2024

Cell-free sensing and recording applications of genetic circuits

BOSTON UNIVERSITY

COLLEGE OF ENGINEERING

Dissertation

CELL-FREE SENSING AND RECORDING APPLICATIONS

OF GENETIC CIRCUITS

by

JINGYAO CHEN

B.S., University of California, San Diego, 2019
M.S., Boston University, 2022

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2024

Approved by

First Reader   _____
Wilson W. Wong, Ph.D.
Associate Professor of Biomedical Engineering

Second Reader   _____
Ahmad S. Khalil, Ph.D.
Professor of Biomedical Engineering

Third Reader   _____
Alexander A. Green, Ph.D.
Associate Professor of Biomedical Engineering

Fourth Reader   _____
Douglas Densmore, Ph.D.
Professor of Electrical and Computer Engineering
Professor of Biomedical Engineering
Professor of Materials Science and Engineering

Fifth Reader   _____
Mary J. Dunlop, Ph.D.
Associate Professor of Biomedical Engineering

DEDICATION

I would like to dedicate this work to the pursuit of knowledge in biology, a journey

fueled by curiosity and tempered by humility. May we forever seek to unravel the

intricate mysteries of life and its complexities.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my Principal Investigator, Wilson, whose passion for science, vast knowledge, and extensive experience have been invaluable throughout this research journey. His guidance and mentorship have played a pivotal role in shaping my understanding of the subject.

A heartfelt thank you goes out to my dedicated lab members. The collaborative working environment fostered within the lab has greatly enriched my research experience. The exchange of ideas, support, and camaraderie have undoubtedly contributed to the success of this dissertation.

I sincerely thank my parents for their unwavering support and boundless care, especially during these challenging times of separation due to the pandemic. Despite the physical distance between us, their mental support has been my guiding light, helping me navigate through the rigorous research and experiments.

I extend my warmest thanks to my friends for their help from all perspectives, from mental support to research advice and even career planning. Their passion for life, work, and research makes my journey memorable and bearable. With family far away from us, we are the strongest support to each other.

This dissertation would not have been possible without the support and contributions of these individuals. Their belief in me and their collective influence has shaped this research into a meaningful and fulfilling endeavor.

CELL-FREE SENSING AND RECORDING APPLICATIONS

OF GENETIC CIRCUITS

JINGYAO CHEN

Boston University College of Engineering, 2024

Major Professor: Wilson W. Wong, Associate Professor of Biomedical Engineering

ABSTRACT

Synthetic genetic circuits have revolutionized numerous fields, ranging from academic research and point-of-care diagnostics to disease therapeutics and industrial biomanufacturing. These circuits provide a powerful tool for precise spatiotemporal control over biological and biochemical interactions, thereby enhancing our understanding of these complex systems and expanding their applicability. The last few decades have witnessed a surge in research efforts, both in cell-free and cellular systems. These endeavors include those to improve the sensitivity and specificity of diagnostics and optimize the safety, efficacy, and tunability of existing treatments. This dissertation delves into the exploration of Boolean logic gates in the cell-free realm: the development of a 'Cell-Free Recombinase Integrated Boolean Operating System' (CRIBOS) for expanding the capabilities of cell-free sensing applications.

Applications of Boolean logic gates have flourished within cellular systems and animal models. However, a persisting gap in the field is in their exploration within the cell-free system. This deficiency has resulted in a constrained toolkit for studying and applying Boolean logic gates in cell-free settings. Recognizing this limitation in the field and aiming to extend the frontiers of genetic circuits beyond traditional boundaries, I introduce

CRIBOS, leveraging the advantages of recombinase, known for its high orthogonality, efficiency, and sensitivity. I designed more than 20 multi-input-multi-output recombinase Boolean logic gates in a cell-free context, from which a set of critical rules crucial for building genetic circuits in the cell-free environment was also established. In addition, integrating allosteric transcription factor (aTF)-based sensors with CRIBOS enabled multiplex environmental sensing within the cell-free environment. Moreover, the CRIBOS system showcased its versatility by facilitating the creation of a biological memory storage device, demonstrating robust functionality with high stability over four months. Implementing CRIBOS not only expands the application of multiplex Boolean logic gates from cellular systems to the cell-free environment but also expands their overall versatility, opening new avenues for the design and application of sophisticated genetic circuits.

CITATIONS TO PREVIOUSLY PUBLISHED WORK

Portions of this dissertation are reproduced from material in the preparation process for publication.

Sections of Chapters 1, 2, and 3 are in review as follows:

Jingyao Chen, Wilson W. Wong. Cell-Free Recombinase Integrated Boolean Operating System (CRIBOS). (In Review).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

aTFs.. Allosteric Transcription Factors

a.u... Arbitrary Unit

BLADE .... Boolean Logic and Arithmetic through DNA Excision

bp...Base Pairs

CID...Chemically Induced Dimerization

Conc ....Concentration

CRIBOS .... Cell-Free Recombinase Integrated Boolean Operating System

CRISPR.................................Clustered Regularly Interspaced Short Palindromic Repeats

CtcS................................allosteric Transcription Factors Responsive to Chlortetracycline

DETECTR.. DNA Endonuclease-targeted CRISPR Trans Reporter

DNA..Deoxyribonucleic acid

E. coli ....Escherichia Coli

FITC..Fluorescein

FPLC... Fast Protein Liquid Chromatography

GFP .... Green Fluorescent Protein

GMOs..Genetically Modified Organisms

HucR ....allosteric Transcription Factors Responsive to Uric Acid

mCh..mCherry Fluorescent Protein

mg .... Milligram

min ....Minute

mL..Milliliter

CHAPTER ONE: Establishing a recombinase-based Boolean logic platform in the cell-free environment

1.1 Abstract

Cell-free gene expression systems, recognized for their potential applications in fundamental research and biomanufacturing, provide a versatile platform for studying gene circuits and bio-computation. We present the Cell-Free Recombinase Integrated Boolean Operating System (CRIBOS), a site-specific recombinase-based multiplex genetic circuit platform designed for cell-free environments. We used CRIBOS to build more than 20 multi-input-multi-output circuits in the cell-free system, including 2-input-1-output genetic circuits and a 2-input-4-output decoder. Combined with allosteric transcription factors (aTFs)-based sensors, the circuits demonstrate multiplex environmental sensing and detection applications. Moreover, CRIBOS demonstrates remarkable portability and stability in a paper-based setting. Utilizing paper-based CRIBOS, we present a biological memory storage device that can preserve DNA-based biological information for over four months with minimal resources, energy costs, and maintenance requirements. The implementation of CRIBOS not only expands the application of multiplex Boolean logic gates from cellular systems to the cell-free environment but also augments their overall versatility, opening new avenues for the design and application of sophisticated genetic circuits.

