, 40(20), e154.

Lindenburg, L. and Hollfelder, F. (2021) '"NAD-display": Ultrahigh-Throughput in Vitro Screening of NAD(H) Dehydrogenases Using Bead Display and Flow Cytometry', *Angew. Chem. Int. Ed.*, 60(16), 9015-9021.

Link, K. H., Guo, L., Ames, T. D., Yen, L., Mulligan, R. C. and Breaker, R. R. (2007) 'Engineering high-speed allosteric hammerhead ribozymes', *Biol. Chem.*, 388(8), 779-786.

Liu, Y., Wang, N., Chan, C. W., Lu, A., Yu, Y., Zhang, G. and Ren, K. (2021) 'The Application of Microfluidic Technologies in Aptamer Selection', *Front. Cell Dev. Biol.*, 9, 730035.

Lorenz, R., Bernhart, S. H., Honer Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F. and Hofacker, I. L. (2011) 'ViennaRNA Package 2.0', *Algorithms Mol. Biol.*, 6, 26.

Lotz, T. S. and Suess, B. (2018) 'Small-Molecule-Binding Riboswitches', *Microbiol. Spectr.*, 6(4), RWR-0025-2018.

Loveday, E. K., Zath, G. K., Bikos, D. A., Jay, Z. J. and Chang, C. B. (2021) 'Screening of Additive Formulations Enables Off-Chip Drop Reverse Transcription Quantitative Polymerase Chain Reaction of Single Influenza A Virus Genomes', *Anal. Chem.*, 93(10), 4365-4373.

Lu, H., Caen, O., Vrignon, J., Zonta, E., El Harrak, Z., Nizard, P., Baret, J. C. and Taly, V. (2017) 'High throughput single cell counting in droplet-based microfluidics', *Sci. Rep.*, 7(1), 1366.

Lu, Y. (2017) 'Cell-free synthetic biology: Engineering in an open world', *Synth. Syst. Biotechnol.*, 2(1), 23-27.

Lutz, B., Faber, M., Verma, A., Klumpp, S. and Schug, A. (2014) 'Differences between cotranscriptional and free riboswitch folding', *Nucleic Acids Res.*, 42(4), 2687-2696.

Lynch, S. A., Desai, S. K., Sajja, H. K. and Gallivan, J. P. (2007) 'A high-throughput screen for synthetic riboswitches reveals mechanistic insights into their function', *Chem. Biol.*, 14(2), 173-184.

Lynch, S. A. and Gallivan, J. P. (2009) 'A flow cytometry-based screen for synthetic riboswitches', *Nucleic Acids Res.*, 37(1), 184-192.

Ma, D., Shen, L., Wu, K., Diehnelt, C. W. and Green, A. A. (2018a) 'Low-cost detection of norovirus using paper-based cell-free systems and synbody-based viral enrichment', *Synth. Biol.*, 3(1),

ysy018.

Ma, G., Wu, G., Li, X., Wang, H. and Zhou, M. (2018b) 'Characterization of Ciprofloxacin Resistance in Laboratory-Derived Mutants of Vibrio parahaemolyticus with qnr Gene', *Foodborne Pathog. Dis.,* 15(11), 711-717.

Machtel, P., Bakowska-Zywicka, K. and Zywicki, M. (2016) 'Emerging applications of riboswitches - from antibacterial targets to molecular tools', *J. Appl. Genet.,* 57(4), 531-541.

Madin, K., Sawasaki, T., Ogasawara, T. and Endo, Y. (2000) 'A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes', *Proc. Natl. Acad. Sci. U.S.A.,* 97(2), 559-564.

Mandal, M. and Breaker, R. R. (2004) 'Adenine riboswitches and gene activation by disruption of a transcription terminator', *Nat. Struct. Mol. Biol.,* 11(1), 29-35.

Mankowska, S. A., Gatti-Lafranconi, P., Chodorge, M., Sridharan, S., Minter, R. R. and Hollfelder, F. (2016) 'A Shorter Route to Antibody Binders via Quantitative in vitro Bead-Display Screening and Consensus Analysis', *Sci. Rep.,* 6, 36391.

Marcoux, P. R., Dupoy, M., Mathey, R., Novelli-Rousseau, A., Heran, V., Morales, S., Rivera, F., Joly, P. L., Moy, J.-P. and Mallard, F. (2011) 'Micro-confinement of bacteria into w/o emulsion droplets for rapid detection and enumeration', *Colloids Surf. A: Physicochem. Eng. Asp.,* 377(1-3), 54-62.

Martin, B. R., Giepmans, B. N. G., Adams, S. R. and Tsien, R. Y. (2005) 'Mammalian cell–based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity', *Nat. Biotechnol.,* 23(10), 1308-1314.

Martini, L. and Mansy, S. S. (2011) 'Cell-like systems with riboswitch controlled gene expression', *Chem. Commun.,* 47(38), 10734-10736.

Martini, L. and Mansy, S. S. (2014) 'Measuring riboswitch activity in vitro and in artificial cells with purified transcription-translation machinery', *Methods Mol. Biol.,* 1111, 153-164.

Mazutis, L., Baret, J. C., Treacy, P., Skhiri, Y., Araghi, A. F., Ryckelynck, M., Taly, V. and Griffiths, A. D. (2009) 'Multi-step microfluidic droplet processing: kinetic analysis of an in vitro translated enzyme', *Lab Chip.,* 9(20), 2902-2908.

McCaskill, J. S. (1990) 'The equilibrium partition function and base pair binding probabilities for RNA secondary structure', *Biopolymers,* 29(6-7), 1105-1119.

McCown, P. J., Corbino, K. A., Stav, S., Sherlock, M. E. and Breaker, R. R. (2017) 'Riboswitch diversity and distribution', *RNA,* 23(7), 995-1011.

Mercier, J. F. and Slater, G. W. (2005) 'Solid phase DNA amplification: a Brownian dynamics study of crowding effects', *Biophys. J.,* 89(1), 32-42.

Michu, E., Mráčková, M., Vyskot, B. and Žlůvová, J. (2010) 'Reduction of heteroduplex formation in PCR amplification', *Biol. Plant.,* 54(1), 173-176.

Miller, M. B. and Bassler, B. L. (2001) 'Quorum sensing in bacteria', *Annu. Rev. Microbiol.,* 55, 165-199.

Mishler, D. M. and Gallivan, J. P. (2014) 'A family of synthetic riboswitches adopts a kinetic trapping mechanism', *Nucleic Acids Res.,* 42(10), 6753-6761.

Morita, E. H., Sawasaki, T., Tanaka, R., Endo, Y. and Kohno, T. (2003) 'A wheat germ cell-free system is a novel way to screen protein folding and function', *Protein Sci.,* 12(6), 1216-1221.

Muller, S., Appel, B., Balke, D., Hieronymus, R. and Nubel, C. (2016) 'Thirty-five years of research into ribozymes and nucleic acid catalysis: where do we stand today?', *F1000Res,* 5, 1511.

Muranaka, N., Sharma, V., Nomura, Y. and Yokobayashi, Y. (2009a) 'Efficient Design Strategy for Whole-Cell and Cell-Free Biosensors based on Engineered Riboswitches', *Anal. Lett.,* 42(1), 108-122.

Muranaka, N., Sharma, V., Nomura, Y. and Yokobayashi, Y. (2009b) 'An efficient platform for genetic selection and screening of gene switches in Escherichia coli', *Nucleic Acids Res.,* 37(5), e39.

Muranaka, N. and Yokobayashi, Y. (2010) 'Posttranscriptional signal integration of engineered riboswitches yields band-pass output', *Angew. Chem. Int. Ed.,* 49(27), 4653-4655.

Mustafina, K., Fukunaga, K. and Yokobayashi, Y. (2020) 'Design of Mammalian ON-Riboswitches Based on Tandemly Fused Aptamer and Ribozyme', *ACS Synth. Biol.,* 9(1), 19-25.

Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L. and Breaker, R. R. (2002) 'Genetic Control by a Metabolite Binding mRNA', *Chem. Biol.,* 9(9), 1043-1049.

Nakano, M., Komatsu, J., Matsuura, S.-i., Takashima, K., Katsura, S. and Mizuno, A. (2003) 'Single-molecule PCR using water-in-oil emulsion', *J. Biotechnol.,* 102(2), 117-124.

Nakao, S., Komagoe, K., Inoue, T. and Katsu, T. (2011) 'Comparative study of the membrane-permeabilizing activities of mastoparans and related histamine-releasing agents in bacteria, erythrocytes, and mast cells', *Biochim. Biophys. Acta,* 1808(1), 490-497.

Narita, A., Ogawa, K., Sando, S. and Aoyama, Y. (2006) 'Visible sensing of nucleic acid sequences with a genetically encodable unmodified RNA probe', *Angew. Chem. Int. Ed.,* 45(18), 2879-2883.

Nevin, D. E. and Pratt, J. M. (1991) 'A coupled in vitro transcription-translation system for the exclusive synthesis of polypeptides expressed from the T7 promoter', *FEBS Lett.,* 291(2), 259-263.

Nieto-Alamilla, G., Marquez-Gomez, R., Garcia-Galvez, A. M., Morales-Figueroa, G. E. and Arias-Montano, J. A. (2016) 'The Histamine H3 Receptor: Structure, Pharmacology, and Function', *Mol. Pharmacol.,* 90(5), 649-673.

Noderer, W. L., Flockhart, R. J., Bhaduri, A., Diaz de Arce, A. J., Zhang, J., Khavari, P. A. and Wang, C. L. (2014) 'Quantitative analysis of mammalian translation initiation sites by FACS-seq', *Mol. Syst. Biol.,* 10, 748.

Noireaux, V., Bar-Ziv, R. and Libchaber, A. (2003) 'Principles of cell-free genetic circuit assembly', *Proc. Natl. Acad. Sci. U.S.A.,* 100(22), 12672-12677.

Nomura, Y., Chien, H. C. and Yokobayashi, Y. (2017) 'Direct screening for ribozyme activity in mammalian cells', *Chem. Commun.,* 53(93), 12540-12543.

Nomura, Y. and Yokobayashi, Y. (2015) 'Aptazyme-based riboswitches and logic gates in mammalian cells', *Methods Mol. Biol.,* 1316, 141-148.

Nomura, Y., Zhou, L., Miu, A. and Yokobayashi, Y. (2013) 'Controlling mammalian gene expression by allosteric hepatitis delta virus ribozymes', *ACS Synth. Biol.,* 2(12), 684-689.

Nord, O., Uhlen, M. and Nygren, P. A. (2003) 'Microbead display of proteins by cell-free expression of anchored DNA', *J. Biotechnol.,* 106(1), 1-13.

Nshogozabahizi, J. C., Aubrey, K. L., Ross, J. A. and Thakor, N. (2019) 'Applications and limitations of regulatory RNA elements in synthetic biology and biotechnology', *J. Appl. Microbiol.,* 127(4), 968-984.

Nudler, E. and Mironov, A. S. (2004) 'The riboswitch control of bacterial metabolism', *Trends Biochem. Sci.,* 29(1), 11-17.

Obexer, R., Godina, A., Garrabou, X., Mittl, P. R., Baker, D., Griffiths, A. D. and Hilvert, D. (2017) 'Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase', *Nat. Chem.,* 9(1), 50-56.

Obexer, R., Pott, M., Zeymer, C., Griffiths, A. D. and Hilvert, D. (2016) 'Efficient laboratory evolution of computationally designed enzymes with low starting activities using fluorescence-activated droplet sorting', *Protein Eng. Des. Sel.,* 29(9), 355-366.

Ogawa, A. (2009) 'Biofunction-assisted sensors based on a new method for converting aptazyme activity into reporter protein expression with high efficiency in wheat germ extract', *ChemBioChem,* 10(15), 2465-2468.

Ogawa, A. (2011) 'Rational design of artificial riboswitches based on ligand-dependent modulation of internal ribosome entry in wheat germ extract and their applications as label-free biosensors', *RNA,* 17(3), 478-488.

Ogawa, A. (2012) 'Rational construction of eukaryotic OFF-riboswitches that downregulate internal ribosome entry site-mediated translation in response to their ligands', *Bioorg. Med. Chem. Lett.,* 22(4), 1639-1642.

Ogawa, A. (2013) 'Ligand-dependent upregulation of ribosomal shunting', *ChemBioChem,* 14(13), 1539-1543.

