PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
Facultad de Ciencias Biológicas
Programa Doctorado en Ciencias Biológicas
Departamento Genética Molecular y Microbiología

Open platform for the implementation of RNA sensing reactions in cell-free systems

by

Anibal Andrés Arce Medina

Thesis submitted to the Pontificia Universidad Católica de Chile,
as one of the requirements to qualify for the academic Doctoral degree in Biological Sciences with mention in Molecular Genetics and Microbiology.

Supervisor	:	Dr. Fernán Federici (Pontificia Universidad Católica de Chile)
Co-Supervisor	:	Dr. Neil Dalchau (Microsoft Research Institute, UK)
Committee	:	Dr. Andreas Schüller (Pontificia Universidad Católica de Chile)
		Dr. Hannetz Roschzttardtz (Pontificia Universidad Católica de Chile)
		Dr. Jim Hasselof (University of Cambridge, UK)

January, 2021

Santiago, Chile

Dedicada a mis abuelos

Jorge y Rebeca

Alicia y Fernando

Agradecimientos

No puedo comenzar estos agradecimientos de otra manera. Gracias a mis padres, Patricio Arce y Maria Consuelo Medina, por darme todo su amor, sus consejos, y apoyo en los momentos de poca luz. Gracias por ser un ejemplo de las cosas buenas que quiero en mi vida, tanto en el mundo de las ciencias como en la vida familiar. Gracias a toda mi familia que me apoyó siempre.

Gracias a todos los docentes del programa de doctorado que me enseñaron sus pasiones y motivaciones. Muy especialmente agradezco a Fernán Federici por su constancia, paciencia y acompañamiento durante mi Tesis. Por contagiar sus ideas que al principio suenan increíbles, pero de a poco las concretizamos. Por enseñar a pensar cómo contribuir a un mundo más colaborativo, mas justo.

Gracias a todos los que contribuyeron con materiales y enseñanzas prácticas. Gracias a Isaac Nuñez y Tamara Matute por apoyar este trabajo desde sus orígenes y enseñarme las buenas prácticas en la bioingeniería. Gracias a los estudiantes de pregrado que trabajaron en sus proyectos con conmigo: Alexandra Bettinelli, y Juan Puig.

Gracias a todos los colaboratores internacionales: Especialmente a Neil Dalchau, quién me recibió en Inglaterra varias veces. Por enseñarme modelización Bayesiana y programación y celebrar mis pequeños exitos. Gracias a Jenny Molloy, Chiara Gandini y Fernando Guzman por trabajar en equipo en varios proyectos durante esta Tesis.

Resumen

La biología sintética busca el desarrollo de funciones programables y predecibles en sistemas biológicos. Uno de los recientes avances en esta área son los *toehold swtiches*: reguladores de la expresión genética diseñados *de-novo* y que permiten la traducción de un gen de forma gatillada por la interacción con un RNA *"trigger"* de secuencia específica.

Por otro lado, los sistemas de expresión genética libre de células (*cell-free*) han sido utilizados durante décadas en el campo de la biología molecular y fueron claves en el descubrimiento del código genético. Sin embargo, sólo muy recientemente han empezado a utilizarse como una herramienta para la bioingeniería. La expresión de *toehold swtiches* en sistemas de libre de células generó una plataforma de diagnóstico muy promisoria ya que los sensores de RNA tipo *toehold* pueden ser liofilizados para su transporte a temperatura ambiente con el objetivo de ser utilizadas en terreno y en zonas remotas. No obstante, a la fecha, estos sensores solamente han sido implementados en sistemas de expresión genetica reconstituidos, del tipo "PURE" (por sus siglas en inglés: *Protein expression Using Reconstituted Elements*). Los sistemas PURE requieren transporte en frío (-80 °C) desde los proveedores en el hemisferio norte, y son prohibitivamente caros para los usuarios en América Latina.

En este trabajo, se describe la implementación de sensores de RNA de tipo *toehold* en sistemas libres de células utilizando extractos celulares de *E. coli* que pueden ser producidos localmente y han reducido los costos en dos ordenes de magnitud, manteniendo un desmpeño comparable al de los sistemas PURE comerciales. Mediante el uso de CRISPRi para silenciar nucleasas en las células antes de lisarlas se logró aumentar la estabilidad de dsDNAs lineales, permitiendo el uso de productos de PCR como sustrato de las reacciones de expressión genética libre de células incluyendo la expresión de sensores de RNA tipo *toehold*.

Como prueba de concepto, utilizamos las herramientas desarrolladas en esta tesis para el prototipado de 8 nuevos sensores de tipo *toehold* para detectar RNA del patógeno viral PVY (*Potato Virus Y*) que afecta dramatcamente la productividad del cultivo de papas (*Solanum tuberosum*) en el mundo.

La implementación local y de bajo costo de sensores de RNA de tipo *toehold* en sistemas de expresión genética libres de células podría permitir escalar la capacidad diagnóstica a nivel regional e impulsar nuevos modelos más descentralizados de monitoreo de ácidos nucleicos y de enfermedades infecciosas.

Abstract

Synthetic biology seeks the development of programmable and predictable functions in biological systems. One of the latest advances in the field are the toehold switches: *de-novo* engineered regulators of gene expression that allow the translation of a gene only after the interaction with a cognate "trigger" RNA of a specific sequence.

On the other hand, cell-free protein expression systems have been used for decades facilitating discoveries in molecular biology, and have played a critical role in the elucidation of the genetic code. More recently, cell free systems have emerged as a tool for the engineering of genetic devices, bioproducts, and biosensors, among many other applications. The expression of programmable toehold switch-based RNA sensors in cell-free systems has generated as promising platform for diagnostics. Cell-free toehold sensors can be freeze-dried for room temperature transport to the point-of-need. These sensors, however, have been implemented using reconstituted PURE (Protein expression Using Reconstituted Elements) cell-free protein expression systems that are difficult to source in Latin America, due to their prohibitively expensive commercial cost, and cold-chain shipping requirements (-80 ^{o}C) from suppliers in the northern hemisphere.

Here, we describe the implementation of RNA toehold sensors using *E. coli*

cell lysate-based cell-free protein expression systems, which can be produced locally and reduce the cost of sensors by two orders of magnitude, while providing sensor performance comparable to commercial PURE cell-free systems. Further optimization of the cell extracts with a CRISPRi strategy enhanced the stability of linear dsDNAs, enabling the use of PCR products as a substrate for cell-free gene expression reactions including toehold sensors. As a proof-of-concept application, we used the tools developed in this thesis to prototype and screen 8 novel RNA toehold sensors for the potato pathogen Potato Virus Y (PVY virus) that dramatically reduces the yield of this important staple crop. The local implementation of low-cost cell-free toehold sensors could enable biosensing capacity at the regional level and lead to more decentralized models for global surveillance of infectious diseases.

Contents

Resumen **3**

Abstract **5**

1 Introduction **9**

1.1 Toehold RNA sensors: *De-novo* designed regulators for Synthetic Biology 9

1.2 Cell-free toehold sensors (CFTS) in PURE systems 12

1.3 Crude extracts as alternatives for PURE systems 13

1.4 Energy metabolism for cell-free crude extracts and cost-effective alternatives 15

1.5 Potato virus Y (PVY) as a target for designing cell-free toehold sensors 19

Hypothesis **22**

General Aims **22**

Specific Aims **22**

2 Materials and Methods **23**

2.1 Rapid DNA preparation using Golden Gate and Gibson assembly methods 23

2.2 Input DNA preparation for cell-free transcription-translation reactions . 31

2.3 Crude extract preparation 31
2.4 Stability of PCR-derived linear DNA in cell-free reactions 33
2.5 Cell-free transcription-translation reaction using home-made cell extracts 34
2.6 Cost breakdown analysis 35
2.7 Cell-free transcription-translation reaction using NEB Purexpress . . . 36
2.8 Trigger RNA preparation 36
2.9 Lyophilization and storage of cell-free reactions 36
2.10 Isothermal NASBA amplification 37
2.11 Selection of conserved sequences across PVY genomes 38
2.12 *In-silico* toehold design 41
2.13 PCR amplification of PVY toehold sensors for rapid screening 49

3 Results from aim 1: To Implement and optimize an in-house low-cost cell-free gene expression system 53

3.1 Optimization of magnesium concentration 54
3.2 Comparison of low and high-cost energy sources for cell free sfGFP expression 56
3.3 Comparison of in-house prepared methods with commercial cell free systems for sfGFP expression 59
3.4 Stability of lyophilized constitutive gene expression reactions 61
3.5 CRISPRi optimization of linear dsDNA stability 63
3.6 Cost breakdown analysis 70

4 Results from aim 2: Benchmarking RNA toehold sensing capacity using low-cost cell-free preparation methods 73

4.1 Selecting reporter genes for cell-free toehold sensors 74
4.2 Performance of RNA sensing capacity vs commercial PURExpress . . . 87

4.3 Cell-free toehold sensors can be lyophilized 91
4.4 Effect of DNA input concentration on the sensing capacity of CFTS . . 93
4.5 NASBA amplification increased sensibility of Cell-free toehold sensors . 95
4.6 The use of linear DNA for CFTS in the CRISPRi strains 96

5 Results from aim 3: To generate novel RNA toehold sensors for the PVY virus. 98

5.1 Search of conserved sequences in PVY genomes corresponding from different strains 98
5.2 Development of software for designing PVY toehold sensors 101
5.3 Prototyping PVY toehold sensors using PCR- derived linear DNA in CRISPRi optimized cell-free reactions 102
5.4 Validation of good performance of the PVY toehold sensor 1 and sensor 7 stored in plasmids 107

6 Discussion 111

6.1 Decentralizing cell-free RNA sensing with the use of low-cost cell extracts 111
6.2 Further optimizations of low-cost cell-free RNA sensing reactions 112
6.3 Improvements in computational design of novel toehold sensors 115
6.4 Understanding batch-to-batch variability in cell-free gene expression reactions and increasing robustness in CFTS 117

7 Conclusions 119

8 Appendix 121

8.1 Breakdown costs of cell-free reactions produced in the UK 121
8.2 Breakdown costs of cell-free reactions produced in Chile 128

1. *Introduction*

1.1 Toehold RNA sensors: *De-novo* designed regulators for Synthetic Biology

The aim of synthetic biology is to engineer programmable and predictable functions in biological systems. This multidisciplinary field is enabling the engineering of artificial genetic systems of unprecedented scale and reliability (e.g. [1]). Synthetic Biology applies classical engineering approaches to biology such as modular design, characterization of core components and mathematical modelling combined with large scale DNA synthesis and assembly [2, 3]. A key aspect of this engineering process is the design of elementary components that can be used to program higher level functions (e.g. artificial transcriptional control for the engineering of logic gates that can be used for higher level circuitry [1]). Therefore, part of the synthetic biology community has focused on identifying steps, processes and components that can be engineered or reprogrammed to create controllable elementary functions, e.g. programmable repressor-promoter pairs [4, 5, 1]) and recombinase-based switches [6].

In this context, the engineering of RNA components has emerged as a promising tool for circuitry design [7, 8, 9, 10]. Green *et al.* have developed a novel type of regulator of gene expression in bacteria called toehold switch [11]. Toehold switches

are *de-novo* designed RNA structures that can be programmed to detect RNAs of arbitrary sequences. A toehold switch works by forming a hairpin at the 5′-end of a reporter gene that prevents RBS-ribosome interactions, thereby stopping translation. This structure is released by toehold-mediated strand displacement by an unstructured RNA 'trigger' with a sequence that is complementary to the switch (Figure 1.1).

Computational and experimental characterization of the first-generation library allowed identifying specific parameters that correlate with the higher specificity and broader dynamic range of the RNA toeholds sensors (optimal length of toehold region and hairpin structures others). Using this information, Green *et al.* [11] recently developed a library of 12 forward-engineered, orthogonal toeholds switches and triggers that displayed dynamic ranges comparable to protein-based regulators of gene expression in *E. coli*.

These forward-engineered designs have been applied to the development of sensors for naturally occurring RNAs in *E. coli* [11]. However, the fixed sequences of potential RNA triggers present significant challenges for effective system activation. For example, unlike synthetic trigger RNAs (designed *de-novo* to be single-stranded), natural RNA targets (for instance, mRNAs, miRNA, or viral RNAs) may display strong secondary structures [12], making the toehold binding capacity unfavourable and affecting the thermodynamical properties that drive the strand displacement process. Also, the toehold sequences defined by the natural trigger mRNA can exhibit base-pairing both internally and with sequences downstream of the hairpin module [11, 13]. Moreover, when used for sensing naturally occurring RNAs, toeholds sensors are prone to off-target miss-activation.

To overcome the issues related to off-target activation, a detailed analysis of the sequences and thermodynamic properties of the toehold sensor devices and trigger RNAs has to be performed. This analysis, along with the genomic complexity of the

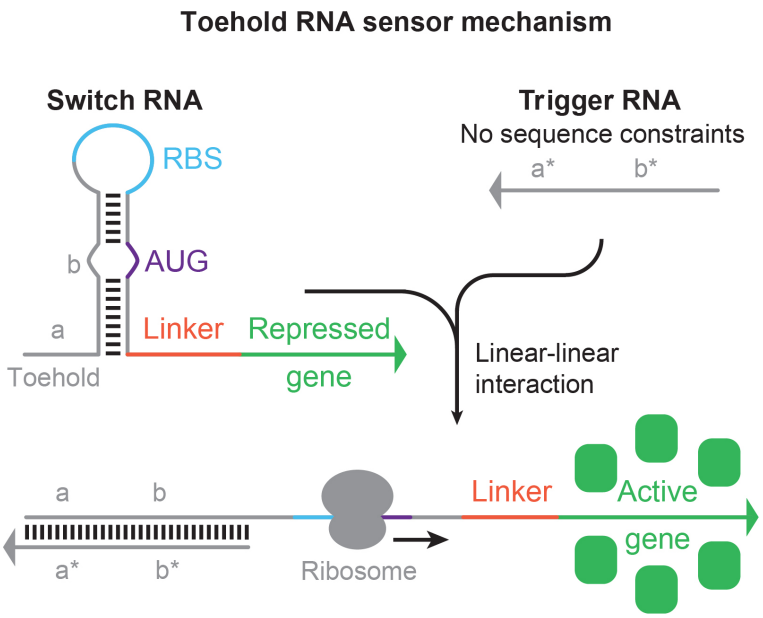

Figure 1.1: Schematic model of the toehold sensor mechanism. Binding of the trigger RNA to the toehold domain (sequences a and a*) facilitates the strand displacement reaction of the rest of the trigger (sequence b*) into the hairpin domain (sequence b). This interaction remodels the structure of the switch RNA, leaving RBS available for interaction with the ribosome, opening the transcript for translation. Image reproduced from Green *et al.* [11]

.

species of the target RNA, has to be taken into consideration for sensor design. To tackle most of the issues related to secondary structures within the switch RNA and complexity of the trigger RNA, toeholds engineered for detection of endogenous RNAs contain some structural modifications including:

1. Increasing toehold domain length (from 12 nt to 24 or 30 nt) [11, 14, 15]. This modification increases the possibility that single-stranded regions from the trigger mRNA seed the interaction with the toehold switch itself.

2. Decreased the size of the loop domain. This modification was shown to decrease leakness (expression of the reporter in absence of the trigger RNA) [15].

Both modifications have improved the capacity of the toehold sensors towards detecting mRNA from endogenous and exogenous genes in *E. coli* [11] and for the recognition of viral sequences from Ebola [14], Zika virus [15] and Norovirus [16].

1.2 Cell-free toehold sensors (CFTS) in PURE systems

A range of RNA toehold sensors [14, 15, 16] have been implemented successfully in a reconstituted cell-free gene expression system, known commercially as PURExpress (NEB), which consists of a complex mixture of the components involved in gene expression from *E. coli*. Although the exact composition of PURExpress has not been made publically available, it is based on the publication from *Shimizu et al.* [17], which developed the PURE system (Protein synthesis Using Recombinant Elements) that combines more than 36 individually purified reagents involved in transcription (T7 RNA polymerase); translation (IF1, IF2, IF3, EF-G, EF-Tu, EF-Ts, RF1, RF3, 20 aminoacyl-tRNA synthetases, methionyl-tRNA transformylase (MTF),

and ribosomes); and energy-recycling (creatine phosphate, creatine kinase, myokinase, nucleoside-diphosphate kinase, and pyrophosphatase). In addition, the system contains 46 tRNAs, NTPs, creatine phosphate, folinic acid, and amino acids.

Cell-free toehold sensors (CFTS) have been shown to work on different substrates such as paper, where they are stable after lyophilization and transportation at room temperature. This feature makes them ideal for in-field deployment where viral RNAs can be detected upon rehydration of freeze-dried CFTS reactions [18].

1.3 Crude extracts as alternatives for PURE systems

Although CFTS are ideal for point-of-care deployment, constraints in the accessibility to PURE systems in South America and the global south impose significant barriers to implementing, protyping and designing *de novo* CFTS locally. PURE systems are not only expensive and proprietary but also difficult to source since they need to be imported from the Northern hemisphere under strict cold-chain shipping (GeneFrontier from Japan, or Creative Biolabs and New England Biolabs from the US) that increases its costs and the risk of activity loss. Alternatively, crude extracts from *E. coli*, which do not require purification of any particular component, could also be used as an alternative substrate for CFTS [19].

Cell-free gene expression systems based on crude extracts are created by mixing the molecular machinery present in a cell lysate with energetic substrates and cofactors to enable transcription-translation coupled reactions *in vitro*. These systems have been extensively used since the 1960s when they played a critical role in the elucidation of the information flux from DNA to proteins [20, 21, 22]. Moreover, the Nobel prize-winning experiments that led Marshall Nirenberg and his colleagues to deciphering

the codons of the genetic code were performed mainly using *E. coli* crude extracts [21, 23, 24].

In subsequent decades, cell-free systems continued to be essential for elucidating molecular mechanisms. More recently, cell-free systems gained popularity as a substrate for engineering biomolecular systems outside living cells [25, 26]. For instance, progress in disulfide bond formation and membrane anchoring has made possible the manufacturing of protein complexes such as monoclonal antibodies [27], ion channels [28], functional the ATP synthase [29] and viable virus-like particles [30, 31]. Moreover, advances in the understanding of endogenous transcription in *E. coli* [32] and the engineering of fluorescent RNA aptamers [33, 34] have facilitated the development of transcriptional cascades [32], light-sensing circuits [35], transcriptional oscillators in cell-free gene expression reactions [36, 37, 38].

The field of biosensing has also benefited recently from cell-free systems, with the bioengineering of novel sensors for small molecules including endocrine disruptors [39], fluoride [40], atrazine [41], and heavy metals [42], allowing the detection of metallic contaminants in water samples [43] or micronutrients in serum [44].

Various protocols have been crafted to simplify extract preparation and optimize its performance, including diverse methods for cell disruptions such as homogenization [45, 46], sonication [47, 48], bead-beating [49], enzymatic lysis [50], or flash freezing [51]. Several protocols for producing cell-free crude extract reactions have shown compatibility with lyophilization and enhanced stability through improved preservation techniques [52, 53, 54, 55]. However, most of the diversity across cell-free preparation methods and the costs associated with reagents arise from the use of different energy sources for ATP regeneration methods.

1.4 Energy metabolism for cell-free crude extracts and cost-effective alternatives

Transcription and translation of a gene involves many steps that consume ATP or its equivalents. The incorporation of a single nucleotide in an mRNA uses the energy equivalent of 2 ATP equivalents per phosphodiester bond formation, while 4 ATP equivalents are needed per new peptide bond formation in an aminoacidic chain (one ATP is hydrolysed to AMP in the aminoacylation of the tRNA reaction; one GTP is hydrolysed to GDP in the aatRNA binding reaction catalysed by EF-Tu; and one GTP is hydrolysed to GDP in the translocation reaction catalysed by EF-G) [56]. Therefore, an adequate and steady supply of ATP is necessary for efficient cell-free gene expression [17, 57, 58, 59].

The most straightforward ATP regeneration systems are achieved by incorporating high-energy phosphate donor molecules to cell-free reactions, often supplemented with purified enzymes and cofactors. Examples are acetyl phosphate plus acetate kinase [60]; phosphoenolpyruvate (PEP) plus pyruvate kinase [61]; creatine phosphate plus creatine kinase [62]; and 3-phosphoglycerate (3-PGA) [63]. However, these phosphate donor molecules are expensive and cold chain dependent, limiting the potential of the cell-free expression technology to the global south.

Another issue of energizing the cell-free reaction using substrate-level phosphorylation is the risk of increasing concentrations of inorganic phosphate (iP) to levels that impair the reaction. This effect is likely due to iP chelating free magnesium ions, which serve as an essential co-factor in many nucleotide-dependent reactions, and whose optimal concentration range in cell-free expression systems is very narrow [49].

Many efforts have been taken to activate the endogenous central metabolic path-

ways present in the crude extracts towards the regeneration of chemical energy cost-effectively and leaving fewer unwanted sub-products. For instance, Jewett *et al.* showed that the oxidative phosphorylation pathway is functional in cell-free extracts, and described an ATP regeneration mechanism that relies on ATP-synthases present in inverted membrane vesicles produced during the cell lysis process [64]. Using these findings, Jewett *et al.* developed an energy system based on glutamate, which achieved a high protein synthesis yield (1.2 mg/ml). Glucose-6-phosphate has also been used as an energy source for cell-free gene expression reactions [58], but initial attempts using glucose were not successful. Calhoun and Swartz showed that glucose produced a rapid decrease in pH (due to the accumulation of organic acids) that inhibits protein synthesis [57]. This issue was partially solved by Kim *et al.*, fortifying the buffering capacity of the reaction mixture, reporting yields up to 0.9 mg/mL [65]. Moreover, Kim *et al.* described the use of S12 extracts, which unlike the typically used S30 extracts, do not require ultra-centrifugation or dialysis steps, leaving residual cofactors available in the reaction media and making available the protocol to the use of conventional bench-top instruments [65].

Later studies showed extended protein production time using polymeric carbohydrates as an energy reservoir that released glucose at a steady rate, avoiding a rapid decrease in pH [66]. Recently, Caschera *et al.* reported a cost-effective method that couples a polyphosphate molecule (HMP) with maltodextrin (Figure 1.2). HMP, releases inorganic phosphates slowly in aqueous solution; maltodextrin, a polysaccharide consisting in D-glucose monomers, can be slowly catabolized for successive ATP generation. Initially, HMP is the molecule phosphate donor for Glucose-1-phosphate (G1P) generation using maltodextrin as substrate in a reaction dependent on maltodextrin phosphorylase, which is present in the crude extract. G1P is subsequently processed in the glycolytic pathway to produce ATP, and reducing power, while lactate and acetate

are wastes along with inorganic phosphate. The latter can now recycle to be used in the generation of G1P. Caschera *et al.* showed that supplementation with NAD and CoA cofactors improves ATP synthesis, iP recycling and lead to yields to yield up to 1.65 mg/mL. The cost-efficiency in Caschera *et al.* protocol [67], combined with the simplicity of the S12 extract [65], makes a suitable alternative for low-cost and local implementation of cell-free reactions in the global south.

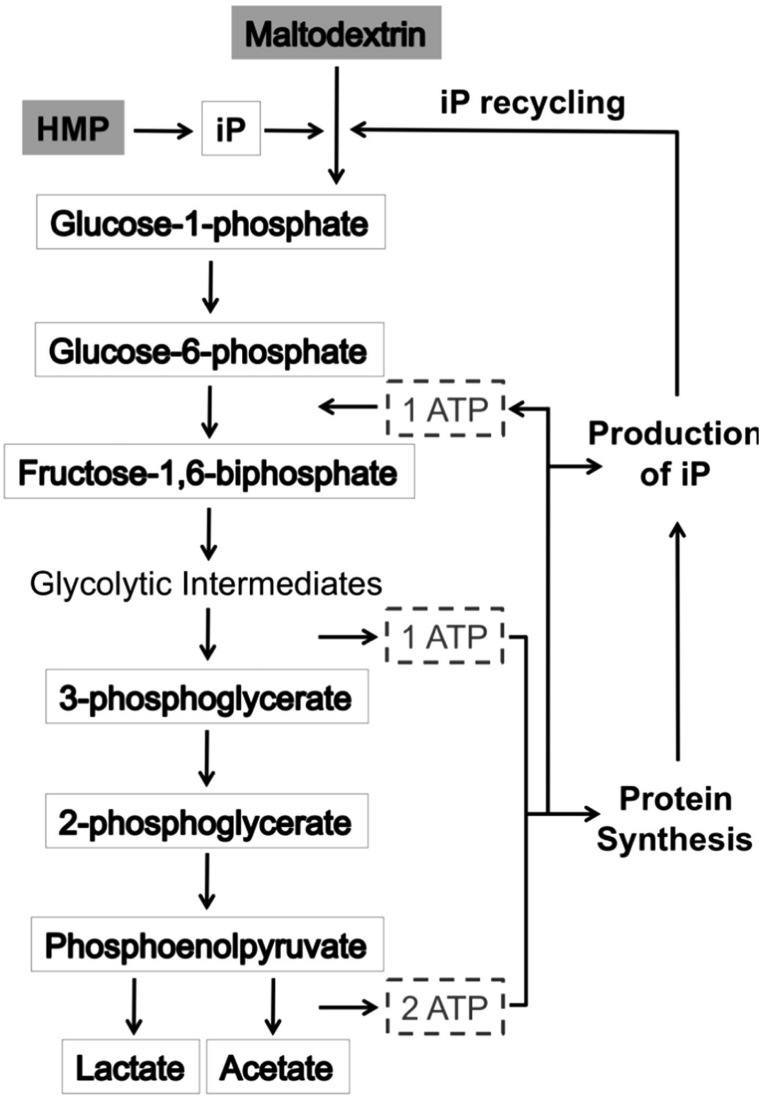

Figure 1.2: Simplified schematic model of metabolism based on maltodextrin and HMP for ATP production and iP recycling. Chemicals shaded in gray are added to the cell-free reaction. Image taken as published in Caschera *et al.* [67].

1.5 Potato virus Y (PVY) as a target for designing cell-free toehold sensors

Potato (*Solanum tuberosum* subsp. tuberosum) is one of the essential staple foods for humanity, being the fourth most cultivated crop behind only rice, wheat, and corn. Potato is propagated vegetatively through tubers, being prone to accumulating and spreading viral diseases. To date, Potato virus Y (PVY) is considered the most dangerous virus pathogen to this crop and is distributed worldwide [68]. Besides potato, PVY virus has a wide range of hosts, including pepper, tomato, tobacco, eggplant, and others. PVY can efficiently be transmitted from one plant to another via animal vectors such as aphids [69] or transmitted mechanically through plant wounds [70]. One strategy to cope with the spread of this virus in the fields was implemented in the US, and consisted on planting virus-free certificated seeds annually, while removing plants originated by uncertificated sources [71].

One of the origins of the potato crop, the south of Chile, possesses ideal conditions to produce virus-free potato seeds due to its land quality, weather conditions, and phytosanitary isolation. However, to maintain this ideal scenario, the early detection of emergent diseases is essential, and the installation of affordable molecular diagnostic capacity directly in the field is an unmet need for potato farmers in Chile, and worldwide. In this thesis, we focus on the PVY virus as a target for the design of cell-free toehold sensors.

PVY virus belongs to the Potyviruses genus, and has a single-stranded, positive-sense RNA genome of about 9.7 kb nucleotides in length which is covalently bound to a VPg protein on the 5' end, and polyadenylated on the 3' end. A single open reading frame is present on the PVY genome, which encodes for a polyprotein (3063 aa) that is

cleaved by three viral proteases (P1, HC-Pro and Nib) to produce ten functional proteins (P1, HC-Pro, P3, 6k1, Cl, 6k2, VPg,NIa, NIb and CP)[68]. PVY is well-known for its genetic diversity and propensity to evolve through two mechanisms: the accumulation point of mutations; and recombination between PVY strains that exchange large (several hundred nt-long) segments of their genome. Phylogenetic analysis suggests that PVY-N and PVY-O are parental strains which genomes can recombine at at recombination junctions (RJ, Figure 1.3) generating novel strains such as PVY-N:O, PVY-N-Wi, PVY-NTN (Figure 1.3) [72, 68, 73, 74, 75].

The different PVY strains are associated with different symptomatic responses on the host plants. For example, the PVY-N strain is described to cause leaf necrosis and mild or even no damage to the tuber [68, 76]; while the other parental strain (PVY-O), provokes mild tuber damage and do not cause leaf necrosis [68]. On the other hand, the PVY-NTN recombinant strain has been strongly associated with the necrotic ringspot disease (PTNRD) [77, 78, 79]. The necrotic disease on the tubers caused by PVY-NTN strains is the main cause of economic loss associated with PVY infections since tubers with PTNRD are discarded. The presence of PVY recombinant and necrotic strains has been confirmed in Africa [80], Asia [78], Europe [81] and America [82, 79] and are becoming the more prevalent strains in many potato production areas of the world [68].

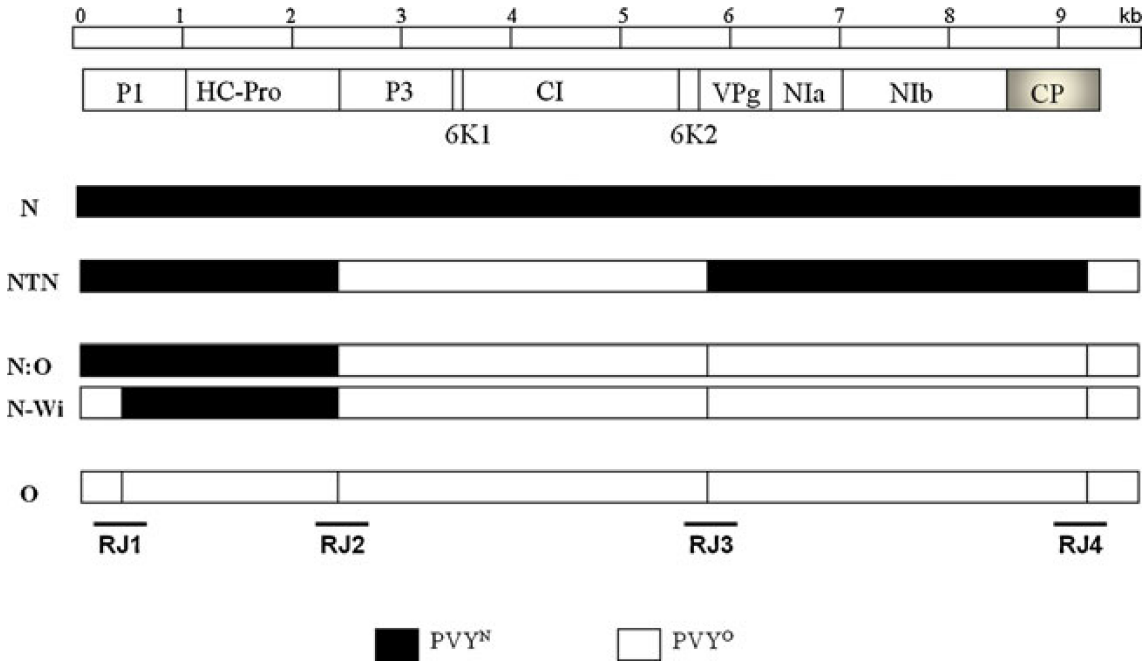

Figure 1.3: Schematic diagram of the main types of PVY recombinants :PVY-NTN, PVY-N:O, and PVY-N-Wi. Different coloring of genomic fragments indicate the parental genome of origin, either from PVY-N (black) or PVY-O (white). Positions of recombinant junctions, RJ 1 to 4, are labeled below the diagram. The ten PVY genes are drawn above the diagram approximately to scale. Image taken as published in Karasev *et al.* [83].

Hypothesis and aims

Hypothesis

De novo designed cell-free toehold RNA sensors for the detection of PVY viral sequence can be implemented in home-made, low-cost *E. coli* cell extracts.

General Aims

To develop a platform for low-cost implementation of cell-free toehold sensors for RNA targets of local interest.

Specific Aims

1. To implement and optimize low-cost cell-free gene expression methods.
2. To benchmark RNA toehold sensing capacity using low-cost cell-free preparation methods.
3. To generate novel RNA toehold sensors for the PVY virus.

2. *Materials and Methods*

2.1 Rapid DNA preparation using Golden Gate and Gibson assembly methods

In order to rapidly produce plasmid DNA with the desired combinations of reporter genes and toehold sensors, two different techniques were used: Golden Gate [84, 85] and Gibson assembly [86]. Golden Gate level 0 parts were domesticated by Gibson assembly into level 0 plasmids. Golden Gate relies upon Type IIS restriction enzymes (such as BsaI) which each recognize a six base pair non-palindromic sequence and cut at a specified distance from that recognition sequence, generating overhanging fusion sites of 4 nt. The particular particular set of four nucleotide sequences on each end of the genetic parts types is known as "grammar", and act as sticky ends when the level 0 plasmids are cut with BsaI, allowing for specific modular cloning [84, 85]. Several repositories of genetic parts (i.e BioBricks, `https://biobricks.org/`; MoClo `http://www.addgene.org/kits/marillonnet-moclo/`) use common grammar sites for their genetic parts. For example, most DNA coding sequences (CDS) are flanked with sites "C" on their 5' and "D" on the 3' end making all CDS parts interchangeable; promoter regions are between cites "A" and "B" normally, while ribosome binding sites (RBS) are flanked with "C" and "D".