1.2 Introduction

The genetic circuit is a foundation tool in synthetic biology with applications in medicine[1,2], diagnostics[3–5], agricultural production[6,7], and environmental surveillance[8,9]. Synthetic gene circuits enable dynamic, precise, interactive, and logical control of cell function, which can facilitate drug discovery by allowing real-time monitoring of diverse molecular events and disease-relevant dynamics[1]. Genetic circuits have also been applied to develop new classes of rapid, cost-effective, deployable diagnostic devices applicable to various diseases and chemicals[4]. Moreover, in chemical manufacturing, engineered cell factories equipped with sensor-regulator systems enable dynamic control of expression pathways and optimal allocation of cellular resources[10]. In agriculture production, yield can be boosted by smart plants incorporated with synthetic circuits capable of sensing and adapting to environmental challenges[11].

In the recent two decades, cell-free systems have been increasingly recognized as valuable platforms for fundamental interrogation and biomanufacturing[12–14]. A cell-free system is a membrane-free, simplified version of cells that operates independently without concern for cell viability. It provides many advantages for developing and characterizing circuits and bio-computation, such as a large dynamic range for chemical measurements, reduced interference from irrelevant cellular pathways, and a high level of control over the composition of the genetic circuits. Moreover, a cell-free environment can detect small molecules that are toxic or impermeable to cells[15,16]. It is also more environmentally friendly and safer by avoiding the use of genetically modified organisms (GMOs). It can provide clearer signals and exclude most of the intrinsic noise caused by complex

organelles and biochemical reactions inside the cell by simplifying the protein composition[17].

With its portability and ease of detection, cell-free has been combined with diverse genetic circuits for point-of-care diagnostics, environmental sensing, and fundamental research on metabolic pathways. Examples include but are not limited to the aTFs-based water contaminant detection system ROSALIND (RNA Output Sensors Activated by Ligand Induction)[18], as well as the CRISPR-based infectious disease diagnostics device SHERLOCK and DETECTR[19]. These examples demonstrated the high sensitivity and robustness of input-output circuits created in the cell-free system and proved it to be an ideal platform for research in synthetic biology and the construction of bio-computing devices. However, although various types of bio-computing devices and Boolean logic gates have been applied to different bacterial[20], yeast[21], mammalian cell strain[22,23], and animal models[24] to enable complex bio-computation, their implementation in the cell-free field lags and circuits with multiple inputs and multiple outputs have remained scarce in the cell-free system, primary due to limited toolkits for building cell-free biocomputing circuits. This limitation necessitates the layering of multiple genetic circuits to construct biocomputing devices and study the complicated metabolic pathways in the cell-free system, requiring extensive construction and characterization for circuit design. The restricted exploration and construction of cell-free Boolean logic gate toolkits largely hinder the cell-free's broader application. Thus, in this project, I aimed to create a cell-free Boolean biocomputing platform for complex multi-input-multi-output genetic circuit constructions, making a streamlined procedure to simplify the construction of cell-free

genetic circuits to facilitate the study of complex metabolic pathways, controlled biologics manufacturing, and multiplex environmental sensing.

Two types of cell-free systems have been developed: the Transcription-Translation (TXTL) and the Protein Synthesis Using Recombinant Elements (PURE). The TXTL system involves extracting components from whole cells by lysing bacteria and purification to remove unnecessary elements[15]. This purified cell lysate forms the basis for conducting cell-free experiments. On the contrary, the PURE system is a mixture of purified proteins[25]. Instead of extracting the whole bacteria cell extract, researchers selectively produce, purify, and mix 30-60 proteins necessary for transcription and translation to form a PURE reaction mixture. Because of the different production methods, compared with the TXTL, the PURE system is much more expensive but offers a higher purity composition, resulting in a stronger signal and greater stability of products. With the motivation to develop cell-free genetic circuits that are highly robust and sensitive, I chose to use the PURE system to develop the Boolean biocomputing platform.

Site-specific DNA recombinases demonstrate a powerful capability for developing multiplex Boolean logic gates in various organisms and models such as bacteria[26], mammalian cells[27,28], and mouse brains[29]. In addition, they perform different types of modification, such as excision and inversion, based on the direction of the recognition sites (Figure 1). Moreover, many recombinases have been developed or discovered, each targeting their corresponding recognition sites with high efficiency and orthogonality, further expanding the pool of usable genetic tools[30]. The success of recombinase-based genetic circuits in various living organisms proves that recombinases can be easily adapted

to different platforms. However, recombinase-based Boolean logic gates have yet to be applied to a cell-free system.