Ogawa, A. (2014) 'Rational design of artificial ON-riboswitches', *Methods Mol. Biol.,* 1111, 165-181.

Ogawa, A. (2015) 'Engineering of ribosomal shunt-modulating eukaryotic ON riboswitches by using a cell-free translation system', *Methods Enzymol.,* 550, 109-128.

Ogawa, A. (2021) 'Suppressor tRNA-based Biosensors for Detecting Analytes', *Anal. Sci.,* 37(3), 407-414.

Ogawa, A. and Itoh, Y. (2020) 'In Vitro Selection of RNA Aptamers Binding to Nanosized DNA for

Constructing Artificial Riboswitches', *ACS Synth. Biol.,* 9(10), 2648-2655.

Ogawa, A. and Maeda, M. (2007) 'Aptazyme-based riboswitches as label-free and detector-free sensors for cofactors', *Bioorg. Med. Chem. Lett.,* 17(11), 3156-3160.

Ogawa, A. and Maeda, M. (2008a) 'An artificial aptazyme-based riboswitch and its cascading system in E. coli', *ChemBioChem,* 9(2), 206-209.

Ogawa, A. and Maeda, M. (2008b) 'A novel label-free biosensor using an aptazyme-suppressor-tRNA conjugate and an amber mutated reporter gene', *ChemBioChem,* 9(14), 2204-2208.

Ogawa, A., Masuoka, H. and Ota, T. (2017a) 'Artificial OFF-Riboswitches That Downregulate Internal Ribosome Entry without Hybridization Switches in a Eukaryotic Cell-Free Translation System', *ACS Synth. Biol.,* 6(9), 1656-1662.

Ogawa, A., Murashige, Y., Tabuchi, J. and Omatsu, T. (2017b) 'Ligand-responsive upregulation of 3' CITE-mediated translation in a wheat germ cell-free expression system', *Mol. Biosyst.,* 13(2), 314-319.

Ogawa, A., Murashige, Y. and Takahashi, H. (2018) 'Canonical translation-modulating OFF-riboswitches with a single aptamer binding to a small molecule that function in a higher eukaryotic cell-free expression system', *Bioorg. Med. Chem. Lett.,* 28(14), 2353-2357.

Ogawa, A. and Tabuchi, J. (2015) 'Biofunction-assisted aptasensors based on ligand-dependent 3' processing of a suppressor tRNA in a wheat germ extract', *Org. Biomol. Chem.,* 13(24), 6681-6685.

Ogawa, A., Tabuchi, J., Doi, Y. and Takamatsu, M. (2016) 'Biofunction-assisted DNA detection through RNase H-enhanced 3' processing of a premature tRNA probe in a wheat germ extract', *Bioorg. Med. Chem. Lett.,* 26(15), 3658-3661.

Ouellet, J. (2016) 'RNA Fluorescence with Light-Up Aptamers', *Front. Chem.,* 4, 29.

Paige, J. S., Wu, K. Y. and Jaffrey, S. R. (2011) 'RNA mimics of green fluorescent protein', *Science,* 333(6042), 642-646.

Pamme, N. (2007) 'Continuous flow separations in microfluidic devices', *Lab Chip.,* 7(12), 1644-1659.

Pandit, K. R., Rueger, P. E., Calabrese, R. V., Raghavan, S. R. and White, I. M. (2015) 'Assessment of surfactants for efficient droplet PCR in mineral oil using the pendant drop technique', *Colloids Surf. B: Biointerfaces,* 126, 489-495.

Panula, P., Chazot, P. L., Cowart, M., Gutzmer, R., Leurs, R., Liu, W. L., Stark, H., Thurmond, R. L. and Haas, H. L. (2015) 'International Union of Basic and Clinical Pharmacology. XCVIII. Histamine Receptors', *Pharmacol. Rev.,* 67(3), 601-655.

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.

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. (2016a) 'Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components', *Cell,* 165(5), 1255-1266.

Pardee, K., Slomovic, S., Nguyen, P. Q., Lee, J. W., Donghia, N., Burrill, D., Ferrante, T., McSorley, F. R., Furuta, Y., Vernet, A., Lewandowski, M., Boddy, C. N., Joshi, N. S. and Collins, J. J. (2016b) 'Portable, On-Demand Biomolecular Manufacturing', *Cell,* 167(1), 248-259.e12.

Paul, S., Stang, A., Lennartz, K., Tenbusch, M. and Überla, K. (2013) 'Selection of a T7 promoter mutant with enhanced in vitro activity by a novel multi-copy bead display approach for in vitro evolution', *Nucleic Acids Res.,* 41(1), e29.

Pavlova, N., Kaloudas, D. and Penchovsky, R. (2019) 'Riboswitch distribution, structure, and function in bacteria', *Gene,* 708, 38-48.

Perez, J. G., Stark, J. C. and Jewett, M. C. (2016) 'Cell-Free Synthetic Biology: Engineering Beyond the Cell', *Cold Spring Harb. Perspect. Biol.,* 8(12), a023853.

Peselis, A. and Serganov, A. (2014) 'Themes and variations in riboswitch structure and function', *Biochim. Biophys. Acta,* 1839(10), 908-918.

Peterman, N., Lavi-Itzkovitz, A. and Levine, E. (2014) 'Large-scale mapping of sequence-function relations in small regulatory RNAs reveals plasticity and modularity', *Nucleic Acids Res.,* 42(19), 12177-12188.

Peterman, N. and Levine, E. (2016) 'Sort-seq under the hood: implications of design choices on large-

scale characterization of sequence-function relations', *BMC Genom.,* 17, 206.

Pooggin, M. M., Ryabova, L. A., He, X., Futterer, J. and Hohn, T. (2006) 'Mechanism of ribosome shunting in Rice tungro bacilliform pararetrovirus', *RNA,* 12(5), 841-850.

Porter, E. B., Polaski, J. T., Morck, M. M. and Batey, R. T. (2017) 'Recurrent RNA motifs as scaffolds for genetically encodable small-molecule biosensors', *Nat. Chem. Biol.,* 13(3), 295-301.

Pothoulakis, G., Ceroni, F., Reeve, B. and Ellis, T. (2014) 'The spinach RNA aptamer as a characterization tool for synthetic biology', *ACS Synth. Biol.,* 3(3), 182-187.

Pratt, J. M. (1984) 'Coupled transcription-translation in prokaryotic cell-free systems', in Hames, B.D. & Higgins, S.J. (eds.) *Transcription and Translation: A Practical Approach.* Oxford: IRL Press, 179-209.

Qian, L. and Winfree, E. (2011) 'Scaling up digital circuit computation with DNA strand displacement cascades', *Science,* 332(6034), 1196-1201.

Quarta, G., Sin, K. and Schlick, T. (2012) 'Dynamic energy landscapes of riboswitches help interpret conformational rearrangements and function', *PLoS Comput. Biol.,* 8(2), e1002368.

Raina, M. and Ibba, M. (2014) 'tRNAs as regulators of biological processes', *Front. Genet.,* 5, 171.

Rehm, C. and Hartig, J. S. (2014) 'In vivo screening for aptazyme-based bacterial riboswitches', *Methods Mol. Biol.,* 1111, 237-249.

Rohlhill, J., Sandoval, N. R. and Papoutsakis, E. T. (2017) 'Sort-Seq Approach to Engineering a Formaldehyde-Inducible Promoter for Dynamically Regulated Escherichia coli Growth on Methanol', *ACS Synth. Biol.,* 6(8), 1584-1595.

Rollin, J. A., Tam, T. K. and Zhang, Y. H. P. (2013) 'New biotechnology paradigm: cell-free biosystems for biomanufacturing', *Green Chem.,* 15(7), 1708-1719.

Rossmanith, J. and Narberhaus, F. (2017) 'Modular arrangement of regulatory RNA elements', *RNA Biol.,* 14(3), 287-292.

Ryabova, L. A. and Hohn, T. (2000) 'Ribosome shunting in the cauliflower mosaic virus 35S RNA leader is a special case of reinitiation of translation functioning in plant and animal systems ', *Genes Dev.,* 14, 817-829

Ryckelynck, M., Baudrey, S., Rick, C., Marin, A., Coldren, F., Westhof, E. and Griffiths, A. D. (2015) 'Using droplet-based microfluidics to improve the catalytic properties of RNA under multiple-turnover conditions', *RNA,* 21(3), 458-469.

Sakatani, Y., Mizuuchi, R. and Ichihashi, N. (2019) 'In vitro evolution of phi29 DNA polymerases through compartmentalized gene expression and rolling-circle replication', *Protein Eng. Des. Sel.,* 32(11), 481-487.

Salimiaghdam, N., Singh, L., Schneider, K., Nalbandian, A., Chwa, M., Atilano, S. R., Bao, A. and Kenney, M. C. (2020) 'Potential adverse effects of ciprofloxacin and tetracycline on ARPE-19 cell lines', *BMJ Open Ophthalmol.,* 5(1), e000458.

Sando, S., Narita, A., Abe, K. and Aoyama, Y. (2005) 'Doubly catalytic sensing of HIV-1-related CCR5 sequence in prokaryotic cell-free translation system using riboregulator-controlled luciferase activity', *J. Am. Chem. Soc.,* 127(15), 5300-5301.

Sawasaki, T., Ogasawara, T., Morishita, R. and Endo, Y. (2002) 'A cell-free protein synthesis system for high-throughput proteomics', *Proc. Natl. Acad. Sci. U.S.A.,* 99(23), 14652-14657.

Schmidt-Puchta, W., Dominguez, D., Lewetag, D. and Hohn, T. (1997) 'Plant ribosome shunting in vitro', *Nucleic Acids Res.,* 25(14), 2854-2860.

Schutz, S. S., Beneyton, T., Baret, J. C. and Schneider, T. M. (2019) 'Rational design of a high-throughput droplet sorter', *Lab Chip.,* 19(13), 2220-2232.

Schutze, T., Rubelt, F., Repkow, J., Greiner, N., Erdmann, V. A., Lehrach, H., Konthur, Z. and Glokler, J. (2011) 'A streamlined protocol for emulsion polymerase chain reaction and subsequent purification', *Anal. Biochem.,* 410(1), 155-157.

Schwarz-Schilling, M., Aufinger, L., Muckl, A. and Simmel, F. C. (2016) 'Chemical communication between bacteria and cell-free gene expression systems within linear chains of emulsion droplets', *Integr. Biol.,* 8(4), 564-570.

Scott, W. G., Horan, L. H. and Martick, M. (2013) 'The hammerhead ribozyme: structure, catalysis, and gene regulation', *Prog. Mol. Biol. Transl. Sci.,* 120, 1-23.

Seelig, G., Soloveichik, D., Zhang, D. Y. and Winfree, E. (2006) 'Enzyme-free nucleic acid logic circuits', *Science,* 314(5805), 1585-1588.

Shanidze, N., Lenkeit, F., Hartig, J. S. and Funck, D. (2020) 'A Theophylline-Responsive Riboswitch Regulates Expression of Nuclear-Encoded Genes', *Plant Physiol.,* 182(1), 123-135.

Sharma, V., Nomura, Y. and Yokobayashi, Y. (2008) 'Engineering complex riboswitch regulation by dual genetic selection', *J. Am. Chem. Soc.,* 130(48), 16310-16315.

Sharon, E., Kalma, Y., Sharp, A., Raveh-Sadka, T., Levo, M., Zeevi, D., Keren, L., Yakhini, Z., Weinberger, A. and Segal, E. (2012) 'Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters', *Nat. Biotechnol.,* 30(6), 521-530.

Sherlock, M. E., Sudarsan, N., Stav, S. and Breaker, R. R. (2018) 'Tandem riboswitches form a natural Boolean logic gate to control purine metabolism in bacteria', *eLife,* 7, e33908.

Shields, C. W. t., Reyes, C. D. and Lopez, G. P. (2015) 'Microfluidic cell sorting: a review of the advances in the separation of cells from debulking to rare cell isolation', *Lab Chip.,* 15(5), 1230-1249.

Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K. and Ueda, T. (2001) 'Cell-free translation reconstituted with purified components', *Nat. Biotechnol.,* 19(8), 751-755.

Shimizu, Y., Kanamori, T. and Ueda, T. (2005) 'Protein synthesis by pure translation systems', *Methods,* 36(3), 299-304.

Shis, D. L. and Bennett, M. R. (2013) 'Library of synthetic transcriptional AND gates built with split T7 RNA polymerase mutants', *Proc. Natl. Acad. Sci. U.S.A.,* 110(13), 5028-5033.