In order to generate toehold sensors in level 1 plasmids without changing the sequence from the original toehold designs from Green *et al.* [11], a slight modification on the B grammar was made, generating B'. This change was necessary in order for the toehold sequence to start right after the +1 transcribed nucleotide of the promoter T7. The final grammar used on this thesis is as follows:

- A: GGAG
- B' AGGG. Instead of TATC, which is B in Iverson *et al.* [84].
- C: AATG.
- D: AGGT
- F: CGCT

Most of the level 0 parts used in this thesis are original (Table 2.1), with a few exceptions that were either part of the CIDAR MOCLO collection purchased by our lab through Addgene [84], or previous work from our research group [87].

The empty level 1 plasmids were also made by Gibson assembly and had two features that are important for efficient cloning. First of all, they allow the constitutive expression of the amilCP chromoprotein [93]. This transcriptional unit (TU) is replaced with the TU of interest in the Golden Gate reaction, enabling visual discrimination of colonies that retained amilCP (blue colonies) with the ones that replaced the amilCP for the transcriptional unit being cloned (white colonies).

biologically inactive unique nucleotide sequences (UNSes) that facilitate accurate ordered assembly.

Also the receiver plasmids have unique nucleotide sequences (UNSes) flanking the TUs being cloned [94]. UNSes are a set of computational designed 40-bp mers of biologically inactive sequences. For instance the 12X destination vector possesses UNS1

Table 2.1: Level 0 genetic parts used in this thesis. Credits are given to whom involved in domestication and references to original report of the sequences

ID	Grammar 5'	Grammar 3'	Domesticated in	References
L0 promoter T7	A	B'	This work	[11, 88, 89]
L0 ZIKV Sensor 27	B'	C	This work	[15]
L0 ZIKV Sensor 8	B'	C	This work	[15]
L0 ZIKV Sensor 32	B'	C	This work	[15]
L0 ZIKV Trigger 8	A	D	This work	[15]
L0 ZIKV Trigger 27	A	D	This work	[15]
L0 Toehold Sensor 1	B'	C	This work	[11]
L0 Trigger 1	A	D	This work	[11]
L0 sfGFP	C	D	Previous work from our group [87]	[87, 90]
L0 deGFP	C	D	This work	[91]
L0 YFP	C	D	Obtained via Addgene	[84]
L0 YeNL	C	D	Colaboration with Tamara Matute	[92]
L0 LacZ	C	D	This work	[15]
L0 LacZα	C	D	This work	[16]
L0 Terminator T7	D	F	This work	[11, 88, 89]

and UNS2 on each side of entry site of the TU being cloned, allowing for very efficient and specific PCR reactions (for checking colonies, and amplification of TUs of interest) using few conserved primers for different DNA constructions.

Level 1 plasmids were done using the Golden Gate protocol in which 15 fmol of each plasmid containing the required level 0 parts and 7.5 fmol of the receiver level 1 plasmid are mixed in a final volume of 5 μL. A mastermix of enzymes is prepared just before use with the following volumes per reaction: 3 μl of dH2O, 1 μl of T4-ligase buffer 10X (NEB, B0202), 0.5 μl of BSA (at 1 mg/mL) 0.25 μl of T4 DNA ligase (at 400 U μL/1, NEB, M0202), and 0.25 μl of BsaI (NEB, R0535). 5 uL of the mastermix are added to the mixture of plasmids, to a final volume of 10 μL. Temperature conditions for Golden Gate consist of 30 cycles at 37 ^{o}C for 3 min, 16 ^{o}C for 4 min, followed by 5 min at 50 ^{o}C and 10 minutes at 80 ^{o}C. When the reaction is finished, 5 μL are used to transform chemically competent TOP10 cells. Next day, white colonies were screened by colony PCR using UNS primers [94], colonies that displayed the expected size of the transcriptional unit, were stored in glycerol and sequenced. A list of all the level 1 plasmids constructed and tested in this thesis is shown below in Table 2.2 which also includes one plasmid from Duo *et al.* which allows the expression of LacZ-Ω peptide [16].

Table 2.2: Level 1 Plasmids used in this thesis. Including references to original description of sequences.

ID	Cloned in	Description	References
L1-12x	This work	Empty level 1 plasmid used for receiving the transcriptional units generated by Golden Gate allowing blue/white discrimination of colonies. Incorporated TU is flanked with UNS1 and UNS2-UNSX respectively.	[93, 94]
L1-23x	This work	Empty level 1 plasmid used for receiving the transcriptional units generated by Golden Gate allowing blue/white discrimination of colonies. Incorporated TU is flanked with UNS2 and UNS3-UNSX.	[93, 94]
12x-sfGFP	This work	T7 transcriptional unit used for constitutively expression of sfGFP.	[90]
L1-12x-LacZ	This work	T7 transcriptional unit used for constitutively expression of LacZ.	[15]
12x-LacZα	This work	T7 transcriptional unit used for constitutively expression of LacZα.	[16]

Table 2.2 (continued)

ID	Cloned in	Description	References
12x-YeNL	This work	T7 transcriptional unit used for constitutively express nanolantern YeNL.	[92]
12x-S1-YFP	This work	T7 transcriptional unit encoding forward engineered synthetic toehold sensor 1 controlling the translation of YFP.	[11]
12x-S1-sfGFP	This work	T7 transcriptional unit encoding forward engineered synthetic toehold sensor 1 controlling the translation of sfGFP.	[11]
12x-S1-deGFP	This work	T7 transcriptional unit encoding forward engineered synthetic toehold sensor 1 controlling the translation of deGFP.	[11]
12x-S8-sfGFP	This work	T7 transcriptional unit encoding ZIKV toehold sensor 8 controlling the translation of sfGFP.	[15]

Table 2.2 (continued)

ID	Cloned in	Description	References
12x-S27-YeNL	This work	T7 transcriptional unit encoding ZIKV toehold sensor 27 controlling the translation of nanolantern YeNL.	[15] [92]
12x-S27-ZFL	This work	T7 transcriptional unit encoding ZIKV toehold sensor 27 controlling the translation of LacZ.	[15]
12x-S27-Zα	This work	T7 transcriptional unit encoding ZIKV toehold sensor 27 controlling the translation of LacZα.	[15]
12x-S8-ZFL	This work	T7 transcriptional unit encoding ZIKV toehold sensor 8 controlling the translation of LacZ.	[15]
12x-S8-Zα	This work	T7 transcriptional unit encoding ZIKV Toehold Sensor 8 controlling the translation of LacZα.	[15]
12x-S32-ZFL	This work	T7 transcriptional unit encoding ZIKV toehold sensor 32 controlling the translation of LacZ.	[15]
23x-T8	This work	T7 transcriptional unit of the ZIKV Trigger 8.	[15]

Table 2.2 (continued)

ID	**Cloned in**	**Description**	**References**
23x-T27	This work	T7 transcriptional unit of the ZIKV Trigger 27.	[15]
pT7-LacZω	Ma *et al.*	T7 transcriptional unit encoding LacZω peptide used for the α complementation experiments	[16]

2.2 Input DNA preparation for cell-free transcription-translation reactions

Once the sequence of the plasmid was verified, DNA input for the cell-free gene expression reactions was produced by midi-prepping (Promega, A2492) an overnight culture of 200 mL LB following the provider's instructions. After elution with ultra-pure water on 800 μL, an equal volume of binding buffer was added and mixed well before continuing the second DNA cleaning protocol using the Wizard-SV PCR cleanup system (Promega, A6754) following the instructions provided by the supplier but adding one extra centrifugation of 5 minutes to eliminate residual ethanol on the column before final elution using ultra-pure water. Plasmid DNA input was used in a set of concentrations ranging from 0.5 nM to 10 nM in cell-free reactions. Preparation of linear DNA for expression in *E. Coli*, includes PCR amplification of desired transcriptional units using UNS1F and UNSXR primers [94]. Amplicon was purified using PCR cleanup system (Promega, A6754) with the incorporation of the extra 5 minutes centrifugation after to eliminate ethanol before eluting with ultra-pure water.

2.3 Crude extract preparation

Where appropriate BL21 DE3 Star cells (Invitrogen, C601003), BL21dLacGold cells (Aligent, 230132) [50], or the CRISPR+, CRISPR- derivatives from this work were grown overnight at 37 °C on 2xYT agar plates with the appropriate antibiotics. A single colony was grown in a 5 mL 2XYT culture overnight at 37 °C. The next day, a 50 mL culture was started with a 1:1.000 dilution overnight at 37 °C. On the fourth day, a total of 2 liters of culture was started and split in 5 2-liter Erlenmeyer, at an initial OD600 0.05 in 2XYT supplemented with 40 mM phosphate dibasic (Sigma,

94046), 22 mM phosphate monobasic (Sigma, 74092) solutions. This culture was grown at 37 °C with vigorous agitation (230 rpm) until OD600 reached 1.5 (approximately 2.5 hours). Cells were then induced for 3 hours at 37°C with 100 mM IPTG before harvesting by centrifugation at 5000 g and 4°C. Pellets were washed twice with cold the S30 buffer (5 mM tris acetate pH 8.2, 14 mM magnesium acetate, 60 mM potassium glutamate, 2 mM dithiothreitol). Pellet was then weighted, and a relationship of 0.9 mL of S30B and 5 g of 0.1 mm diameter silica beads (Biospec, 11079101) were added respectively per gram of pellet obtained. 2.5 mL bead-beading tubes and cups were filled with the viscous solution formed with cells and glass beads, without generating bubbles, then sealed and beaten for 30 s using a homogenizer (MP Biomedicals, FastPrep-24 5G). To remove the beads from extracts, processed samples were placed on the top of a micro-chromatography column (Biorad, 7326204) elution end down, into an empty bead-beading tube placed into a 15 mL Falcon tube. This complex was centrifuged at 6000 g for 5 minutes at 4 °C. Properly beat extracts were clear, and two distinct layers were observed. Supernatant for all tubes was pooled and let it at 37 °C with agitation inside a 5 mL unsealed tube for one hour for runoff. After the runoff reactions were completed, the tubes were placed on ice and re centrifuged at 6000g for 10 minutes to poll down the debris generated during the run-off reaction. The supernatant from this centrifugation, the crude extract, was aliquoted, frozen, and stored at -80 °C until use. A link to the full protocol of crude extract and cell-free reaction preparation can be found here: `https://www.protocols.io/view/preparation-of-cell-free-rnapt7-reactions-kz2cx8e`.

2.4 Stability of PCR-derived linear DNA in cell-free reactions

Linear DNA fragments of 1.5 kb and 500 bp were prepared via PCR, for its use as linear DNA input and standards, respectively. Those were purified by wizard gel and PCR clean-up kit (Promega), so that a final concentration of 200 ng/μL was achieved. 600 ng of the 1.5 kb linear DNA were added to 20 μL cell-free reactions and incubated at 29 °C for 0, 30, 60 or 120 minutes and quickly submerged in liquid nitrogen to stop the reaction. Frozen samples were put on ice and immediately, a volume of binding buffer from PCR clean-up kit (Promega, A9282) was added followed with 2 μL (300 ng) of the 500 bp linear DNA that serves as an internal standard for the purification. The samples were then purified following standard supplier's protocol and eluted with 30 μL of ultra-pure water. Then 0.3 volumes of loading buffer was added to the purified samples and 10 μL were run on a 1 % agarose gel electrophoresis containing 1X SYBR-Safe for DNA visualization. 1.5 kb and 500 bp DNA fragments were placed as well in two independent lines along with another line for 1 kb Ladder (NEB, N232). Once bands were clearly identifiable, the gel was photographed on a blue LED light transilluminator with a blue-light filter. The intensity of the 1.5 kb bands in each sample was integrated using ImageJ and normalized with the intensity of the 500 bp band of the same lane.

2.5 Cell-free transcription-translation reaction using home-made cell extracts

In-house cell-free reactions' volume was composed with 45% crude lysate and 40% reaction buffer. The remaining 15% included DNA input solution, and in the case of the RNA sensing reactions, trigger RNA from *in vitro* transcription and chlorophenol red-$beta-D-galactopyranoside (CRPG, purchased from Sigma 59767 at final concentration of 0.6 mg/ml). A typical reaction is composed of 50 mM HEPES pH 8, 1.5 mM ATP and GTP, 1.4 mM CTP and UTP as triphosphate ribonucleotides, 0.2 mg/ml tRNA (Roche, 01109541001), 0.26 mM Coenzyme A, 0.33 mM NAD, 0.756 mM cAMP, 0.0068 mM folinic acid, 0.112 mg/mL spermidine, 2% PEG-8000, 3.4 mM of each of the 20 amino acids (glutamate is also added in excess in potassium and magnesium salts), as energy source it was used 12 mg/ml maltodextrin and 0.6 mg/ml sodium hexametaphosphate (Sigma, 305553) or 30 mM 3-PGA. Cell-free reactions were mounted on ice in a final volume of 10 μL in 1.5 ml Eppendorf tubes and vortexed for 30 seconds before sampling 5 μL per reaction that was loaded in V-bottom 96 well plates (Corning, CLS3957) or rounded 384 well-plates (Corning, CLS3540) for experiments of GFP and LacZ outputs respectively. All the reactions were incubated at 29 °C in the Synergy HTX plate reader (BioTek).

When LacZ-α was used as reporter, a previous expression of LacZ-Ω was taken place 2 hours before, and diluted 1:10 in S30 buffer, from which 1 μL was used to supplement 20 μL cell-free reactions.

2.6 Cost breakdown analysis

A comprehensive list of all the reagents needed for the preparation of crude extracts and energy buffers in Chile and in the UK can be found in sections 8.2 8.1 respectively. These list also indicate the individual cost of reagents and quantities purchased.

The following assumptions have been made in order to compute the cost per reaction:

- 1 liter of cell culture in the experimental conditions described yields 6g of biomass.
- 8 micro-chromatography columns are needed to process 6 grams of biomass.
- 6 grams of processed biomass yield 3.75 mL of crude extract.
- 3.75 mL of crude extract is sufficient to prepare 1667 cell-free reactions (5 μL).
- No loses in reagents during manipulation, or buffer preparation.

The purchased quantity of each reagent is used to compute the maximum reactions that could be performed with that particular reagent according with our assumptions. The limiting reagent was then identified for each method and geographic place (3-PGA energy performed in Chile, 3-PGA energy performed in the UK, maltodextrin energy source performed in Chile, maltodextrin energy source performed in the UK). The number of reactions of the limiting reagent was used to compute the cost per reaction and the contribution of each reagent to this cost. Exchange rates as per in the 8th of June 2020 were used to convert local currency to USD.

2.7 Cell-free transcription-translation reaction using NEB Purexpress

The cell-free reactions that used PURExpress consisted of: NEB Solution A (40%) and B (30%), chlorophenol red-β-D-galactopyranoside (0.6 mg/ml), and DNA constructs encoding the toehold sensors. When LacZ-α was used as reporter, a previous expression of LacZ-Ω was taken place 2 hours before, and diluted 1:10 in ultra-pure water, from which 1 μl was used to supplement 20 μL cell-free reactions. Paper-based NEB reactions were done adding 5 μL reactions on paper discs (2 mm) (Whatman, 1442-042) that were blocked in 5% BSA overnight prior to use.

2.8 Trigger RNA preparation

PCR products of Trigger 8 and Trigger 27 were made by amplifying the respectives plasmids using primers UNS2F and UNSX. Amplificons containing the T7 transcriptional units were then cleaned by PCR cleanup kit (Promega, A6754) and 250 ng of DNA were used for the *in vitro* transcription reaction using the Hi scribe T7 kit (Promega, E2040S) in 40 μL final volume and incubated at 37 °C overnight. 1 μL of DNAse I (NEB M0303S) is added to each tube and incubated one hour at 37 °C. After that, standard RNA clean up was performed according to the RNeasy MinElute kit (Qiagen, 74204) before quantification.

2.9 Lyophilization and storage of cell-free reactions

Lyophilization was performed by assembling 9 μL cell-extract and 8 μL of reaction buffer in a PCR tube. In a different tube, 3 μL of the corresponding plasmid was

transferred, when LacZ reporter was used, 1 μL of substrate CPRG (9 mg/ml) was added too. Assembled reaction tubes were closed with adhesive aluminum film and punctured with a 16-gauge needle to create a hole. The tubes were placed in FreeZone 2.5 L Triad Benchtop Freeze Dryer (Labconco), which was previously pre-chilled at -75 °C (shelf temperature) for overnight freeze-drying of two segments. The first segment was at -75 °C (shelf temperature) with a temperature collector of -80 °C for 12 hours and 0.04 mbar of pressure. The second segment was performed at -20 °C for 4 h with the same pressure. Samples were stored at room temperature in open plastic bags over 30 g of silica inside a closed Tupperware. Plasmid samples were rehydrated with 20 μL of water, then total volume was used to rehydrate the reaction buffer and then the crude extract. pT7:sfGFP was used at 6 nM final concentration while ZIKV Toehold 27 was used in 10 nM final concentration. Trigger 27 RNA was added at a final concentration of 300 nM.

2.10 Isothermal NASBA amplification

NASBA lyophilized kit was purchased from Life Sciences and was used following the provider's instructions. Briefly, negative controls and RNA dilution in the range 27 mM to 0.2 fM were prepared just before use in ultra-pure water.

Lyophilized reaction buffer containing nucleotide mix (Life Sciences, LRB-10) was reconstituted with DRB-10 diluent (Life Sciences, DRB-10) and heated at 41 °C for 5 minutes then kept at room temperature. The initial mixture consisted of 10 μL of the reconstituted reaction buffer, 1 μL RNasin (Promega, N2111), 380 mM of each DNA primer, 2 μL ultra-pure water and 1 μL RNA dilution were assembled at 4 °C and incubated at 65 °C for 2 min, followed by a 10 min annealing at 41 °C. During the annealing, the lyophilized enzyme mix, consistent of AMV RT, RNaseH, T7 RNAP, and

High molecular weight sugar mix (Life Sciences, LEM-10), was reconstituted with cold D-PDG (Life Sciences, D-PDG-10) and placed on ice. Immediately after the annealing, tubes were spin down and 5 μL of the dissolved enzyme mix was added to the reaction. The final reaction (20 μL) was incubated at 41 °C for 2 h with thermocycler with the lead at 98 °C, followed by a 10 minute inactivation of enzymes at 70 °C before opening and processing the samples. The amplified product was then cleaned up with the RNeasy Mini elute kit (Qiagen, 74204).

2.11 Selection of conserved sequences across PVY genomes

Table 2.3 shows the list of GenBank entries that were included in a multiple sequence alignment (MSA) to identify common regions across different PVY strains. MSA was performed using MUSCLE algorithm (`https://www.ebi.ac.uk/Tools/msa/muscle/`) and lead the identification of a region of interest (Figure 5.1). From this region, two sequences of different lengths called "PVY long" (Table 2.4) and "PVY short" (Table 2.5) were used for screening PVY toehold sensors.

Table 2.3: GenBank accessions of PVY complete genomes used to identify conserved sequences across different PVY strains.

GenBank ID	Strain	Length (pb)	Country of origin
AY166866	PVY-NTN	9700	Canada
MT264736	PVY-NTN	9701	Ireland
MT264732	PVY-NTN	9694	Ireland
EF026075	PVY-N	9694	USA
MK572827	PVY-N:O	9618	Colombia
JQ924285	PVY-O	9649	Brazil
EF026074	PVY-O	9666	USA
MT264738	PVY-O	9687	Ireland
AJ890349	PVY-O	9699	Poland
MT264737	PVY-Nwi	9697	Ireland

Table 2.4: "PVY long" DNA and RNA sequences of the used for the screening PVY sensors 1 to 4. Promoter T7 sequence is underlined.

Description	Sequence
DNA encoding the PVY long Direct RNA	GATCGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGGAGTT TGGGTTATGATGGATGGAGATGAACAAGTCGAATACCCACTGAAACCA ATCGTTGAGAATGCAAAACCAACACTTAGGCAAATCATGGCACATTTC TCAGATGTTGCAGAAGCGTATATAGAAATGCGCAACAAAAAGGAACCA TATATGCCACGATATGGTTTAGTTCGTAATCTGCGCGATGGAAGTTTG GCTCGCTATGCTTTTGACTTTTATGAAGTT
PVY long Direct RNA	GAGUUUGGGUUAUGAUGGAUGGAGAUGAACAAGUCGAAUACCCACUGA AACCAAUCGUUGAGAAUGCAAAACCAACACUUAGGCAAAUCAUGGCAC AUUUCUCAGAUGUUGCAGAAGCGUAUAUAGAAAUGCGCAACAAAAAGG AACCAUAUAUGCCACGAUAUGGUUUAGUUCGUAAUCUGCGCGAUGGAA GUUUGGCUCGCUAUGCUUUUGACUUUUAUGAAGUU

Table 2.5: "PVY short" DNA and RNA sequences of the used for the screening PVY sensors 5 to 8. Promoter T7 sequence is underlined.

Description	**Sequence**
DNA encoding the PVY short	GATCGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGCTTAG GCAAATCATGGCACATTTCTCAGATGTTGCAGAAGCGTATATAGAAAT
Direct RNA	GCGCAACAAAAAGGAACCATATATGCCACGATATGGTTTA
PVY short Direct RNA	CUUAGGCAAAUCAUGGCACAUUUCUCAGAUGUUGCAGAAGCGUAUAUA GAAAUGCGCAACAAAAAGGAACCAUAUAUGCCACGAUAUGGUUUA

2.12 *In-silico* toehold design

With the aim of designing novel toehold sensors, we have implemented a jupyter notebook pipeline called NupackSensors, which is based on an algorithm described by Green *et al.* [16] (`http://www.github.com/elanibal/NupackSensors`). To identify possible triggers from a target sequence, the algorithm first determines all contiguous 36 nucleotide sub-sequences. Each possible trigger defines one specific toehold sensor. After filtering out the sensors that contained stop codons, potential structural defects of each design were calculated, along with a set of thermodynamic parameters (Table 2.6). The following score function was implemented, so one score value was assigned to each of the possible designs [16]:

$$Score = -71\delta_{Sensor} - 49.1\delta_{ActiveSensor} - 22.6\delta_{BindingSite} + 54.3 \tag{2.1}$$

Here, δ_{Sensor} represents the average number of incorrectly paired nucleotides respecting the ideal toehold sensor structure (see Figure 2.1 for schematics of the ideal design structures and Table 2.5 for their dot-bracket notation) ; similarly, the parameter $\delta_{ActiveSensor}$ represents the average number of incorrectly paired nucleotides respecting the ideal activated sensor structure (Table 2.5), and $\delta_{BindingSite}$ represents the deviations from the perfect single-stranded binding site in the toehold domain, which is essential for the initial trigger binding and subsequent strand-displacement reaction. After assigning one score to each toehold sensor design, they are ranked. According to this ranking, sixteen sensors were selected for experimental screening (Table 2.8).

Table 2.6: Definition of each parameter evaluated over the PVY toehold sensors.

Parameter	Definition	Notes
ΔG_{Switch}	The minimum free energy secondary structure, of the toehold sensor.	Calculated at 29°C using the *mfe* function in NUPACK. Values are represented in Kcal/mol units.
$\Delta G_{Trigger}$	The minimum free energy secondary structure, of the trigger RNA.	Calculated at 29°C using the *mfe* function in NUPACK. Values are represented in Kcal/mol units.
ΔG_{Sensor}	The minimum free energy secondary structure, of the trigger RNA.	Calculated at 29°C using the *mfe* function in NUPACK. Values are represented in Kcal/mol units.
$Net\Delta G_{Complex}$	$Net\Delta G$ of the toehold sensor and trigger RNA interaction.	Net $\Delta G_{Complex} formation = \Delta G_{Complex} - (\Delta G_{Trigger} + \Delta G_{Switch})$.
$SS_{Trigger}$	The average probability of the nucleotides in the trigger RNA to be unpaired.	Calculated at 29°C using the *pairs* function in NUPACK.

Table 2.6 (continued)

Parameter	Definition	Notes
$\Delta G_{RBS-Linker}$	The minimum free energy secondary structure, of the sequence between the first nucleotide of the RBS until the end of the toehold sensor.	Calculated at 29°C using the *mfe* function in NUPACK, considering the toehold sensor positions 49 until 96. Values are represented in Kcal/mol units.
δ_{Sensor}	The average number of incorrectly paired nucleotides in the equilibrium respecting the ideal sensor structure.	Calculated at 29°C using the `complexdefect` function in NUPACK.
$\delta_{Complex}$	The average number of incorrectly paired nucleotides in the equilibrium respecting the ideal activated toehold-trigger complex.	Calculated at 29°C using the `complexdefect` function in NUPACK.
$SS_{ActiveSensor}$	The average probability of the nucleotides in the active toehold sensor RNA to be unpaired.	Calculated at 29°C using the pairs function in NUPACK, considering toehold sensor positions 37 until 96.

Table 2.6 (continued)

Parameter	**Definition**	**Notes**
$\delta_{BindingSite}$	The average number of incorrectly paired nucleotides in the equilibrium respecting the ideal activated trigger binding site.	Calculated at 29°C using the `complexdefect` function in NUPACK, considering toehold sensor sequence from 1 until position 26.
$\delta_{ActiveSensor}$	The average number of incorrectly paired nucleotides in the equilibrium respecting the ideal acivated toehold sensor.	Calculated at 29°C using the `complexdefect` function in NUPACK, considering the sequence of the toehold sensor positions 37 until 96.
Score	Metric for scoring toehold sensors design on the *in-silico* screening.	
ON/OFF endpoint measurement	Experimental value that measures the toehold sensor performance.	

Table 2.7: Definition of the specified secondary structures designs for toehold sensors. Secondary structures is shown in dot-bracket notation

Structure ID	**Specified secondary structure for design**
Toehold sensor	.....((((((((((...(((((.....)))))...))))))))))..... .....
Active sensor	...(((((.....))))).....
Active complex	((((((((((((((((((((((((((((((...(((((.....)))))..... .....+))))))))))))))))))))))))))))))

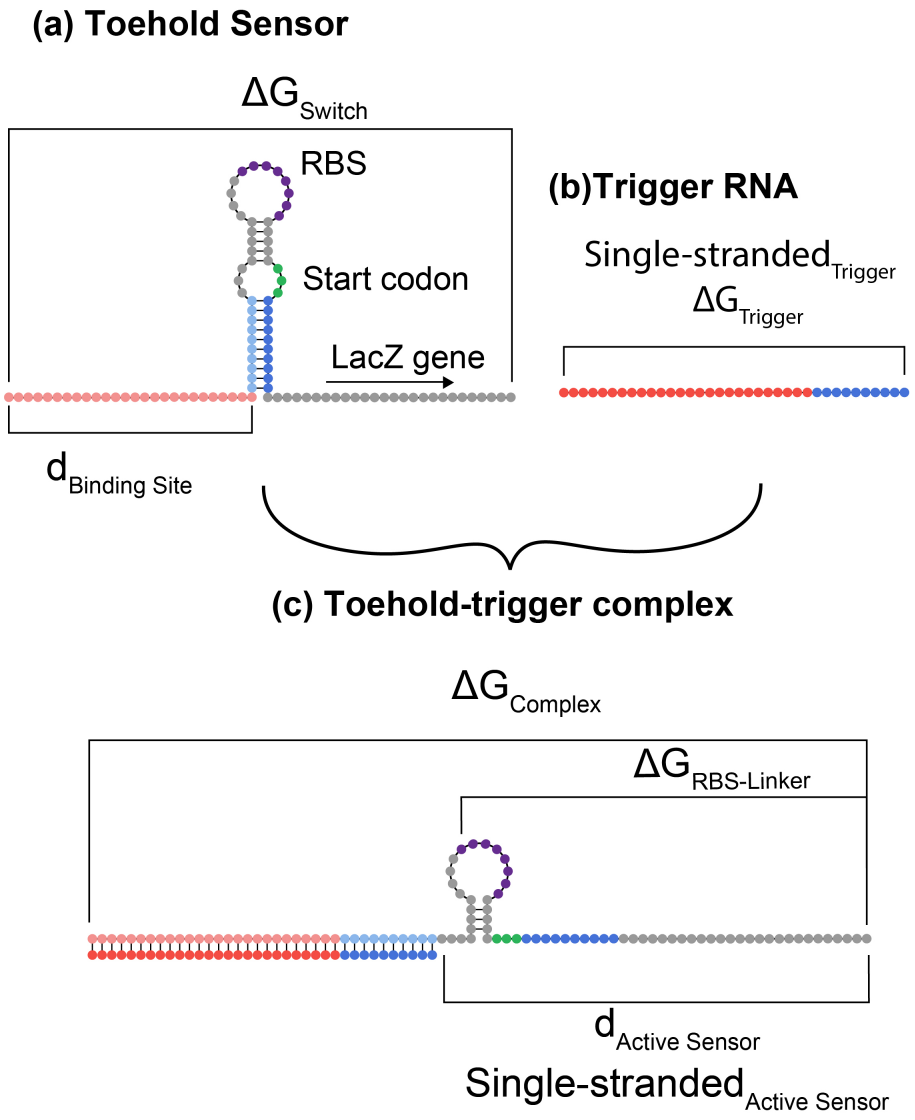

Figure 2.1: Ideal structures of (a) toehold sensor RNA, (b) trigger RNA, and the (c) toehold-trigger complex. Trigger RNA (b) is complementary to the binding site of the toehold sensor (a) and its first 10 nucleotides of the hairpin structure. Strand-displacement reactions between trigger RNA and toehold sensor causes the formation of the toehold-trigger complex (c). Each circle represents one ribonucleotide. Complementary sequences are shown in the same color but different intensities. The sequence used to define each parameter is also shown.

Table 2.8: Toehold and trigger sequences for PVY toehold sensors

Toehold sensor ID	Toehold sequence	Trigger sequence
Sensor 1	CGCAGAUUACGAACUAAACCAUAUCGUGGCA UAUAUGGACUUUAGAACAGAGGAGAUAAAGA UGAUAUAUGCCACACCUGGCGGCAGCGCAAG AAG	AUAUAUGCCACGAUA UGGUUUAGUUCGUAA UCUGCG
Sensor 2	UCAUAAAAGUCAAAAGCAUAGCGAGCCAAAC UUCCAGGACUUUAGAACAGAGGAGAUAAAGA UGUGGAAGUUUGGACCUGGCGGCAGCGCAAG AAG	UGGAAGUUUGGCUCG CUAUGCUUUUGACUU UUAUGA
Sensor 3	CGCGCAGAUUACGAACUAAACCAUAUCGUGG CAUAUGGACUUUAGAACAGAGGAGAUAAAGA UGAUAUGCCACGAACCUGGCGGCAGCGCAAG AAG	AUAUGCCACGAUAUG GUUUAGUUCGUAAUC UGCGCG
Sensor 4	UCCAUCGCGCAGAUUACGAACUAAACCAUAU CGUGGGGACUUUAGAACAGAGGAGAUAAAGA UGCCACGAUAUGGACCUGGCGGCAGCGCAAG AAG	CCACGAUAUGGUUU AGUUCGUAAUCUGCG CGAUGGA
Sensor 5	AUACGCUUCUGCAACAUCUGAGAAAUGUGCC AUGAUGGACUUUAGAACAGAGGAGAUAAAGA UGAUCAUGGCACAACCUGGCGGCAGCGCAAG AAG	AUCAUGGCACAUUUC UCAGAUGUUGCAGAA GCGUAU

Table 2.8 (continued)

Toehold sensor ID	**Toehold sequence**	**Trigger sequence**
Sensor 6	UCGUGGCAUAUAUGGUUCCUUUUUGUUGCGC AUUUCGGACUUUAGAACAGAGGAGAUAAAGA UGGAAAUGCGCAAACCUGGCGGCAGCGCAAG AAG	GAAAUGCGCAACAAA AAGGAACCAUAUAUG CCACGA
Sensor 7	UGGCAUAUAUGGUUCCUUUUUGUUGCGCAUU UCUAUGGACUUUAGAACAGAGGAGAUAAAGA UGAUAGAAAUGCGACCUGGCGGCAGCGCAAG AAG	AUAGAAAUGCGCAAC AAAAAGGAACCAUAU AUGCCA
Sensor 8	UAAACCAUAUCGUGGCAUAUAUGGUUCCUUU UUGUUGGACUUUAGAACAGAGGAGAUAAAGA UGAACAAAAAGGAACCUGGCGGCAGCGCAAG AAG	AACAAAAAGGAACCA UAUAUGCCACGAUAU GGUUUA

2.13 PCR amplification of PVY toehold sensors for rapid screening

Linear DNA encoding the PVY toehold sensors was produced by PCR amplification using Phusion high fidelity DNA polymerase (ThermoFisher, F530L). 50 uL PCR reactions were prepared according provider‘s instructions, using LacZ as template.