Therefore, to expand the tool kit for genetic circuits, I developed a cell-free recombinase platform and demonstrated that I could create the same high-complexity circuits originally developed in cells (Fig. 2). Although recombinase has been widely studied and well-characterized in a cellular system, transitioning from cell to cell-free is not a simple conversion. Thus, to apply the recombinase circuits to cell-free, we performed large-scale circuit design characterization and established design rules applicable to cell-free genetic circuit development.
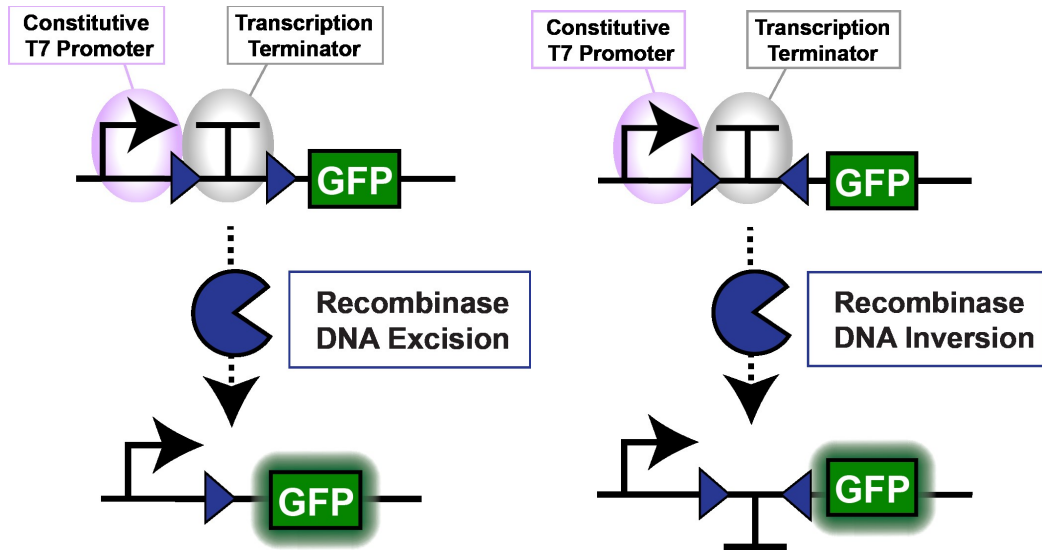
Figure 1: Mechanism of SSRs excision and inversion

By binding the two recognition sites (blue triangles) on either side of the target sequence ("T" indicates transcription terminator), recombinase can either excise/invert the target sequence based on the orientation of the two recognition sites.

A

Inputs

Antibiotics,
Heavy metals,
Light, etc.

**CELL-FREE
SYSTEM**

Synthetic gene circuits
(e.g. logic & signal process.)

Sensor

Executor

*Outputs
and Memory*

Cell-Free SSR-based Biocomputing Platform Design

ORTHOGONAL	MULTIPLEX	VERSATILE	BIOSAFE
programmable, orthog. sensing	*multi-input, multi-output sensing*	Cytotoxic / Cell-Impermeable — *sense wide spectrum of chemicals*	GMO FREE — *minimize genetically modified organisms*

Figure 2: Cell-free recombinase-based biocomputing platform design

A cell-free recombinase-based biocomputing platform allows for the construction of orthogonal, multiplex, versatile, and biosafe genetic circuits. Such a platform has potential applications in multiplex environmental sensing, biological information storage, and disease diagnostics.

1.3 Recombinase excision reporter characterization in a cell-free system

Although recombinase genetic circuits have been widely applied to different cellular and animal models, exploration of their application in cell-free conditions is rare. As such, we have limited knowledge of their design rules. To address this issue, we characterized Cre recombinase excision circuits to establish the potential of recombinase in a cell-free reaction.

The excision reporter comprises a T7 promoter, a ribosome binding site (RBS), a terminator flanked by two Cre recognition loxP sites, and a mCherry reporter gene (Fig. 3A). The results of arbitrary mCherry fluorescence signal were normalized to the nanomolar equivalent of standard Texas Red dye and indicated as normalized translational output (NTO). For the starting reporter design, we used a combination of rrnB T1 and T7Te terminator (Term 1), a double terminator combination verified in a bacterial recombinase system in the previous study, to suppress reporter expression before recombination[31]. In addition, the expected reporter plasmid post-excision, referred to as a LoxP-mCherry plasmid, was also cloned and tested. This serves as a positive control to ensure the successful expression of the reporter plasmid after excision by Cre (Fig. 3A). With the expression and activity of Cre recombinase, we expect the expression of the mCherry reporter gene to go from OFF to ON. At the same time, the LoxP-mCherry should constitutively produce high reporter output. However, opposite results were observed. The reporter-alone condition showed a basal output signal that is a 27-fold change over the blank condition. In addition, an 82% and a 55% decrease in NTO were observed for reporter and LoxP-mCherry, correspondingly, when incubated with Cre-expressing

plasmids (Fig. 3B). These results suggest multiple issues of the current circuit design. First, the high reporter basal expression indicates that the terminator between loxP sites was too weak and failed to terminate the T7 RNA polymerase activity. Second, the decreased reporter expression in the Reporter+Cre condition indicates that adding the Cre protein expression negatively impacts reporter expression, with the underlying mechanism remaining unclear.

We first focused on lowering the basal mCherry expression by implementing strong terminators to shut down the T7RNAP readthrough. We chose the T7 terminator, an engineered terminator for the T7RNAP and expected to prevent T7RNAP readthrough efficiently, as the first candidate (Term 2). However, the results show that the T7 terminator failed to thoroughly terminate the reporter transcription pre-excision (Fig. 3C). Our result agrees with a previous study of synthetic terminator design, in which the T7 terminator only showed 80% termination efficiency[32]. In addition, the previous research also screened an array of terminator designs and created a synthetic terminator (Term 3 - synTerm) that can prevent 99% T7RNAP readthrough. The synTerm is composed of a synthetic T7 terminator, rrnBT1 terminator, and T7 terminator (Fig. 3C)[32]. Replacing the terminator in our original reporter design with the synTerm reduced reporter leaky expression by>78%. Adding one extra copy of synTerm in the reporter (Term 4 - 2XsynTerm) can further lower 94% reporter basal expression to a fluorescence level similar to the blank condition (Fig. 3C).