Siu, R. H. P., Liu, Y., Chan, K. H. Y., Ridzewski, C., Slaughter, L. S. and Wu, A. R. (2021) 'Optimization of on-bead emulsion polymerase chain reaction based on single particle analysis', *Talanta,* 221, 121593.

Smith, J. M., Chowdhry, R. and Booth, M. J. (2021) 'Controlling Synthetic Cell-Cell Communication', *Front. Mol. Biosci.,* 8, 809945.

Smith, M. T., Berkheimer, S. D., Werner, C. J. and Bundy, B. C. (2014) 'Lyophilized Escherichia coli-based cell-free systems for robust, high-density, long-term storage', *BioTechniques,* 56(4), 186-193.

Soleja, N., Manzoor, O., Khan, I., Ahmad, A. and Mohsin, M. (2018) 'Role of green fluorescent proteins and their variants in development of FRET-based sensors', *J. Biosci.,* 43(4), 763-784.

Soukup, G. A. and Breaker, R. R. (1999) 'Engineering precision RNA molecular switches', *Proc. Natl. Acad. Sci. U.S.A.,* 96(7), 3584-3589.

Soukup, G. A., Emilsson, G. A. and Breaker, R. R. (2000) 'Altering molecular recognition of RNA aptamers by allosteric selection', *J. Mol. Biol.,* 298(4), 623-632.

Soukup, J. K. and Soukup, G. A. (2004) 'Riboswitches exert genetic control through metabolite-induced conformational change', *Curr. Opin. Struct. Biol.,* 14(3), 344-349.

Spencer, A. C., Torre, P. and Mansy, S. S. (2013) 'The encapsulation of cell-free transcription and translation machinery in vesicles for the construction of cellular mimics', *J. Vis. Exp.,* 80, e51304.

Stoltenburg, R., Reinemann, C. and Strehlitz, B. (2007) 'SELEX-A (r)evolutionary method to generate high-affinity nucleic acid ligands', *Biomol. Eng.,* 24(4), 381-403.

Strobel, B., Sporing, M., Klein, H., Blazevic, D., Rust, W., Sayols, S., Hartig, J. S. and Kreuz, S. (2020) 'High-throughput identification of synthetic riboswitches by barcode-free amplicon-sequencing in human cells', *Nat. Commun.,* 11(1), 714.

Sudarsan, N., Barrick, J. E. and Breaker, R. R. (2003) 'Metabolite-binding RNA domains are present in the genes of eukaryotes', *RNA,* 9(6), 644-647.

Suess, B., Hanson, S., Berens, C., Fink, B., Schroeder, R. and Hillen, W. (2003) 'Conditional gene expression by controlling translation with tetracycline-binding aptamers', *Nucleic Acids Res.,* 31(7), 1853-1858.

Suess, B. and Weigand, J. E. (2008) 'Engineered riboswitches: overview, problems and trends', *RNA Biol.,* 5(1), 24-29.

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 Synth. Biol.,* 3(6), 387-397.

Sunami, T., Matsuura, T., Suzuki, H. and Yomo, T. (2010) 'Synthesis of functional proteins within liposomes', *Methods Mol. Biol.,* 607, 243-256.

Swartz, J. (2006) 'Developing cell-free biology for industrial applications', *J. Ind. Microbiol. Biotechnol.,* 33(7), 476-485.

Swartz, J. R., Jewett, M. C. and Woodrow, K. A. (2004) 'Cell-free protein synthesis with prokaryotic combined transcription-translation', *Methods Mol. Biol.,* 267, 169-182.

Szeto, K., Latulippe, D. R., Ozer, A., Pagano, J. M., White, B. S., Shalloway, D., Lis, J. T. and Craighead, H. G. (2013) 'RAPID-SELEX for RNA aptamers', *PLoS One,* 8(12), e82667.

Tabuchi, T. and Yokobayashi, Y. (2021) 'Cell-free riboswitches', *RSC Chem. Biol.,* 2(5), 1430-1440.

Tabuchi, T. and Yokobayashi, Y. (2022) 'High-throughput screening of cell-free riboswitches by fluorescence-activated droplet sorting', *Nucleic Acids Res.,* 50(6), 3535-3550.

Takahashi, K. and Yokobayashi, Y. (2019) 'Reversible Gene Regulation in Mammalian Cells Using Riboswitch-Engineered Vesicular Stomatitis Virus Vector', *ACS Synth. Biol.,* 8(9), 1976-1982.

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. (2015) 'Rapidly characterizing the fast dynamics of RNA genetic circuitry with cell-free transcription-translation (TX-TL) systems', *ACS Synth. Biol.,* 4(5), 503-515.

Takahashi, M. K., Tan, X., Dy, A. J., Braff, D., Akana, R. T., Furuta, Y., Donghia, N., Ananthakrishnan, A. and Collins, J. J. (2018) 'A low-cost paper-based synthetic biology platform for analyzing gut microbiota and host biomarkers', *Nat. Commun.,* 9(1), 3347.

Taylor, G. J. and Sarles, S. A. (2015) 'Using Cell-Free Expression to Create Light-Activated Proteins In Situ in Droplet Interface Bilayer Networks', *Mater. Res. Soc. Symp. Proc.,* 1720, 39–45.

Terekhov, S. S., Smirnov, I. V., Stepanova, A. V., Bobik, T. V., Mokrushina, Y. A., Ponomarenko, N. A., Belogurov, A. A., Jr., Rubtsova, M. P., Kartseva, O. V., Gomzikova, M. O., Moskovtsev, A. A., Bukatin, A. S., Dubina, M. V., Kostryukova, E. S., Babenko, V. V., Vakhitova, M. T., Manolov, A. I., Malakhova, M. V., Kornienko, M. A., Tyakht, A. V., Vanyushkina, A. A., Ilina, E. N., Masson, P., Gabibov, A. G. and Altman, S. (2017) 'Microfluidic droplet platform for ultrahigh-throughput single-cell screening of biodiversity', *Proc. Natl. Acad. Sci. U.S.A.,* 114(10), 2550-2555.

Thavarajah, W., Silverman, A. D., Verosloff, M. S., Kelley-Loughnane, N., Jewett, M. C. and Lucks, J. B. (2020) 'Point-of-Use Detection of Environmental Fluoride via a Cell-Free Riboswitch-Based Biosensor', *ACS Synth. Biol.,* 9(1), 10-18.

Thompson, J. R., Marcelino, L. A. and Polz, M. F. (2002) 'Heteroduplexes in mixed-template amplifications: formation, consequence and elimination by 'reconditioning PCR", *Nucleic Acids Res.,* 30(9), 2083-2088.

Tickner, Z. J. and Farzan, M. (2021) 'Riboswitches for Controlled Expression of Therapeutic Transgenes Delivered by Adeno-Associated Viral Vectors', *Pharmaceuticals,* 14(6), 554.

Topp, S. and Gallivan, J. P. (2008) 'Random walks to synthetic riboswitches—a high-throughput selection based on cell motility', *ChemBioChem,* 9(2), 210-213.

Tucker, B. J. and Breaker, R. R. (2005) 'Riboswitches as versatile gene control elements', *Curr. Opin. Struct. Biol.,* 15(3), 342-348.

Tuerk, C. and Gold, L. (1990) 'Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase', *Science,* 249(4968), 505-510.

Vega Laso, M. R., Zhu, D., Sagliocco, F., Brown, A. J., Tuite, M. F. and McCarthy, J. E. (1993) 'Inhibition of translational initiation in the yeast Saccharomyces cerevisiae as a function of the stability and position of hairpin structures in the mRNA leader', *J. Biol. Chem.,* 268(9), 6453-6462.

Vitreschak, A. G., Rodionov, D. A., Mironov, A. A. and Gelfand, M. S. (2004) 'Riboswitches: the oldest mechanism for the regulation of gene expression?', *Trends Genet.,* 20(1), 44-50.

Vyawahare, S., Brundage, M., Kijac, A., Gutierrez, M., de Geus, M., Sinha, S. and Homyk, A. (2021) 'Sorting droplets into many outlets', *Lab Chip.,* 21(21), 4262-4273.

Wachsmuth, M., Domin, G., Lorenz, R., Serfling, R., Findeiß, S., Stadler, P. F. and Mörl, M. (2015) 'Design criteria for synthetic riboswitches acting on transcription', *RNA Biol.,* 12(2), 221-231.

Wachsmuth, M., Findeiß, S., Weissheimer, N., Stadler, P. F. and Mörl, M. (2013) 'De novo design of a synthetic riboswitch that regulates transcription termination', *Nucleic Acids Res.,* 41(4), 2541-2551.

Wachter, A. (2010) 'Riboswitch-mediated control of gene expression in eukaryotes', *RNA Biol.,* 7(1),

67-76.

Wan, X., Marsafari, M. and Xu, P. (2019) 'Engineering metabolite-responsive transcriptional factors to sense small molecules in eukaryotes: current state and perspectives', *Microb. Cell Fact.,* 18(1), 61.

Wang, B., Kitney, R. I., Joly, N. and Buck, M. (2011) 'Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology', *Nat. Commun.,* 2, 508.

Wang, J., Gong, Q., Maheshwari, N., Eisenstein, M., Arcila, M. L., Kosik, K. S. and Soh, H. T. (2014) 'Particle display: a quantitative screening method for generating high-affinity aptamers', *Angew. Chem. Int. Ed.,* 53(19), 4796-4801.

Wang, P., Kojima, T., Kobayashi, T. and Nakano, H. (2012) 'Comprehensive analysis of the DNA-binding specificity of an Aspergillus nidulans transcription factor, AmyR, using a bead display system', *Biosci. Biotechnol. Biochem.,* 76(6), 1128-1134.

Warner, K. D., Chen, M. C., Song, W., Strack, R. L., Thorn, A., Jaffrey, S. R. and Ferre-D'Amare, A. R. (2014) 'Structural basis for activity of highly efficient RNA mimics of green fluorescent protein', *Nat. Struct. Mol. Biol.,* 21(8), 658-663.

Weigand, J. E., Sanchez, M., Gunnesch, E. B., Zeiher, S., Schroeder, R. and Suess, B. (2008) 'Screening for engineered neomycin riboswitches that control translation initiation', *RNA,* 14(1), 89-97.

Werstuck, G. and Green, M. R. (1998) 'Controlling gene expression in living cells through small molecule-RNA interactions', *Science,* 282(5387), 296-298.

Wick, S., Walsh, D. I., Bobrow, J., Hamad-Schifferli, K., Kong, D. S., Thorsen, T., Mroszczyk, K. and Carr, P. A. (2019) 'PERSIA for Direct Fluorescence Measurements of Transcription, Translation, and Enzyme Activity in Cell-Free Systems', *ACS Synth. Biol.,* 8(5), 1010-1025.

Wickiser, J. K., Cheah, M. T., Breaker, R. R. and Crothers, D. M. (2005a) 'The kinetics of ligand binding by an adenine-sensing riboswitch', *Biochem.,* 44(40), 13404-13414.

Wickiser, J. K., Winkler, W. C., Breaker, R. R. and Crothers, D. M. (2005b) 'The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch', *Mol. Cell,* 18(1), 49-60.

Wieland, M., Benz, A., Klauser, B. and Hartig, J. S. (2009) 'Artificial ribozyme switches containing natural riboswitch aptamer domains', *Angew. Chem. Int. Ed.,* 48(15), 2715-2718.

Wieland, M. and Hartig, J. S. (2008) 'Improved aptazyme design and in vivo screening enable riboswitching in bacteria', *Angew. Chem. Int. Ed.,* 47(14), 2604-2607.

Williams, R., Peisajovich, S. G., Miller, O. J., Magdassi, S., Tawfik, D. S. and Griffiths, A. D. (2006) 'Amplification of complex gene libraries by emulsion PCR', *Nat. Methods,* 3(7), 545-550.

Winkler, W. C. and Breaker, R. R. (2003) 'Genetic control by metabolite-binding riboswitches', *ChemBioChem,* 4(10), 1024-1032.

Winkler, W. C. and Breaker, R. R. (2005) 'Regulation of bacterial gene expression by riboswitches', *Annu. Rev. Microbiol.,* 59, 487-517.

Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. and Breaker, R. R. (2004) 'Control of gene expression by a natural metabolite-responsive ribozyme', *Nature,* 428(6980), 281-286.