In order to produce the LacZ-α toehold sensors the corresponding forward primer containing the promoter T7 and the toehold was used in combination with the primer "Sensor-α-Rv. The generation of PVY toehold sensors controlling LacZ the UNSXR primer was used (Table 2.9). Amplicon was purified using the Wizard gel and PCR clean-up kit (Promega) and used to perform cell-free RNA sensing reactions.

Table 2.9: List of primers used to generate linear dsDNA used in cell-free gene expression reactions. T7 promoter is underlined

Primer Name	**Sequence**	**References and notes**
PVY-1-Fw	CGTTTCTACGGTAGCCGGGCG CTAATACGACTCACTATAGGGCGCAGAT TACGAACTAAACCATATCGTGGCATATA TGGACTTTAGAACAGAGGAGATAAAGAT GATATATGCCACACCTGGCGGCAGCGCA AGAAGAATGACCATGATTACGGATTCAC TGGCC	This work. Used for the generation of the linear PVY toehold sensor 1.
PVY-2-Fw	CGTTTCTACGGTAGCCGGGCG CTAATACGACTCACTATAGGGTCATAAA AGTCAAAAGCATAGCGAGCCAAACTTCC AGGACTTTAGAACAGAGGAGATAAAGAT GTGGAAGTTTGGACCTGGCGGCAGCGCA AGAAGAATGACCATGATTACGGATTCAC TGGCC	This work. Used for the amplification of linear PVY toehold sensor 2.
PVY-3-Fw	CGTTTCTACGGTAGCCGGGCG CTAATACGACTCACTATAGGGCGCGCAG ATTACGAACTAAACCATATCGTGGCATA TGGACTTTAGAACAGAGGAGATAAAGAT GATATGCCACGAACCTGGCGGCAGCGCA AGAAGAATGACCATGATTACGGATTCAC TGGCC	This work. Used for the amplification of linear PVY toehold sensor 3.

Table 2.9 (continued)

Primer Name	**Sequence**	**References and notes**
PVY-4-Fw	CGTTTCTACGGTAGCCGGGCG CTAATACGACTCACTATAGGGTCCATCG CGCAGATTACGAACTAAACCATATCGTG GGGACTTTAGAACAGAGGAGATAAAGAT GCCACGATATGGACCTGGCGGCAGCGCA AGAAGAATGACCATGATTACGGATTCAC TGGCC	This work. Used for the generation of the linear PVY toehold sensor 4.
PVY-5-Fw	CGTTTCTACGGTAGCCGGGCG CTAATACGACTCACTATAGGGTGGCATA TATGGTTCCTTTTTGTTGCGCATTTCTA TGGACTTTAGAACAGAGGAGATAAAGAT GATAGAAATGCGACCTGGCGGCAGCGCA AGAAGAATGACCATGATTACGGATTCAC TGGCC	This work. Used for the generation of the linear PVY toehold sensor 5.
PVY-6-Fw	CGTTTCTACGGTAGCCGGGCG CTAATACGACTCACTATAGGGTCGTGGC ATATATGGTTCCTTTTTGTTGCGCATTT CGGACTTTAGAACAGAGGAGATAAAGAT GGAAATGCGCAAACCTGGCGGCAGCGCA AGAAGAATGACCATGATTACGGATTCAC TGGCC	This work. Used for the generation of the linear PVY toehold sensor 6.

Table 2.9 (continued)

Primer Name	**Sequence**	**References and notes**
PVY-7-Fw	CGTTTCTACGGTAGCCGGGCG CTAATACGACTCACTATAGGGCAACAAA AAGGAACCATATATGCCACGATATGGTT TGGACTTTAGAACAGAGGAGATAAAGAT GAAACCATATCGACCTGGCGGCAGCGCA AGAAGAATGACCATGATTACGGATTCAC TGGCC	This work. Used for the generation of the linear PVY toehold sensor 7.
PVY-8-Fw	CGTTTCTACGGTAGCCGGGCG CTAATACGACTCACTATAGGGTAAACCA TATCGTGGCATATATGGTTCCTTTTTGT TGGACTTTAGAACAGAGGAGATAAAGAT GAACAAAAAGGAACCTGGCGGCAGCGCA AGAAGAATGACCATGATTACGGATTCAC TGGCC	This work. Used for the generation of the linear PVY toehold sensor 8.

3. *Results from aim 1: To Implement and optimize an in-house low-cost cell-free gene expression system*

In this aim, we addressed the local production of cell-free gene expression systems that were simple to implement in Chile at a low cost. To pursue this goal we have combined advances from different published methods: the use of S12 crude extracts, which does not require ultra-centrifugation nor dialysis [65]; a highly concentrated and stable amino acid mixture [95]; and cost-effective energy source based on maltodextrin and polyphosphates that do not require cold-chain transportation as other energy sources (e.g. 3-PGA) [96]. In order to produce the cell extract, the BL21 DE3 STAR *E. coli* strain was chosen because it allows the expression of the T7 RNA polymerase upon induction with IPTG, and harbours a mutation in the RNAseE gene that enhances shelf-stability of RNAs up to two folds *in-vivo* [97]. Using this method we have generated 4 crude extracts (named Batch A, B, C and D), that were further used for characterization and optimization.

3.1 Optimization of magnesium concentration

Several authors have indicated that Mg^{2+} concentration is a key factor for the performance of cell-free gene expression reactions [32, 49, 98]. Mg^{2+} is an essential cofactor for transcription [99], and also important for maintaining the structure of ribosomes [53, 100, 101].

In cell free, Mg^{2+} availability could be modified by precipitation with pyrophosphates (derived from transcription), and inorganic phosphates (derived from the metabolism of maltodextrin and HMP). Early *in vitro* studies of ribosomes have shown that 10 mM Mg^{2+} is an optimal concentration for the poly(U)-dependent poly(Phe) synthesis reaction [102]. While Sun *et al.* showed an increase of two folds on GFP endpoint yields when optimizing Mg^{2+} concentrations in cell-free reactions, which optimal levels were found at about 8.6 nM [49].

The four extracts prepared in this work (Batches A-D) contained 7 mM Mg^{2+} derived only from the cell lysis buffer. We studied the effects of supplementing them with different concentrations of Mg^{2+} on the endpoint sfGFP expression. 15 nM of plasmidic DNA that allows constitutive T7 polymerase-dependant transcription of sfGFP was mixed with crude extracts (Batches A-D) supplemented with maltodextrin and different concentrations of magnesium glutamate. Poor sfGFP signal was detected at 7 nM and above 11.5 nM Mg^{2+}, while an optimal range between 8.5 nM and 10 nM was found for all the extracts (Figure 3.1), which is consistent with previous reports [49].

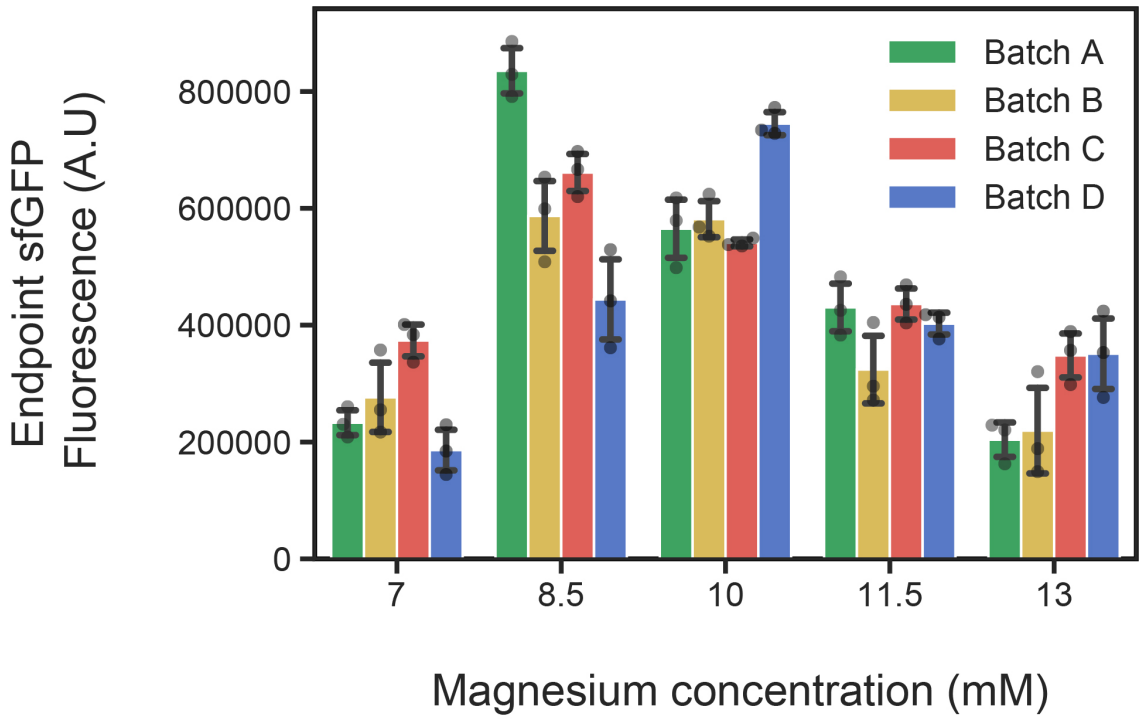

Figure 3.1: Effect of magnesium concentration on sfGFP expression. Endpoint sfGFP fluorescence measurements in cell-free reactions performed in four different crude extracts (Batch A, B, C, D) supplemented with maltodextrin energy buffers containing a range of magnesium concentrations. Optimal levels for magnesium were found between 8.5-10 mM for all extracts. Data points are shown in light grey circles. Error bars represent standard deviations of three independent trials.

3.2 Comparison of low and high-cost energy sources for cell free sfGFP expression

We prepared the 3-phosphoglyceric acid (3-PGA) energy supply according to [49] and compared the performance of sfGFP expression on Batches A-D using 3-PGA or low-cost maltodextrin energy solutions at the optimal magnesium concentration found previously. The endpoint measurements were significantly higher for the cost-effective maltodextrin-based energy source in the four batches tested (Figure 3.2 and 3.3). Therefore, cell-free preparations based on maltodextrin were used for subsequent tests and comparisons.

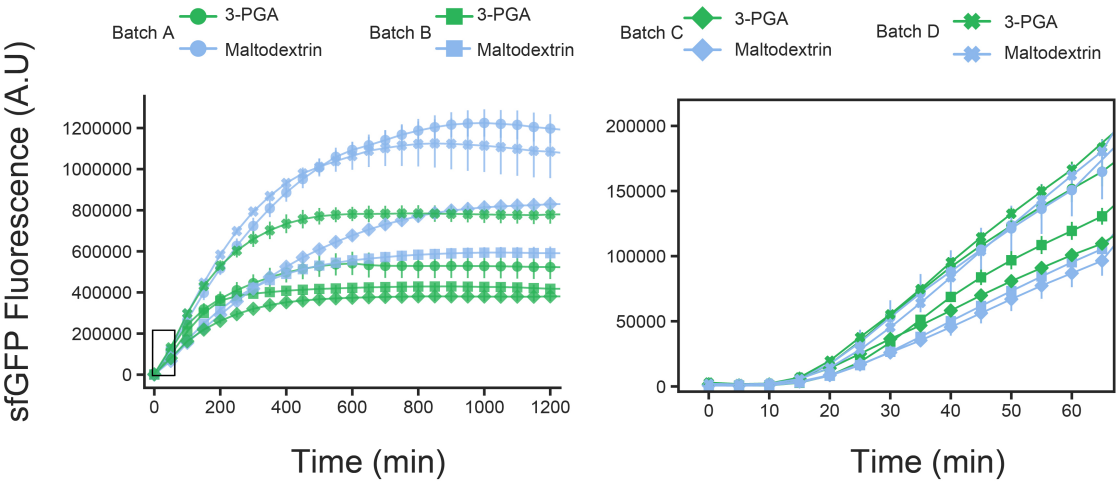

Figure 3.2: Constitutive sfGFP production on batches A-D supplemented with maltodextrin (light blue) or 3-PGA (green) energy buffers. Left: Dynamic of the fluorescence intensity on the constitutive sfGFP expression reactions. Right: initial increment on sfGFP fluorescence intensity during the first hour.

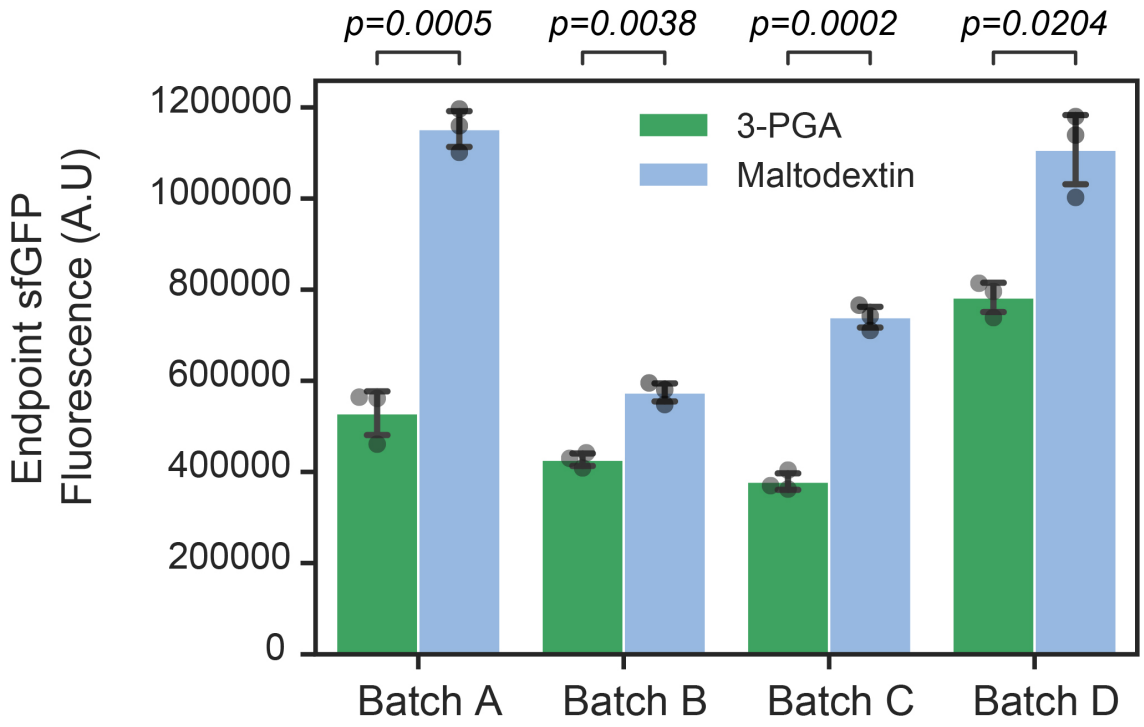

Figure 3.3: Endpoint sfGFP fluorescence measured in four different cell extracts (Batch A, B, C, D) fuelled with either maltodextrin plus polyphosphates (light blue) or 3-PGA (green) as the energy source. Data points are shown in light grey circles; error bars represent standard deviations of three independent trials. Independent sample t-test with Bonferroni correction for multiple comparisons was done to compare sfGFP outputs between the two energy sources within the same batch, p-values shown above each comparison.

3.3 Comparison of in-house prepared methods with commercial cell free systems for sfGFP expression

In order to assess the performance of in-house prepared cell free reactions, we compared the expression of sfGFP in our system to that of commercial alternatives (PURExpress, from New England Biolabs; S30 T7 High-Yield Protein Expression System, from Promega). We measured sfGFP constitutive expression in a plate reader to visualize the temporal dynamics. Despite the variability observed in the rates and endpoint measurements of the in-house cell-free reactions, all the batches reached higher endpoint fluorescent values than the commercial alternatives (Figure 3.4).

Although absolute concentrations of sfGFP production were not available, the results from fluorescent measurements are consistent with the yields reported in the litarature. For example PURExpress's manual claims a range of yields of 10–200 μg/ml, while S30 T7 High-Yield Protein Expression System claims yields up to 500 μg/ml of protein. On the other hand, Caschera *et al.* reported yields up to 1.65 mg/ml using the maltodextrin energy system [96] while Shin *et al.* reported 0.65 mg/ml using 3-PGA as energy source [32].

These results demonstrated that our low-cost and in-house prepared system for cell free reactions is comparable in performance to the commercial kits, being an appropriate system for the study of CFTS in low-cost reactions in the next chapters.

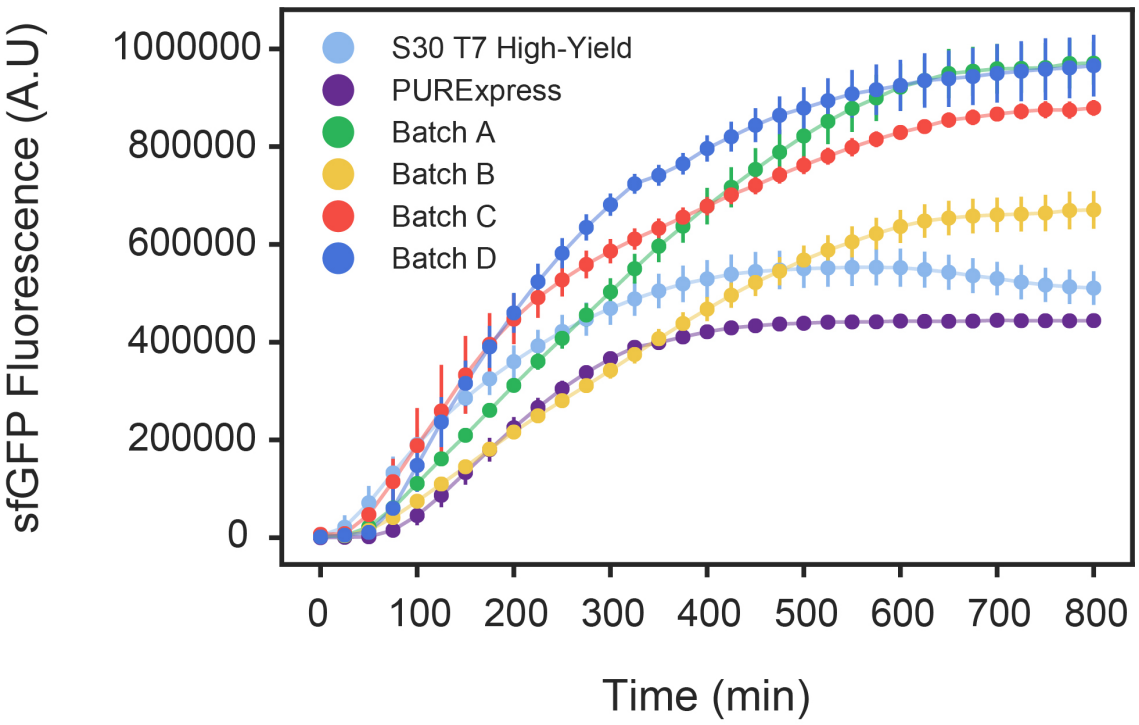

Figure 3.4: Comparison of in-house prepared cell free reactions to NEB PURExpress and Promega S30 T7 High Yield commercial kits. We tested the four optimized batches (Batch A, B, C, D) using maltodextrin plus polyphosphates as the energy source. Dots are centered at the arithmetic mean at each time point. Error bars represent the standard deviations in three independent trials

3.4 Stability of lyophilized constitutive gene expression reactions

Next, we assessed the performance of the in-house prepared cell free system upon lyophilization. Lyophilization permits the preservation and transport of cell-free gene expression reactions at room temperature, which is critical for field deployment outside laboratory conditions where CFTS can be used upon rehydration. In this work, we have measured the stability of the freeze-dried cell-free reactions based on maltodextrin energy solution for up to 90 days. We observed more than 60% recovery of the endpoint sfGFP expression 7 days after lyophilization, which decreased to 17% after 90 days at room temperature (Figure 3.5). These results demonstrated that, although there is room for optimization, the low-cost system is able to withstand preservation at room temperature upon lyophilization.

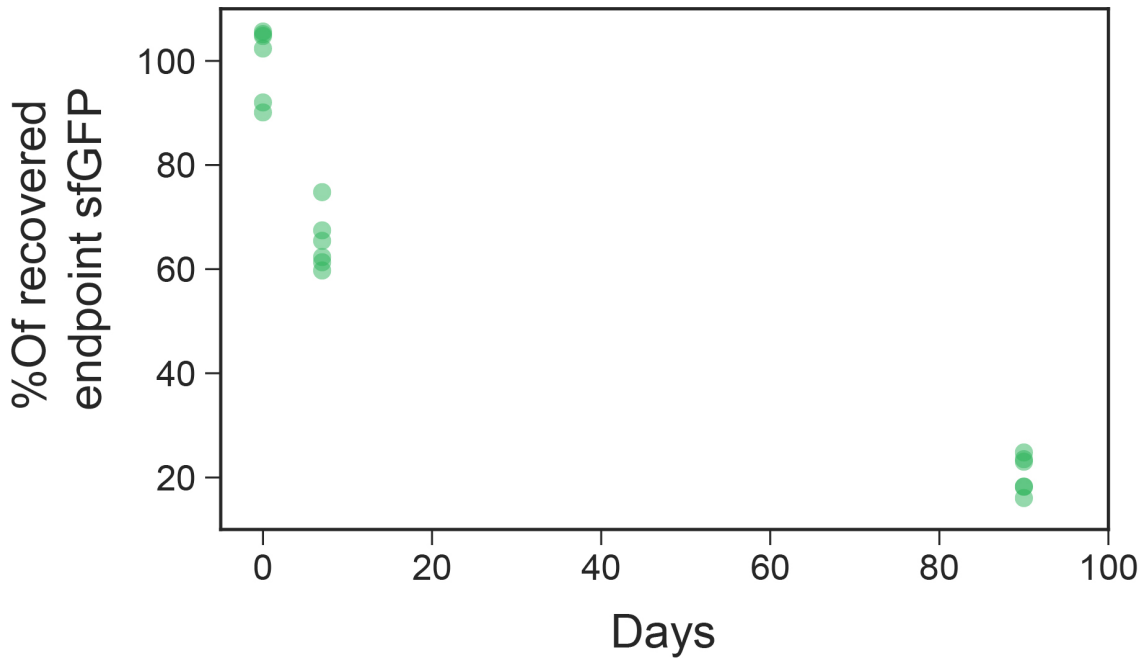

Figure 3.5: Shelf stability of the lyophilized cell-free reactions over three months. Reactions were lyophilized and stored in a pot containing desiccants for up to 90 days at room temperature. Endpoint sfGFP production was compared to a fresh sample at 7 and 90 days after lyophilization, which recovered 60% and 17% of the sfGFP production on average.

3.5 CRISPRi optimization of linear dsDNA stability

Next, we sought to explore the optimization of extracts for their use with linear DNA as input for the cell free reactions. The use of linear DNA for cell-free gene expression offers several advantages compared to plasmid DNA that has to be extracted and concentrated from bacterial cultures [103, 47, 104, 105, 106, 107]. Linear DNA can be produced by PCR amplification, enabling the rapid production of DNA designs with different primers and templates. The use of linear dsDNA in cell-free extracts, however, is affected by DNA degradation primarily due to the action of the Exonuclease V, which is encoded in the RecBCD operon and participates in DNA repair [103, 106, 108]; and Endonuclease I (encoded by endA), which is the dominant source of endonuclease activity in *E.coli* [109]. Previous studies have increased the efficiency of cell-free protein synthesis from PCR products by several methods including the supplementation with GamS, a viral protein that inhibits Exonuclease V [104, 110]; the incorporation of protective modified nucleotides downstream of PCR amplification [111]; the incorporation of competitive DNA strands containing χ-sites [112]; supplementation of dsDNA binding proteins [35]; depletion of Exonuclease V in the crude extracts [105]; genetic replacement of RecBCD with the lambda phage red recombinase [107]; and generation of strains lacking endA [107, 113].

Here, we explored a different approach, based on CRISPR interference (CRISPRi). CRISPRi is a genetic technique that allows for sequence-specific and reversible repression of gene expression in bacteria at the transcription level by blocking either transcriptional initiation or elongation [114, 115]. In order to increase the stability of PCR-derived linear DNA in our extracts, a CRISPRi-based strategy was implemented

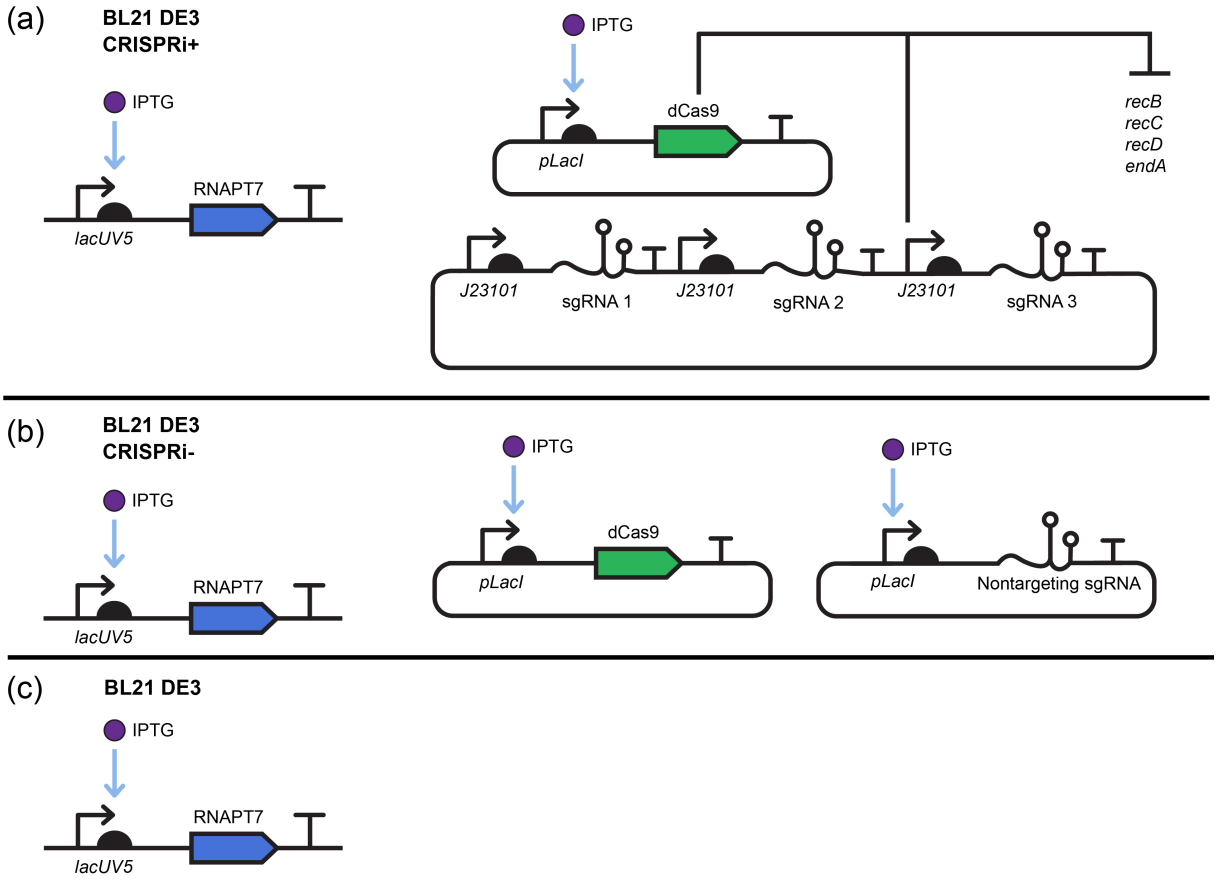

Figure 3.6: Schematic representation of CRISPRi genetic circuits used for nuclease silencing. a: In the CRISPRi+ strain, IPTG induces the expression of RNAPT7 and dCas9, and a complementary plasmid enables the constitutive expression of three sgRNAs that target recB, recC, recD, and endA genes. b: In the control CRISPRi- strain, IPTG induces the expression of dCas9 in addition to the RNAPT7, a complementary plasmid supports the expression of the "non-targeting" sgRNA that does not have any target in the genome. c: In the basal genotype strain, IPTG only induces the production of RNAPT7 before harvesting the cells.

to control (before harvesting) the genes that encode for Exonuclease I and V. This work was part of the undergraduate thesis dissertation of Juan Manuel Puig, under my supervision (Biochemistry thesis 2020).

The genetic devices utilized to pursue this goal were the following:

- A single plasmid allowing the *in-vivo* transcription of three sgRNAs targeting the 5' ends of recB, recC:recD, and end A (recC and recD are transcribed from the same promoter at the 5' end of recC [116]).
- A plasmid for IPTG-inducible dCas9 expression *in-vivo*.
- A control plasmid harbouring an sgRNA that does not have a known target in the *E.coli* genome (Nontargeting sgRNA) previously used by Nuñez *et al.* [117].

Accordingly, four strains were prepared: CRISPRi+ (BL21 DE3 STAR), CRISPRi- (BL21 DE3 STAR) and BL21 DE3 STAR (Figure 3.6), and the stability of DNA was quantified on a lapse of 2 hours (Figure 3.7 and Figure 3.8).

The following exponential degradation model was fitted to the data:

$$D(t) = D_0.e^{-K_{deg}t}, \tag{3.1}$$

where D_0 is the initial concentration of DNA, $D(t)$ is the remaining concentration of DNA ([DNA]) at a given time t, and K_{deg} is the degradation rate of the DNA. The exponential degradation model assumes that the loss of DNA is simply proportional to its current concentration and that the lower plateau if [DNA] is equal to zero:

$$\frac{d[DNA]}{dt} = K_{deg}.[DNA] \tag{3.2}$$

The ordinary least squares approach was used for parameter estimation. Model screening indicated no statistical evidence of the parameters D_0 being different across

the data sets evaluated (F-test for extra sum of squares, $p = 0.9285$), finding the best fit value at $D_0 = 0.989$. Statistical analysis of the K_{deg} parameters showed significant differences between cell extract genotypes (F-test for extra sum of squares, $p < 0.001$). Therefore, the CRISPRi strategy had a positive effect on DNA stability, due to a significant decrease in the degradation rates of the CRISPRi-optimized extracts (K_{deg}= 0.008/min), with respect to the two controls (K_{deg}= 0.022/min for the CRISPRi-negative control, and K_{deg}= 0.040/min for the BL21-derived crude extract).

Protein production from PCR-derived linear dsDNA was also measured. The CRISPRi optimized extracts significantly increased protein production capacity from PCR products, leading to a two-fold increase in the sfGFP endpoint signal (Figure 3.9) respcting the BL21 control. These results demonstrated that our system is suitable for linear, PCR-amplified DNAs, which is required for rapid prototyping of novel toehold sensors in following chapters.

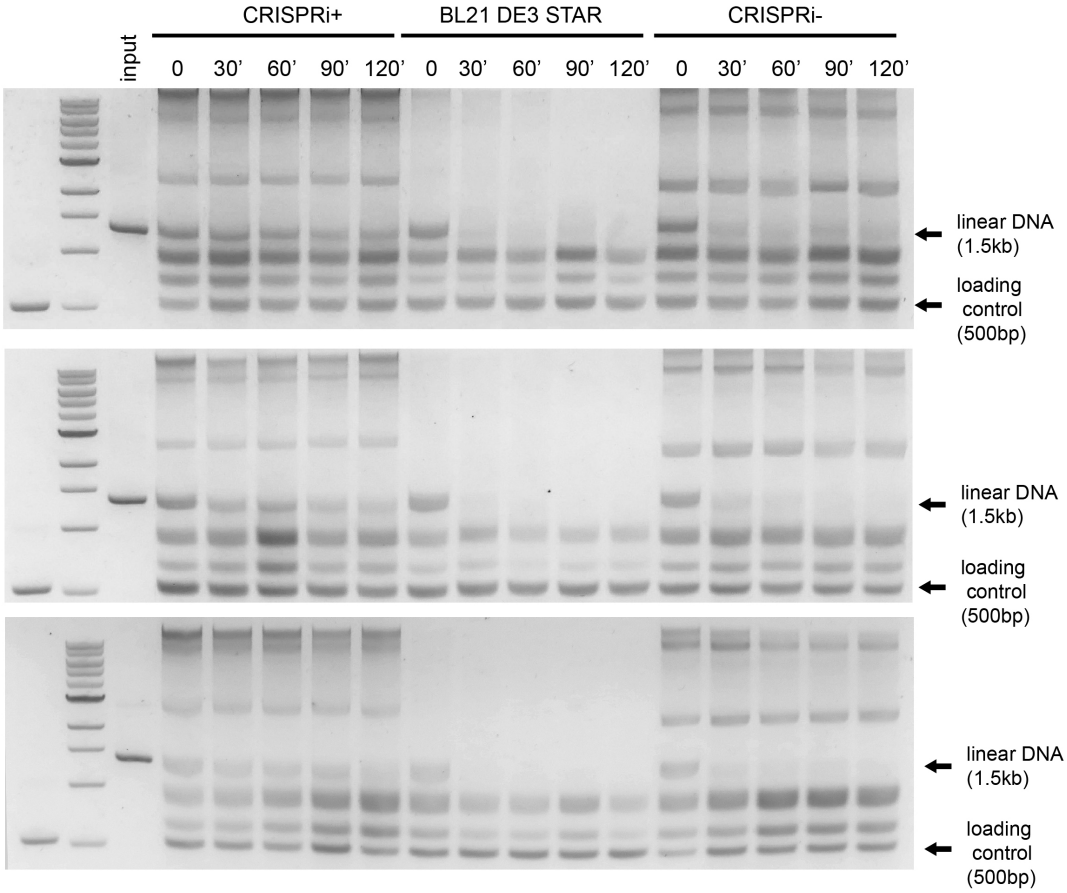

Figure 3.7: Increased stability of PCR-derived linear DNA in CRISPRi optimized cell-free extracts. Three independent experiments measuring the stability of a linear product of 1.5 kb that was incubated in cell-free reactions during a range of 120 minutes. Agarose gels were used for resolving the linear DNA product of interest and integration of its signal, 500pb was used as a loading control for quantification. Experiments were performed in three different genotypes: CRISPR+, CRISPR- and BL21 DE3 STAR.