With the basal expression issue solved, we next investigated the mechanism of the decreased reporter expression in the Reporter+Cre condition. We hypothesized that the Cre

binds to the loxP on the reporter post-recombination, thus blocking the T7 transcription. One loxP site will be left in the reporter plasmid after excision, therefore serving as a repressive operator site for Cre to inhibit the promoter activity. To decrease the Cre binding to loxP sites post-excision, the loxP sites were replaced by the asymmetric, mutant loxP (mutLoxP) sites lox66/lox72. The lox66/72 pairs carry 5bp of mutation on the left or right end of the lox66 or lox72 sites, respectively, resulting in 10bp mutation on the recombined loxP sites post-excision (Fig. 4A)[33]. We hypothesize that the 5bp sequence mutation isn't strong enough to prevent Cre recombinase from binding and modifying the DNA but will result in a 10-bp mutation post-excision, which can effectively prompt Cre recombinase to dissociate from the reporter plasmid. We created a mutLoxP-mCherry plasmid to test this hypothesis by replacing the original loxP site with the mutLoxP site. Both LoxP-mCherry and mutLoxP-mCherry plasmids were then exposed to Cre recombinase to assess the impact of Cre expression on constitutive mCherry expression. Implementing the mutated lox66/72 sites efficiently decreases Cre binding on the recombined recognition sites and recovers over 39% of reporter expression in the LoxP-mCherry plasmids (Fig. 4B).

The success of recovering output signal with the mutLoxP site suggests potential interference between T7 RNA polymerase and Cre binding to DNA. Recognizing the possibility of competition between these enzymes for binding, we hypothesize that the insert length, or distance between the promoter and the loxP sites, could influence the performance of the excision circuit. Thus, we also varied insert lengths to elucidate their impact on circuit performance. Within all the insert lengths that were tested, 10bp provides the most significant fold change of reporter expression with and without Cre plasmid (Fig.

4C). We rationalize that it might be due to the competitive binding between the Cre recombinase and the T7RNAP – the two enzymes might interfere the other's binding efficiency and activity if their binding sites are too close. However, ribosome binding sites should not be too far downstream to the T7 promoter since a longer sequence behind the transcription starting site will result in a longer non-coding mRNA sequence with a more complex secondary structure that can negatively affect translation efficiency. Thus, finding an optimal distance between promoter and recognition sites is critical, and 10bp is determined to be the insert length that can maximize the translation efficiency and T7RNAP activity in the recombinase excision circuits.

A cell-free system is supplemented with limited energy and enzymes. Thus, to better leverage the limited sources in the genetic circuits, we fine-tuned the ratio of reporter and recombinase plasmids for the excision circuits (Fig. 5A). The dose-response curve of recombinase plasmids demonstrates that while a sufficient amount of recombinase plasmids is needed to trigger reporter signals efficiently, high-level of recombinase plasmids will also lead to decrease in reporter fluorescence. We hypothesize that this is due to the rapid depletion of energy and resources for recombinase expression, leaving insufficient resources for reporter fluorescence expression post recombinase excision.

With the characterization of the terminator sequence, recognition sites, insert length, and plasmid dosage, we present a cell-free Cre excision circuit that originally only had a 0.2-fold change to 41-fold change, resulting in more than 200 folds improvement of the excision reporter output signal (Fig. 4D). This streamlined characterization procedure was also applied to excision circuits of five other recombinases in the cell-free

environment. These recombinases have been proven to function in bacterial systems and can trigger above 20-fold change of reporter output signal with high efficiency and orthogonality in cell-free (Fig. 5B).
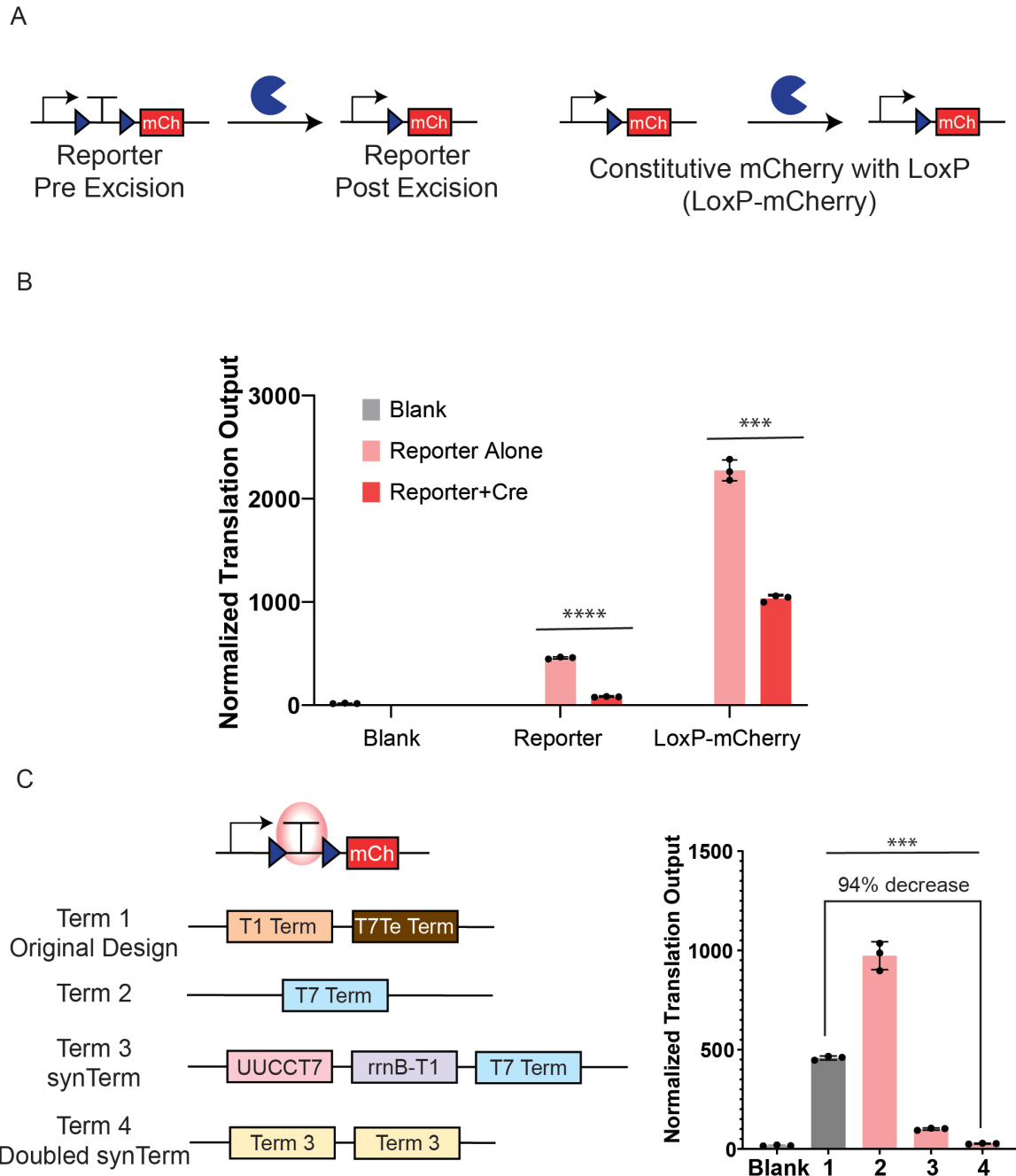
Figure 3: Cell-free recombinase excision reporter terminator characterization