Witt, M., Phung, N. L., Stalke, A., Walter, J.-G., Stahl, F., von Neuhoff, N. and Scheper, T. (2017) 'Comparing two conventional methods of emulsion PCR and optimizing of Tegosoft-based emulsion PCR', *Eng. Life Sci.,* 17(8), 953-958.

Wittmann, A. and Suess, B. (2011) 'Selection of tetracycline inducible self-cleaving ribozymes as synthetic devices for gene regulation in yeast', *Mol. Biosyst.,* 7(8), 2419-2427.

Wolfinger, M. T., Flamm, C. and Hofacker, I. L. (2018) 'Efficient computation of co-transcriptional RNA-ligand interaction dynamics', *Methods,* 143, 70-76.

Wouters, M. M., Vicario, M. and Santos, J. (2016) 'The role of mast cells in functional GI disorders', *Gut,* 65(1), 155-168.

Xayaphoummine, A., Bucher, T. and Isambert, H. (2005) 'Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots', *Nucleic Acids Res.,* 33(Web Server issue), W605-10.

Xiang, J. S., Kaplan, M., Dykstra, P., Hinks, M., McKeague, M. and Smolke, C. D. (2019) 'Massively parallel RNA device engineering in mammalian cells with RNA-Seq', *Nat. Commun.,* 10(1), 4327.

Xu, B., Nguyen, N.-T. and Neng Wong, T. (2011) 'Droplet Coalescence in Microfluidic Systems', *Micro Nanosyst.,* 3(2), 131-136.

Xu, C., Hu, S. and Chen, X. (2016) 'Artificial cells: from basic science to applications', *Mater. Today,* 19(9), 516-532.

Yokobayashi, Y. (2019a) 'Applications of high-throughput sequencing to analyze and engineer ribozymes', *Methods,* 161, 41-45.

Yokobayashi, Y. (2019b) 'Aptamer-based and aptazyme-based riboswitches in mammalian cells', *Curr. Opin. Chem. Biol.,* 52, 72-78.

Zawada, J. F., Yin, G., Steiner, A. R., Yang, J., Naresh, A., Roy, S. M., Gold, D. S., Heinsohn, H. G. and Murray, C. J. (2011) 'Microscale to manufacturing scale-up of cell-free cytokine production--a new approach for shortening protein production development timelines', *Biotechnol. Bioeng.,* 108(7), 1570-1578.

Zhang, Y., Minagawa, Y., Kizoe, H., Miyazaki, K., Iino, R., Ueno, H., Tabata, K. V., Shimane, Y. and Noji, H. (2019) 'Accurate high-throughput screening based on digital protein synthesis in a massively parallel femtoliter droplet array', *Sci. Adv.,* 5(8), eaav8185.

Zhang, Y. and Noji, H. (2017) 'Digital Bioassays: Theory, Applications, and Perspectives', *Anal. Chem.,* 89(1), 92-101.

Zhang, Y., Zheng, T., Wang, L., Feng, L., Wang, M., Zhang, Z. and Feng, H. (2021) 'From passive to active sorting in microfluidics: A review', *Rev. Adv. Mater. Sci.,* 60(1), 313-324.

Zhong, G., Wang, H., Bailey, C. C., Gao, G. and Farzan, M. (2016) 'Rational design of aptazyme riboswitches for efficient control of gene expression in mammalian cells', *eLife,* 5, e18858.

Zhou, J., Arugula, M. A., Halamek, J., Pita, M. and Katz, E. (2009) 'Enzyme-based NAND and NOR logic gates with modular design', *J. Phys. Chem. B.,* 113(49), 16065-16070.

Zhou, T., Ji, X., Shi, L., Zhang, X. and Joo, S. W. (2020) 'Droplet fusion by the interplay of electric potential and converging–diverging geometry in micro-channels', *J. Chem. Technol. Biotechnol.,* 96(2), 448-453.

Zhou, X., Wu, H., Cui, M., Lai, S. N. and Zheng, B. (2018) 'Long-lived protein expression in hydrogel particles: towards artificial cells', *Chem. Sci.,* 9(18), 4275-4279.

Zhuo, Z., Yu, Y., Wang, M., Li, J., Zhang, Z., Liu, J., Wu, X., Lu, A., Zhang, G. and Zhang, B. (2017) 'Recent Advances in SELEX Technology and Aptamer Applications in Biomedicine', *Int. J. Mol. Sci.,* 18(10), 2142.

Zubaite, G., Simutis, K., Galinis, R., Milkus, V., Kiseliovas, V. and Mazutis, L. (2017) 'Droplet Microfluidics Approach for Single-DNA Molecule Amplification and Condensation into DNA-Magnesium-Pyrophosphate Particles', *Micromachines,* 8(2), 62.

Zuker, M. (2003) 'Mfold web server for nucleic acid folding and hybridization prediction', *Nucleic Acids Res.,* 31(13), 3406-3415.

Zuker, M. and Stiegler, P. (1981) 'Optimal Computer Folding of Large Rna Sequences Using Thermodynamics and Auxiliary Information', *Nucleic Acids Res.,* 9(1), 133-148.

APPENDICES

APPENDIX 1: LIST OF PRIMERS

Table 4. List of primers and oligonucleotides. 2-Bio: dual biotin, iSp18: 18-atom hexa-ethyleneglycol spacer, TexRd: Texas Red or sulforhodamine 101 (NHS ester), 2-OMe: 2'-O-methyl-RNA(ss), mA/mU/mC/mG: 2'-O-methyl-ribonucleotides, ROX: Rhodamine Red-X or 6-carboxy-X-rhodamine (NHS ester), BHQ-2: Black Hole Quencher-2, UDI: Illumina unique dual index (**Table 6**).

Name	Sequence (5'–3')
P0L (Anchor)	(5' 2-Bio + iSp18)TGGCGAAAGGTGCTTGTTGTAATACGACTC
P3L (Forward)	GGCGAAAGGTGCTTGTTGTAATACGACTC
P4L (Reverse)	(5' TexRd)TAGTACGATGCCAGCAGGTCTTATGTAATCC
P-eGFP-R	TAGTACGATGCCAGCAGGTCTCACTTG
P-GFP1.10-R	GCCAAGCTGGAGACCGTTTAAACTCA
ROX-OMB (2-OMe)	(5' ROX)mCmCmUmGmUmAmCmGmAmUmGmCmCmAmGmCmAmGmG(3' BHQ-2)
P-HA-Nova-T1-#	ACACTCTTTCCCTACACGACGCTCTTCCGATCT(Barcode_#)GGCGAAAGGTGCTTGTTGTAATACG
P-HA-Nova-B1-#	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT(Barcode_#)AGGACCATGTGGTCACGCAT
TruSeq-i5-UDI00##	AATGATACGGCGACCACCGAGATCTACAC(UDI00##)ACACTCTTTCCCTACACGACGC
TruSeq-i7-UDI00##	CAAGCAGAAGACGGCATACGAGAT(UDI00##)GTGACTGGAGTTCAGACGTGTG

Table 5. Custom barcodes used in the P-HA-Nova-T1 and P-HA-Nova-B1 primers (**Table 4**). The different lengths is to increase the nucleotide diversity of the base calling during the sequencing run.

Name	Sequence (5'–3')
…T1-1 (cycle 0)	TATCTC
…T1-2 (cycle 1)	GTATCA
…T1-3 (cycle 2)	AGAGTT
…T1-4 (cycle 3)	CATGCTA
…T1-5 (cycle 4)	TCTCAGAT
…T1-6 (cycle 5)	CTGTAATC
…T1-7	GACTTCACT
…T1-8	ACGTCTAGA
…B1-1	ACTGCT
…B1-2	GACTATC

Table 6. Illumina unique dual indexes (UDI) used in the TruSeq-i5 and TruSeq-i7 primers (**Table 4**).

Name	Sequence (5'–3')	
	i5	i7
UDI0001	AACCGCGG	AGCGCTAG
UDI0002	GGTTATAA	GATATCGA
UDI0003	CCAAGTCC	CGCAGACG
UDI0004	TTGGACTT	TATGAGTA
UDI0005	CAGTGGAT	AGGTGCGT

APPENDIX 2: SEQUENCES OF THE CONSTRUCTS

Table 7. Sequences of the constructs. To keep this table concise, only the variable or unique sequences are shown, while the shared constant regions have been abbreviated as Leader (5' end) and Trailer (3' end).

- All GFP11 constructs have the structure: (5') [Leader #1]…Variable Sequence…[Trailer #1] (GFP11) (3').
- All eGFP constructs have the structure: (5') [Leader #2]…Variable Sequence…[Trailer #2] (eGFP) (3').
- All GFP1–10 constructs have the structure: (5') [Leader #2]…Variable Sequence…[Trailer #3] (GFP1–10) (3').
- All HDC constructs have the structure: (5') [Leader #1]…Variable Sequence…[Trailer #4] (HDC) (3').

The "tag 1" and "tag 2" sequences are underlined. T7 promoter is shown in UPPERCASE. Aptamers are shown in **BOLD UPPERCASE**. The coding sequences (ORFs) are shown in *italicized lowercase*. The start and the stop codons are shown in *ITALICIZED UPPERCASE*. Restriction sites are shown in **bold lowercase**.

Description	Sequence (5'–3')
Leader #1 (tag 1 + P_T7)	ggcgaaaggtgcttgttgTAATACGACTCACTATAg…
Leader #2 (P_T7)	(ga)ttgTAATACGACTCACTATAg…
Trailer #1 (*gfp11* + tag 2)	…*ATG*cgtgaccacatggtccttcatgagtatgtaaatgctgctgggattaca*TAA*gacctgctggcatcgtacta
Trailer #2 (*egfp* + tag 2)	…*ATG*cgt*ATG*gtgagcaagggcgaggagctgttcaccggggtggtgcccatcctggtcgagctggacggcgacgtaaacg gccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacggcaagctgaccctgaagttcatctgcaccacc ggcaagctgcccgtgccctggcccaccctcgtgaccaccctgacctacggcgtgcagtgcttcagccgctaccccgacca catgaagcagcacgacttcttcaagtccgccatgcccgaaggctacgtccaggagcgcaccatcttcttcaaggacgacg gcaactacaagacccgcgccgaggtgaagttcgagggcgacaccctggtgaaccgcatcgagctgaagggcatcgacttc aaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggccgacaagca gaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcgccgaccactaccagc agaacacccccatcggcgacggccccgtgctgctgcccgacaaccactacctgagcacccagtccgccctgagcaaagac cccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgccgggatcactctcggcatggacgagctgta caag*TGA*gacctgctggcatcgtacta
Trailer #3 (*gfp(1–10)*-6xHis)	…*ATG*cgtatgagcaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgttaatgggc acaaattttctgtcagaggagagggtgaaggtgatgctacaatcggaaaactcacccttaaatttatttgcactactgga aaactacctgttccatggccaacacttgtcactactctgacctatggtgttcaatgcttttcccgttatccggatcacat gaaaaggcatgactttttcaagagtgccatgcccgaaggttatgtacaggaacgcactatatctttcaaagatgacggga aatacaagacgcgtgctgtagtcaagtttgaaggtgatacccttgttaatcgtatcgagttaaagggtactgatttttaa gaagatggaaacattctcggacacaaacttgagtacaactttaactcacacaatgtatacatcacggcagacaaacaaaa gaatggaatcaaagctaacttcacagttcgccacaacgttgaagatggttccgttcaactagcagaccattatcaacaaa atactccaattggcgatggccctgtcctttaccagacaaccattacctgtcgacacaaactgtcctttcgaaagatccc aacgaaaagTccggcggtggaggtctcgagcatcatcatcatcatcat*TGA*gtttaaacggtctccagcttggc
Trailer #4 (*hdc*-6xHis + tag 2)	…*ATG*acattatcaatctcagatcagaacaagttagatagcttctggtcatattgtgtaaagaaccagtatttttaatatcg gataccctgaatccgcggattttgattatactattctggagcgctttatgcgtttcagtattaataactgcggtgactgg ggcgagtattgcaactaccttcttaattctttcgattttgagaaggaggttatggagtatttcgcacaattattcaagat tccattcgaggagtcctgggggttacgtaaccaatggtggtacggaaggcaatatgttcggatgttacctgggacgtgaga tcttccctaatggaactctttactacagcaaagacacccattacagcgtagcaaaaattgtcaaactcttacgtattaag tccacgttggtcgagagccagcctaacggtgagatggattatgccgaccttatcaagaagatcaaggcagacaacgaaaa acaccctattatcttcgccaatatcggaaccactgtgcgtggagcaattgacaatattgcgattattcagcaatcgattt cggaactgggaatcgaacgtaaggattactatttacacgcggatgccgcattgtcagggatgattcttccatttgtggat aaccctcaacgtttaatttcgcggacgggattgatagtatcggagtctctgggcataagatgatcggaagtcccattcc ttgcgggattgtcgtcgctaaaaagaagaatgtagatcgtatttccgttgaaatcgattacatttccgcgcatgaccaaaa cgatctcaggatcgcgcaatggtcataccccattaatgatgtgggaagcaatccgttctccatagttgggaagaatggcgt cgtcgcattgagcgttctttaaacatggcgcaatatgcagtggatcgtttccaaagcgcgggaattgatgcctggcgcaa taagaactcgattacggttgtatttccttgcccttccgaggcggtttggaaaaagcactgcttggcgacatctggagata tcgctcatcttatcgcaacggctcaccaccttgattcgtccaagatcgatgctttaatcgacgatgtgattgcggacttg aagaaacaagctgcctcaggtggtggcggattagaacaccaccaccatcaccat*TAA*tggattacataagacctgctggc atcgtacta
P_T7-RBS-GFP11 ("Strong" RBS)	[Leader #1]…**gaattc**taaggaggtaaatta…[Trailer #1]
P_T7-wRBS-GFP11 ("Weak" RBS)	[Leader #1]…**gagctc**tcacacaggac…[Trailer #1]
P_T7-RBS-eGFP	[Leader #2]…ggttaaggaggtaaatta…[Trailer #2]
P_T7-RBS-GFP(1–10)	[Leader #2]…ggaacagaaggaggtaaatta…[Trailer #3]
P_trc-*gfp(1–10)*-6xHis	CTGTTGACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTgaattaagctttaaggaggat ttagt…[Trailer #3]
HA-C1g (N6) [GFP11]	[Leader #1]…ggNNNNNN**CCAGTGGGGTTGAAGGAAAGTAACAGA**cgcattaatttaaggaggtaaatta …[Trailer #1]
HA-C1g (N5) [GFP11]	[Leader #1]…ggNNNNN**CCAGTGGGGTTGAAGGAAAGTAACAGA**cgcattaatttaaggaggtaaatta …[Trailer #1]