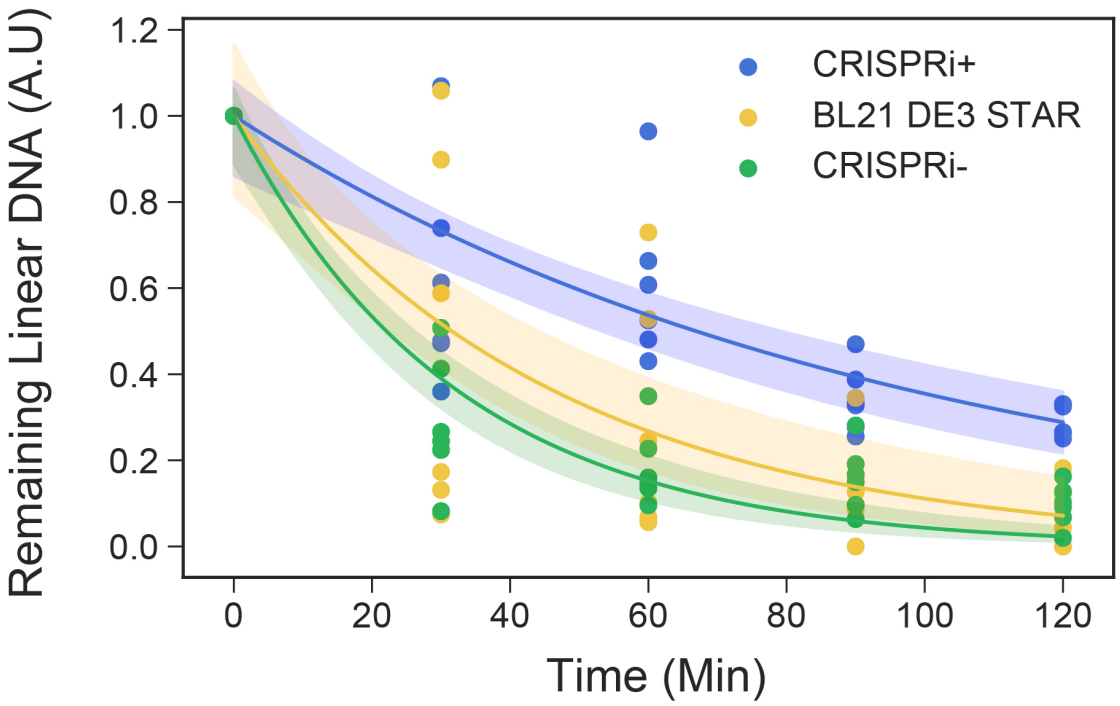

Figure 3.8: Increased stability of PCR-derived linear DNA in CRISPRi optimized cell-free extracts. Exponential DNA degradation models for each genotype evaluated. The colored areas indicate 95% confidence intervals for the model, which was approximated by sampling values of D_0 and K_{deg} from a multivariate Gaussian distribution with mean and covariance identified from inferences performed per genotype. Degradation rate K_{deg}, decreased significantly in the CRISPRi-optimized extracts (K_{deg} = 0.010/min; R^2=0.73) with respect to the controls (K_{deg}= 0.021/min; R^2=0.70 for the CRISPRi-negative control, and K_{deg}= 0.031/min; R^2=0.90 for the BL21-derived crude extract). F-test for extra sum of squares, p <0.001.

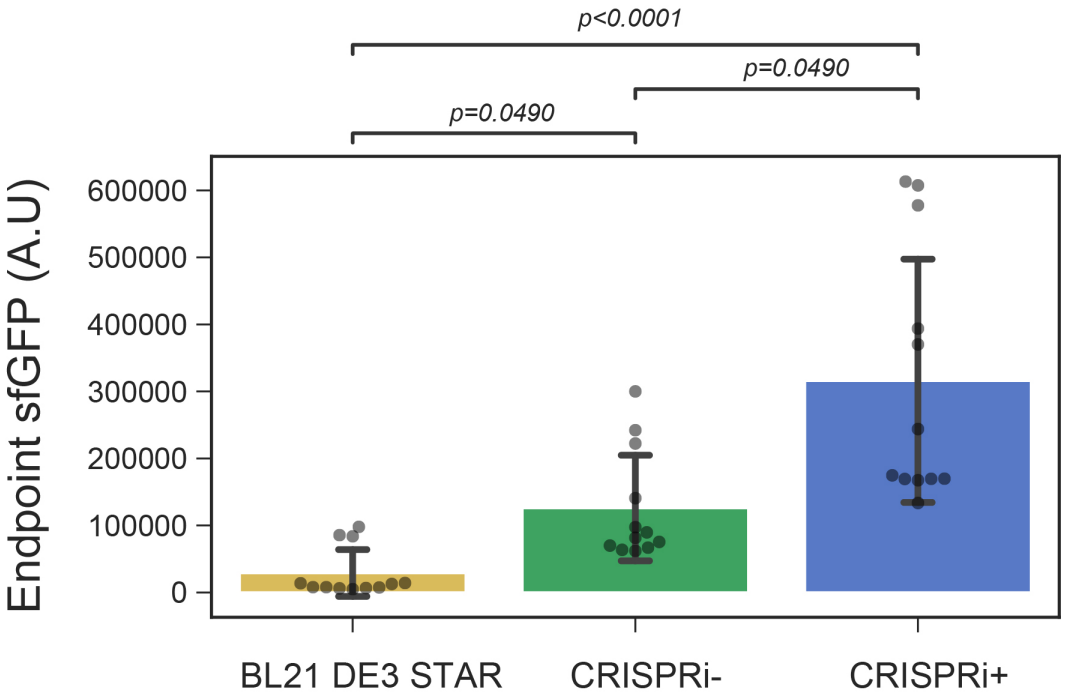

Figure 3.9: Increased endpoint sfGFP production from constitutive expression of linear DNA in CRISPRi optimized cell-free extracts. End-point sfGFP expression from PCR products in crude extracts derived from CRISPRi-optimized (CRISPR+), negative control for CRISPRi (CRISPR-), and the starting genotype (BL21 DE3 STAR) as control. Kruskal-Wallis test was performed followed by Dunn's multiple comparisons obtaining significant differences in all groups.

3.6 Cost breakdown analysis

A cost-breakdown analysis was performed (see section 2.6) comparing reagent prices available to the collaborating institutions in Chile and the UK (Sections 8.2 and 8.1 respectively).

We estimated the cost per reaction being 0.069 USD for the maltodextrin-based preparation produced in Chile, versus 7.76 USD of the commercial alternative PURExpress bought in Chile, representing a reduction of two orders of magnitude in costs. Producing reactions based on maltodextrin was also more than 20% cheaper than for the 3-PGA alternative. Notably, the home-made preparations were about two times more costly to produce in Chile than in the UK (Table 3.1), highlighting the barriers for developing low-cost alternatives in the global south.

Regarding the composition of the costs in individual reactions, the most notable difference appears in the energy source employed. 3-PGA energy solution contributes 23% of the cost per reaction, regardless of the geographic location. The use of maltodextrin decreases the contribution of the energy source to the total cost of the reaction to bellow 1% (Figure 3.10).

Regarding the costs of the crude extract preparation, two factors represent the major contribution: IPTG and the micro-chromatography columns. IPTG is used for induction of the cell culture in order to produce RNA polymerase T7 (and trigger the CRISPRi system in the case of the linear dsDNA enhanced stability extracts) before harvesting the cells; while the micro-chromatography columns are used after the bead-beating lysis step in order to separate the beads from the crude extract.The most expensive item per reaction are the ribo-NTPs used for providing substrate for efficient transcription, and the Coenzyme A (Figure 3.10).

Table 3.1: Cost breakdown analysis of cell-free reactions performed in Chile and UK (Cambridge University). All the reagents needed for the production of maltodextrin (MDX) or 3-PGA methods, along with the reagents needed for extract production were used to calculate a cost per reaction (considering 5 μl reaction volume). Exchange rates as the 8th of June 2020 were used to convert local currency to USD. This value was used to compare cost reductions with respect to the commercial alternative PURExpress (also considering 5 μl reaction volume) as well as for cost differences between places and methods.

Place	**Method**	**Cost in US dollars**	**% cost reduction from the value above**	**Times the costs in UK**	**Times the costs in Chile**
Chile, PUC	PURExpress	7.76		1.2	NA
	3-PGA	0.091	98.8	2.03	NA
	MDX	0.069	24.2	2	NA
UK, UniCam	PURExpress	6.5		NA	0.83
	3-PGA	0.045	99	NA	0.49
	MDX	0.036	23	NA	0.5

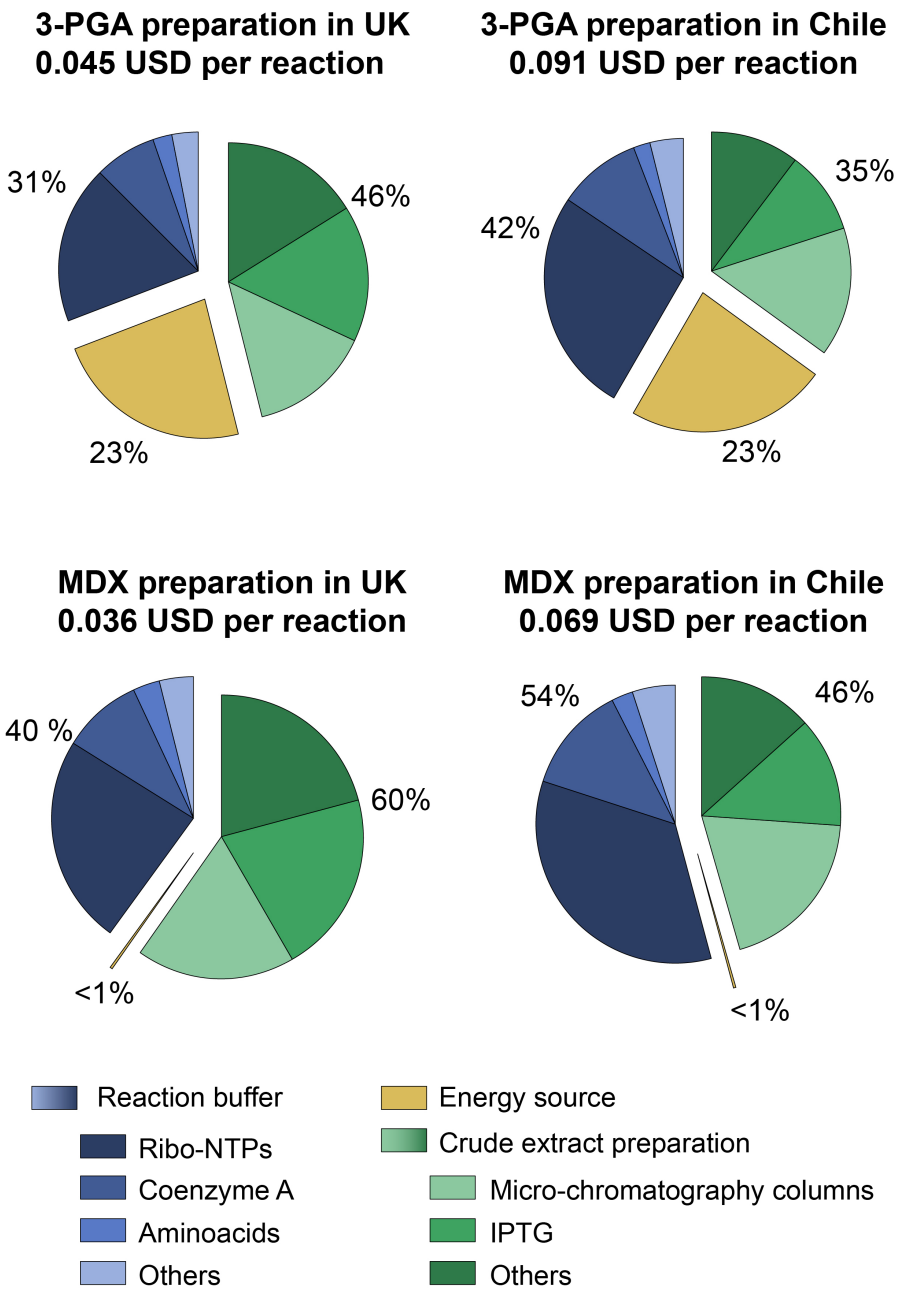

Figure 3.10: Composition of the costs of cell-free reactions prepared in Chile and in the UK using 3-PGA or Maltodextrin (MDX) preparation protocols. Fraction of the costs representing energy source are shown in yellow, crude extract preparation reagents in green and reaction buffer in blue. We assumed 1 liter of culture yields 6 g of biomass, needing 8 micro-chromatography columns to be processed, representing 3.75 ml of crude extract per liter of culture.

4. *Results from aim 2: Benchmarking RNA toehold sensing capacity using low-cost cell-free preparation methods*

Commercial cell-free gene expression reactions have been used for the implementation of RNA sensors with the possibility of developing point-of-care diagnostics [14, 18, 16]. However, PURExpress -the commercial cell-free gene expression system used for the implementation of the CFTS- is prohibitively expensive and have cold chain shipping constrains.

In Chapter3, we have shown the establishment and validation of an in-house and low-cost cell-free preparation method that produced similar yields of sfGFP expression (Figure 3.4) by a fraction of the cost of the commercial alternatives. In particular we show a decrease of two orders of magnitude in costs respecting PURExpress (Figure 3.10, Table 3.1). In this chapter, we sought to test CFTS in the low-cost cell extracts to replace the expensive PURE systems.

4.1 Selecting reporter genes for cell-free toehold sensors

We tested the performance of fluorescent and enzymatic reporters in the low cost cell free system introduced in Chapter 3. In order to aid the testing of different combinations of toehold sensors and reporters, a Golden Gate strategy for modular molecular cloning was implemented (see Section 2.1). We generated level 0 parts (promoter T7, toehold sensors, ribosome binding sites, reporter genes, and terminator T7) as well as level 1 transcriptional units (Figure 4.1a). First, we used constitutively-expressed super folder GFP (sfGFP) to test how fluorescent output signal changed in response to plasmid concentration ranging from 7 nM to 7 pM. The fluorescence intensity of sfGFP was found to display a strong signal at 7 nM of input plasmid concentration. The signal decreased significantly at lower concentrations than 700 pM, being indistinguishable from the background fluorescence at concentrations below 7 pM of DNA input (Figure 4.1b).

Next, we tested constitutively-expressed enzymatic reporters, since they can amplify the signal produced per unit of functionally expressed protein. Luciferases [42, 118, 119] are among the most widely used in cell-free systems. Luciferases are enzymes that produce light when they oxidize their substrate. The latest advances in engineering luciferase reporters reduced the protein size while increasing the luminescence intensity to more than 150 times than the original firefly and renilla luciferases [120]. In this work, we cloned one of the latest versions of nano-lantern (YeNL), a small luciferase that used fumazine as the substrate for the ATP-independent light emission reaction (Figure 4.2a), and prepared a constitutive expression plasmid to test its signal in response to DNA input concentration in cell-free reactions (Figure 4.2b). The

luminescent signal was distinguishable from the background in the first 20 minutes of reaction and achieved a quick peak during the first 40-60 minutes. Unlike the sfGFP reporter, whose signal was only different from the background in concentrations above 7 pM, a luminescent signal was observed for the complete range of input DNA tested (7 nM to 0.7 pM) (Figure 4.2c).

Next, we tested enzymatic reporters based on the β-galactosidase enzyme, which are widely used for cell-free sensors, including CFTS [16, 18, 44, 15]. LacZ encodes for a β-galactosidase enzyme that can cleave the yellow substrate chlorophenol red-β-D-galactopyranoside generating chlorophenol red as a product (Figure 4.3a). This property has been used for the development of colorimetric tests used in CFTS [15, 16]. LacZ enzyme can be split into two peptides named alpha (α, Figure 4.3b) and omega (ω, Figure 4.3c). The lacZ α-peptide consists of the first 59 amino acids from the N terminus of lacZ full-length, while the ω-peptide (lacZω) comprises the remaining 970 lacZ residues. The complete lacZ must form a tetramer before it becomes catalytically active (Figure 4.3b,c). In this thesis, LacZ full-length and LacZα genes were cloned into plasmids for constitutive expression. lacZ α-reactions were supplemented with the ω-peptide previously expressed in a cell-free reaction. Similar results were obtained using LacZ full-length (Figure 4.4a) or lacZ α (Figure 4.4b) as reporters, observing enzymatic activity using concentrations as low as 0.7 pM of input plasmid concentration (Figure 4.4a).

Once the concentration and signal of each reporter was characterized from constitutive expression in cell-free reactions, we tested them as reporters for toehold sensors. To this end, we first used toehold sensor designs from the literature [11, 15]. The rationale behind this approach was to use pairs of toehold trigger-switches already shown to work (in commercial cell-free systems). The first experiments included Sensor 1, a forward engineered toehold sensor developed previously [11]. We have domesticated

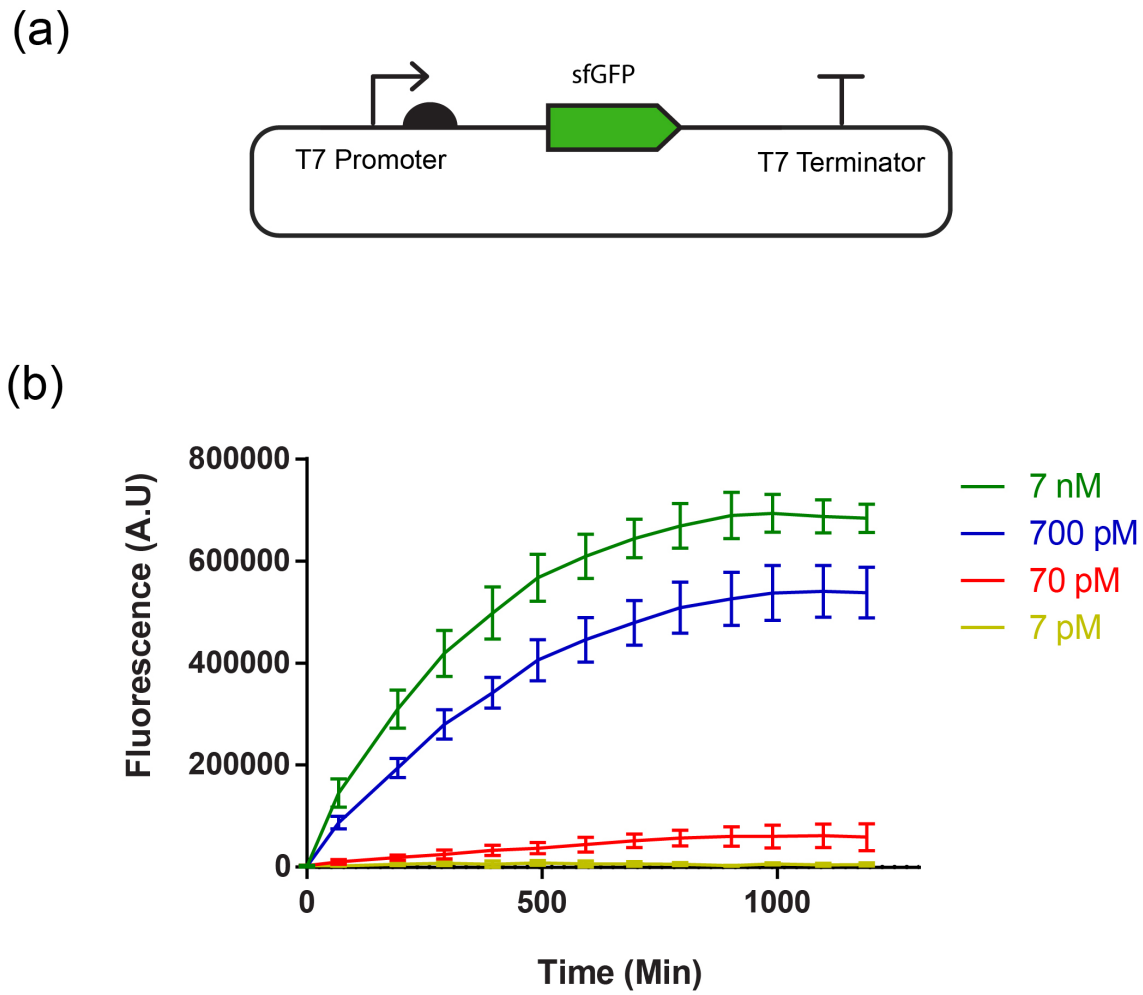

Figure 4.1: Effect of different input DNA concentration to the fluorescence intensity response. a: Scheme of the plasmid input DNA that allows constitutive sfGFP expression in cell-free reactions. b: Fluorescence intensity response to different input DNA concentrations in the range 7 nM to 7 pM.

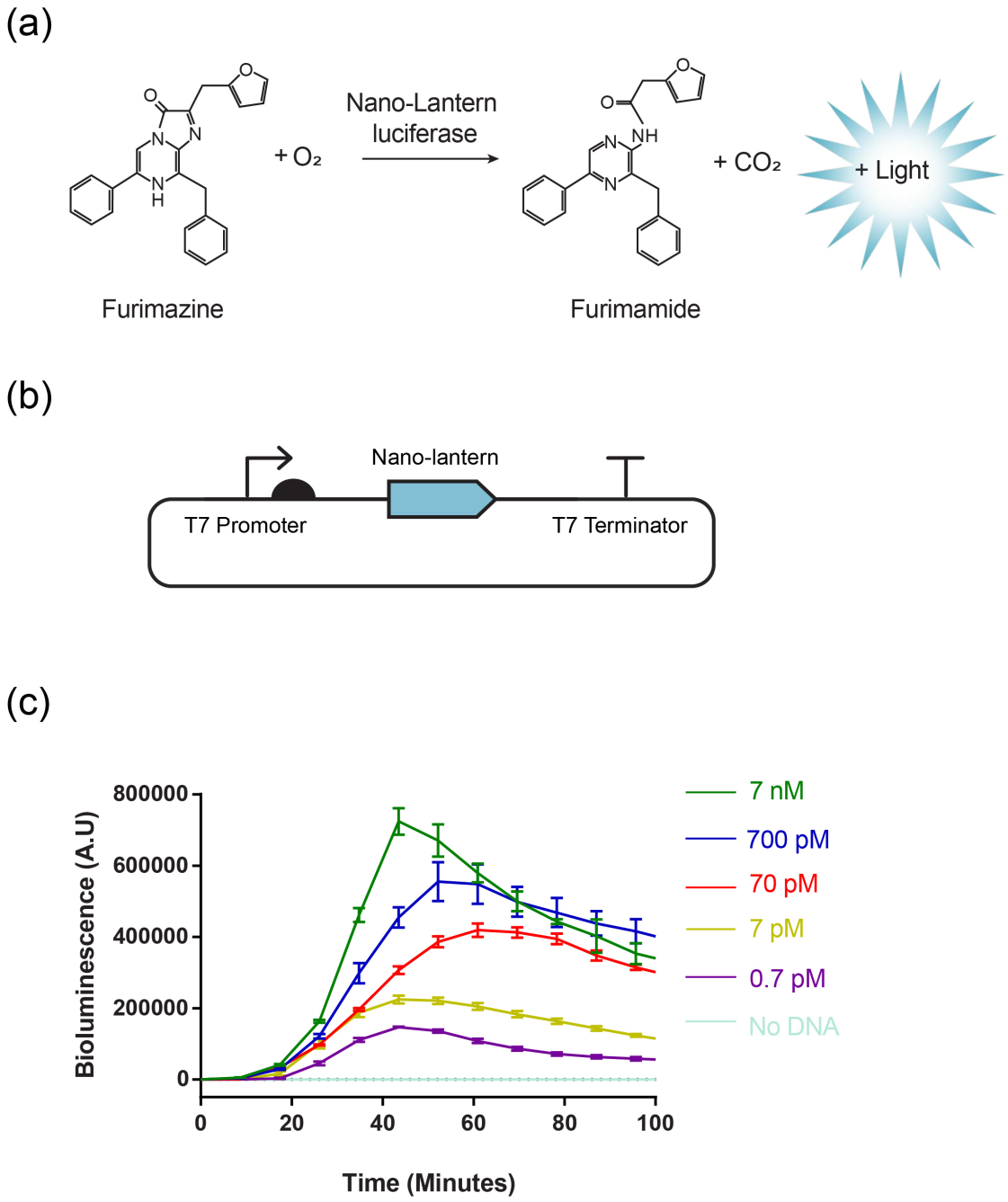

Figure 4.2: Effect of different input DNA concentration on the luminescence response. a: Scheme of the chemical reaction catalyzed by the nano-lantern luciferase to produce light and furimamide using furimazine as substrate. b: Scheme of the plasmid input DNA that allows constitutive nano-lantern expression in cell-free reactions. This plasmid was used to test input vs. signal relationship. c: luminescence response to different input DNA concentrations in the range 7 nM to 0.7 pM.

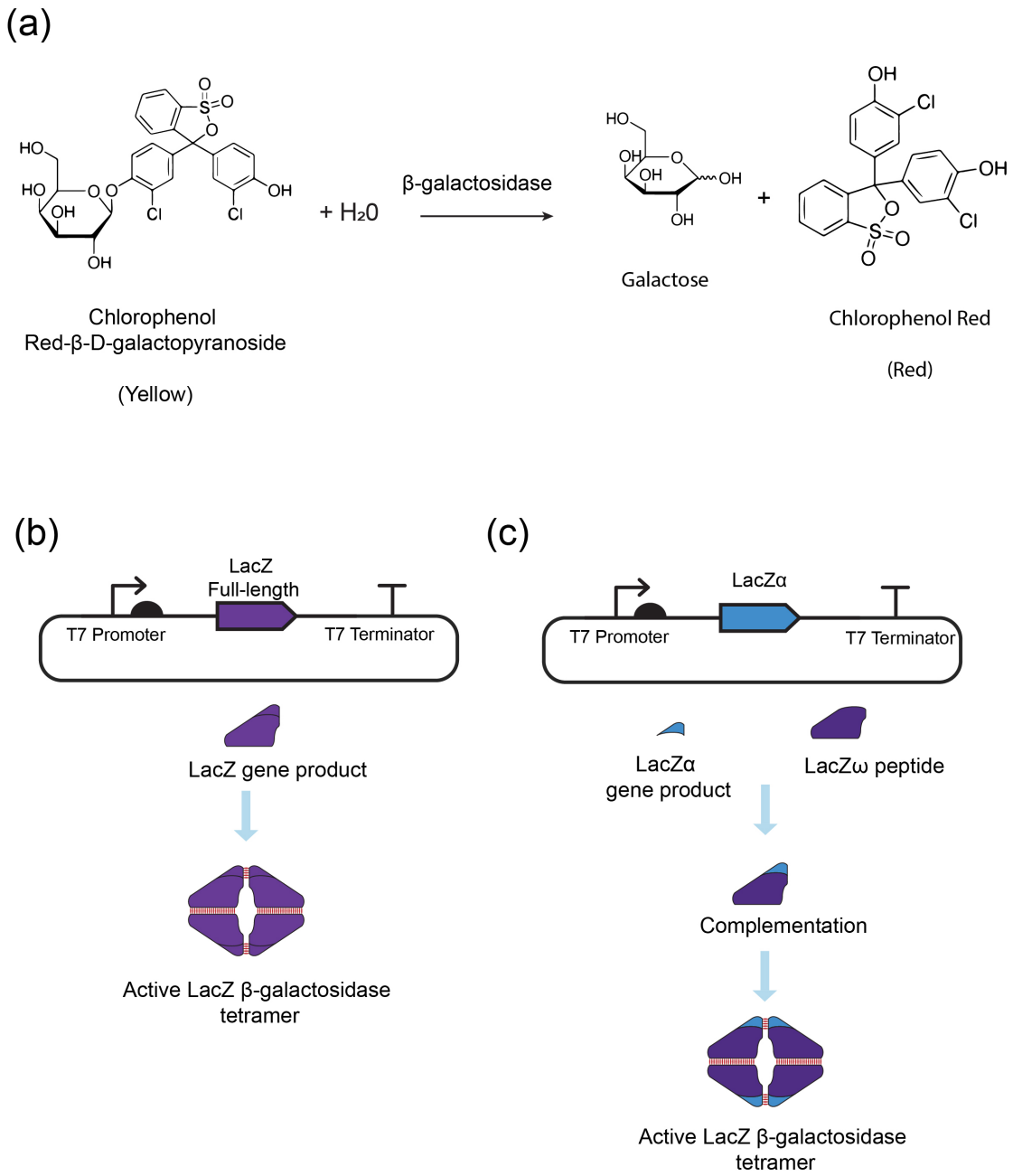

Figure 4.3: Enzymatic reaction catalyzed by LacZ β-galactosidase and plasmids used for constitutive expression of LacZ reporters. a: Scheme of the chemical reaction catalyzed by β-galactosidase that claves the yellow substrate Chlorophenol red-β-D-galactopyranoside generating chlorophenol red as a product. b: Schematic representation of the plasmid allowing the expression of LacZ full-length that forms the active LacZ tetramer and c: Schematic representation of the plasmid allowing the expression of lacZ α that in combination with the ω-peptide can forms the active LacZ tetramer.

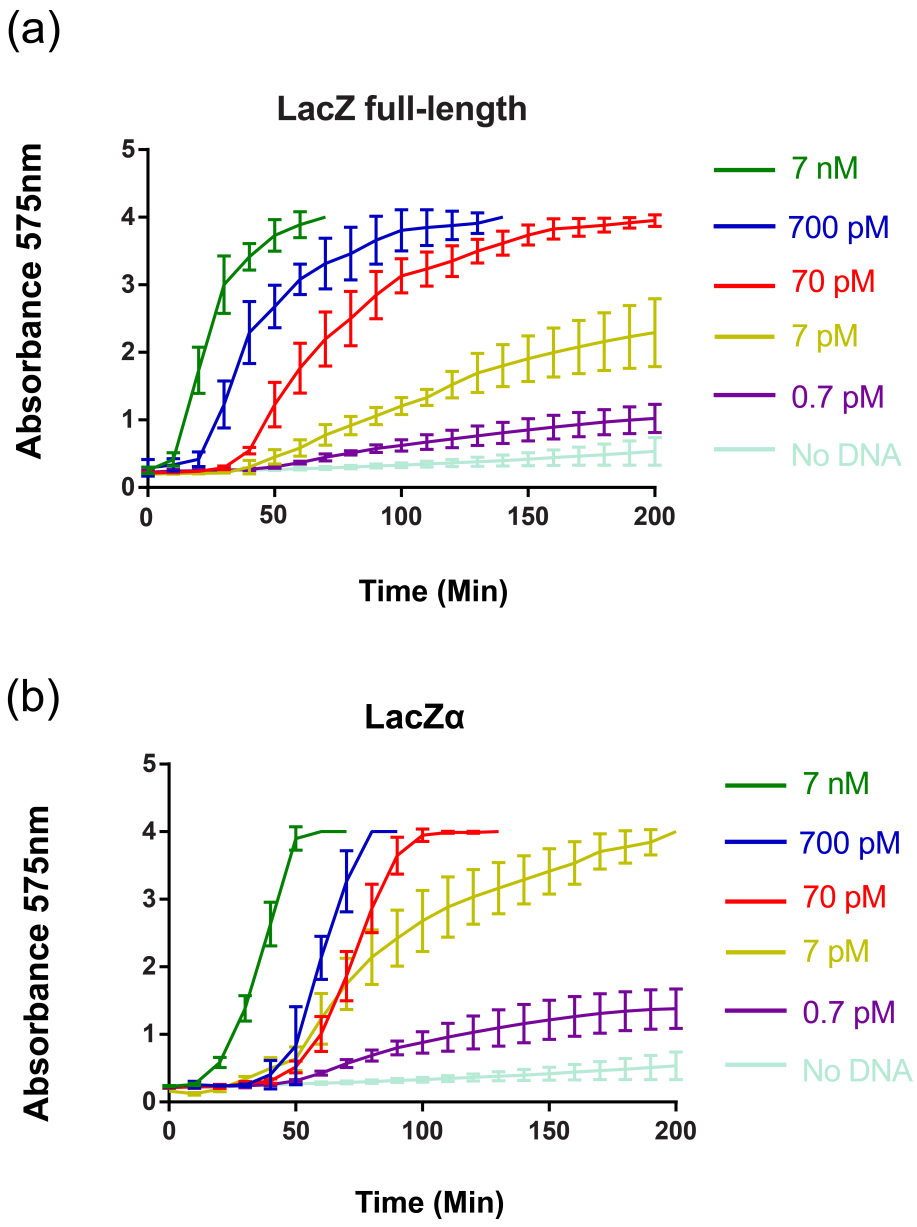

Figure 4.4: Effect of different input DNA concentration to the colorimetric change from yellow to red, measured as Absorbance at 575 nm. Constitutive expression plasmids of LacZ α (a) or LacZ full-length (b) reporters were used. Input DNA concentrations were in the range 7 nM to 0.7 pM.

the sequence of Sensor 1 in a level 0 plasmid, from which the development of several level 1 constructions was possible using Golden Gate (see section 2.1). In particular we have made different plasmids in which this toehold sensor controlled the translation of different fluorescent reporters (deGFP, YFP, and sfGFP). The plasmids encoding for each transcriptional unit were incubated along with an appropriate trigger molecule (Trigger 1), as RNA or ssDNA at different concentrations. In this setting, our initial attempts to activate the expression of fluorescent reporters using Trigger 1 RNA from *in vitro* transcription or co-incubation with a plasmid encoding for Trigger 1 were not successful. However, we observed the system's activation using ssDNA as a trigger at 2.5 μM and higher concentrations (Figure 4.5). The highest performance was achieved with the sfGFP reporter, followed by YFP, with a poorer response from deGFP. Besides the observed activation, a highly concentrated trigger ssDNA was needed to activate the fluorescent reporters. The highest fluorescent signals observed were at least one order of magnitude lower than what we found for constitutively expressed fluorescent reporters.