(A) Excision reporter and LoxP-mCherry plasmids design. Before Cre excision, the terminator after the promoter suppresses the mCherry reporter gene transcription. When Cre is expressed, the recombinase will excise the terminator and allow for transcription, thus triggering mCherry fluorescent output. LoxP-mCherry plasmid, an expected reporter sequence post-excision, is cloned

as a positive control for excision circuit testing. Blue triangle: loxP recognition site for Cre recombinase. T: transcription terminator.

(B) Results of the original reporter design. The high NTO of the Reporter alone condition indicates leaky expression of the reporter pre-excision, and the low NTO of the Reporter+Cre condition indicates inhibition of expression by Cre recombinase expression post-excision.

(C) Terminator optimization to prevent leaky reporter expression. To determine the leaky reporter expression pre-excision, reporter plasmid alone was incubated in a PURE mixture, and mCherry reporter expression was detected by a plate reader. By improving the terminator design, we can shut off 94% of leaky mCherry expression pre-excision using terminator design 4.

Data shown are the means of technical triplicate samples with error bars indicating +1 standard deviation. P-values were calculated as described in the methods.
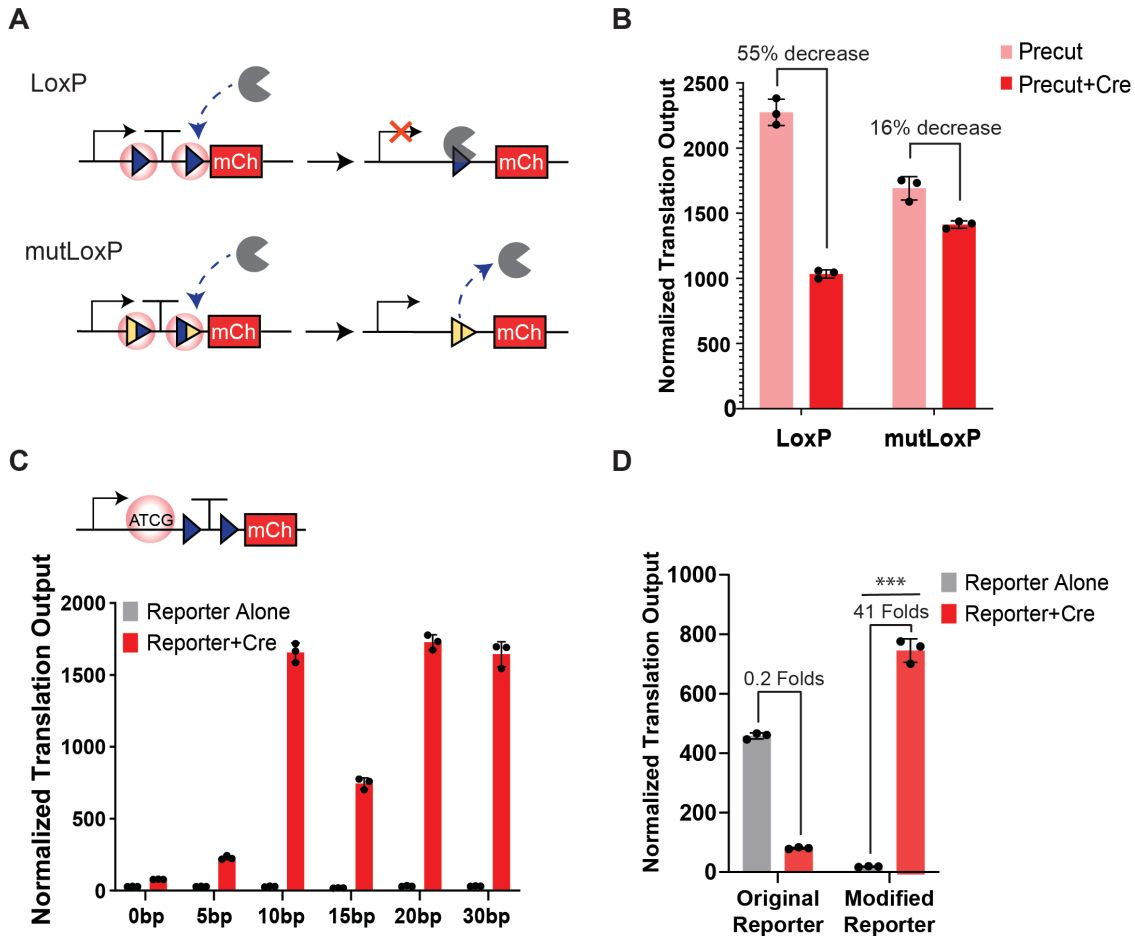
Figure 4: Cell-free recombinase excision reporter optimization

(A) Schematic for comparison and mechanism of loxP and mutLoxP sites. Compared with loxP, mutLoxP sites have 5bp mutation at the outer bound side of each recognition site (mutated and non-mutated sequence indicated as yellow and blue, correspondingly). An identical loxP site will be left in the reporter plasmid post-excision. For mutLoxP sites, a new sequence with 10bp of mutation will be left in the reporter plasmids post-excision.