HA-C1g (N4) [GFP11]	[Leader #1]…ggNNN**NCCAGTGGGTTGAAGGAAAGTAACAGA**cgcattaatttaaggaggtaaatta …[Trailer #1]
HA-C1g-3 [GFP(1–10)]	[Leader #2]…ggaggcg**TCCAGTGGGTTGAAGGAAAGTAACAGA**cgcattaatttaaggaggtaaatta*ATGcgt* …[Trailer #3]
HA-C1g-10 [GFP(1–10)]	[Leader #2]…ggaagcg**TCCAGTGGGTTGAAGGAAAGTAACAGA**cgcattaatttaaggaggtaaatta*ATGcgt* …[Trailer #3]
HA-C1g-19 [GFP11/eGFP]	[Leader #1/#2]…ggatgcg**TCCAGTGGGTTGAAGGAAAGTAACAGA**cgcattaatttaaggaggtaaatta …[Trailer #1/#2]
HA-C1g-19/OFF [eGFP]	[Leader #2]…ggatagg**TCCAGTGGGTTGAAGGAAAGTAACAGA**cgcattaatttaaggaggtaaatta …[Trailer #2]
HA-C1g-19/ON [eGFP]	[Leader #2]…ggaaattaatgcg**TCCAGTGGGTTGAAGGAAAGTAACAGA**cgcattaatttaaggaggtaaatta …[Trailer #2]
HA-C1g-19/MM [eGFP]	[Leader #2]…ggatgcg**TCCAGTGGGTTGAAGGAAAGTAACAGA**acacataatttaaggaggtaaatta …[Trailer #2]
HA-OFF4-a9 (WT) [GFP11/eGFP]	[Leader #1/#2]…ggcgcattaatttactttct**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta …[Trailer #1/#2]
HA-OFF4-a9 (N6) [GFP11]	[Leader #1]…ggcgcattNNNNNNctttct**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta …[Trailer #1]
HA-OFF4-a9 (N5) [GFP11]	[Leader #1]…ggcgcattNNNNNctttct**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta …[Trailer #1]
HA-OFF4-a9 (N4) [GFP11]	[Leader #1]…ggcgcattNNNNctttct**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta …[Trailer #1]
HA-OFF4-a9 (N3) [GFP11]	[Leader #1]…ggcgcattNNNctttct**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta …[Trailer #1]
HA-OFF4-a9-1 [GFP11/eGFP]	[Leader #1/#2]…ggcgcatttctgacctttct**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta …[Trailer #1/#2]
HA-OFF4-a9-3 [GFP11/eGFP]	[Leader #1/#2]…ggcgcatttttcggctttct**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta …[Trailer #1/#2]
HA-OFF4-a9-13 [GFP11/eGFP]	[Leader #1/#2]…ggcgcattgcttgactttct**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta …[Trailer #1/#2]
HA-OFF4-a9-13 [GFP(1–10)]	[Leader #2]…ggcgcattgcttgactttct**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta*ATGcgt* …[Trailer #3]
HA-OFF4-a9-13-mA [GFP11]	[Leader #1]…ggcgcattcgttgacttact**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta …[Trailer #1]
HA-OFF4-a9-13-mA [GFP(1–10)]	[Leader #2]…ggcgcattcgttgacttact**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta*ATGcgt* …[Trailer #3]
HA-OFF4-a9-13-mB [GFP(1–10)]	[Leader #2]…ggcgcattcgttgactttat**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta*ATGcgt* …[Trailer #3]
HA-OFF4-a9-13-mC [GFP(1–10)]	[Leader #2]…ggcgcattcgttgactttct**TTCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta*ATGcgt* …[Trailer #3]
HA-OFF4-a9-13-mD [GFP(1–10)]	[Leader #2]…ggcgcattcgttgactttct**TCCAGCGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta*ATGcgt* …[Trailer #3]
HA-OFF4-a9-13-mE [GFP(1–10)]	[Leader #2]…ggcgcattcgttgactttct**TCCAGTGGGTTGAAGGATAGTAACAGA**aggaggtaaatta*ATGcgt* …[Trailer #3]
HA-OFF4-a9-13-mF [GFP(1–10)]	[Leader #2]…ggcgcattcgttgactttct**TCCAGCGGGTTGAAGGATAGTAACAGA**aggaggtaaatta*ATGcgt* …[Trailer #3]
HA-OFF4-a9-14 [GFP11/eGFP]	[Leader #1/#2]…ggcgcatttctcacctttct**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta …[Trailer #1/#2]
HA-OFF4-a9-18 [GFP11/eGFP]	[Leader #1/#2]…ggcgcatttttgacctttct**TCCAGTGGGTTGAAGGAAAGTAACAGA**aggaggtaaatta …[Trailer #1/#2]
CFX-a1 (WT) [GFP11/eGFP]	[Leader #1/#2]…gggaccaa**CGCAACAGACTAGGTTGTGACTGCTTAGGCAGTTGTGGACGGCTAAGCCCACCAGAGGTCCGTAACTAGTCG**cgcattaatttaaggaggtaaatta…[Trailer #1/#2]
CFX-a1 (N6) [GFP11]	[Leader #1]…gggaccaa**CGCAACAGACTAGGTTGTGACTGCTTAGGCAGTTGTGGACGGCTAAGCCCACCAGAGGTCCGTAACTAGTCG**cgNNNNNNtttaaggaggtaaatta…[Trailer #1]
CFX-a1-sr5 [GFP11/eGFP]	[Leader #1/#2]…gggaccaa**CGCAACAGACTAGGTTGTGACTGCTTAGGCAGTTGTGGACGGCTAAGCCCACCAGAGGTCCGTAACTAGTCG**cgcgttaatttaaggaggtaaatta…[Trailer #1/#2]
CFX-a1-sr5 (N6) [GFP11]	[Leader #1]…gggaNNNNN**NCAACAGACTAGGTTGTGACTGCTTAGGCAGTTGTGGACGGCTAAGCCCACCAGAGGTCCGTAACTAGTCG**cgcgttaatttaaggaggtaaatta…[Trailer #1]
CFX-a1-sr5-2 [GFP11/eGFP]	[Leader #1/#2]…gggacgtg**CGCAACAGACTAGGTTGTGACTGCTTAGGCAGTTGTGGACGGCTAAGCCCACCAGAGGTCCGTAACTAGTCG**cgcgttaatttaaggaggtaaatta…[Trailer #1/#2]
CFX-a1-sr5-19 [GFP11/eGFP]	[Leader #1/#2]…gggagtcg**CGCAACAGACTAGGTTGTGACTGCTTAGGCAGTTGTGGACGGCTAAGCCCACCAGAGGTCCGTAACTAGTCG**cgcgttaatttaaggaggtaaatta…[Trailer #1/#2]
P_{T7}-*hdc*-6xHis	[Leader #1]…ggagatttgtttaactttaaggaggtaattttt…[Trailer #4]
P_{T7}-CFXsr5.19-*hdc*	[Leader #1]…gggagtcg**CGCAACAGACTAGGTTGTGACTGCTTAGGCAGTTGTGGACGGCTAAGCCCACCAGAGGTCCGTAACTAGTCG**cgcgttaatttaaggaggtaaatta*ATGcgt*…[Trailer #4]
P_{T7}-HA-C1g-3-*hdc*	[Leader #1]…ggaggcg**TCCAGTGGGTTGAAGGAAAGTAACAGA**cgcattaatttaaggaggtaaatta*ATGcgt* …[Trailer #4]

APPENDIX 3: SUMMARY OF THE SEQUENCING RUNS

Table 8. Summary of the high-throughput sequencing runs. max.: maximum, avg.: average, mdn.: median.

HA-C1g (N4, N5, N6)

Sequencing platform	Illumina MiSeq v3
Read ends	Single-read
Read length	150 bp
Run statistics	
Density (K/mm²)	1127 ± 42
PF Clusters (M)	25.97
% PF Clusters	93.43
Yield (G)	3.9
% >= Q30 bases	94.08
Mean Quality	36.28

Sequencing data	Total	Cycle 0	Cycle 1	Cycle 2	Cycle 3	Cycle 4	
Total raw reads	22813551						
Sorted reads	22499883	11178476	5709002	3279952	1553967	778486	
Processed reads	14819186	6713618	3863911	2436060	1181679	623918	
Unique variants (>0 reads)	5376	5376	5376	5374	5341	5167	
		100.0%	100.0%	100.0%	99.3%	96.1%	
Unique variants (>0.01% cycle reads)		3854	3166	3206	2220	1953	
		71.7%	58.9%	59.6%	41.3%	36.3%	
Reads per variant (max.)		8759	14420	11192	41091	17426	
Reads per variant (avg.)		1249	719	453	220	116	
Reads per variant (mdn.)		995	480	316	80	22	

HA-OFF4-a9 (N3, N4, N5, N6)

Sequencing platform	Illumina NovaSeq 6000
Read ends	Paired-end
Read length	150 bp
Run statistics	
% of lane	12.40
PF Clusters (M)	60.97
% PF Clusters	100.0
Yield (G)	18.4
% >= Q30 bases	94.97
Mean Quality	36.16

Sequencing data	Total	Cycle 0	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5
Total raw reads	60973370						
Sorted reads	59511474	9197153	10207304	9940022	10067480	10016114	10083401
Processed reads	56590426	8693158	9820521	9534414	9588368	9528074	9425891
Unique variants (>0 reads)	5440	5440	5440	5440	5440	5440	5440
		100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Unique variants (>0.01% cycle reads)		4938	2179	2268	1532	1396	1065
		90.8%	40.1%	41.7%	28.2%	25.7%	19.6%
Reads per variant (max.)		10728	23539	22690	61720	88566	183589
Reads per variant (avg.)		1598	1805	1753	1763	1751	1733
Reads per variant (mdn.)		1464	282	335	74	68	71

CFX-a1 (N6) Run 1

Sequencing platform	Illumina MiSeq v2 nano
Read ends	Single-read
Read length	150 bp
Run statistics	
Density (K/mm²)	1045
PF Clusters (M)	1.32
% PF Clusters	86.61
Yield (G)	0.26
% >= Q30 bases	94.14
Mean Quality	36.40