Next, we tested a toehold sensor for ZIKV (Sensor 8; Pardee *et al.* [15]). A level 1 plasmid was constructed to enable ZIKV Sensor 8 controlling the expression of sfGFP, and this was incubated with Trigger 8 (T8) from *in vitro* transcribed RNA or ssDNA in cell-free gene expression reactions. In the first case, a slight but low activation was observed using concentrated RNA (500 nM) (Figure 4.6a). Two ssDNA triggers were generated considering different lengths: a minimal T8 trigger, which is 36 nt long (T8 short, 36 nt); and a more extended version consistent with the minimal trigger but extended by fifty nucleotides on each side with the corresponding ZIKV genomic region, to achieve a 136 nt long trigger (T8 long, 136 nt; Table). At the same concentration (10 μM), endpoint expression (18 h post-induction) of sfGFP was almost twice as high for T8 long ssDNA as it was for T8 short ssDNA, possibly due to enhanced half-life

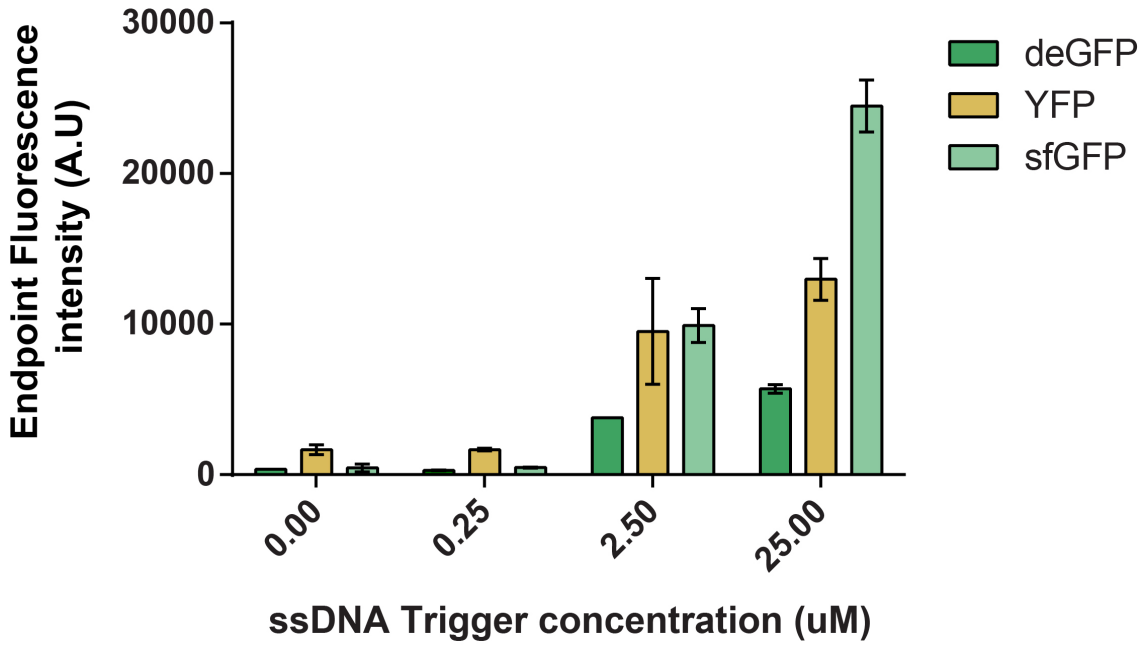

Figure 4.5: Effect of different trigger concentration (ssDNA) over the Toehold Sensor 1 activation using different fluorescent reporters as outputs. Endpoint fluorescence intensity was measured after 18 hours incubation in home-made cell-free gene expression system. Error bars represent standard deviation from three independent experiments

of the ssDNA (Figure 4.6b). Although fluorescent reporter activation was achieved upon toehold trigger incubation, it demanded high ssDNA concentrations as a trigger. Furthermore, the signal outputs obtained were low in comparison to constitutively expressed reporters. These results indicated that although fluorescent proteins work well for constitutive expression, their output signal are not optimum for CFTS in low cost extracts.

Next, we tested enzymatic reporters as outputs of toehold RNA sensing reactions. First, ZIKV toehold Sensor 8 was built to control the translation of the nano-lantern reporter, and incubated either with its cognate Trigger 8 RNA or with RNA Trigger 27 (negative control) in cell-free gene expression reactions. The rapid decrease in the luminescence signal observed using the constitutive nano-lantern reporters could be attributed to substrate depletion 4.2. Therefore, in one set of RNA sensing reactions, the substrate furimazine was added 10 hours after the reaction began (Figure 4.7a) to ensure sufficient expression of the reporter before interaction with the substrate; in the other set, the substrate was added in the initial mix when the reaction started similarly to previous settings (Figure 4.7b). In both cases, a clear activation signal was observed. When furimazine was added in the initial reaction (Figure 4.7b), the signal was observable from 20 minutes of reaction and lasted for more than 200 minutes. During the first 200 minutes of the reaction, the negative control (Trigger 27) showed a low but constant background, while the positive reaction (Trigger 8) achieved maximum luminescence signal around 150 minutes. On the other hand, a much brighter signal was observed when the substrate furimazine was added once the reaction had ended (after 10 hours of reaction). However, this rapid increase in signal came along with a fast decay, that after 15 minutes became indistinguishable from the negative control (Figure 4.7a).

Finally, LacZ was tested as an output for CFTS reactions. ZIKV Sensor 27 was

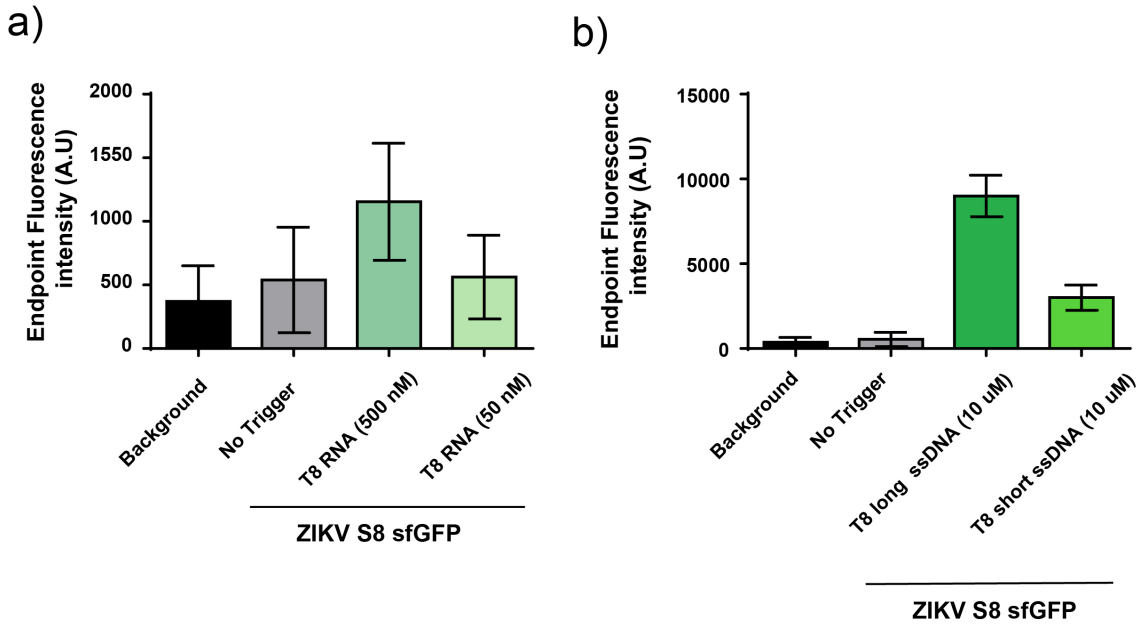

Figure 4.6: Activation of ZIKV Sensor 8 controlling the expression of sfGFP a: Effect of different RNA Trigger 8 concentrations over ZIKV Sensor 8 activating sfGFP traduction. b: Effect of different lengths of ssDNA trigger 8 over ZIKV Sensor 8 activating sfGFP traduction. Endpoint fluorescence intensity was measured after 18 hours incubation in home-made cell-free gene expression system. Error bars represent standard deviation from three independent experiments.

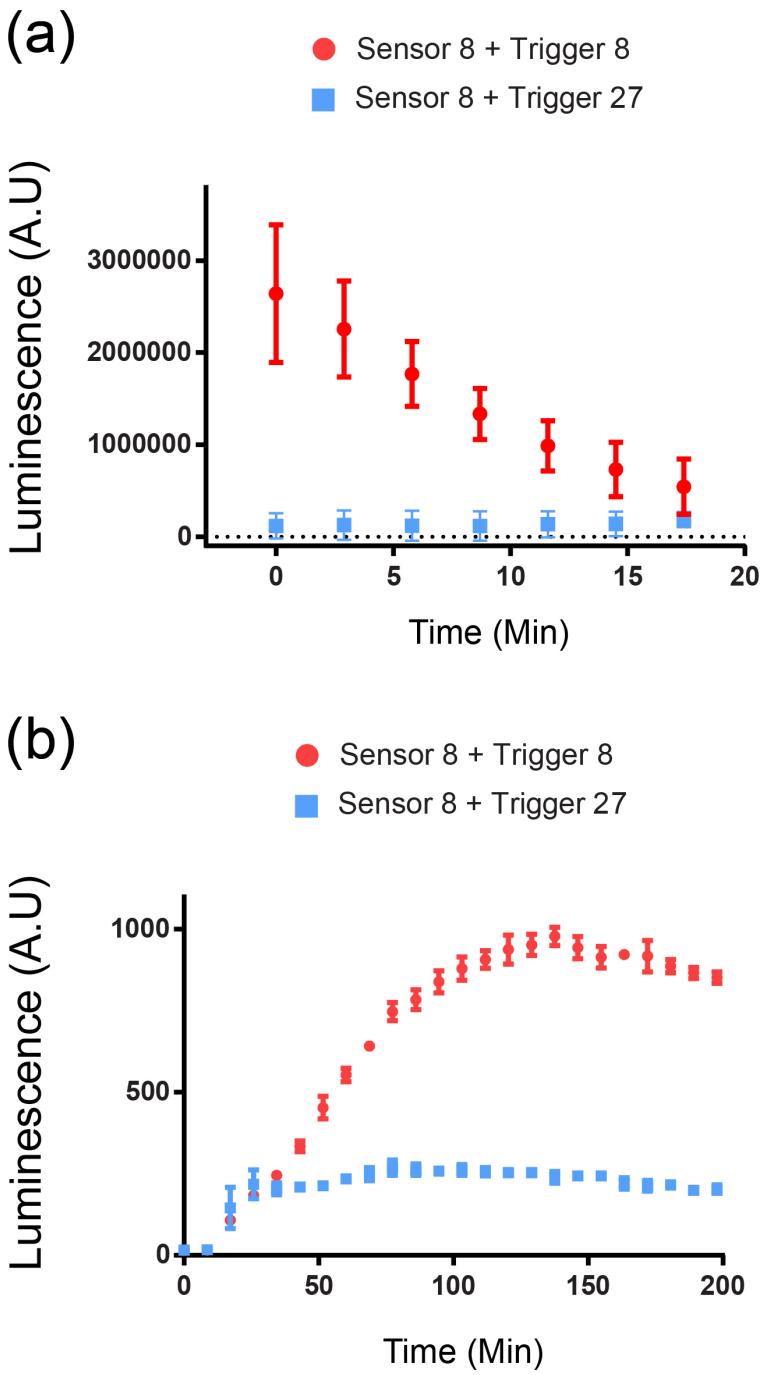

Figure 4.7: Activation of ZIKV Sensor 27 controlling the expression of nano-lantern. a: Reactions were prepared incubating Sensor 8 plasmid DNA along with RNA Trigger 8 or RNA Trigger 27. Substrate furimazine was added after 10 hours of reaction and recorded immediately. b: Substrate furimazine was added in the initial reaction.

cloned to control the translation of full-length LacZ in cell-free gene expression reactions (Figure 4.8). This plasmid was incubated with either its cognate RNA Trigger 27, or RNA Trigger 8 as a negative control. The change in absorbance was observed during the first 90 minutes and its dynamics were clearly different from the negative control (Figure 4.8a). After 200 minutes, the color of reactions containing Sensor 27 with the Trigger 27 changed completely from yellow to red (Figure 4.8b). Although some background was observed in the untriggered reactions (where Sensor 27 was incubated with RNA trigger 8 or without trigger), the color change resulting from the cognate trigger reaction was clearly different to the controls (Figure 4.8b).

To summarize these results, both enzymatic reporters tested, nano-lantern and LacZ, are suitable outputs for toehold RNA sensing reactions in a sequence-specific manner. We decided to continue with LacZ reporters for further experiments as they produced a more stable signal and they were previously used for cell-free toehold sensors.

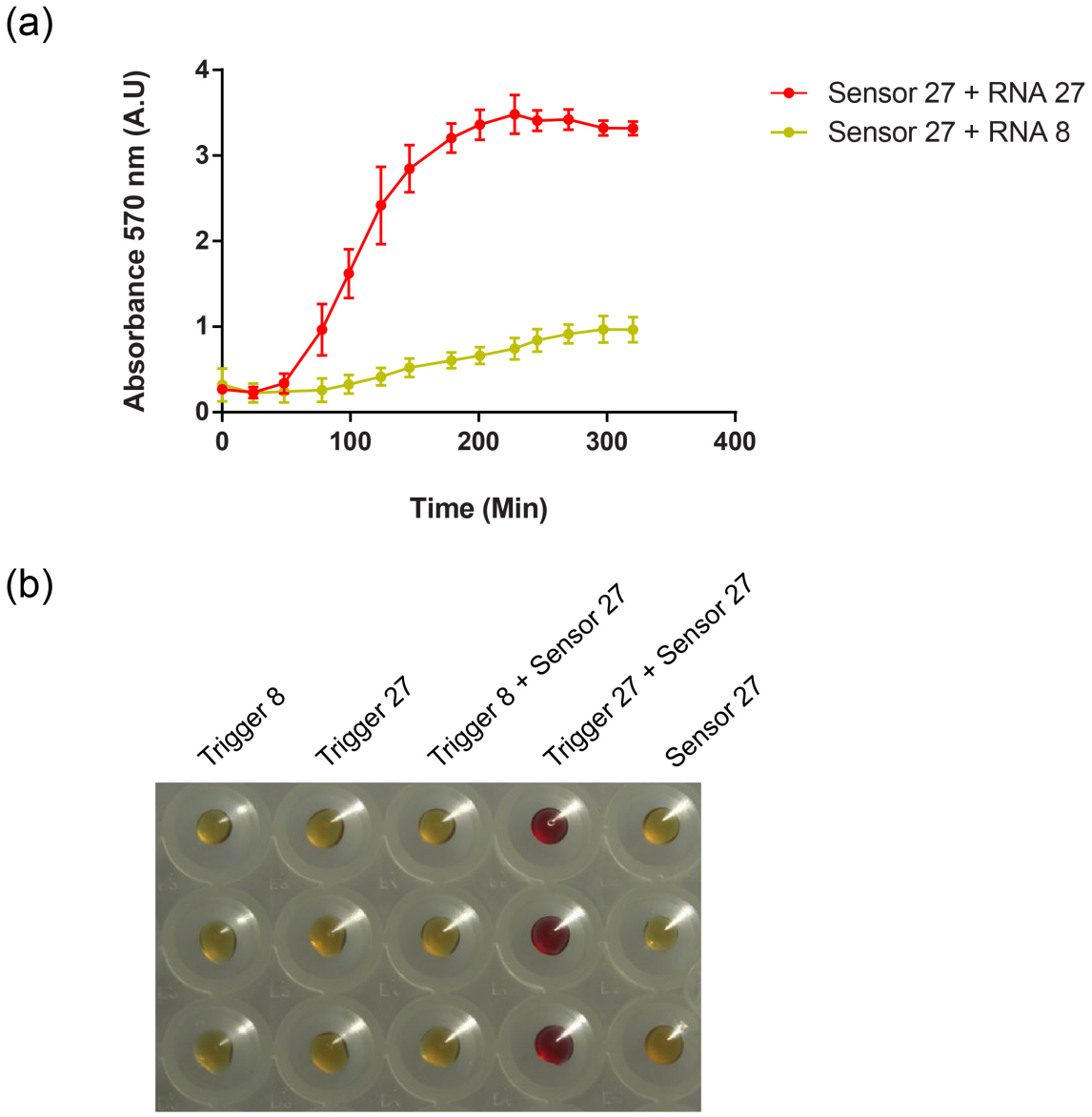

Figure 4.8: Activation of ZIKV Sensor 27 in cell-free reactions. a: Dynamics of the RNA sensing reactions performed with ZIKV toehold Sensor 27 incubated with RNA Trigger 8 or RNA Trigger 27. b: Colorimetric change after 200 minutes of incubation at 29 °C.

4.2 Performance of RNA sensing capacity vs commercial PURExpress

In order to compare the performance of CFTS in in-house produced extracts vs. commercial systems, we established a collaboration with the Pardee laboratory (University of Toronto). Crude extracts were prepared in Chile and delivered to the Pardee laboratory, where comparisons were performed with the PURExpress commerical system. We selected ZIKV toehold sensors 8 and 27 and their respective triggers due to their high dynamic range, orthogonality, and previous successful demonstration with PURExpress [18]. These toehold sensors, controlling LacZ expression, were incubated with each RNA trigger independently (Figure 4.9a). Sensor 8 behaved similarly in both reaction conditions, while sensor 27 exhibited a higher ON state in PURExpress while maintaining very low-background in both (Figure 4.9b,c). Besides these differences, both systems succeeded in sensing target RNAs in a sequence-specific manner.

Equivalent experiments were performed using LacZ-α as the reporter (Figure 4.10). Toehold sensors employing full-length LacZ reporter exhibited equal or higher ON/OFF endpoint absorbance values compared to LacZ-α in both systems tested. This can be explained by the higher OFF state observed on both toehold sensors when using LacZ-alpha as a reporter (Figure 4.11). These results demonstrated that low-cost in-house prepared CFTS are comparable in performance to PURE-based CFTS reactions.

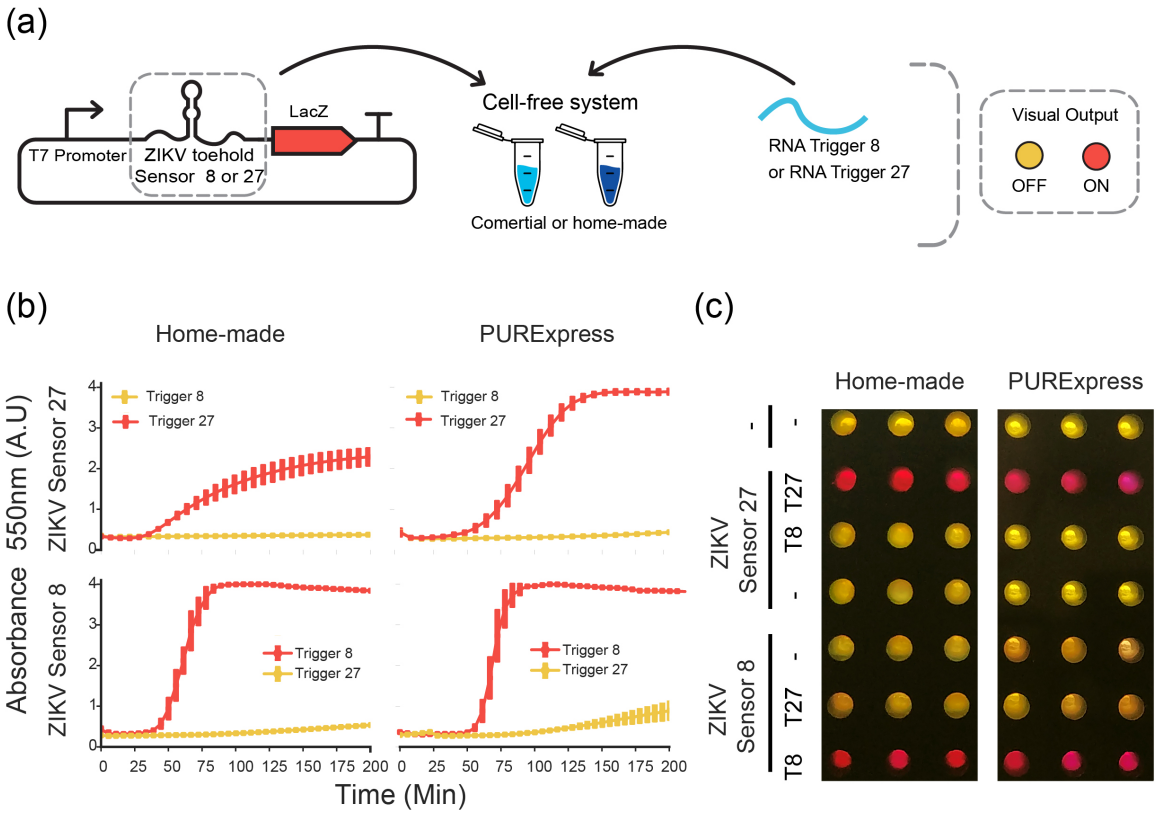

Figure 4.9: Performance of low-cost ZIKV toehold sensors. a: Schematic representation of toehold-mediated RNA sensing. b: Dynamics of the RNA sensing reactions performed with ZIKV toehold sensor 8 and 27, regulating the expression of the full-length LacZ in home-made cell extracts and PURExpress cell-free reactions. c: End-point visualization of the experiments after 200 minutes of incubation at 29°C.

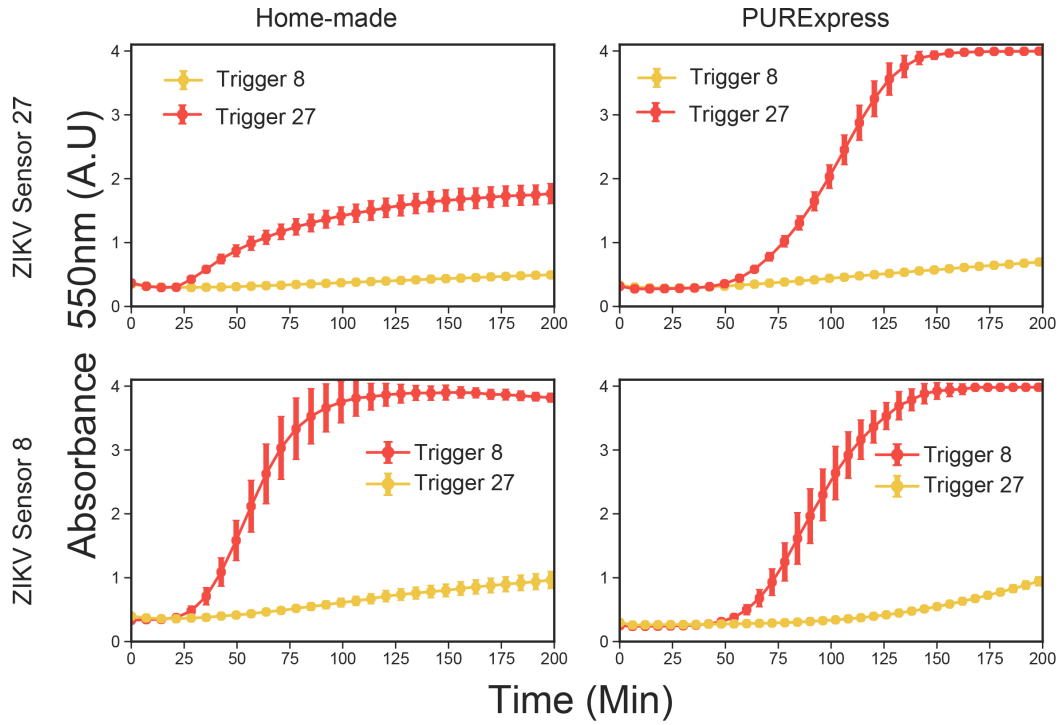

Figure 4.10: Comparison of RNA sensing reactions performed on in-house cell-free preparations and commercial PURExpress using LacZ-alpha as a reporter. Dynamics of RNA sensing reactions performed with the ZIKV toehold sensor 8 and 27, regulating the expression of the LacZ-alpha. These reactions were performed in home-made cell-free preparations (left) or commercial PURExpress (right). These reactions were supplemented with the pre-synthesized LacZ-omega peptide.

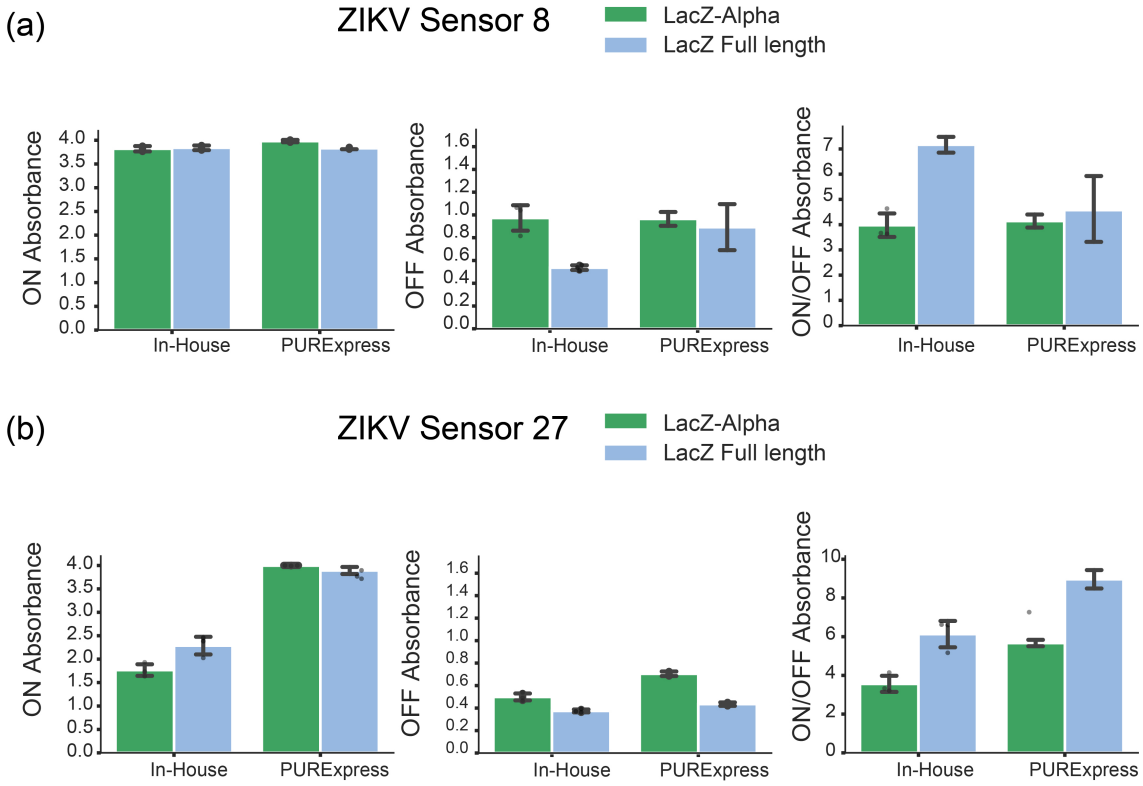

Figure 4.11: Performance comparison of the ZIKV toehold sensor 8 and 27 using LacZ-alpha or full-length as a reporter. Reporters LacZ-alpha (green) and full-length LacZ (light blue) were compared at endpoint absorbance (at 550 nm) for the trigger-activated (ON), untriggered (OFF), and ON/OFF measurements of ZIKV Sensor 8 (a) or ZIKV Sensor 27 (b) after 200 minutes of incubation at 29°C.

4.3 Cell-free toehold sensors can be lyophilized

Deployment of RNA sensors in point-of-care settings is facilitated by the lyophilization of reactions, enabling room temperature transportation and usage upon rehydration. It was already shown in Chapter 3 that cell-free reactions support the expression of sfGFP for a period of 90 days after lyophilization, with a considerable time-dependent loss in activity. In order to test whether home-made CFTS can also be freeze-dried and stored at room-temperature, the ZIKV Sensor 27 was lyophilized in home-made cell extracts and challenged with trigger RNA after one and seven days after lyophilization (Figure 4.12). Besides the evident decrease in dynamic range, especially after seven days, these reactions retained the ability to detect its target trigger, while maintaining target specificity. These results showed that, although there is room for improvement, low-cost CFTS can be lyophilized and stored for several days.

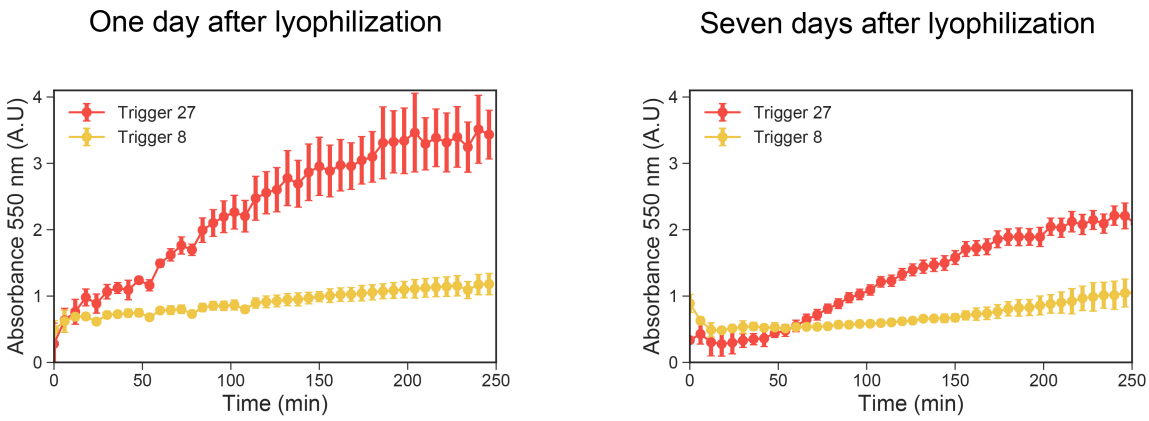

Figure 4.12: Cell-free extracts were lyophilized and stored at room temperature for up to seven days. Reactions containing the ZIKV Sensor 27 were prepared using the freeze dried extracts and challenged with Tigger 27 or Trffer 8 on days one (a) and seven (b) after lyophilization.

4.4 Effect of DNA input concentration on the sensing capacity of CFTS

We studied the effect of DNA input concentration on the leaky expression and saturation of CFTS reporter signals. We tested three ZIKV toehold sensors controlling LacZ expression (Sensor 8, 27, and 32) in concentrations ranging from 0.5 nM to 10 nM (Figure 4.13). Every sensor tested increased its OFF state signal as a function of the input DNA concentration. Sensor 27 exhibited the highest dynamic range due to a low OFF state, followed by Sensor 32. For Sensor 27, the highest dynamic range was observed when input DNA ranged between 1.5 nM and 4 nM. Sensor 8, on the other hand, showed a higher OFF state and a lower dynamic range as expected since it belongs to a less optimized series of toehold sensors (Figure 4.13) [15]. These results demonstrated that although an effective range is found between 0.5 and 4 nM, the optimal DNA input concentration should be individually evaluated for each sensor.