(B) By implementing mutLoxP sites, we can recover 39% precut NTO when co-incubated with Cre plasmids.

(C) Optimize the insert length between promoter and recognition sites. Too short distances between promoters and recognition sites result in lower reporter mCherry expression. An insert of at least 10bp is needed to optimize excision reporter output.

(D) The optimized excision reporter showed minimum leaky expression and high NTO post-excision, resulting in an excision circuit that can generate a 41-fold change of NTO between pre- and post-excision.

Data shown are the means of technical triplicate samples with error bars indicating +1 standard deviation. P-values were calculated as described in the methods.
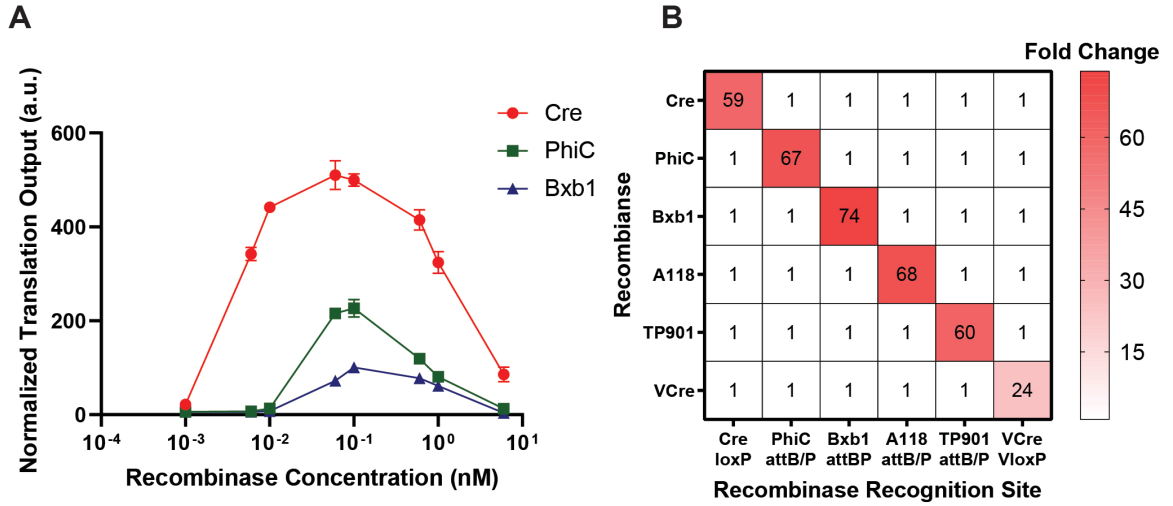
Figure 5: Cell-free recombinase excision genetic circuit fine-tuning

(A) The dose-response curve for recombinase plasmid concentration showed that the optimal concentration is around 6nM. Too high or too low a concentration will decrease reporter gene expression.
(B) Six different recombinase excision circuits can function in cell-free conditions and generate a higher than 20-fold change of NTO between with and without recombinases.

Data shown are the means of technical triplicate samples with error bars indicating +1 standard deviation. P-values were calculated as described in the methods.

1.4 Serine integrase inversion reporter characterization in a cell-free system

Following the successful construction of excision circuits, the subsequent exploration involved the development of recombinase inversion circuits, aiming to further diversify the toolkit for cell-free Boolean logic gate constructions. In contrast to excision, inversion presents advantages such as reduced basal expression of the reporter and a broader spectrum of potential genetic circuits to be engineered.

Recombinases can be classified into two categories: one is tyrosine recombinase (such as Cre and Flp), and the other one is serine integrase (such as PhiC and Bxb1)[34]. Tyrosine recombinase and serine integrase differ in their recombination mechanism and directionality. Tyrosine recombinases operate by breaking and rejoining single strands, forming intermediate Holliday junction, while serine integrase induces double-stranded breaks and reconnects the DNA strands[34,35]. Unlike tyrosine recombinase, whose recombination can be reversible, serine integrase is highly directional and creates irreversible recombination[35]. Such characteristics make serine integrase a better option for inversion circuits than tyrosine recombinase because tyrosine recombinase might keep flipping the target sequence back and forth while serine integrase won't. Therefore, in the project, I mainly focused on inversion genetic circuits of the serine integrase PhiC.

We test three variations of inversion reporters: promoter inversion, terminator inversion, and reporter gene inversion. In the design of the promoter inversion reporter, the T7 promoter is oriented in the opposite direction to the target gene. In the absence of PhiC recombination, the original reporter leads to the transcription of a random DNA sequence. Upon PhiC recombination, the reporter is inverted back to the correct direction, facilitating

the transcription and translation of the reporter gene (Fig. 6A). Similar to the terminator excision reporter, the terminator inversion reporter incorporates a terminator before the reporter gene to suppress reporter transcription. However, in contrast to the excision reporter, where recognition sites point in the same direction, the inversion reporter's recognition sites point in opposite directions, resulting in the inversion of the terminator sequence and triggering transcription and translation of the reporter signal (Fig. 6B). The target gene inversion reporter operates similarly to the promoter inversion reporter, but instead of inverting the promoter, this design inverts the target gene to initiate transcription and translation of the correct reporter signal (Fig. 6C).

In the process of optimizing the reporter design for the inversion circuit, I extended the same streamlined approach for excision reporter characterization. This procedure focuses on three key elements: recombinase recognition sites, the insert length between the promoter and recognition sites, and the strong terminator sequence to suppress leaky reporter expression. Different from Cre loxP sites, which result in an identical loxP site post-recombination, PhiC attB/attP sites produce a different pair of attL/attR recombination sites. This eliminates concerns about recombinase binding post-recombination when designing the PhiC reporter. Using 2XsynTerm in the terminator excision design mitigates pre-inversion reporter transcription, minimizing leaky reporter expression. However, potential competitive binding between T7RNAP and the recombinase remains a concern for the reporter output signal. Therefore, to optimize the inversion reporter design, I explored different insert lengths between the promoter and recombinase recognition sites, ensuring interference-free bindings of these two enzymes to

their target sequence.