Sequencing data	Total	Cycle 0			Cycle 3		Cycle 5
Total raw reads	1170256						
Sorted reads	1149031	434375			404283		310373
Processed reads	518214	247175			154447		116592
Unique variants (>0 reads)	4096	4096			3285		3068
		100.0%			80.2%		74.9%
Unique variants (>0.01% cycle reads)		3603			663		540
		88.0%			16.2%		13.2%
Reads per variant (max.)		408			6743		13523
Reads per variant (avg.)		60			38		28
Reads per variant (mdn.)		51			2		2

CFX-a1 (N6) Run 2

Sequencing platform	Illumina NovaSeq 6000
Read ends	Paired-end
Read length	150 bp
Run statistics	
% of lane	9.22
PF Clusters (M)	51.15
% PF Clusters	100.0
Yield (G)	15.4
% >= Q30 bases	96.40
Mean Quality	36.54

Sequencing data	Total	Cycle 0	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5
Total raw reads	50618895						
Sorted reads	50056790	12678321	4525667	4217190	17091491	4419733	7124388
Processed reads	35427450	11271195	3519599	3414738	10092592	3385559	3743767
Unique variants (>0 reads)	4096	4096	4096	4096	4096	4096	4096
		100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Unique variants (>0.01% cycle reads)		3579	1026	1123	718	658	632
		87.4%	25.0%	27.4%	17.5%	16.1%	15.4%
Reads per variant (max.)		41824	49916	34033	363582	196501	165315
Reads per variant (avg.)		2752	859	834	2464	827	914
Reads per variant (mdn.)		2254	89	64	204	56	85

CFX-a1-sr5 (N6)

Sequencing platform	Illumina NovaSeq 6000
Read ends	Paired-end
Read length	150 bp
Run statistics	
% of lane	11.05
PF Clusters (M)	54.34
% PF Clusters	100.0
Yield (G)	16.4
% >= Q30 bases	95.26
Mean Quality	36.19

Sequencing data	Total	Cycle 0	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5
Total raw reads	54340974						
Sorted reads	45113130	9495092	8770285	9265975	5799156	5788918	5993704
Processed reads	41796625	8858803	8131638	8604911	5302821	5381831	5516621
Unique variants (>0 reads)	4096	4096	4096	4093	4090	4093	4093
		100.0%	100.0%	99.9%	99.9%	99.9%	99.9%
Unique variants (>0.01% cycle reads)		3223	2012	1891	1509	949	906
		78.7%	49.1%	46.2%	36.8%	23.2%	22.1%
Reads per variant (max.)		16169	25619	46445	48031	106391	163758
Reads per variant (avg.)		2163	1985	2101	1295	1314	1347
Reads per variant (mdn.)		1675	763	627	75	61	64

APPENDIX 4: PROTOCOLS

A. BUFFER SOLUTIONS AND MEDIA

Table 9. Composition of the buffer solutions and media employed.

Name	Composition	Protocol
Oil (droplets)	5% (w/w) Pico-Surf 1 in HFE-7500	**B.1**
Sheath fluid	0.1% (w/w) Pico-Surf 1 in HFE-7500	**F.1**
Binding buffer	5 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 1 M NaCl	**C.1**
TK buffer	20 mM Tris-HCl (pH 8.4), 50 mM KCl	**C.1, C.3, G.1**
TE buffer	10 mM Tris-HCl (pH 7.5), 1 mM EDTA	**G.2**
TEK buffer	15 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 25 mM KCl	**C.3**
TENa buffer	10 mM Tris-HCl (pH 7.5), 5 mM EDTA, 200 mM NaCl	**C.3**
Storage buffer	10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 20 mM KCl	**D.4**
Breaking buffer	15 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 50 mM NaCl, 25 mM KCl, 0.5% (w/v) Triton X-100, 0.5% (w/v) SDS	**C.3, G.1**
Lysis buffer	50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 10 mM imidazole, 0.05% (w/v) Tween 20	**D.2**
Wash buffer	50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 20 mM imidazole, 0.05% (w/v) Tween 20	**D.3**
Elution buffer	50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 250 mM imidazole, 0.05% (w/v) Tween 20	**D.3**
LB medium	10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, pH 7.0	**D.1**

B. DROPLET GENERATION

The water-in-oil emulsion droplets were generated by On-chip Droplet Generator (On-chip Biotechnologies) in 2D chip-800DG microfluidics chips (material: COP) at 4°C. As the oil phase, 5% (w/w) Pico-Surf 1 (Sphere Fluidics) in HFE-7500 3M Novec Engineered Fluid (Fluorochem) was used, and the aqueous phase was either a PCR mix (Q5 High-Fidelity 2X Master Mix, NEB) or a cell-free protein synthesis (CFPS) reaction mix (PURE*frex* 1.0, Gene Frontier).
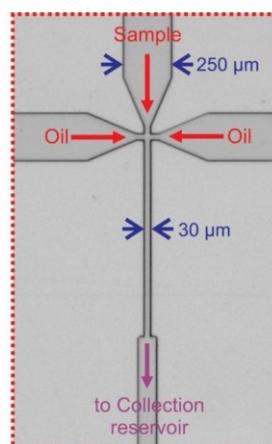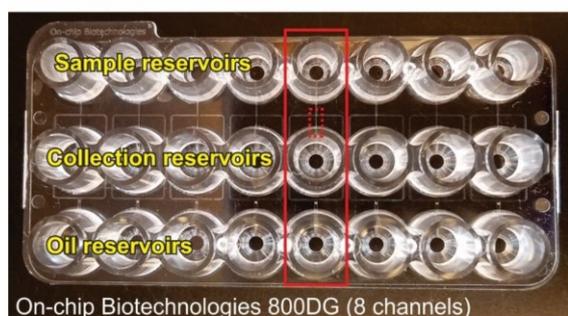
Figure 40. 8-channel 2D chip-800DG microfluidics chip (On-chip Biotechnologies) for droplet generation.

B.1. PREPARING THE CHIP AND SETTING UP THE DROPLET GENERATOR MACHINE	1. Turn on the On-chip Droplet Generator (On-chip Biotechnologies) and the temperature control unit, if needed. It takes 5–10 minutes to reach 4°C or 50°C starting from room temperature.

2. Set up the pressures as following:

Channel	Pressure
Sample	58 kPa
Oil	79 kPa

3. Fill an oil reservoir for each sample of an 2D chip-800DG microfluidics chip (8-channels) with 5% (w/w) Pico-Surf 1 in HFE-7500.
4. Insert the chip into the chip holder and fasten the lid.
5. Open the sample and oil valves of the corresponding channels.
6. Briefly turn on the machine to prime the channels.

B.2. DROPLET GENERATION	7. Take the chip out and add the samples to the corresponding sample reservoirs.
	8. Insert the chip into the chip holder and fasten the lid.
	9. Within the software start the droplet generation.

- The throughput at these settings is ~2.2×10^6 droplets per 10 minutes.
- The droplets produced are ~26 µm (25.9 ± 0.4 µm) in diameter.
- Be careful of the oil levels. If the oil reservoir is completely filled, it takes about 12 minutes to become depleted and to overflow the collection reservoir.

10. Turn off the machine once the droplets are done.

11. Take the chip out. Retrieve the emulsion from the collection reservoir and transfer it to a new tube.

C. SINGLE-TEMPLATE AMPLIFICATION ON MAGNETIC BEADS (BEAMING)

The template DNAs encoding the riboswitch variants were clonally amplified by emulsion PCR (ePCR) on magnetic beads according to the BEAMing protocol (Diehl et al., 2006) with few modifications. The magnetic beads were always collected using a home-made neodymium magnet stand.

C.1. BINDING OF PRIMERS TO BEADS

1. Thoroughly resuspend the streptavidin-conjugated magnetic beads (Dynabeads MyOne Streptavidin C1, Invitrogen).

2. Transfer 30 µl of beads (1 volume) to a clean low-binding tube.

3. Place the tube in a magnetic rack and wait for ~1 minute. Carefully remove the liquid without disturbing the pellet.

4. Wash the beads 2 times with 1 volume (30 µl) of TK buffer. Pellet the beads with a magnet, then carefully remove the liquid.

5. Resuspend the beads in 1 volume (30 µl) of Binding buffer.

6. Add 1/3 volume (10 µl) of biotinylated anchor primer P0L (100 µM) and vortex immediately. The anchor primer was modified at the 5' end with a dual-biotin tag connected via an 18-atom hexa-ethyleneglycol spacer (iSp18) synthesized by IDT.

7. Incubate at room temperature for 45 minutes. Ensure to keep the beads suspended by mixing them every 10–15 minutes or using a rotating device.

8. Pellet the beads with a magnet, then carefully remove the liquid.

9. Wash the beads 3 times with 1 volume (30 µl) of TK buffer. Pellet the beads with a magnet, carefully remove the liquid.

10. Resuspend in 1 volume (30 µl) of TK buffer.

11. Store at 4°C or continue working.

C.2. SINGLE TEMPLATE AMPLIFICATION ON BEADS (BEAMING)

12. On ice, prepare the PCR master mix as following:

Component	20 µl reaction	Final Conc.
Nuclease-free water	5.8 µl	
P0L-coated beads (7–12×10^6/µl)	0.8 µl	3–5×10^5/µl
F primer: P3L (1.25 µM)	0.8 µl	50 nM
R primer: P4L (100 µM)	1.6 µl	4000 nM
Template (1 pM)*	1.0 µl	50 fM
Q5 High-Fidelity 2X Master Mix (NEB)	10.0 µl	1X

* Start the dilution of the template only after adding all the other component except Q5 to the PCR mix. Use a new tip for each step of a serial dilution. DNA tends to stick to the plastic tube and tips, mix by pipetting, do NOT vortex.

13. Generate the droplets as described before (see **B.1–B.2**, steps **1–11**, page **VI**).

14. Transfer the emulsion to clean 0.2 ml PCR tubes. Remove as much oil (HFE-7500) as possible from the bottom. Gently spin-down the liquid.

15. Place the tubes in a thermal cycler and run a PCR reaction according to the following settings:

Step	Temp.	Time
Initial Denaturation	98°C	1 minute
35 cycles	98°C	15 seconds
	72°C	25 seconds
Final Extension	72°C	2 minutes
Hold	4°C	

C.3. BREAKING EMULSION AND WASHING THE BEADS

Scale up or down as necessary. For a 20 μl BEAMing reaction:

16. (Optional) Add 1 volume (20 μl) of TENa buffer. EDTA will prevent unwanted amplification during the emulsion breaking.

17. Break the emulsion with an anti-static gun (MILTY Zerostat 3). Spin-down the liquid. Using chemical demulsifiers such as perfluorooctanol (Pico-Break 1, Sphere Fluidics) or chloroform is not recommended as they can interfere with downstream reactions (PCR and CFPS).

18. Place the tube in a magnetic rack and wait for ~1 minute. Carefully remove ~80–90% of the liquid without disturbing the pellet. Optional: Repeat Step 15–17 to break the remaining emulsion.

19. Resuspend in 2 volumes (40 μl) of Breaking buffer. Mix well then spin-down. Pellet the beads with a magnet, then carefully remove the liquid.

20. Wash the beads 2 times with 2 volumes (40 μl) of TK buffer. Pellet the beads with a magnet, then carefully remove the liquid.

21. Resuspend the beads in 2 volumes (40 μl) of TK buffer. Mix well then spin-down.

22. Transfer to a new tube. Pellet the beads with a magnet, carefully remove the liquid.

23. Wash the beads 2 times with 2 volumes (40 μl) of TK buffer. Pellet the beads with a magnet, then carefully remove the liquid.

24. Resuspend the beads in 2 volumes (40 μl) of TEK buffer.

25. Store at 4°C or continue working.

D. PRODUCTION OF RECOMBINANT GFP1–10 PROTEIN

To produce the GFP1–10 protein fragment, *Escherichia coli* BL21 (DE3) was transformed with the expression plasmid pTrcHis-*gfp(1–10)*-6xHis. The His-tagged GFP1–10 was purified using Ni-NTA magnetic agarose beads (QIAGEN) following the manufacturer's protocol.

D.1. BACTERIAL CULTURE

1. In a test tube with 1–5 ml of LB medium supplemented with 100 μg/ml of ampicillin, seed the *E. coli* strain transformed with the GFP1–10 plasmid.