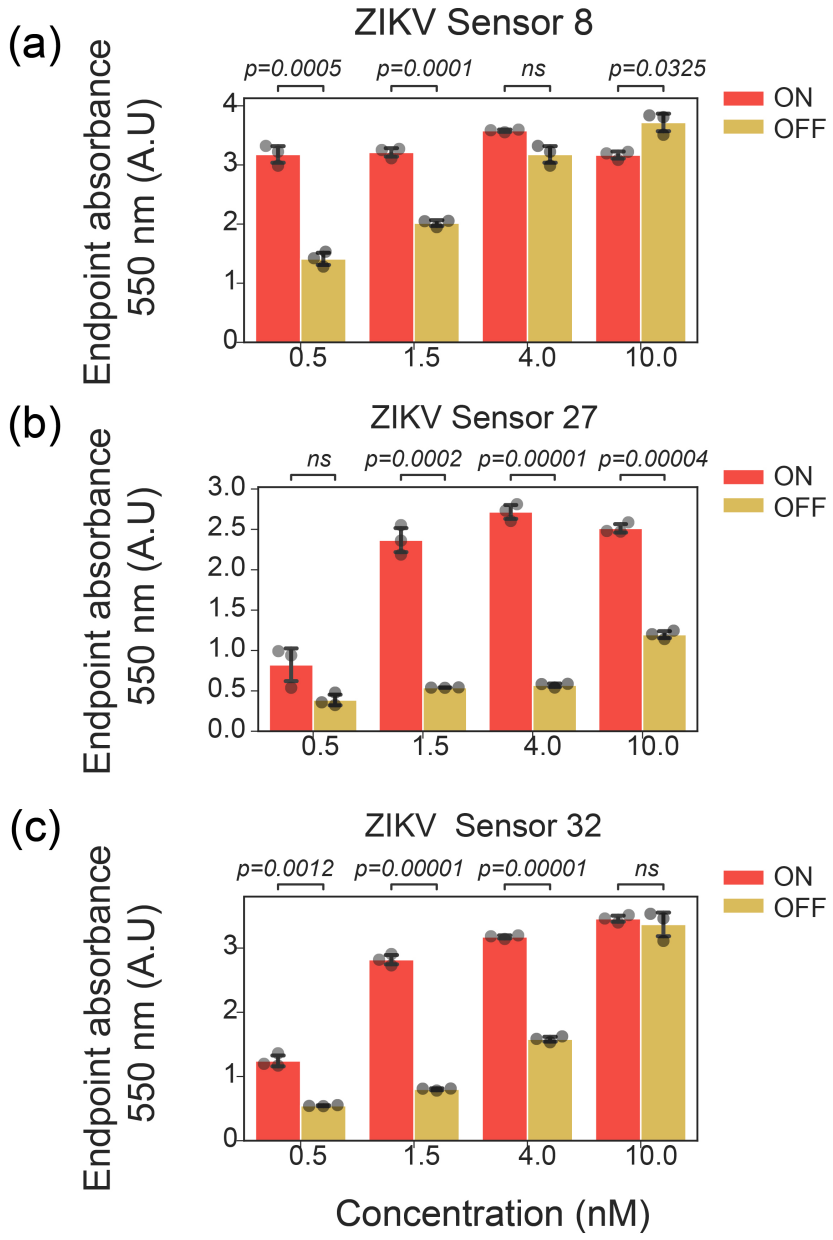

Figure 4.13: Plasmid concentration effect on the response dynamic range of ZIKV toehold sensors 8, 27, and 32. Endpoint absorbance of RNA sensing reactions were performed with the ZIKV toehold Sensor 8 (a), Sensor 27 (b), and Sensor 32 (c) at different concentrations of starting plasmid DNA, ranging from 0.5 nM to 10 nM. Reactions in the presence of the corresponding trigger are shown in red (ON), while reactions without a trigger are shown in yellow (OFF). Independent sample t-test with Bonferroni correction for multiple comparisons was performed to compare ON vs. OFF states within the same concentration for each sensor. Significant p-values are shown above each comparison or "ns", where no significant difference was observed.

4.5 NASBA amplification increased sensibility of Cell-free toehold sensors

To increase the sensitivity of CFTS in home-made cell-free preparations, we performed NASBA amplification of the target RNA before adding it to the reactions containing the sensors. We chose ZIKV toehold sensor 27, due to the high dynamic range found previously and performed NASBA reactions of the Trigger 27 using the primers described on the literature [15]. While unamplified RNA was detected in the range 274 nM to 2.7 nM, NASBA-amplified RNAs were detected at concentrations as low as 2.7 fM, representing a sensitivity increment of six orders of magnitude (Figure 4.14).

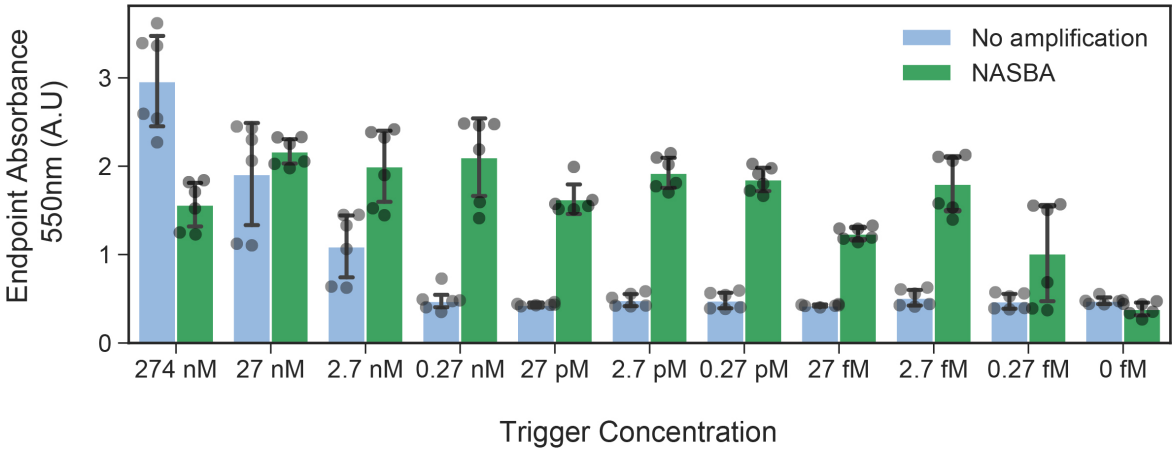

Figure 4.14: NASBA enhances the sensitivity of home-made CFTS. End-point measurement of RNA sensing reactions performed with ZIKV sensor 27 with Trigger 27 in a range of concentrations with and without NASBA isothermal amplification

4.6 The use of linear DNA for CFTS in the CRISPRi strains

Following the promising results of increased stability of constitutive gene expression from linear DNA in a CRISPRi strain (Chapter 3), we tested linear CFTS in CRISPRi-optimized extracts. We prepared CRISPRi-optimized extracts from the BL21dLacGold strain, due to its compatibility to LacZ reporters, and tested the RNA sensing capability of ZIKV toehold sensors expressed from PCR linear products. Using a concentration of linear DNA of approximately 17 nM, were able to detect RNA with both sensors 8 and 27 in a sequence-specific manner (Figure 4.15). The temporal dynamics were similar to those of plasmid derived sensors (shown in Figures 4.9b and 4.10). These results demonstrated that linear DNA can be used for prototyping CFTS in home-made extracts.

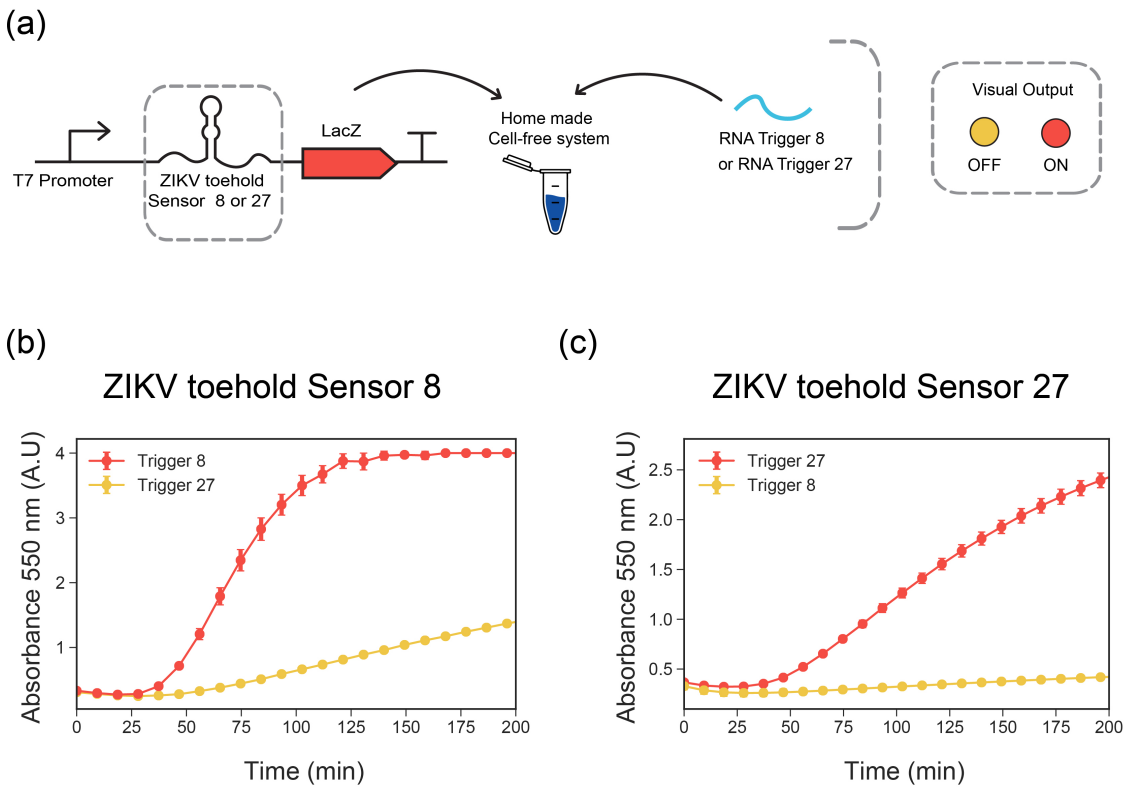

Figure 4.15: Performance of RNA sensing reactions using PCR-derived linear DNA encoding for ZIKV toehold sensors.a: Scheme of the experiment performed.b: Dynamical profile of RNA sensing reactions using ZIKV toehold Sensor 8 and Sensor 27(c).

5. *Results from aim 3: To generate novel RNA toehold sensors for the PVY virus.*

In previous chapters, we have established and validated a method for cell-free gene expression reactions (Chapter 3) that is suitable for cell-free toehold RNA sensing reactions (Chapter 4).

In this chapter, we sought the development of novel cell-free toehold sensors specific for the potyvirus potato Y (PVY) that affect potato farmers and seed makers around the globe. The challenges faced on this objective include for the identification of specific genetic regions that are conserved across PVY strains, and the computer-aided design, and experimental validation of the toehold sensors for this sequences.

5.1 Search of conserved sequences in PVY genomes corresponding from different strains

In order to design a cell-free toehold sensor that is useful for detecting PVY viruses, a multiple sequence alignment was performed, considering a diversity of PVY strains. Ten PVY complete genome annotations were included in this analysis that correspond to different PVY strains (Table 2.3).

Few regions showed high degree of conservation across the full alignment, for in-

stance the longest string of 100% conservation found was only 87 nucleotides long and correspond to a section that encodes the capsid protein (Cp). A longer sequence of 227 nucleotides, that flanked this section of 100% conservation, displayed high degree of conservation across the full alignment, however some punctual nucleotide changes in the necrotic (PVY-NTN) strains were observed (Figure 5.1). Both consensus sequences (named "PVY short" and "PVY long" respectively) were chosen for analysis for designing cell-free toehold sensors.

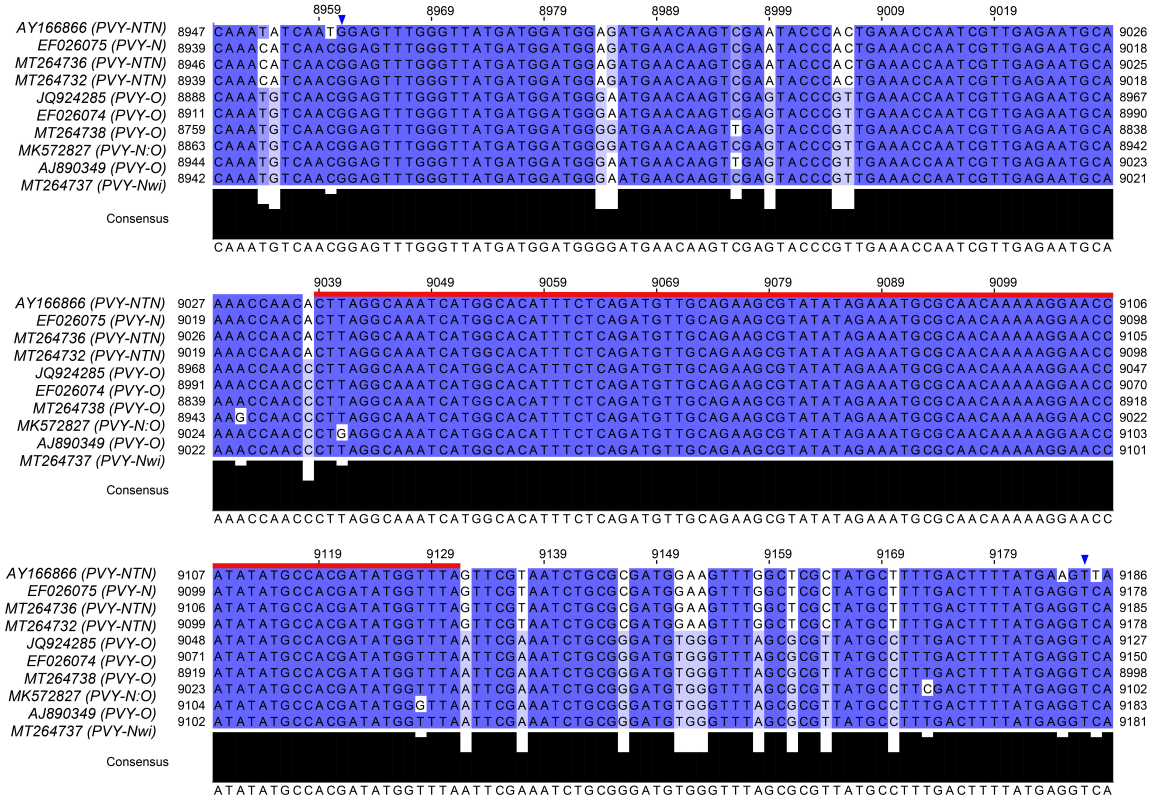

Figure 5.1: Conserved region across several sequences of the PVY genotypes. The consensus contained between blue triangles marks was named "Long" and was used for screening PVY toehold sensors 1 to 4; while the consensus region highlighted in red line was named "Short" and was used in the screening the toehold 5 to 8 , it contains the longest string of 100% conservation across all the genomes analyzed.

5.2 Development of software for designing PVY toehold sensors

The majority of published toehold sensors have been designed with NUPACK, a software tool for modeling the thermodynamics properties of DNA and RNA molecules [11, 15, 16]. In the first publication where toehold switches were shown, Green *et al.* [11] described explicitly the algorithm used for designing novel toehold sensors. However, a software implementation of this algorithm was not published alongside the research findings, making it difficult for readers to reproduce their results. This was addressed by a team of researchers participating in the international genetically engineered machines (iGEM) 2017 competition, based at EPFL. The team developed an on-line software tool that implements the algorithm described in [11] and published as a GitHub repository (`https://github.com/EPFLliGem/Toehold-Designer`). Although the application is useful, it has not been updated to include the published scoring functions from more recent works [16]. Therefore, in this thesis, we developed a new implementation of the toehold design algorithm and shared in a new GitHub repository named NupackSensors (`https://github.com/elanibal/NupackSensors`). Included is an easy-to-use Jupyter notebook that steps through the algorithm with a running example. The algorithm works as follows:

1. Division of the target sequence in windows of the trigger size (36 nt).

2. For each possible trigger, define a toehold sensor sequence according to the ideal designs (Figure 2.1).

3. Filter out sequences that contain stop codons.

4. For each design, evaluate a set of thermodynamic parameters, and sequence

properties using NUPACK.

5. Rank the possible toehold designs according to a scoring function.

After filtering out sequences that contained stop codons, the two selected sequences PVY Long, and PVY Short, delivered two hundred eighty eight (288) and ninety five (95) possible toeholds sensors respectively. For each target sequence, the toehold sensors were ranked according to Ma*et al.* scoring function [16], and finally 8 sensors were selected for experimental screening. From them, four were designed to be triggered by the short sequence (sensor 1 to 4), although the trigger sequence is also present in the "PVY long"". On the other hand, sensor 5 to 8 were able to be triggered by the long sequence.

5.3 Prototyping PVY toehold sensors using PCR-derived linear DNA in CRISPRi optimized cell-free reactions

A rapid screening strategy was implemented in order to test whether the *in-silico* designed PVY toehold sensors were functional using in-house prepared cell-free reactions. The CRISPRi-optimized strain was used with the aim of testing PVY toehold sensor directly from PCR amplification. Individual ultramer-primers containing the T7 promoter and sequence corresponding to each toehold sensor were designed (Table 2.9). Each sensor was then prepared by PCR amplification, purified, and incubated with the corresponding synthetic trigger RNA, or its reverse complement sequence as a control (Figure 5.2A). An untriggered control without RNA addition was prepared too; this control was used as a baseline in the ON/OFF endpoint calculations (at 200

minutes). All of the sensors tested displayed no interaction with the reverse complementary trigger RNA respecting the untriggered reaction. Among the 8 sensors tested, two of them achieved ON/OFF endpoint measurements greater than 2.3 (Figure 5.2B,C). Individual performance on this screening of each PVY sensor is shown in Figure 5.3 (for sensors 1 to 4) and Figure 5.4 (for sensors 5 to 8).

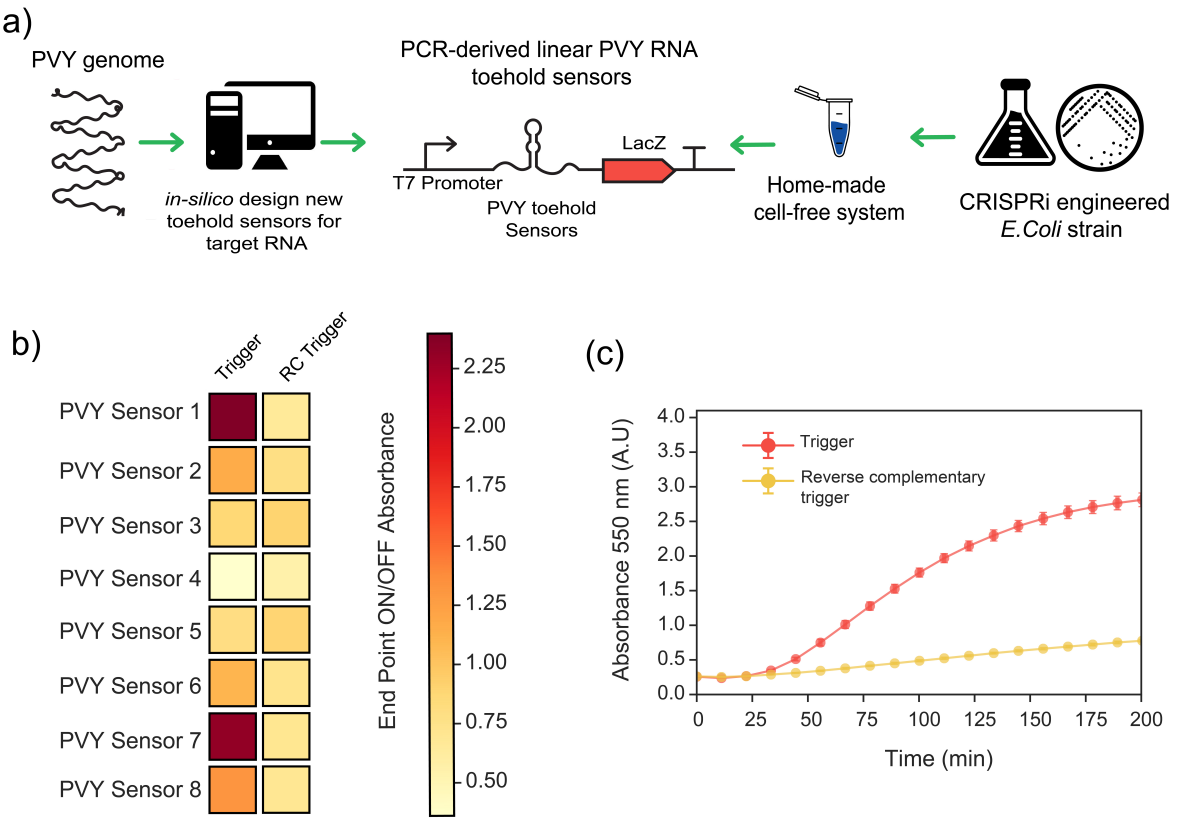

Figure 5.2: (a) Scheme of the global strategy for fast prototyping *de-novo* designed RNA toeholds sensors against synthetic fragments of PVY virus. PCR products encoding for each of the 8 PVY toehold sensors tested were incubated in a cell-free reaction and supplemented with the corresponding RNA trigger or its reverse complementary sequence (negative control) (b) PCR-purified transcriptional units (at 10 nM final concentration) encoding for PVY RNA toehold sensors were incubated with synthetic trigger RNA or its reverse complementary as a control. RNA was included at 300 nM final concentration. ON/OFF absorbance was measured after 200 minutes with respect to the untriggered control (the toehold sensor incubated in cell-free reaction without any trigger, at the same concentration than the triggered reactions). (c) Dynamics of the PVY toehold sensor 1, that achieved the highest endpoint ON/OFF absorbance in this fast screening.

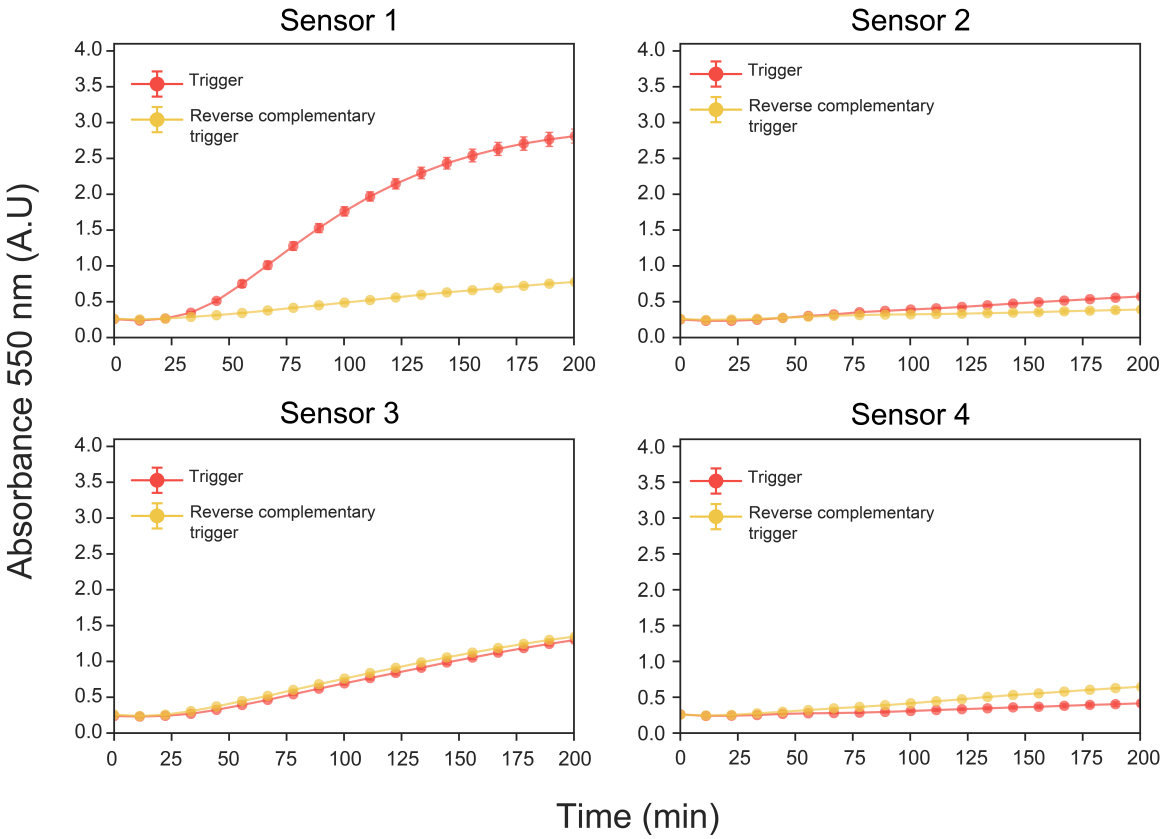

Figure 5.3: Dynamics of the fast prototyping reactions of PVY toehold sensors 1 to 4 in CRISPIRi optimized cell extracts. Linear dsDNA of each of the PVY cell-free sensors 1 to 4 was prepared individually by PCR and incubated with the corresponding RNA trigger (red) or its reverse complementary sequence (yellow) in cell-free gene expression reactions using the CRISPIRi optimized cell lysates. Absorbance at was measured in a plate reader at 29 °C.

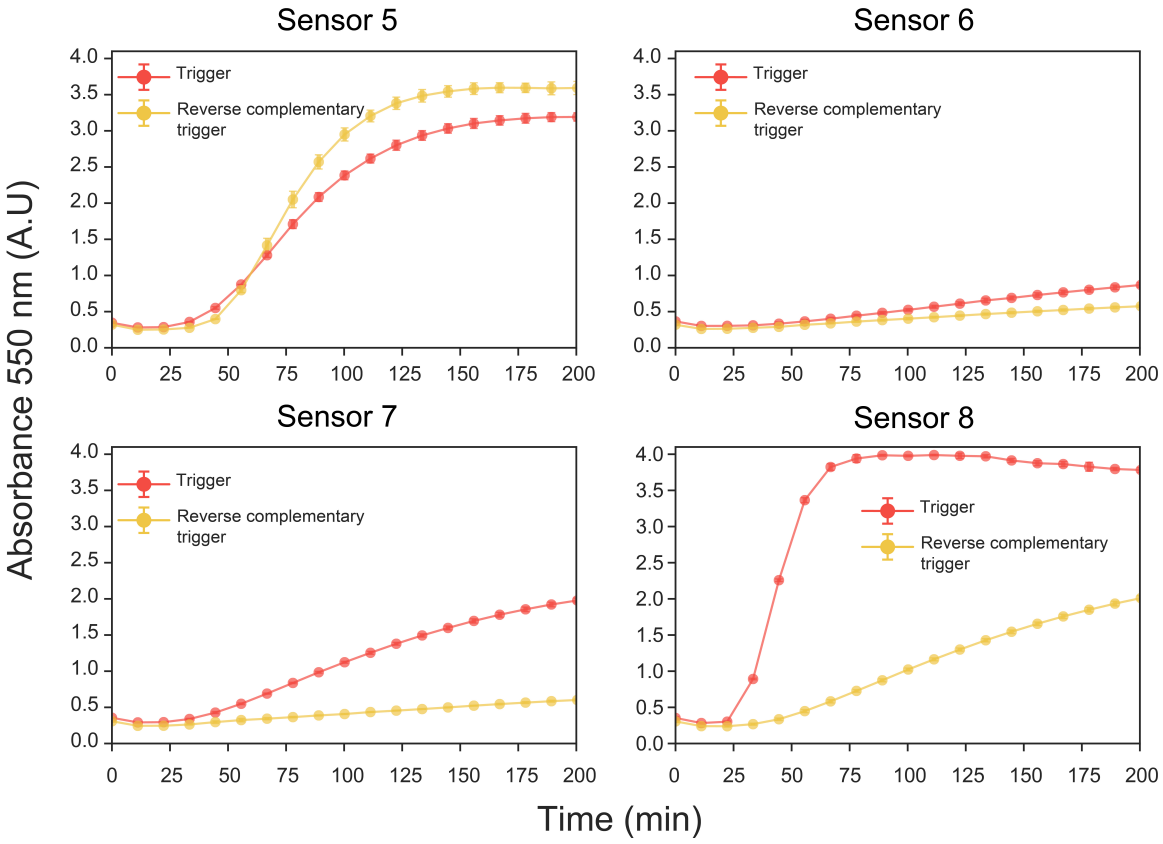

Figure 5.4: Dynamics of the fast prototyping reactions of PVY toehold sensors 5 to 8 in CRISPIRi optimized cell extracts. Linear dsDNA of each of the PVY cell-free sensors 1 to 4 was prepared individually by PCR and incubated with the corresponding RNA trigger (red) or its reverse complementary sequence (yellow) in cell-free gene expression reactions using the CRISPIRi optimized cell lysates. Absorbance at was measured in a plate reader at 29 °C.

5.4 Validation of good performance of the PVY toehold sensor 1 and sensor 7 stored in plasmids

In order to validate the fast characterization of PVY toehold sensors using linear DNA, two of the good performing designs (sensor 1 and sensor 7) and two of the poorer designs (sensor 4 and sensor 5) were chosen to be cloned into plasmid for further assays. After sequence verification, 1 nM of plasmid DNA was incubated with the corresponding trigger RNA or a control corresponding to the ZIKV trigger 27 in order to test the sequence-specificity of the activation, also one control without the addition of RNA was placed (Figure 5.5).

Consistently with the linear DNA screening, sensor 4 and sensor 5 displayed poor sensing quality using plasmidic DNA. Sensor 5 showed high background in the unspecific and untriggered reactions. Similarly, sensor 4 showed unspecific activation in the negative controls to a lesser endpoint absorbance value than sensor 5, but without displaying any sequence-specific activation (Figure 5.5).

On the other hand, the sensors that achieved good results using linear DNA, behaved consistently when cloned as pladmids. For instance sensor 1 and sensor 7 showed specific activation when the corresponding trigger RNA was used. At 1 nM sensor 7 showed high specific activation (ON/OFF $\sim$3.5) while sensor 1 displayed a slightly poorer activation (ON/OFF $\sim$ 2) (Figure 5.5).

Because the PVY toehold sensor 1 exhibited low specific activation along with little background, it was assumed that was being used at sub-optimal concentration. Therefore, sensor 1 was also tested at higher DNA input concentrations (4 nM, Figure 5.6a ; and 2 nM Figure 5.6b). In this sets of experiments, the PVY sensor 1 was

compared side by side with ZIKV Sensor 27 that achieved higher dynamic ranges at higher input DNA concentrations than the other sets of ZIKV sensors tested in this thesis (Figure 4.13). At 4 nM, both PVY and ZIKV sensors achieved activation near saturation values with their respective triggers (Figure 5.6a), however a higher background was observed in the PVY sensor 1, compared with the background on the ZIKV sensor 27. The background observed in sensor 1 was lower at 2 nM input DNA (Figure 5.6b). This result indicates that PVY sensor 1 has a lower optimal input DNA concentration than ZIKV sensor 27.

Summarizing, in this chapter the use of PCR-derived linear DNAs in the CRISPRi-engineered lysates was validated for rapid screening and characterization of toehold sensors. Moreover, it was shown that PVY sensor 1 and sensor 7 display sequence-specific activation triggered with a conserved sequence of the PVY viruses.

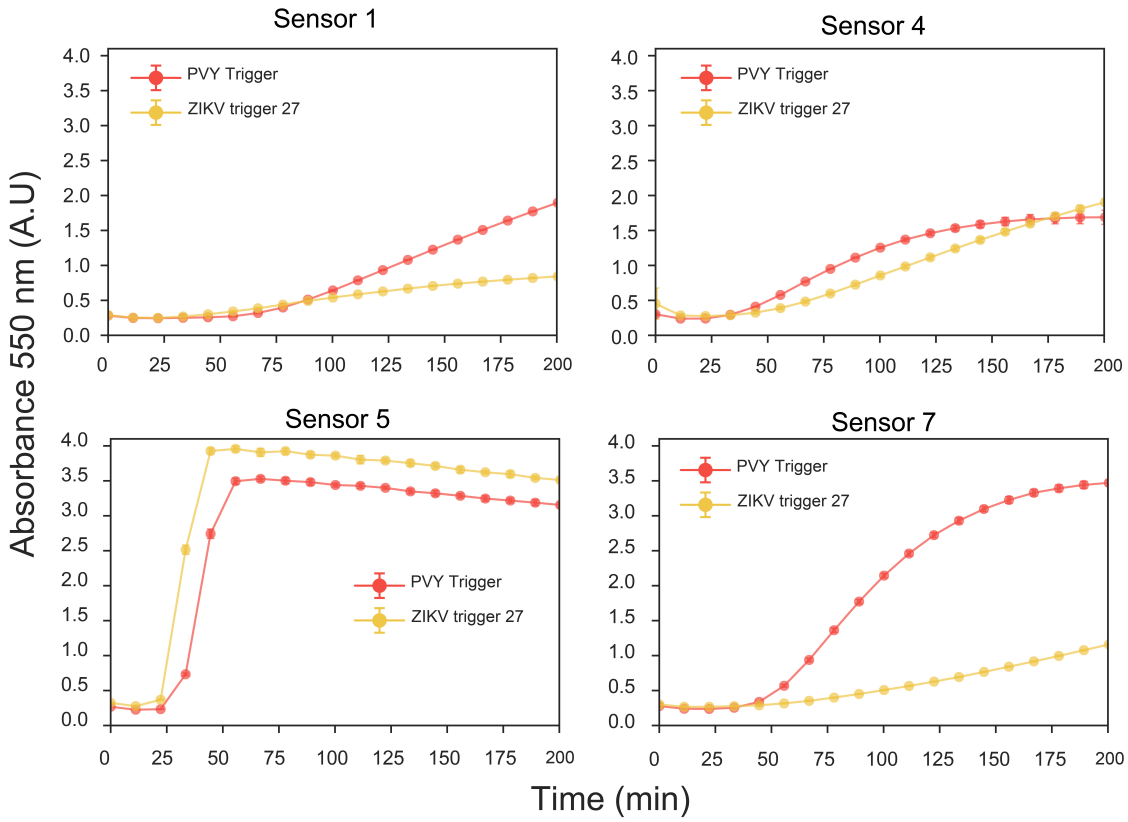

Figure 5.5: Validation of the screening process testing two good performing PVY toehold sensors (sensor 1 and sensor 7) and two poor performing toehold designs (sensor 4 and sensor 5) using plasmidic DNA. 1 nM input plasmidic DNA encoding each toehold sensor was incubated with its corresponding trigger RNA (at 300 nM) or with the unrelated ZIKV sensor 27, in order to test the sequence-specificity on their activation. Sensors that were identified as good performing in the linear DNA screening show sequence-specific activation using plasmidic DNA, while the sensor 4 and sensor 5 exhibit non specific activation similarly to what was observed in the linear screening.