Incremental inserts of 0, 5, 10, 15, and 20 base pairs (bp) of DNA were introduced into the promoter, terminator, and reporter gene inversion reporters to refine the insert length between recognition sites and the promoter. Optimal output signals were observed for the promoter and target gene inversion reporters at an insert length of 15 bp, while the terminator inversion reporter exhibited optimal performance with a 30 bp insert (Fig. 6A-C). Deviations from this optimal range, either too long or too short, resulted in a significant decline in the output signal.

It is observed that the output signal increases with a longer insert length. This experimental phenomenon substantiates our initial hypothesis concerning potential interference between T7RNAP and recombinase binding. The proximity of their binding sites could lead to mutual hindrance, negatively affecting the reporter output expression. Importantly, expanding the distance between these binding sites appears to significantly mitigate this adverse effect. Moreover, it is also observed that the output signal decreases when the insert length exceeds the optimal range, suggesting that an overly extended insert may harm circuit performance. This could be attributed to the complex secondary structure of the mRNA transcribed from the random insert sequence before the target gene, causing extra challenges for the target gene translation.

The optimized designs for the promoter, terminator, and target gene inversion reporters yielded fold changes of 13, 6.2, and 3.8, respectively (Fig. 6D). The successful development of inversion genetic circuits significantly broadens the spectrum of potential complex Boolean logic gates that can be constructed in a cell-free system. These outcomes

mark a substantial advancement in the versatility and applicability of engineered genetic

circuits within a controlled environment.

21
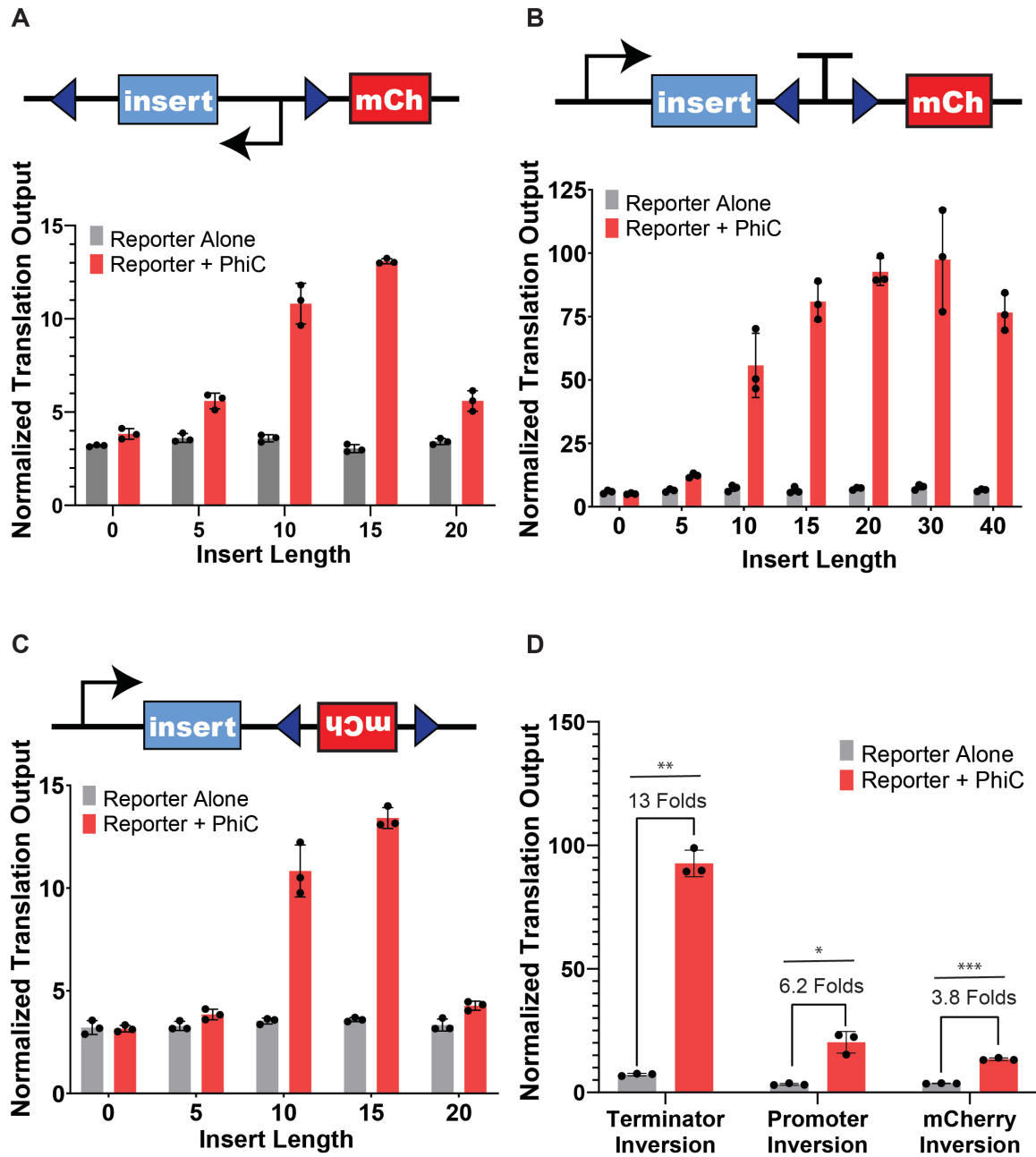
Figure 6: PhiC inversion genetic circuit fine-tuning

Recognition sites facing opposite directions are placed on either end of (A) the Inverted promoter, (B) the Terminator, and (C) the Inverted target gene. Inserts with 0-40bp were inserted between the recognition sites and the promoter. All three graphs show a trend of increasing NTO before the optimal value and starting to decrease after passing the optimal value, indicating that an insert sequence that is too long or too short will negatively impact the reporter output signal expression.

(D) Optimized inversion reporters function properly in a cell-free system and present fold change of 13, 6.2, and 3.8 for terminator inversion, promoter inversion, and mCherry inversion circuits, correspondingly, between conditions with and without PhiC recombinase.