2. Incubate overnight at 30–37°C in an orbital shaker (250 rpm).

3. Transfer ~100 μl of the overnight culture (starter) to 5 ml of fresh LB medium supplemented with 100 μg/ml of ampicillin. Scale up as necessary.

4. Incubate at 30°C in an orbital shaker (250 rpm) until the bacterial culture reach an OD_{600}~0.6 (~1 hour).

5. Add 25 μl of 100 mM IPTG to each 5 ml culture (0.5 mM final concentration).

6. Incubate at 37°C for 3 hours in an orbital shaker (250 rpm).

7. Dispense the culture into 1.5 ml microcentrifuge tubes (~1–1.5 ml aliquots). Harvest the cells by centrifugation (4,000–8,000×g for 2–5 minutes). Discard the supernatant.

8. Store the cell pellets at -80°C or keep on ice and continue working.

D.2. CELL LYSIS

9. Freeze the pellets (-20°C) for at least 30 minutes. Then thaw and resuspend in 1 ml of Lysis buffer and freshly added lysozyme (~1 μg/μl final concentration).

10. Incubate on ice (4°C) for 30 minutes.

11. (Optional) Perform 3–5 freeze–thaw cycles with liquid nitrogen or dry ice.

	12. On ice, lyse the cells by sonication (Ultrasonic Disruptor UD-100, Tomy: 50% output, 6 cycles of 10 seconds ON and 5 seconds OFF). The cell suspension should become translucent after this step. Repeat if necessary.
	13. Centrifuge at 10,000×g for 30 minutes at 4°C to clarify the lysates. Collect the supernatant (~1 ml) and transfer to a clean tube. Discard the cell debris (pellet).
D.3. HIS-TAGGED GFP1–10 PROTEIN PURIFICATION	14. Resuspend the Ni-NTA magnetic agarose beads (QIAGEN) and immediately add 200 μl to each 1 ml cleared lysate. Mix thoroughly by pipetting.
	15. (Optional) If the lysate is viscous, add DNase I (~5 units/ml final concentration).
	16. Incubate at room temperature for 1 hour (or 4°C for 2 hours). Ensure to keep the beads suspended by mixing them every 10–15 minutes or using a rotating device.
	17. Spin-down and place the tubes in a magnetic rack. Wait for ~1 minute, then carefully remove the liquid without disturbing the pellet.
	18. Wash the beads 3 times with 500 μl of Wash buffer. Pellet the beads with a magnet, then carefully remove the liquid.
	19. Resuspend the beads in 50–100 μl of Elution buffer. Incubate at room temperature for ~1 minute.
	20. Place the tubes in a magnetic rack and wait for ~1 minute. Carefully collect the liquid without disturbing the pellet and transfer to a clean tube. Repeat one more time to remove any carry-over beads. Beads can be regenerated and reused.
	21. (Optional) Repeat steps 20–22 to maximize the recovery.
D.4. DESALTING (IMIDAZOLE REMOVAL)	22. To desalt and concentrate the protein, transfer the solution to an Amicon Ultra-0.5 ml (MWCO 3 kDa or 10 kDa) centrifugal filter (Merck Millipore).
	23. Fill the filter column with Storage buffer (~400 μl). Centrifuge at 15,000×g for 15–30 minutes. Discard the flow-through. Repeat 3–5 times until the traces of imidazole are negligible.
	24. Place the filter upside-down in a clean collection tube. Centrifuge at 1,000×g for 2 minutes to recover the protein solution.
	25. Measure the protein concentration and store at -80°C.

E. CELL-FREE PROTEIN SYNTHESIS (CFPS) IN DROPLETS

The cell-free protein synthesis reaction was performed in droplets using PUREfrex 1.0 (Gene Frontier). The reaction parameters were set according to the manufacturer's recommendations.

E.1. CELL-FREE PROTEIN SYNTHESIS (CFPS) REACTION IN EMULSION	1. On ice, prepare the CFPS reaction mix on ice as following:

Component	20 μl reaction	Final Conc.
Beads from BEAMing*	(6–8 μl)	
Nuclease-free water**	5 μl	
PURE*frex* 1.0 Sol. 1 (2X)	10 μl	1X
PURE*frex* 1.0 Sol. 2 (10X)	1 μl	1X
PURE*frex* 1.0 Sol. 3 (10X)	1 μl	1X
GFP1–10 protein (1 mg/ml)	1 μl	50 μg/ml (~26 μM)
ROX-OMB (10 μM)	1 μl	0.5 μM
Ligand (20X)***	1 μl	1X

* The concentration of the beads has to be empirically adjusted to yield ~10–30% transcriptionally active droplets (ROX+). Under the current protocol conditions, 6–8 μl is the recommended volume.

** Beads tend to stick together, to reduce agglomeration and promote dispersion, wash the beads once with nuclease-free water, pellet them with a magnet, and remove the liquid. Then first add the nuclease-free water (5 μl) to the beads and thoroughly resuspend by pipetting before adding the rest of components.

*** For the no-ligand condition, replace the ligand volume with nuclease-free water or the dilution buffer of the ligand.

2. Generate the droplets as described before (see **B.1–B.2**, steps **1–11**, page **VI**).
3. Incubate the emulsion at 37°C for 4 hours, then transfer to ice.
4. Immediately proceed to the droplet sorting (see **F.1–F.2**, page **X**).

F. FLUORESCENCE-ACTIVATED DROPLET SORTING (FADS)

The droplets containing the CFPS reaction mix were sorted by fluorescence-activated droplet sorting (FADS) using a commercially available On-chip Sort instrument (On-chip Biotechnologies) and Chip-Z1001 microfluidics chips (material: COP).
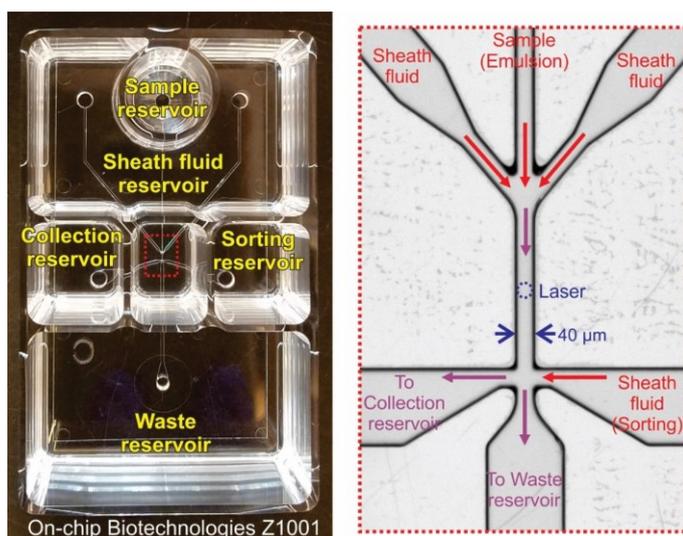
Figure 41. Chip-Z1001 microfluidics chip (On-chip Biotechnologies) for droplet sorting.

F.1. PREPARING THE CHIP AND SETTING UP THE SORTER MACHINE	

1. Add sheath fluid (0.1% (w/w) Pico-Surf 1 in HFE-7500) to each reservoir of the Chip-Z1001 microfluidics chip (On-chip Biotechnologies).
2. Carefully insert the chip in the chip holder and secure the lid.
3. Pull out the chip stage of the On-chip Sort instrument, and carefully insert the chip holder into it. Secure the lid and carefully reintroduce the chip stage into the machine.
4. Prime the channels of the chip.
5. After finish priming, empty the sample reservoir and waste reservoir.
6. Fill the main reservoir and sorting reservoir with sheath fluid, visually confirm that the fluid level in both reservoirs is equal.
7. Remove the sheath fluid from the collection reservoir leaving a thin layer (~100–200 μl) of sheath fluid.
8. Add ~400–500 μl of low-density mineral oil (Sigma) to the collection reservoir. Visually confirm that an upward meniscus (convex) is formed between the sheath fluid (bottom) and the mineral oil (top). The meniscus will trap the sorted droplets in the center of the reservoir and facilitate their recovery. Be careful when adding the mineral oil as it is not compatible with the material of the chips (cyclo olefin polymer, COP).
9. Add the sample (droplets) directly from the CFPS reaction to the sample reservoir. If possible, remove excess of sheath fluid from the bottom of the reservoir.

F.2. FLUORESCENCE-ACTIVATED DROPLET SORTING (FADS)

10. Carefully insert the chip with the sample in the chip holder and secure the lid.
11. Pull out the chip stage of the On-chip Sort instrument, and carefully insert the chip holder into it. Secure the lid and carefully reintroduce the chip stage into the machine.
12. Align the chip stage. Always perform the alignment before starting any run. Also verify that the lasers are properly aligned.

13. Set up the filter parameters and selection regions and gates.

Channel	Purpose
FSC (front scatter) SSC (side scatter)	To select the proper droplet size. More than 95% of the droplets should be clustered in one spot.
FL-2 (ex. 488 nm, em. 543±22 nm)	To measure GFP fluorescence and detect expression of the reporter gene.
FL-5 (ex. 561 nm, em. 676±37 nm)	To measure ROX fluorescence and detect the presence of transcriptional activity.

14. Pre-run the sample without sorting (flow cytometry only). Each sample is slightly different; thus, the regions and gates need to be adjusted accordingly.

15. Choose the desired region or gate and apply it for sorting/selection.

16. Start the sorting run. The run can be stopped and restarted at any moment.

- Remember to re-align the chip stage before starting each run.

- Selected droplets should be collected in the collection reservoir, everything else will go to the waste reservoir.

- At the current settings, the throughput of the On-chip Sort instrument is approximately 6.3×10^5 droplets/h for a droplet size of ~26 μm diameter.

- Periodically verify the sheath fluid levels inside the reservoirs. When necessary, fill the main reservoir and sorting reservoir with sheath fluid and empty the waste reservoir. If the collection reservoir fills up, carefully remove some of the sheath fluid by slowly pipetting up with a fine tip from the bottom of the reservoir without disturbing the droplets.

- Droplets tend to float and concentrate at the top of the sample reservoir, sometimes this generates artifacts during the sorting run after a while. If that happens, stop the run, gently stir the droplets inside the sample reservoir, then restart the sorting.

G. BEADS RECOVERY AND LIBRARY REGENERATION

The beads from the sorted droplets were carefully recovered and a standard PCR was performed to regenerate the DNA library. The new DNA templates were directly used for the next cycle of BEAMing.

G.1. BEADS RECOVERY AND WASHING

1. Carefully remove most of the sheet fluid (HFE-7500) with a fine pipette tip, suctioning from the bottom of the reservoir without disturbing the droplets. Leave ~50–80 μl of the sheath fluid in the reservoir.

2. Carefully retrieve the remaining sheath fluid containing the droplets from the collection reservoir and gently transfer it to a clean 0.2 ml PCR tube. Collecting a little of mineral oil is not a problem.

3. (Optional) Wash the walls of the used tip by pulling ~50–80 μl of sheath fluid from the bottom of the tube, and then slowly pushing the liquid out.

4. Add 30 μl of TK buffer and 0.5–1 μl of blank (non-coated) magnetic beads.

5. Break the emulsion with an anti-static gun (MILTY Zerostat 3).

6. Spin-down the liquid. Place the tube in a magnetic rack and wait for ~1 minute. Carefully remove as much liquid as possible without disturbing the pellet. Remove mineral oil (top) first, then HFE-7500 (bottom), and finally aqueous phase (mid).

7. Resuspend the beads in 100 μl of Breaking buffer. Mix well, then spin-down. Pellet the beads with a magnet, then carefully remove the liquid. Try to remove all the remaining oil and HFE-7500 (white emulsion on the top and bottom) at this stage.

8. Wash the beads 3 times with 50 μl of TK buffer. Pellet the beads with a magnet, carefully remove all the liquid.

G.2. LIBRARY REGENERATION THROUGH PCR

9. On ice, prepare the PCR mix in a clean tube as follow:

Component	12 µl reaction	Final Conc.
F primer: P3L (10 µM)	0.6 µl	0.5 µM
R primer: P4L (10 µM)	0.6 µl	0.5 µM
Q5 High-Fidelity 2X Master Mix (NEB)	6.0 µl	1X
Nuclease-free water	4.8 µl	

10. Resuspend the washed beads with 12 µl of the PCR mix. Add the PCR mix to the wall of the tube where the pellet is located to catch as many beads sticked to the wall as possible.