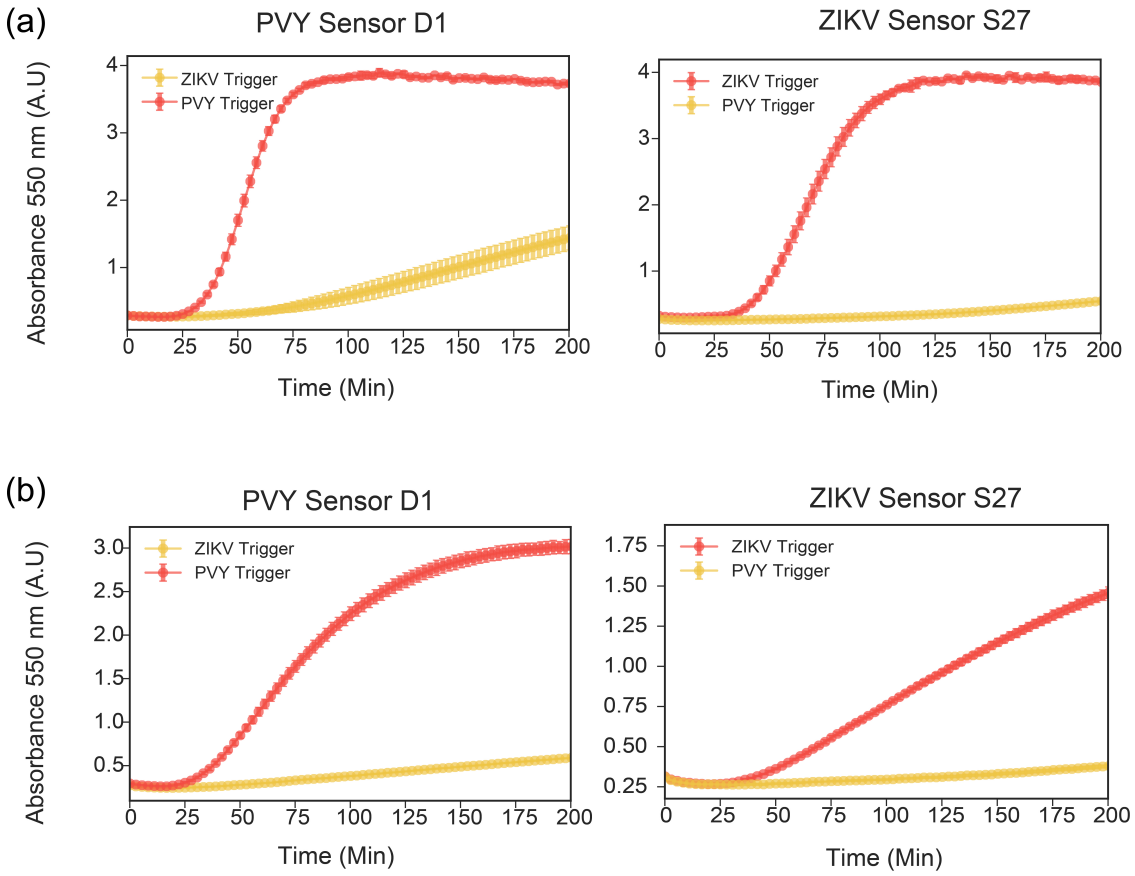

Figure 5.6: Validation of PVY sensor 1 as a functional, and sequence specific RNA toehold sensor. PVY toehold sensor 1 was cloned and sequence verified. Plasmids encoding PVY toehold 1 and ZIKV Sensor 27 were incubated with either RNA PVY trigger ("PVY Short") or ZIKV RNA (Trigger 27). Reactions were performed using 4 nM input DNA (a) or 2 nM input DNA (b)

.

6. *Discussion*

6.1 Decentralizing cell-free RNA sensing with the use of low-cost cell extracts

Cell-free gene expression reactions have played a significant role in biological research. They were crucial for understanding the genetic code and the elucidation of the molecular mechanisms that govern transcription and translation. The use of cell-free *in-vitro* biology to facilitate the application of engineering principles to genetic circuit design through more open and controllable experimental setups has recently gained global momentum [30, 32, 121, 122]. Cell-free systems have also gained popularity in the emerging area of field-deployable *in-vitro* biosensors [16, 18, 25, 26, 123]. One of the most prominent applications is the use of cell free for toehold gene expression regulators, which activate the expression of reporter genes when a trigger RNA with a cognate sequence (defined by the user) is present in the reaction. To date, several toehold RNA sensors have been designed and implemented with the aim of detecting infectious agents such as Zika virus and Norovirus. Coupled with preparatory steps, these RNA sensors can detect clinically relevant concentrations of these viruses [16, 18]. Cell-free toehold sensors are GMO-free, which eases biocontainment restrictions, and can be lyophilized on paper or plastic for room temperature storage and

field-deployment upon rehydration. These features make CFTS a promising tool for decentralized strategies of nucleic acids monitoring.

In Latin America, however, the high cost and shipping constraints of the commercial *in-vitro* transcription and translation systems (e.g. PURExpress, and myTXTL) have limited the use and development of toehold RNA sensors. These current constraints are leading to fragile structures that rely on global supply chains and technological dependencies prone to failure upon the stress faced in pandemic scenarios, as evidenced by the recently disrupted supply chain of diagnostic reagents for SARS-CoV-2 testing. Moreover, the increasing risk of novel infectious outbreaks highlights the need for more decentralized and distributed strategies that critically rely on building up local manufacturing and surveillance capabilities.

The work presented in this thesis seeks to contribute to this goal by implementing a low-cost cell-free RNA sensing capacity in locally produced extracts. We have applied this platform to the development of novel CFTS for PVY virus as a proof-of-concept. Despite the immediate contribution to fighting PVY agricultural threat, we expect these resources to be used as a platform for prototyping and engineering novel RNA toehold sensors for other genetic targets of local interest.

6.2 Further optimizations of low-cost cell-free RNA sensing reactions

Future and ongoing research are focused on two fronts. Firstly, increasing cost-efficiency of cell-free reactions. The cost-breakdown analysis performed in this thesis has identified clear targets for future improvements (Section 3.6). For instance, the use of cheaper alternative methods for induction such as auto-induced media or opto-genetic regulation; and the use of different lysis methods that do not require micro-

chromatography columns such as autolysis, sonication, and cell pressing are possible targets. One of the most important features of cell-free RNA sensors is the possibility of lyophilization for storage, transportation and use upon rehydration on the field. Our findings revealed a fast decrease in gene expression efficiency and in RNA sensing capacity after lyophilization, probably due to insufficient stabilization of the cell-free reactions in the storage conditions. Ongoing research projects established from our group with collaborators at Cambridge university include improving the atmospheric conditions by the addition of desiccants,oxygen absorbers and argon gas on the lyophilized cell-free reactions.

Another point of improvement is evaluating whether alternative energy sources are less prone to activity-loss than maltodextrin after lyophilization, the underlying hypothesis being that maltodextrin requires a rather complex set of enzymatic reactions and inverted vesicles to achieve energy recycling as compared to a less complex energy source such as phosphoenolpyruvate.

A second target for further optimization is related to the pre-amplification step. For cell-free toehold sensors to reach detection limits of clinically or agronomically relevant RNA concentrations, it is necessary to perform a prior RNA amplification step. Isothermal nucleic acid amplification techniques such as NASBA (Nucleic Acid Sequence-based Amplification) [124] or RPA (Recombinant polymerase amplification) that do not require sophisticated equipment have been used for this purpose.

NASBA reaction is based on three enzymes: RNA polymerase T7, RNAseH, and Reverse transcriptase; it exponentially amplifies RNA using a dsDNA intermediate step. Pardee and colleagues have successfully used NASBA, to amplify the trigger RNAs from plasma samples derived from macaque infected with Zika virus. Then, the output of this amplification was detected in a sequence-specific manner using the cell-free RNA sensing reaction. NASBA is commercially available by various sup-

pliers including a lyophilized version that was used in this thesis (purchased from Life Sciences, USA). Individual enzymes are also available by commercial suppliers from the northern hemisphere, such as New England Biolabs (USA), Takara (Japan), and others. On the other hand, RPA employs three core enzymes; a recombinase (e.g. uvsX and uvsY), a single-stranded binding protein (SSB; e.g. gp32), and a strand displacement polymerase (e.g. Bsu). SSB stabilizes single-stranded DNA generated by the strand exchange reaction catalyzed by the recombinase and specific primers that invade the dsDNA. The strand-displacement polymerase uses the seeding primers for generating dsDNA that is subject to further invasions in the following rounds of amplification. Adding RNA polymerase T7 and reverse transcriptase to the RPA mix can be used to generate RNA amplification from a DNA or RNA sample which can be used on the cell-free toehold sensing reaction [16]. RPA has gained global attention since it has been used as an amplification step for CRISPR-based diagnostics and implemented for Sars-CoV-2 [125, 126]. RPA reactions have been recently patented (`https://patentimages.storage.googleapis.com/16/bc/ad/8f7e43a556aeb4/US10329603.pdf`) and currently there is only one supplier on the market, Twist Dx (based in Cambridge, UK), which is experiencing significant delays in delivery (i.e. up to 6 months).

One solution for lowering the costs of RNA sensing and alleviate supply chain stress is the preparation of home-made NASBA or RPA kits from off-patent enzymes. The Open Bioeconomy lab has curated a repository of DNA sequences that encode for Open Enzyme encoding collection, a variety of more than 42 different enzymes widely used in molecular biology and diagnostics, that are IP-free and have been recently made available at zero cost for the end-user, as a diagnostics tool. Moreover, the Reclone community, a collaboration of the Open Bioeconomy lab, FreeGenes, Ginkgo Bioworks and our group, has created a Golden Gate-compatible toolkit for these enzymes (`https:`

`//openbioeconomy.org/projects/open-enzyme-collections/`), along with affinity tags, vectors, and promoters. The toolkit is boosting collaboration in diagnostics worldwide via resource and protocols sharing. One promising task would be the development of the locally made NASBA and RPA reactions as a previous step of the cell-free RNA sensing reactions, using purified open enzymes.

6.3 Improvements in computational design of novel toehold sensors

The first library of toehold switches consisted of 168 first-generation devices designed *de novo* by Green *et al.* [11]. These designs were characterized in vivo, in E.coli using GFP as reporter. Only a small fraction displayed high ON/OFF values from all those designs (i.e. only 20 toehold switches showed $100 <$ ON/OFF). Heuristic design principles allowed the authors to generate a forward-engineered library of 12 switches that exhibit average ON/OFF ratios exceeding 400 (a dynamic range typically reserved for protein regulators). The cell-free toehold sensors were made in even further design-test-evaluate cycles [14]. For instance, during the development of the ZIKV toehold sensors two types of designs were tested (Series A and B). Both designs differ from the original toehold sensor design presented in [11]: Series A sensors reduced the size of the loop domain from 18 nts to 11 nts to prevent untriggered-ribosome interaction with the loop domain and included a refolding domain that stabilized the active sensor; Series B designs, also reduced the loop domain (to 12-nt in this case), and incorporated a more thermodynamically stable stem, stabilizing the OFF state of the sensor. The experimental evaluation concluded that series B achieved overall lower OFF states than series A, which is a desired attribution. After each round of toehold sensor design, the scoring function could be reevaluated and improved considering the novel data [16]. In

this thesis, a library of 8 toehold sensors was screened using the CRISPRi optimized lyzates in in-house prepared cell-free reactions 5.3). From them, sensor 1 and sensor 7 were cloned into plasmid, and their functionality were confirmed (Figure 5.5 and Figure 5.6) and other promising PVY toehold sensors are in process of being cloned.

The overall process of designing, screening, and selecting a suitable performance toehold switch is considered by many a bottleneck in the usability of toehold switches for their many possible applications (including cell-free toehold sensors [14, 18], *in-vivo* gene regulation in bacteria [11], biocomputation [74], detection of mutations [127], etc.). Predictability of the performance of the toehold switches is limited by the small data sets available and lack of understanding of optimal design rules. To address this, two papers have recently shown the development of machine learning models for sequence-to-function prediction of the behavior of toehold switches using large experimental data sets consisting of more than 91,000 novel toehold switches [128, 129]. Although the deep neural networks (DNNs) trained on nucleotide sequences outperform previous state-of-the-art thermodynamic and kinetic models, there are limitations on using these models for general design or evaluation of cell-free toehold sensors. The data set in which the models were trained only spanned first-generation designs. When challenged to a different set of toehold designs, such as ZIKV series B, the machine learning models could not accurately predict the ranking according to sensor performance (R^2 around 0.2 for the best model)[128, 129], making them of poor practical utility for designing cell-free toehold sensors. Transfer learning techniques and layer re-training using data from forward engineered designs could improve the predictability for these designs.

Another important consideration for *in-silico* design of functional RNA sensors is the pre-isothermal amplification step which is needed to increase sensitivity in cell-free RNA sensing reactions. Possible constrains in RNA structure of the NASBA or RPA amplicons along with optimal primer designs should be considered in the screening

process of cell-free toehold sensors, as the output of these isothermal amplifications is the actual input of the sensing reaction.

The evaluation of the PVY sensors with NASBA amplifications is part of my ongoing research in collaboration with Professor Marlene Rosales from the Agronomy Faculty and the use of Open Enzymes to develop a local version of NASBA using purified enzymes. The incorporation of a pipeline for NASBA optimal primer identification is part of a collaboration with Ignacio Ibarra (Helmholtz Zentrum, München), which seeks to screen primers for COVID-19 with stable single-stranded RNA structure (`https://github.com/ilibarra/primer_design_covid19`).

6.4 Understanding batch-to-batch variability in cell-free gene expression reactions and increasing robustness in CFTS

Batch-to-batch variability is an ongoing problem in cell-free gene expression reactions [49, 130, 131, 132]. This issue has also been evidenced in the work presented in this thesis. For example, Figure 3.4 shows the diversity in dynamics of sfGFP expression in 4 different cell extract. Understanding the sources of variability will help to focus efforts towards optimizing particular processes of the overall reaction (i.e transcription, translation, energy regeneration, etc). In order to understand better this issue, we have constructed and tested several ratiometric reporters in cell-free batches. In those DNA constructions, the expression of two fluorescent proteins is controlled differently, allowing us to observe competition per resources effects and variability on their gene expression dynamics across batches. Using Bayesian modelling techniques, several hypothesis have been tested so far. However, our first attempts have not been

able to discriminate between transcription and translation processes, as both hypotheses achieved similar performance when comparing their models with the experimental data. This led us to consider further DNA constructions that may help to discriminate between these processes. We have built genetic constructs that allow constitutive expression of the fluorescent RNA aptamers called spinach RNA [133]. This is allowing us to identify transcription rates and RNA degradation rates across batches, and therefore help to distinguish whether transcriptional or translational resource is a dominating source of variability.

7. *Conclusions*

- The development of low-cost in-house cell-free preparations that are suitable for the implementation toehold RNA sensing reactions was achieved.
- Although fluorescent proteins work well for constitutive expression, they are not optimal for low-cost CFTS reactions.
- Enzymatic reporters such as LacZ derivatives or Nanolantern are suitable for low-cost CFTS reactions.
- Toehold sensor performance is affected by input DNA concentration.
- In-house CFTS can be combined with preparatory isothermal amplification reactions to increase sensitivity.
- In-house CFTS can be lyophilized. However a rapid decay on activity was observed.
- CRISPRi optimized cell lysates enhance linear dsDNA stability and protein synthesis from linear templates than controls.
- CRISPRi optimized cell lysates allow the use of PCR-derived products for characterization of cell-free toehold sensors.

- The rapid prototyping strategy allowed the development of two novel PVY toe-hold sensors.

8. *Appendix*

8.1 Breakdown costs of cell-free reactions produced in the UK

Table 8.1: Cost breakdown analysis of cell extract preparation including culture, induction, and lysis as costs listing in U. of Cambridge (UK).

Item and category	**Supplier**	**Cat number**	**Q.ty**	**cost (£)**	**% cost per rxn**
2xYTG medium (YT powder)	MP Biomedicals	11357669	454g	50.16	14.57
2xYTG medium (Potassium Phosphate Dibasic 1M - K_2HPO_4)	Sigma	P8584	1L	34.8	5.92
2xYTG medium (Potassium Phosphate Monobasic 1M - KH_2PO_4)	Sigma	P8709	1L	33.92	3.17
2xYTG medium (D-glucose)	Sigma	G8270	1000g	27.14	2.08
Induction (IPTG)	Thermo Scientific	15763552	1g	34.22	34.69

Table 8.1 (continued)

Item and category	**Supplier**	**Catalog ID**	**Quantity**	**Individual cost (GBP)**	**% cost rxn**
Cell extract separation (Micro Bio-Spin Chromatography Columns)	BioRad	15763552	100	97.11	30.25
S30B Buffer (Dithiothreitol)	Sigma	D9779	5g	90.96	5.37
S30B Buffer (L-Glutamic acid hemimagnesium salt)	Sigma/Merck	49605	250g	32.81	0.1
S30B Buffer (L-Glutamic acid potassium salt)	Sigma/Merck	G1149	100g	33.15	3.81
S30B Buffer (TRIS)	Fisher Scientific	BP1521	1000g	31.83	0.04
TOTAL				**466.1**	**100**

Table 8.2: Cost breakdown analysis of cell-free reactions produced in UK using 3-PGA as energy source. 3-PGA reagent is shown highlighted in yellow.

Item	Supplier	Catalog ID	Quantity (g)	Individual cost (GBP)	% cost per rxn	Category
3-PGA	Sigma/Merck	P8877	1	205.33	23.16	Energy source
CTP	Sigma/Merck	C1506	0.1	84.15	10.15	Rxn buffer, NTPs
Coenzyme A	Sigma/Merck	C4282	0.1	220.8	7.20	Rxn buffer, Coenzyme A
GTP	Roche	10106399001	0.25	102.85	5.28	Rxn buffer, NTPs
UTP	Sigma/Merck	94370-250MG	0.25	43.43	2.33	Rxn buffer, NTPs
L-Amino acids kit	Sigma/Merck	LAA21	20	300	2.30	Rxn buffer, Aminoacids
ATP	Sigma/Merck	A8937	1	42.3	0.61	Rxn buffer, NTPs
AMPc	Sigma/Merck	A9501	1	146.2	0.59	Rxn buffer,others
Spermidine	Sigma/Merck	85558	5	126	0.49	Rxn buffer,others
NAD	Sigma/Merck	N6522	0.25	33.2	0.49	Rxn buffer,others
HEPES	Sigma/Merck	H6147	25	53.47	0.42	Rxn buffer,others

Table 8.2 (continued)

Item	**Supplier**	**Catalog ID**	**Quantity (g)**	**Individual cost (GBP)**	**% cost per rxn**	**Category**
tRNAs from *E.Coli*	Roche	10109541001	0.1	119.85	0.39	Rxn buffer,others
Folinic Acid	Sigma/Merck	F7878	0.1	63.75	0.36	Rxn buffer,others
L-Glutamic acid potassium salt	Sigma/Merck	G1149	100	33.15	0.06	Rxn buffer,others
L-Histidine	Sigma/Merck	53319	25	39.69	0.01	Rxn buffer, Aminoacids
L-Cysteine	Sigma/Merck	30089	25	43.02	0.01	Rxn buffer, Aminoacids
L-Arginine	Sigma/Merck	11009	25	25.82	0.01	Rxn buffer, Aminoacids
L-Glutamic acid hemimagnesium salt	Sigma/Merck	49605	250	32.8	0.002	Rxn buffer,others
PEG 8000	Promega	V3011	500	53.25	0	Rxn buffer,others
Crude extract preparation	In house. See Table 8.1			466	46.11	Crude extract, total
TOTAL				**2235.06**	**100**	

Table 8.3: Cost breakdown analysis of cell-free reactions produced in UK using maltodextrin and polyphosphates as energy source. Reagents specific to the maltodextrin energy solution are highlighted in yellow.

Item	**Supplier**	**Catalog ID**	**Quantity (g)**	**Individual cost (GBP)**	**% cost per rxn**	**Category**
Coenzyme A	Sigma/Merck	C4282	0.1	220.8	9.367	Rxn buffer, Coenzyme A
NAD	Sigma/Merck	N6522	0.25	33.2	0.637	Rxn buffer, others
L-amino acid kit	Sigma/Merck	LAA21	20	300	2.992	Rxn buffer, Aminoacids
L-Arginine	Sigma/Merck	11009	25	25.82	0.013	Rxn buffer, Aminoacids
L-Cysteine	Sigma/Merck	30089	25	43.02	0.015	Rxn buffer, Aminoacids
L-Histidine	Sigma/Merck	53319	25	39.69	0.018	Rxn buffer, Aminoacids
L-Glutamic acid hemimagnesium salt	Sigma/Merck	49605	250	32.81	0.003	Rxn buffer, others
L-Glutamic acid potassium salt	Sigma/Merck	G1149	100	33.15	0.08	Rxn buffer, others
Folinic Acid	Sigma/Merck	F7878	0.1	63.75	0.471	Rxn buffer, others

Table 8.3 (continued)

Item	**Supplier**	**Catalog ID**	**Quantity (g)**	**Individual cost (GBP)**	**% cost per rxn**	**Category**
GTP	Roche	10106399001	0.25	102.85	6.863	Rxn buffer, NTPs
ATP	Sigma/Merck	A8937	1	42.3	0.787	Rxn buffer, NTPs
CTP	Sigma/Merck	C1506	0.1	84.15	13.202	Rxn buffer, NTPs
AMPc	Sigma/Merck	A9501	1	146.2	0.774	Rxn buffer, others
tRNAs from *E.Coli*	Roche	10109541001	0.1	119.85	0.51	Rxn buffer, others
Sodium HMP	Sigma/Merck	305553	25	29.33	0.015	Energy source
Maltodextrin	Sigma/Merck	419672	100	30.24	0.078	Energy source
UTP	Sigma/Merck	94370	0.25	43.43	3.03	Rxn buffer, NTPs
Spermidine	Sigma/Merck	85558	5	126	0.643	Rxn buffer, others
HEPES	Sigma/Merck	H6147	25	53.47	0.542	Rxn buffer, others
PEG 8000	Promega	V3011	500	53.25	0	Rxn buffer, others

Table 8.3 (continued)

Item	**Supplier**	**Catalog ID**	**Quantity (g)**	**Individual cost (GBP)**	**% cost per rxn**	**Category**
Crude Extract	In house. See Table 8.1	NA	NA	466	59.961	Crude extract, total
TOTAL				**2089.31**	**100**	

8.2 Breakdown costs of cell-free reactions produced in Chile

Table 8.4: Cost breakdown analysis of cell extract preparation including culture, induction, and lysis as costs listing in Chile

Item and category	**Supplier**	**Cat number**	**Q.ty (g)**	**initial cost (CLP)**	**% cost per rxn**
2xYTG medium (YT powder)	MP Miomedicals	113012022	454g	56,000	11.22
2xYTG medium (Potassium Phosphate Dibasic 1M - K_2HPO_4)	Sigma	P8584-1L	1L	36,700	4.31
2xYTG medium (Potassium Phosphate Monobasic 1M - KH_2PO_4)	Sigma	P8709-1L	1L	42,400	2.74
2xYTG medium (D-glucose)	Sigma	G5767	500g	50,900	5.38
Induction (IPTG)	Promega	V3951	5g	199,500	27.89
Cell extract separation (Micro Bio-Spin Chromatography Columns)	Bio Rad	7326204	100	198,000	42.54
S30B Buffer (Dithiothreitol)	Promega	V3155	25g	310,000	2.53
S30B Buffer (L-Glutamicacid hemimagnesium salt)	Santa Cruz	SC228394	150g	47,500	0.16
S30B Buffer (L-Glutamicacid potassium salt)	Santa Cruz	SC250217	100g	39,500	3.13

Table 8.4 (continued)

Item and category	**Supplier**	**Catalog ID**	**Quantity**	**initial cost (CLP)**	% **cost per crude extract**
S30B Buffer (TRIS)	Bio Rad	BP 1 52 1	500g	67,000	0.11
TOTAL				**1,047,500**	**100**

Table 8.5: Cost breakdown analysis of cell-free reactions produced in Chile using 3-PGA as energy source. 3-PGA is highlighted in yellow.

Item and category	**Supplier**	**Cat number**	**Q.ty**	**cost (£)**	**% cost per rxn**
Coenzyme A	Sigma/Merck	C4282	0.035	194,000	9.396
NAD	Sigma/Merck	N6522	1	27,904	0.053
L-Amino acid kit	Sigma/Merck	LAA21-1KT	20	491,000	1.957
L-Cysteine	Sigma/Merck	1028380100	100	57,206	0.002
L-Histidine	Sigma/Merck	1043510100	100	103,447	0.005
L-glutamic acid hemimagnesium salt	Santa Cruz	SC228394	150	47,500	0.003
L-glutamic acid potassium salt	Santa Cruz	SC250217	100	39,500	0.072
Folinic Acid	Sigma/Merck	F7878	0.1	71,000	0.21
GTP	Sigma/Merck	A2383	0.25	136,000	3.626
ATP	Sigma/Merck	A8937	1	80,000	5.946
CTP	Affymetrix/USB	14121	0.1	195,468	12.253
AMPc	Sigma/Merck	A9501	0.1	108,000	2.283
tRNAs from *E.Coli*	Roche	MRE600	0.1	143,000	0.243
3-PGA	Sigma/Merck	P8877	1	394,000	23.095
UTP	Affymetrix/USB	23160	0.1	75,000	5.228
Spermidine	Sigma/Merck	85558	1	81,000	0.826
HEPES	Sigma/Merck	H6147	100	61,250	0.062
PEG 8000	Promega	V30111	500	59,000	0

Table 8.1 (continued)

Item and category	**Supplier**	**Catalog ID**	**Quantity**	**Individual cost (GBP)**	% **cost rxn**
Crude extract	In-house (see Table 8.4)	NA	NA	1,047,500	34.74
TOTAL				**3,411,775**	**100**

Table 8.6: Cost breakdown analysis of cell-free reactions produced in Chile using maltodextrin and polyphosphates as energy source.Reagents specific to the maltodextrin energy solution are highlighted in yellow.

Item	**Supplier**	**Cat number**	**Q.ty (g)**	**initial cost (CLP)**	**% cost per rxn**
Coenzyme A	Sigma/Merck	C4282	0.035	194,000	12.396
NAD	Sigma/Merck	N6522	1	27,904	0.071
L-Amino acid kit	Sigma/Merck	LAA21-1KT	20	491,000	2.582
L-Cysteine	Sigma/Merck	1028380100	100	57,206	0.003
L-Histidine	Sigma/Merck	1043510100	100	103,447	0.006
L-glutamic acid hemimagnesium salt	Santa Cruz	SC228394	150	47,500	0.002
L-glutamic acid potassium salt	Santa Cruz	SC250217	100	39,500	0.05
Folinic Acid	Sigma/Merck	F7878	0.1	71,000	0.277
GTP	Sigma/Merck	A2383	0.25	136,000	4.784
ATP	Sigma/Merck	A8937	1	80,000	7.845
CTP	Affymetrix/USB	14121	0.1	195,468	16.165
AMPc	Sigma/Merck	A9501	0.1	108,000	3.012
tRNAs from *E.Coli*	Roche	MRE600	0.1	143,000	0.321
Sodium HMP	Sigma/Merck	305553	25	30,553	0.008
Maltodextrin	Sigma/Merck	419672	100	55,000	0.075
UTP	Affymetrix/USB	23160	0.1	58,733	5.401

Table 8.6 (continued)

Item	**Supplier**	**Catalog ID**	**Quantity**	**initial cost (CLP)**	**% cost per rxn**
Spermidine	Sigma/Merck	85558	1	81,000	1.089
HEPES	Sigma/Merck	H6147	100	61,250	0.082
PEG 8000	Promega	V30111	500	59,000	0
Crude extract	In-house (see Table 8.4)	NA	NA	1,047,500	45.83
TOTAL				**3,087,061**	**100**

Bibliography

[1] Alec A. K. Nielsen, Bryan S. Der, Jonghyeon Shin, Prashant Vaidyanathan, Vanya Paralanov, Elizabeth A. Strychalski, David Ross, Douglas Densmore, and Christopher A. Voigt. Genetic circuit design automation. *Science*, 352(6281):aac7341, 4 2016.

[2] Drew Endy. Foundations for engineering biology. *Nature*, 438(7067):449–453, 2005.

[3] Jeffrey C. Way, James J. Collins, Jay D. Keasling, and Pamela A. Silver. Integrating biological redesign: Where synthetic biology came from and where it needs to go. *Cell*, 157(1):151–161, 2020/11/23 2014.

[4] M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–8, 1 2000.

[5] T. S. Gardner, C. R. Cantor, and J. J. Collins. Construction of a genetic toggle switch in escherichia coli. *Nature*, 403(6767):339–42, 2000.

[6] Jerome Bonnet, Peter Yin, Monica E. Ortiz, Pakpoom Subsoontorn, and Drew Endy. Amplifying genetic logic gates. *Science*, 340(6132):599–603, 2013.

[7] Maung Nyan Win and Christina D. Smolke. Higher-order cellular information processing with synthetic rna devices. *Science*, 322(5900):456–460, 2008.

[8] Keller Rinaudo, Leonidas Bleris, Rohan Maddamsetti, Sairam Subramanian, Ron Weiss, and Yaakov Benenson. A universal rnai-based logic evaluator that operates in mammalian cells. *Nature Biotechnology*, 25(7):795–801, 2007.

[9] Alexander A. Green, Jongmin Kim, Duo Ma, Pamela A. Silver, James J. Collins, and Peng Yin. Complex cellular logic computation using ribocomputing devices. *Nature*, 548(7665):117–121, 2017.

[10] James Chappell, Melissa K. Takahashi, Sarai Meyer, David Loughrey, Kyle E. Watters, and Julius Lucks. The centrality of rna for engineering gene expression. *Biotechnology Journal*, 8(12):1379–1395, 2013.

[11] Alexander A. Green, Pamela A. Silver, James J. Collins, and Peng Yin. Toehold Switches: De-Novo-Designed Regulators of Gene Expression. *Cell*, 159(4):925–939, November 2014.

[12] Kyle E Watters, Eric J Strobel, Angela M Yu, John T Lis, and Julius B Lucks. Cotranscriptional folding of a riboswitch at nucleotide resolution. *Nat Struct Mol Biol*, 23(12):1124–1131, December 2016.

[13] Joseph N. Zadeh, Conrad D. Steenberg, Justin S. Bois, Brian R. Wolfe, Marshall B. Pierce, Asif R. Khan, Robert M. Dirks, and Niles A. Pierce. NUPACK: Analysis and design of nucleic acid systems. *J. Comput. Chem.*, 32(1):170–173, January 2011.

[14] Keith Pardee, Alexander A. Green, Tom Ferrante, D. Ewen Cameron, Ajay DaleyKeyser, Peng Yin, and James J. Collins. Paper-Based Synthetic Gene Networks. *Cell*, 159(4):940–954, November 2014.

[15] Keith Pardee, Alexander A. Green, Melissa K. Takahashi, Dana Braff, Guillaume Lambert, Jeong Wook Lee, Tom Ferrante, Duo Ma, Nina Donghia, Melina Fan, Nichole M. Daringer, Irene Bosch, Dawn M. Dudley, David H. O'Connor, Lee Gehrke, and James J. Collins. Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell*, 165(5):1255–1266, May 2016.

[16] Duo Ma, Luhui Shen, Kaiyue Wu, Chris W Diehnelt, and Alexander A Green. Low-cost detection of norovirus using paper-based cell-free systems and synbody-based viral enrichment. *Synthetic Biology*, 3(1), January 2018.