11. Place the tubes in a thermal cycler and run a PCR reaction as following:

Step	Temp.	Time
Initial Denaturation	98°C	1 minute
30 cycles	98°C	10 seconds
	72°C	10 seconds
Final Extension	72°C	2 minutes
Hold	4°C	

12. Purify the PCR product with silica columns or agarose gel electrophoresis. Elute with nuclease-free water, or TE buffer for long term storage. Measure the DNA concentration.

13. Store the purified DNA at -20°C. Use this regenerated library for the next cycle of BEAMing-FADS.

14. (Optional) Perform a CFPS reaction in solution of the regenerated library to assess the enrichment of riboswitches (see **I.1**, steps **1–4**, page **XIV**). This step can be performed in bulk after all the sorting cycles are done.

H. HIGH-THROUGHPUT SEQUENCING AND DATA PROCESSING

To link each library with its corresponding sorting cycle, a custom DNA barcode (6–8 nt) was inserted to each library with PCR. The final sequencing libraries were producing by adding the Illumina sequencing adapters with a second PCR. The sequencing was performed using Illumina MiSeq or Illumina NovaSeq, and the data was processed using a custom script.

H.1. INSERTION OF CUSTOM DNA BARCODES

1. For each DNA library and sorting cycle, prepare a separate PCR mix as follow:

Component	20 µl reaction	Final Conc.
DNA library	(variable)	25 nM
F primer: P-HA-Nova-T1-# (10 µM)*	1 µl	0.5 µM
R primer: P-HA-Nova-B1-# (10 µM)*	1 µl	0.5 µM
Q5 High-Fidelity 2X Master Mix (NEB)	10 µl	1X
Nuclease-free water	to 20 µl	

* "#" denotes the custom barcodes used to identify each sorting cycle (see **Table 5**, page **II**).

2. Place the tubes in a thermal cycler and run a PCR reaction as following:

Step	Temp.	Time
Initial Denaturation	98°C	1 minute
6 cycles	98°C	10 seconds
	69°C	10 seconds
	72°C	15 seconds
Final Extension	72°C	2 minutes
Hold	4°C	

3. Purify the PCR product with silica columns. Elution can be done with nuclease-free water. Measure the DNA concentration.

H.2. FINAL SEQUENCING LIBRARY PREPARATION

4. Adjust the concentration of each library to be the same. Pool them together.

5. To add the Illumina sequencing adapters, prepare a PCR mix as follow:

Component	20 µl reaction	Final Conc.
Barcoded DNA libraries	(0.5 ng each)	
F primer: TruSeq-i5-UDI00## (10 µM)*	1 µl	0.5 µM
R primer: TruSeq-i7-UDI00## (10 µM)*	1 µl	0.5 µM
Q5 High-Fidelity 2X Master Mix (NEB)	10 µl	1X
Nuclease-free water	to 20 µl	

* "##" denotes a UDI pair (see **Table 6**, page **II**). It is possible to multiplex several samples and libraries by using different UDI pairs for each sample.

6. Place the tubes in a thermal cycler and run a PCR reaction according to the following settings:

Step	Temp.	Time
Initial Denaturation	98°C	1 minute
	98°C	10 seconds
12 cycles	69°C	10 seconds
	72°C	15 seconds
Final Extension	72°C	2 minutes
Hold	4°C	

7. Purify the PCR product by agarose gel electrophoresis. Elution can be done with nuclease-free water. Measure the DNA concentration.

H.3. HIGH-THROUGHPUT SEQUENCING

8. If not working with multiplexed libraries, add 10–15% Phi-X control to increase the base-call diversity.

9. Perform the sequencing with Illumina MiSeq or NovaSeq. 150–200 bp single reads or 150 bp paired-end reads is usually enough.

H.4. SEQUENCING DATA PROCESSING

The data was processed with custom shell scripts (an example file can be found in the folder of this thesis at the university archives as *Shell_Script_Example.sh*).

10. If paired end reads are used, merge the raw reads pairs.

11. Sort the raw reads according to the corresponding sorting cycle based on the custom barcode introduced during the sequencing library preparation.

12. Filter out reads containing low quality base calls and sequence errors.

13. Identify and isolate the randomized region within the reads.

14. Count the number of reads for each variant in the libraries.

15. Calculate the abundance (as a percentage or fraction) of each variant among the total number of reads within each cycle, using the formula:

$$a_{i,c} = \frac{r_{i,c}}{\sum_{i=1}^{m} r_{i,c}} \cdot 100$$

where i is a unique riboswitch variant among a total number of variants m; $r_{i,c}$ is the read count (number of reads) of variant i within the cycle c, with the initial library (pre-sorting) deemed as cycle $c = 0$; and $a_{i,c}$ is the abundance of variant i expressed as a percentage of the total number of reads in cycle c.

16. Calculate the "enrichment trend" of each variant, using the formula:

$$T_i = \frac{n \sum_{c=0}^{n-1} (c \cdot a_{i,c}) - \left(\sum_{c=0}^{n-1} c\right)\left(\sum_{c=0}^{n-1} a_{i,c}\right)}{n \sum_{c=0}^{n-1} c^2 - \left(\sum_{c=0}^{n-1} c\right)^2}$$

where i is a unique riboswitch variant among a total number of variants m; $a_{i,c}$ is the abundance of the variant i within the cycle c; T_i is the "enrichment trend" of the variant i; and n is the total number of sorting cycles analyzed starting from cycle $c = 0$ (initial library).

17. Rank the variants according to their "enrichment trend".

I. RIBOSWITCH EVALUATION THROUGH CFPS REACTION IN SOLUTION.

The riboswitch constructs (including promising riboswitch candidates and initial designs) were individually synthetized and tested in solution using PURE*frex* 1.0 (Gene Frontier).

I.1. CELL-FREE PROTEIN SYNTHESIS (CFPS) REACTION IN SOLUTION	

1. On ice, for each sample prepare two reaction mixes (with and without ligand) in 0.2 ml PCR tubes as follow:

Component	8 µl reaction	Final Conc.
DNA template (*gfp11*)*	(variable)	20 nM
PURE*frex* 1.0 Sol. 1 (2X)	4 µl	1X
PURE*frex* 1.0 Sol. 2 (10X)	0.4 µl	1X
PURE*frex* 1.0 Sol. 3 (10X)	0.4 µl	1X
GFP1–10 protein (1 mg/ml)**	0.4 µl	50 µg/ml (~26 µM)
ROX-OMB (10 µM)	0.4 µl	0.5 µM
Ligand (20X)***	0.4 µl	1X
Nuclease-free water	to 8 µl	

* Include an additional reaction with a blank. To prepare the blank, replace the DNA template with nuclease-free water.

** When eGFP constructs are used, replace GFP1–10 protein with nuclease-free water. For GFP11 constructs, the GFP1–10 protein can be replaced with a *gfp1–10* DNA template at the same molarity as the *gfp11* template (20 nM).

*** For the no-ligand condition, replace the ligand volume with nuclease-free water or the dilution buffer of the ligand.

2. Incubate the tubes at 37°C for 4 hours.
3. Transfer the reactions to a 384-well microplate (Greiner Bio-One, black, non-binding). Be sure that the liquid is uniformly distributed at the bottom of the well and there is no bubbles, as this may affect the fluorescence measurements.
4. Measure the GFP fluorescence (ex. 484 nm, em. 510 nm) and ROX fluorescence (ex. 587 nm, em. 599 nm) with a plate reader.

J. RESTRICTION DIGESTION ANALYSIS

To estimate the abundance and ratio of the different templates ("Strong" or "Weak" RBS) in the mock libraries, a PAGE-based restriction enzyme test was performed. A conditioning PCR was included prior the restriction digestion analysis to reduce the presence of heteroduplexes generated during the library amplification through standard PCR.

J.1. CONDITIONING PCR	

1. On ice, prepare the PCR mix as following:

Component	5 µl reaction	Final Conc.
DNA template	(variable)	0.5 nM
F primer: P3L (10 µM)	0.25 µl	0.5 µM
R primer: P4L (10 µM)	0.25 µl	0.5 µM
Q5 High-Fidelity 2X Master Mix (NEB)	2.50 µl	1X
Nuclease-free water	to 5 µl	

2. Place the tubes in a thermal cycler and run a PCR reaction as following:

Step	Temp.	Time
Initial Denaturation	98°C	1 minute
10 cycles	98°C	10 seconds
	72°C	10 seconds
Final Extension	72°C	2 minutes
Hold	4°C	

3. Dilute the PCR product by 10-fold with fresh PCR reaction mix of the same composition except with 2.5X DNA polymerase:

Component	50 µl reaction	Final Conc.
First PCR product (unpurified)	5.0 µl	1/10X
F primer: P3L (10 µM)	2.5 µl	0.5 µM
R primer: P4L (10 µM)	2.5 µl	0.5 µM
Q5 High-Fidelity 2X Master Mix (NEB)	25 µl	1X
Q5 DNA polymerase (2 units/µl, NEB)	1 µl	0.04 units/µl
Nuclease-free water	14 µl	

4. Place the tubes in a thermal cycler and incubate the reaction as following:

Step	Temp.	Time
Denaturation	98°C	2 minutes
Final Extension	72°C	2 minutes
Hold	4°C	

5. Purify the PCR product with silica columns. Elution can be done with nuclease-free water. Measure the DNA concentration.

J.2. RESTRICTION ENZYME DIGESTION

6. On ice, for each sample prepare one reaction mix per restriction enzyme:

Component	5 µl reaction	Final Conc.
DNA sample	(20 ng)	
EcoRI-HF or SacI-HF (20 units/µl, NEB)	0.2 µl	2 units/µl
Cutsmart 10X buffer (NEB)	0.5 µl	1X
Nuclease-free water	to 5 µl	

7. Incubate the reaction at 37°C for 1.5 hours.

8. (Optional) Inactivate the enzymes by a heat treatment (65°C for 20 minutes).

J.3. POLY-ACRYLAMIDE GEL ELECTRO-PHORESIS (PAGE)

9. Add 1 µl of 6X loading dye containing SDS to the digestion reaction. Load the whole volume (6 µl) onto a native polyacrylamide gel (8%).

10. Perform the electrophoresis at 200 V for 15 minutes.

11. Stain the gel for 10 minutes submerged in a SYBR gold (Invitrogen) solution.

12. Using a blue-light transilluminator/scanner, take high-quality pictures of gel.

13. From the pictures, estimate the relative abundance of each construct based on their corresponding bands intensities and fragment size.

K. HISTAMINE CONCENTRATION MEASUREMENT

The absolute histamine concentration in the samples was estimated using a "Histamine Test" colorimetric kit (Kikkoman Biochemifa). The volumes have been scaled down from the original manufacturer's protocol.

K.1. COLORIMETRIC MEASUREMENT OF HISTAMINE

1. Prepare the colorimetric reagent and enzyme solutions according to the manufacturer's instructions.

2. Dilute the samples with distilled water by a dilution factor (DF) of at least 25-folds or more (e.g., 10 µl sample + 240 µl water).

3. In a 96-well transparent plate prepare the reactions as following:

Component	For each sample		Only once	
	(E_s)	(E_b)	(E_{std})	(E_c)
Diluted sample	50 µl	50 µl	-	-
Distilled water	-	-	-	50 µl
Kit's histamine standard (4 mg/L)	-	-	50 µl	-
Kit's colorimetric reagent	50 µl	50 µl	50 µl	50 µl
Kit's enzyme solution	50 µl	-	50 µl	-
Kit's buffer solution	-	50 µl	-	50 µl

4. Mix well by pipetting, then incubate in the dark at 37°C for 15 minutes.

5. Measure absorbance Abs_{470nm} in a plate reader (Infinite M1000 Pro, Tecan).

   - Before measuring, remove any bubble formed inside the wells.
   - If $Abs_{470nm} > 1.1$, samples need to be diluted more.
   - If standard spectrophotometer is used, scale up the volume according to the recommended volume of the cuvette. Alternatively, after the incubation dilute the reaction with water and note the dilution factor.
   - Do not use nanodrop as the microvolumes gives inconsistent measurements.

6. Estimate the histamine concentration with the formula:

$$HA_s = \frac{E_s - E_b}{E_{std} - E_c} \times \frac{4 \times DF}{111.15}$$

where HA_X is the histamine concentration in mM of sample s ; E_s (Sample), E_b (Blank), E_{std} (Standard), and E_c (negative control) are the measured Abs_{470nm} values; DF is the dilution factor of the sample.