[17] Yoshihiro Shimizu, Akio Inoue, Yukihide Tomari, Tsutomu Suzuki, Takashi Yokogawa, Kazuya Nishikawa, and Takuya Ueda. Cell-free translation reconstituted with purified components. *Nat Biotechnol*, 19(8):751–755, August 2001.

[18] Keith Pardee, Shimyn Slomovic, Peter Q. Nguyen, Jeong Wook Lee, Nina Donghia, Devin Burrill, Tom Ferrante, Fern R. McSorley, Yoshikazu Furuta, Andyna Vernet, Michael Lewandowski, Christopher N. Boddy, Neel S. Joshi, and James J. Collins. Portable, On-Demand Biomolecular Manufacturing. *Cell*, 167(1):248–259.e12, September 2016.

[19] Adam D. Silverman, Nancy Kelley-Loughnane, Julius B. Lucks, and Michael C. Jewett. Deconstructing Cell-Free Extract Preparation for *in Vitro* Activation of Transcriptional Genetic Circuitry. *ACS Synth. Biol.*, 8(2):403–414, February 2019.

[20] Marvin R. Lamborg and Paul C. Zamecnik. Amino acid incorporation into protein by extracts of E. coli. *Biochimica et Biophysica Acta*, 42:206–211, January 1960.

[21] Marshall W. Nirenberg and J. Heinrich Matthaei. The Dependence of Cell- Free Protein Synthesis in E. coli upon Naturally Occurring or Synthetic Polyribonucleotides. *Proceedings of the National Academy of Sciences of the United States of America*, 47(10):1588–1602, 1961.

[22] Muriel Lederman and Geoffrey Zubay. DNA-directed peptide synthesis I. A comparison of T2 and Escherichia coli DNA-directed peptide synthesis in two cell-free systems. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis*, 149(1):253–258, November 1967.

[23] M. Nirenberg and P. Leder. RNA Codewords and Protein Synthesis: The Effect of Trinucleotides upon the Binding of sRNA to Ribosomes. *Science*, 145(3639):1399–1407, September 1964.

[24] M. Nirenberg, T. Caskey, R. Marshall, R. Brimacombe, D. Kellogg, B. Doctor, D. Hatfield, J. Levin, F. Rottman, S. Pestka, M. Wilcox, and F. Anderson. The RNA Code and Protein Synthesis. *Cold Spring Harbor Symposia on Quantitative Biology*, 31(0):11–24, January 1966.

[25] Adam D. Silverman, Ashty S. Karim, and Michael C. Jewett. Cell-free gene expression: An expanded repertoire of applications. *Nat Rev Genet*, 21(3):151–170, March 2020.

[26] Nadanai Laohakunakorn, Laura Grasemann, Barbora Lavickova, Grégoire Michielin, Amir Shahein, Zoe Swank, and Sebastian J. Maerkl. Bottom-Up Construction of Complex Biomolecular Systems With Cell-Free Synthetic Biology. *Front. Bioeng. Biotechnol.*, 8:213, March 2020.

[27] Aaron R. Goerke and James R. Swartz. Development of cell-free protein synthesis platforms for disulfide bonded proteins. *Biotechnol. Bioeng.*, 99(2):351–367, February 2008.

[28] Jessica J. Wuu and James R. Swartz. High yield cell-free production of integral membrane proteins without refolding or detergents. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1778(5):1237–1250, May 2008.

[29] Doreen Matthies, Stefan Haberstock, Friederike Joos, Volker Dötsch, Janet Vonck, Frank Bernhard, and Thomas Meier. Cell-Free Expression and Assembly of ATP Synthase. *Journal of Molecular Biology*, 413(3):593–603, October 2011.

[30] Bradley C. Bundy, Marc J. Franciszkowicz, and James R. Swartz. Escherichia coli-based cell-free synthesis of virus-like particles. *Biotechnol. Bioeng.*, 100(1):28–37, May 2008.

[31] Mark Rustad, Allen Eastlund, Paul Jardine, and Vincent Noireaux. Cell-free TXTL synthesis of infectious bacteriophage T4 in a single test tube reaction. *Synthetic Biology*, 3(1), January 2018.

[32] Jonghyeon Shin and Vincent Noireaux. An *E. coli* Cell-Free Expression Toolbox: Application to Synthetic Gene Circuits and Artificial Cells. *ACS Synth. Biol.*, 1(1):29–41, January 2012.

[33] Razvan Nutiu and Yingfu Li. Aptamers with fluorescence-signaling properties. *Methods*, 37(1):16–25, September 2005.

[34] Jeremy R. Babendure, Stephen R. Adams, and Roger Y. Tsien. Aptamers Switch on Fluorescence of Triphenylmethane Dyes. *J. Am. Chem. Soc.*, 125(48):14716–14717, December 2003.

[35] Bo Zhu, Rui Gan, Maria D. Cabezas, Takaaki Kojima, Robert Nicol, Michael C. Jewett, and Hideo Nakano. Increasing cell-free gene expression yields from linear templates in *Escherichia coli* and *Vibrio natriegens* extracts by using DNA-binding proteins. *Biotechnology and Bioengineering*, page bit.27538, August 2020.

[36] Maaruthy Yelleswarapu, Ardjan J. van der Linden, Bob van Sluijs, Pascal A. Pieters, Emilien Dubuc, Tom F. A. de Greef, and Wilhelm T. S. Huck. Sigma Factor-Mediated Tuning of Bacterial Cell-Free Synthetic Genetic Oscillators. *ACS Synth. Biol.*, 7(12):2879–2887, December 2018.

[37] Alexandra M. Tayar, Eyal Karzbrun, Vincent Noireaux, and Roy H. Bar-Ziv. Synchrony and pattern formation of coupled genetic oscillators on a chip of artificial cells. *Proc Natl Acad Sci USA*, 114(44):11609–11614, October 2017.

[38] Henrike Niederholtmeyer, Zachary Z Sun, Yutaka Hori, Enoch Yeung, Amanda Verpoorte, Richard M Murray, and Sebastian J Maerkl. Rapid cell-free forward engineering of novel genetic ring oscillators. *eLife*, 4:e09771, October 2015.

[39] Amin S.M. Salehi, Seung Ook Yang, Conner C. Earl, Miriam J. Shakalli Tang, J. Porter Hunt, Mark T. Smith, David W. Wood, and Bradley C. Bundy. Biosensing estrogenic endocrine disruptors in human blood and urine: A RAPID cell-free protein synthesis approach. *Toxicology and Applied Pharmacology*, 345:19–25, April 2018.

[40] Walter Thavarajah, Adam D. Silverman, Matthew S. Verosloff, Nancy Kelley-Loughnane, Michael C. Jewett, and Julius B. Lucks. Point-of-Use Detection of Environmental Fluoride *via* a Cell-Free Riboswitch-Based Biosensor. *ACS Synth. Biol.*, 9(1):10–18, January 2020.

[41] Adam D. Silverman, Umut Akova, Khalid K. Alam, Michael C. Jewett, and Julius B. Lucks. Design and Optimization of a Cell-Free Atrazine Biosensor. *ACS Synth. Biol.*, 9(3):671–677, March 2020.

[42] Sisko Tauriainen, Marko Virta, Wei Chang, and Matti Karp. Measurement of Firefly Luciferase Reporter Gene Activity from Cells and Lysates Using Escherichia coli Arsenite and Mercury Sensors. *Analytical Biochemistry*, 272(2):191–198, August 1999.

[43] Jaeyoung K. Jung, Khalid K. Alam, Matthew S. Verosloff, Daiana A. Capdevila, Morgane Desmau, Phillip R. Clauer, Jeong Wook Lee, Peter Q. Nguyen, Pablo A. Pastén, Sandrine J. Matiasek, Jean-François Gaillard, David P. Giedroc, James J. Collins, and Julius B. Lucks. Cell-free biosensors for rapid detection of water contaminants. *Nat Biotechnol*, July 2020.

[44] Monica P. McNerney, Yan Zhang, Paige Steppe, Adam D. Silverman, Michael C. Jewett, and Mark P. Styczynski. Point-of-care biomarker quantification enabled by sample-specific calibration. *Sci. Adv.*, 5(9):eaax4473, September 2019.

[45] Filippo Caschera and Vincent Noireaux. Synthesis of 2.3 mg/ml of protein with an all Escherichia coli cell-free transcription–translation system. *Biochimie*, 99:162–168, April 2014.

[46] David V. Liu, James F. Zawada, and James R. Swartz. Streamlining Escherichia Coli S30 Extract Preparation for Economical Cell-Free Protein Synthesis. *Biotechnol Progress*, 21(2):460–465, September 2008.

[47] Prashanta Shrestha, Troy Michael Holland, and Bradley Charles Bundy. Streamlined extract preparation for *Escherichia coli* -based cell-free protein synthesis by sonication or bead vortex mixing. *BioTechniques*, 53(3):163–174, September 2012.

[48] Yong-Chan Kwon and Michael C. Jewett. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Sci Rep*, 5(1):8663, August 2015.

[49] Zachary Z. Sun, Clarmyra A. Hayes, Jonghyeon Shin, Filippo Caschera, Richard M. Murray, and Vincent Noireaux. Protocols for Implementing an Escherichia coli Based TX-TL Cell-Free Expression System for Synthetic Biology. *JoVE*, 79:e50762, September 2013.

[50] Andriy Didovyk, Taishi Tonooka, Lev Tsimring, and Jeff Hasty. Rapid and Scalable Preparation of Bacterial Lysates for Cell-Free Gene Expression. *ACS Synth. Biol.*, 6(12):2198–2208, December 2017.

[51] Kei Fujiwara and Nobuhide Doi. Biochemical Preparation of Cell Extract for Cell-Free Protein Synthesis without Physical Disruption. *PLoS ONE*, 11(4):e0154614, April 2016.

[52] Kristen M. Wilding, Emily Long Zhao, Conner C. Earl, and Bradley C. Bundy. Thermostable lyoprotectant-enhanced cell-free protein synthesis for on-demand endotoxin-free therapeutic production. *New Biotechnology*, 53:73–80, November 2019.

[53] Xiaocui Guo, Yi Zhu, Lihui Bai, and Dayong Yang. The Protection Role of Magnesium Ions on Coupled Transcription and Translation in Lyophilized Cell-Free System. *ACS Synth. Biol.*, 9(4):856–863, April 2020.

[54] Mark Thomas Smith and Scott D. Berkheimer. Lyophilized *Escherichia coli* -based cell-free systems for robust, high-density, long-term storage. *BioTechniques*, 56(4), April 2014.

[55] Quentin M. Dudley, Kim C. Anderson, and Michael C. Jewett. Cell-Free Mixing of *Escherichia coli* Crude Extracts to Prototype and Rationally Engineer High-Titer Mevalonate Synthesis. *ACS Synth. Biol.*, 5(12):1578–1588, December 2016.

[56] Berg JM, Tymoczko JL, Stryer L. *Biochemistry.* W. H. Freeman, 5th edition edition, 2002.

[57] Kara A. Calhoun and James R. Swartz. Energizing cell-free protein synthesis with glucose metabolism. *Biotechnol. Bioeng.*, 90(5):606–613, June 2005.

[58] Tae-Wan Kim, Jung-Won Keum, In-Seok Oh, Cha-Yong Choi, Chang-Gil Park, and Dong-Myung Kim. Simple procedures for the construction of a robust and cost-effective cell-free protein synthesis system. *Journal of Biotechnology*, 126(4):554–561, December 2006.

[59] Ho-Cheol Kim and Dong-Myung Kim. Methods for energizing cell-free protein synthesis. *Journal of Bioscience and Bioengineering*, 108(1):1–4, July 2009.

[60] H Kondo, I Tomioka, H Nakajima, and K Imahori. Construction of a system for the regeneration of adenosine 5'-triphosphate, which supplies energy to bioreactor. *J Appl Biochem*, 6(1-2):29–38, 1984.

[61] Debbie C. Crans, Romas J. Kazlauskas, Bernard L. Hirschbein, Chi-Huey Wong, Obsidiana Abril, and George M. Whitesides. Enzymatic regeneration of adenosine 5′-triphosphate: Acetyl phosphate, phosphoenolpyruvate, methoxycarbonyl phosphate, dihydroxyacetone phosphate, 5-phospho-α-d-ribosyl pyrophosphate, uridine-5′-diphosphoglucose. In *Methods in Enzymology*, volume 136 of *Immobilized Enzymes and Cells, Part C*, pages 263–280. Academic Press, January 1987.

[62] Yen-Shiang Shih and George M. Whitesides. Large-scale ATP-requiring enzymic phosphorylation of creatine can be driven by enzymic ATP regeneration. *J. Org. Chem.*, 42(25):4165–4166, December 1977.

[63] Kalavathy Sitaraman, Dominic Esposito, George Klarmann, Stuart F Le Grice, James L Hartley, and Deb K Chatterjee. A novel cell-free protein synthesis system. *Journal of Biotechnology*, 110(3):257–263, June 2004.

[64] Michael C Jewett, Kara A Calhoun, Alexei Voloshin, Jessica J Wuu, and James R Swartz. An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol Syst Biol*, 4(1):220, January 2008.

[65] Tae-Wan Kim, Ho-Cheol Kim, In-Seok Oh, and Dong-Myung Kim. A highly efficient and economical cell-free protein synthesis system using the S12 extract of Escherichia coli. *Biotechnol Bioproc E*, 13(4):464–469, August 2008.

[66] Ho-Cheol Kim, Tae-Wan Kim, and Dong-Myung Kim. Prolonged production of proteins in a cell-free protein synthesis system using polymeric carbohydrates as an energy source. *Process Biochemistry*, 46(6):1366–1369, June 2011.

[67] Filippo Caschera and Vincent Noireaux. A cost-effective polyphosphate-based metabolism fuels an all E. coli cell-free expression system. *Metabolic Engineering*, 27:29–37, January 2015.

[68] Alexander V. Karasev and Stewart M. Gray. Continuous and Emerging Challenges of *Potato virus Y* in Potato. *Annual Review of Phytopathology*, 51(1):571–586, August 2013.

[69] Edward B. Radcliffe and David W. Ragsdale. Aphid-transmitted potato viruses: The importance of understanding vector biology. *American Journal of Potato Research*, 79(5):353–386, September 2002.

[70] Manphool Fageria, Xianzhou Nie, Angela Gallagher, and Mathuresh Singh. Mechanical Transmission of Potato Virus Y (PVY) Through Seed Cutting and Plant Wounding. *American Journal of Potato Research*, 92(1):143–147, February 2015.

[71] Dennis Halterman, Amy Charkowski, and Jeanmarie Verchot. Potato, Viruses, and Seed Certification in the USA to Provide Healthy Propagated Tubers. *Pest Technology*, 6(1):1–14, 2012.

[72] L. Glais, M. Tribodet, and C. Kerlan. Genomic variability in Potato potyvirus Y (PVY): Evidence that PVY N W and PVY NTN variants are single to multiple recombinants between PVY O and PVY N isolates. *Archives of Virology*, 147(2):363–378, February 2002.

[73] Kelsie J. Green, Celeste J. Brown, Stewart M. Gray, and Alexander V. Karasev. Phylogenetic study of recombinant strains of Potato virus Y. *Virology*, 507:40–52, July 2017.

[74] Alexander A. Green, Jongmin Kim, Duo Ma, Pamela A. Silver, James J. Collins, and Peng Yin. Complex cellular logic computation using ribocomputing devices. *Nature*, 548(7665):117–121, August 2017.

[75] Kelsie J. Green, Celeste J. Brown, and Alexander V. Karasev. Genetic diversity of potato virus Y (PVY): Sequence analyses reveal ten novel PVY recombinant structures. *Archives of Virology*, 163(1):23–32, January 2018.

[76] James M. Crosslin. PVY: An Old Enemy and A Continuing Challenge. *American Journal of Potato Research*, 90(1):2–6, February 2013.

[77] M. Chikh Ali, T. Maoka, and K. T. Natsuaki. Whole Genome Sequence and Characterization of a Novel Isolate of PVY Inducing Tuber Necrotic Ringspot in Potato and Leaf Mosaic in Tobacco. *Journal of Phytopathology*, 156(7-8):413–418, August 2008.

[78] Kazusato Ohshima, Kazuya Sako, Chikako Hiraishi, Akio Nakagawa, Kazutoshi Matsuo, Tetsuji Ogawa, Eishiro Shikata, and Nobumichi Sako. Potato Tuber Necrotic Ringspot Disease Occurring in Japan: Its Association with *Potato virus Y* Necrotic Strain. *Plant Disease*, 84(10):1109–1115, October 2000.

[79] J G McDonald and R P Singh. RESPONSE OF POTATO CULTIVARS T O N O R T H AMERIC˜kN ISOLATES OF PVYNTN! *American Potato Journal*, 73:7, 1996.

[80] H. K. Were, J. N. Kabira, Z. M. Kinyua, F. M. Olubayo, J. K. Karinga, J. Aura, A. K. Lees, G. H. Cowan, and L. Torrance. Occurrence and Distribution of Potato Pests and Diseases in Kenya. *Potato Research*, 56(4):325–342, December 2013.

[81] M. Chrzanowska. New isolates of the necrotic strain of potato virus Y (PVYN) found recently in Poland. *Potato Research*, 34(2):179–182, June 1991.

[82] Victoriano Ramírez-Rodríguez, Katia Aviña-Padilla, Gustavo Frías-Treviño, Laura Silva-Rosales, and Juan Martínez-Soriano. Presence of necrotic strains of Potato virus Y in Mexican potatoes. *Virology Journal*, 6(1):48, 2009.

[83] Alexander V. Karasev and Stewart M. Gray. Genetic Diversity of Potato virus Y Complex. *American Journal of Potato Research*, 90(1):7–13, February 2013.

[84] Sonya V. Iverson, Traci L. Haddock, Jacob Beal, and Douglas M. Densmore. CIDAR MoClo: Improved MoClo Assembly Standard and New *E. coli* Part Library Enable Rapid Combinatorial

Design for Synthetic and Traditional Biology. *ACS Synthetic Biology*, 5(1):99–103, January 2016.

[85] A. Sarrion-Perdigones, M. Vazquez-Vilar, J. Palaci, B. Castelijns, J. Forment, P. Ziarsolo, J. Blanca, A. Granell, and D. Orzaez. GoldenBraid 2.0: A Comprehensive DNA Assembly Framework for Plant Synthetic Biology. *PLANT PHYSIOLOGY*, 162(3):1618–1631, July 2013.

[86] Daniel G Gibson, Lei Young, Ray-Yuan Chuang, J Craig Venter, Clyde A Hutchison, and Hamilton O Smith. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*, 6(5):343–345, May 2009.

[87] Isaac Nuñez, Tamara Matute, Roberto Herrera, Juan Keymer, Timothy Marzullo, Timothy Rudge, and Fernán Federici. Low cost and open source multi-fluorescence imaging system for teaching and research in biology and bioengineering. *PLOS ONE*, 12(11):e0187163, November 2017.

[88] R. A. Ikeda and C. C. Richardson. Interactions of the RNA polymerase of bacteriophage T7 with its promoter during binding and initiation of transcription. *Proceedings of the National Academy of Sciences*, 83(11):3614–3618, June 1986.

[89] Karsten Temme, Rena Hill, Thomas H. Segall-Shapiro, Felix Moser, and Christopher A. Voigt. Modular control of multiple pathways using engineered orthogonal T7 polymerases. *Nucleic Acids Research*, 40(17):8773–8781, September 2012.

[90] Jean-Denis Pédelacq, Stéphanie Cabantous, Timothy Tran, Thomas C Terwilliger, and Geoffrey S Waldo. Engineering and characterization of a superfolder green fluorescent protein. *Nature Biotechnology*, 24(1):79–88, January 2006.

[91] Jonghyeon Shin and Vincent Noireaux. Efficient cell-free expression with the endogenous E. Coli RNA polymerase and sigma factor 70. *Journal of Biological Engineering*, 4(1):1–9, 2010.

[92] Akira Takai, Masahiro Nakano, Kenta Saito, Remi Haruno, Tomonobu M. Watanabe, Tatsuya Ohyanagi, Takashi Jin, Yasushi Okada, and Takeharu Nagai. Expanded palette of Nano-lanterns for real-time multicolor luminescence imaging. *Proceedings of the National Academy of Sciences*, 112(14):4352–4356, April 2015.

[93] Josefine Liljeruhm, Saskia K. Funk, Sandra Tietscher, Anders D. Edlund, Sabri Jamal, Pikkei Wistrand-Yuen, Karl Dyrhage, Arvid Gynnå, Katarina Ivermark, Jessica Lövgren, Viktor Törnblom, Anders Virtanen, Erik R. Lundin, Erik Wistrand-Yuen, and Anthony C. Forster. Engineering a palette of eukaryotic chromoproteins for bacterial synthetic biology. *Journal of Biological Engineering*, 12(1):8, December 2018.

[94] Joseph P Torella, Florian Lienert, Christian R Boehm, Jan-Hung Chen, Jeffrey C Way, and Pamela A Silver. Unique nucleotide sequence–guided assembly of repetitive DNA parts for synthetic biology applications. *Nature Protocols*, 9(9):2075–2089, September 2014.

[95] Filippo Caschera and Vincent Noireaux. Preparation of amino acid mixtures for cell-free expression systems. *BioTechniques*, 58(1), January 2015.

[96] Filippo Caschera and Vincent Noireaux. A cost-effective polyphosphate-based metabolism fuels an all E. coli cell-free expression system. *Metabolic Engineering*, 27:29–37, January 2015.

[97] Pascal J. Lopez, Isabelle Marchand, Susan A. Joyce, and Marc Dreyfus. The C-terminal half of RNase E, which organizes the Escherichia coli degradosome, participates in mRNA degradation but not rRNA processing in vivo. *Mol Microbiol*, 33(1):188–199, July 1999.

[98] Vijayalakshmi H Nagaraj, James M Greene, Anirvan M Sengupta, and Eduardo D Sontag. Translation inhibition and resource balance in the TX-TL cell-free gene expression system. *Synthetic Biology*, 2(1):ysx005, January 2017.

[99] Michael Chamberlin and Janet Ring. Characterization of T7-specific Ribonucleic Acid Polymerase. *Int. J. Biol. Chem.*, September 1972.

[100] Raymond F. Gesteland. Unfolding of Escherichia coli ribosomes by removal of magnesium. *Journal of Molecular Biology*, 18(2):356–IN14, July 1966.

[101] Anton S. Petrov, Chad R. Bernier, Chiaolong Hsiao, C. Denise Okafor, Emmanuel Tannenbaum, Joshua Stern, Eric Gaucher, Dana Schneider, Nicholas V. Hud, Stephen C. Harvey, and Loren Dean Williams. RNA–Magnesium–Protein Interactions in Large Ribosomal Subunit. *J. Phys. Chem. B*, 116(28):8113–8120, July 2012.

[102] Julian Gordon and Fritz Lipmann. Role of divalent ions in poly U-directed phenylalanine polymerization. *Journal of Molecular Biology*, 23(1):23–33, January 1967.

[103] H. L. Yang, L. Ivashkiv, H. Z. Chen, G. Zubay, and M. Cashel. Cell-free coupled transcription-translation system for investigation of linear DNA segments. *Proceedings of the National Academy of Sciences*, 77(12):7029–7033, December 1980.

[104] Zachary Z. Sun, Enoch Yeung, Clarmyra A. Hayes, Vincent Noireaux, and Richard M. Murray. 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, June 2014.

[105] Eiko Seki, Natsuko Matsuda, Shigeyuki Yokoyama, and Takanori Kigawa. Cell-free protein synthesis system from Escherichia coli cells cultured at decreased temperatures improves productivity by decreasing DNA template degradation. *Analytical Biochemistry*, 377(2):156–161, June 2008.

[106] C L Bassett and J R Rawson. In vitro coupled transcription-translation of linear DNA fragments in a lysate derived from a recB rna pnp strain of Escherichia coli. *Journal of Bacteriology*, 156(3):1359–1362, 1983.

[107] Nathalie Michel-Reydellet, Kim Woodrow, and James Swartz. Increasing PCR Fragment Stability and Protein Yields in a Cell-Free System with Genetically Modified *Escherichia coli* Extracts. *J Mol Microbiol Biotechnol*, 9(1):26–34, 2005.

[108] Mark S. Dillingham and Stephen C. Kowalczykowski. RecBCD Enzyme and the Repair of Double-Stranded DNA Breaks. *MMBR*, 72(4):642–671, December 2008.

[109] I. R. Lehman, G. G. Roussos, and E. A. Pratt. The Deoxyribonucleases of Escherichia coli. *Int. J. Biol. Chem.*, 233(3), 1962.

[110] Kalavathy Sitaraman, Dominic Esposito, George Klarmann, Stuart F Le Grice, James L Hartley, and Deb K Chatterjee. A novel cell-free protein synthesis system. *Journal of Biotechnology*, 110(3):257–263, June 2004.

[111] Thomas Hoffmann, Cordula Nemetz, Regina Schweizer, Wolfgang Mutter, and Mandred Watzele. High-Level Cell-Free Protein Expression from PCR-Generated DNA Templates. In *Cell-Free Translation Systems*. Spirin A.S. eds, Springer, Berlin, Heidelberg, December 2001.

[112] Ryan Marshall, Colin S. Maxwell, Scott P. Collins, Chase L. Beisel, and Vincent Noireaux. Short DNA containing χ sites enhances DNA stability and gene expression in *E. coli* cell-free transcription-translation systems: Enhancing TXTL-Based Expression With χ-Site DNA. *Biotechnol. Bioeng.*, 114(9):2137–2141, September 2017.

[113] Seok Hoon Hong, Yong-Chan Kwon, Rey W. Martin, Benjamin J. Des Soye, Alexandra M. de Paz, Kirsten N. Swonger, Ioanna Ntai, Neil L. Kelleher, and Michael C. Jewett. Improving Cell-Free Protein Synthesis through Genome Engineering of *Escherichia coli* Lacking Release Factor 1. *ChemBioChem*, 16(5):844–853, March 2015.

[114] Lei S. Qi, Matthew H. Larson, Luke A. Gilbert, Jennifer A. Doudna, Jonathan S. Weissman, Adam P. Arkin, and Wendell A. Lim. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*, 152(5):1173–1183, February 2013.

[115] Matthew H Larson, Luke A Gilbert, Xiaowo Wang, Wendell A Lim, Jonathan S Weissman, and Lei S Qi. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc*, 8(11):2180–2196, November 2013.

[116] Oytun Portakal and Pakize Dogan. Construction of recBrecD genetic fusion and functional analysis of RecBDC fusion enzyme in Escherichia coli. *BMC Biochem*, 9(1):27, 2008.

[117] Isaac N. Nuñez, Tamara F. Matute, Ilenne D. Del Valle, Anton Kan, Atri Choksi, Drew Endy, Jim Haseloff, Timothy J. Rudge, and Fernan Federici. Artificial Symmetry-Breaking for Morphogenetic Engineering Bacterial Colonies. *ACS Synth. Biol.*, 6(2):256–265, February 2017.

[118] Teijo Pellinen, Tuomas Huovinen, and Matti Karp. A cell-free biosensor for the detection of transcriptional inducers using firefly luciferase as a reporter. *Analytical Biochemistry*, 330(1):52–57, July 2004.

[119] Ju-Young Byun, Kyung-Ho Lee, Yong-Beom Shin, and Dong-Myung Kim. Cascading Amplification of Immunoassay Signal by Cell-Free Expression of Firefly Luciferase from Detection Antibody-Conjugated DNA in an *Escherichia coli* Extract. *ACS Sens.*, 4(1):93–99, January 2019.

[120] Kazushi Suzuki, Taichi Kimura, Hajime Shinoda, Guirong Bai, Matthew J. Daniels, Yoshiyuki Arai, Masahiro Nakano, and Takeharu Nagai. Five colour variants of bright luminescent protein for real-time multicolour bioimaging. *Nat Commun*, 7(1):13718, December 2016.

[121] Luis E. Contreras-Llano, Conary Meyer, Yao Liu, Mridul Sarker, Sierin Lim, Marjorie L. Longo, and Cheemeng Tan. Holistic engineering of cell-free systems through proteome-reprogramming synthetic circuits. *Nat Commun*, 11(1):3138, December 2020.

[122] Barbora Lavickova and Sebastian J. Maerkl. A Simple, Robust, and Low-Cost Method To Produce the PURE Cell-Free System. *ACS Synth. Biol.*, 8(2):455–462, February 2019.

[123] Jaeyoung K. Jung, Khalid K. Alam, Matthew S. Verosloff, Daiana A. Capdevila, Morgane Desmau, Phillip R. Clauer, Jeong Wook Lee, Peter Q. Nguyen, Pablo A. Pastén, Sandrine J. Matiasek, Jean-François Gaillard, David P. Giedroc, James J. Collins, and Julius B. Lucks. Cell-free biosensors for rapid detection of water contaminants. *Nat Biotechnol*, July 2020.

[124] J. Compton. Nucleic acid sequence-based amplification. *Nature*, 350(6313):91–92, 1991.

[125] James P. Broughton, Xianding Deng, Guixia Yu, Clare L. Fasching, Venice Servellita, Jasmeet Singh, Xin Miao, Jessica A. Streithorst, Andrea Granados, Alicia Sotomayor-Gonzalez, Kelsey Zorn, Allan Gopez, Elaine Hsu, Wei Gu, Steve Miller, Chao-Yang Pan, Hugo Guevara, Debra A. Wadford, Janice S. Chen, and Charles Y. Chiu. CRISPR–Cas12-based detection of SARS-CoV-2. *Nature Biotechnology*, 38(7):870–874, July 2020.

[126] Maturada Patchsung, Krittapas Jantarug, Archiraya Pattama, Kanokpol Aphicho, Surased Suraritdechachai, Piyachat Meesawat, Khomkrit Sappakhaw, Nattawat Leelahakorn, Theerawat Ruenkam, Thanakrit Wongsatit, Niracha Athipanyasilp, Bhumrapee Eiamthong, Benya Lakkanasirorat, Thitima Phoodokmai, Nootaree Niljianskul, Danaya Pakotiprapha, Sittinan Chanarat, Aimorn Homchan, Ruchanok Tinikul, Philaiwarong Kamutira, Kochakorn Phiwkaow, Sahachat Soithongcharoen, Chadaporn Kantiwiriyawanitch, Vinutsada Pongsupasa, Duangthip Trisrivirat, Juthamas Jaroensuk, Thanyaporn Wongnate, Somchart Maenpuen, Pimchai Chaiyen, Sirichai Kamnerdnakta, Jirawat Swangsri, Suebwong Chuthapisith, Yongyut Sirivatanauksorn, Chutikarn Chaimayo, Ruengpung Sutthent, Wannee Kantakamalakul, Julia Joung, Alim Ladha, Xin Jin, Jonathan S. Gootenberg, Omar O. Abudayyeh, Feng Zhang,

Navin Horthongkham, and Chayasith Uttamapinant. Clinical validation of a Cas13-based assay for the detection of SARS-CoV-2 RNA. *Nature Biomedical Engineering*, August 2020.

[127] Fan Hong, Duo Ma, Kaiyue Wu, Lida A. Mina, Rebecca C. Luiten, Yan Liu, Hao Yan, and Alexander A. Green. Precise and Programmable Detection of Mutations Using Ultraspecific Riboregulators. *Cell*, 180(5):1018–1032.e16, March 2020.

[128] Jacqueline A. Valeri, Katherine M. Collins, Pradeep Ramesh, Miguel A. Alcantar, Bianca A. Lepe, Timothy K. Lu, and Diogo M. Camacho. Sequence-to-function deep learning frameworks for engineered riboregulators. *Nature Communications*, 11(1):5058, December 2020.

[129] Nicolaas M. Angenent-Mari, Alexander S. Garruss, Luis R. Soenksen, George Church, and James J. Collins. A deep learning approach to programmable RNA switches. *Nature Communications*, 11(1):5057, December 2020.

[130] Fabio Chizzolini, Michele Forlin, Noël Yeh Martín, Giuliano Berloffa, Dario Cecchi, and Sheref S. Mansy. Cell-Free Translation Is More Variable than Transcription. *ACS Synthetic Biology*, 6(4):638–647, April 2017.

[131] Stephanie D. Cole, Kathryn Beabout, Kendrick B. Turner, Zachary K. Smith, Vanessa L. Funk, Svetlana V. Harbaugh, Alvin T. Liem, Pierce A. Roth, Brian A. Geier, Peter A. Emanuel, Scott A. Walper, Jorge L. Chávez, and Matthew W. Lux. Quantification of Interlaboratory Cell-Free Protein Synthesis Variability. *ACS Synthetic Biology*, 8(9):2080–2091, September 2019.

[132] Dominic J.B. Hunter, Akshay Bhumkar, Nichole Giles, Emma Sierecki, and Yann Gambin. Unexpected instabilities explain batch-to-batch variability in cell-free protein expression systems. *Biotechnology and Bioengineering*, 115(8):1904–1914, August 2018.

[133] J. S. Paige, K. Y. Wu, and S. R. Jaffrey. RNA Mimics of Green Fluorescent Protein. *Science*, 333(6042):642–646, July 2011.