Data shown are the means of technical triplicate samples with error bars indicating +1 standard deviation. P-values were calculated as described in the methods.

1.5 Construction and characterization of cell-free 2-input-1-output logic gates with
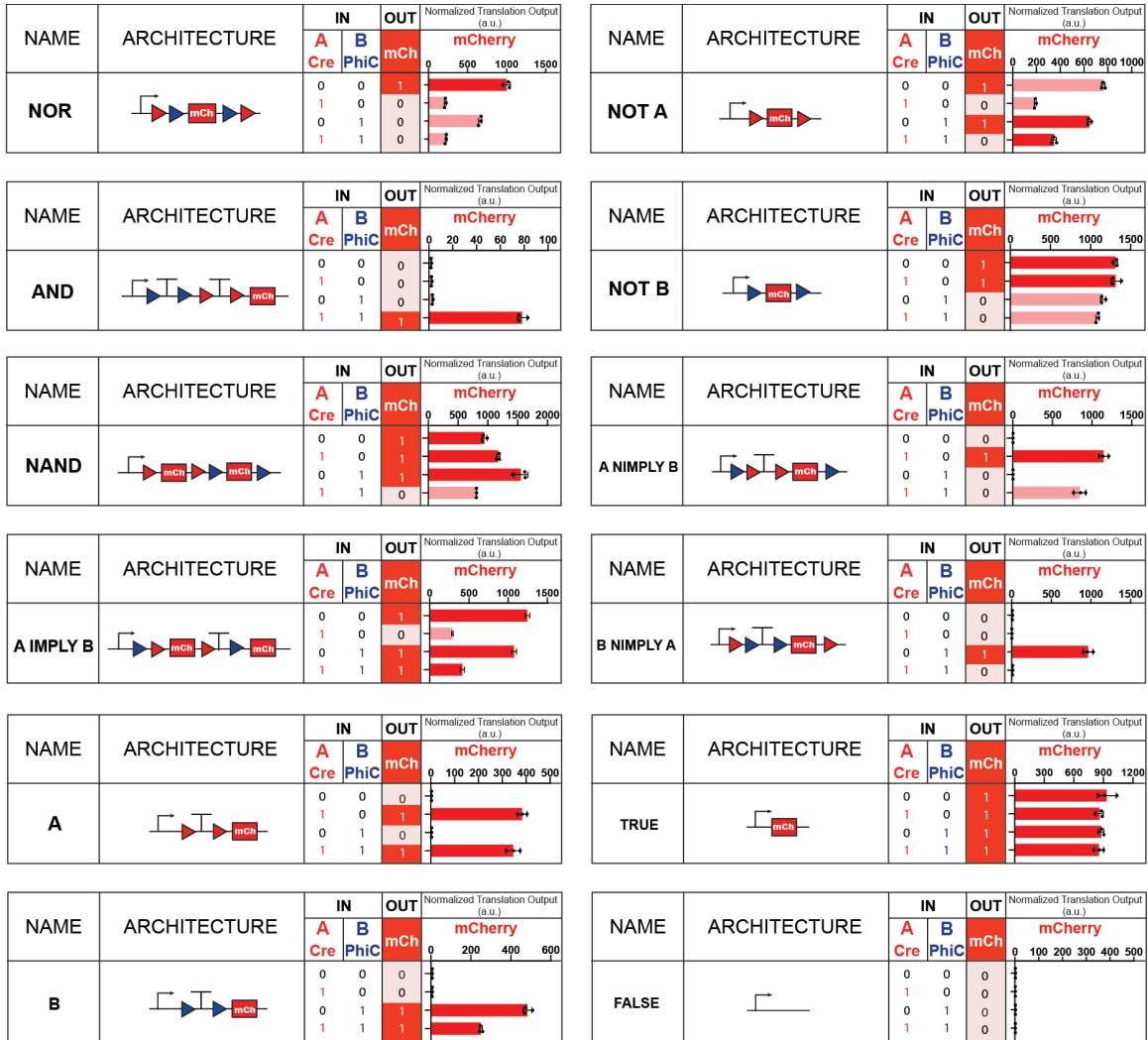
Cre and PhiC recombinases

The high orthogonality exhibited by recombinases and their recognition sequences paves the way for the development of more intricate recombinase genetic circuits featuring multiple inputs and outputs within cell-free conditions. Leveraging the six recombinases established in the cell-free system, I further increase the cell-free Boolean logic gates' complexity beyond simple excision and inversion. One notable demonstration is the creation of 2-input-1-output Boolean logic gates. Prior research has explored the construction of 2-input-1-output Boolean logic gates in cell-free conditions using Toehold-mediated strand displacement circuits at the RNA level, but the development of such logic gates at the DNA level within cell-free systems has been limited[36,37].

To address the existing gap in the field, I constructed 12 distinct 2-input-1-output Boolean logic gates by strategically incorporating Cre and PhiC recombination sites, namely Cre lox66/72 and PhiC attB/attP, on either side of terminator sequences along with a mCherry reporter gene. In the original experimental design, the reporters and recombinase plasmids were premixed and added to the cell-free reaction before incubation. With this design, some circuits do not behave as expected in the logic table, which is mainly due to the delayed shutoff of the reporter fluorescent expression by PhiC recombinase (Fig. 7). We reason that this is because of the weak activity of the PhiC recombinase. While Cre is very efficient and active and thus can promptly shut the reporter expression, the PhiC is much weaker and takes longer to excise the reporter gene. During this time, the reporter simultaneously expresses and accumulates in the reaction, leading to a leaky output signal

for the genetic circuits. Therefore, we hypothesize that rapid recombination by PhiC is the key to decreasing the leaky signals, and pre-expression of the recombinases might help to achieve this goal.

In the following experimental design, instead of adding recombinase and reporter plasmids simultaneously to the cell-free reaction, we first pre-incubated Cre and PhiC recombinase plasmids in the cell-free reaction for 30 minutes. Then, we added the reporter plasmids to finish the circuit setup. The 2-input-1-output logic gates yielded intended behaviors (Fig. 8). It demonstrates that timing is also a critical factor in developing and optimizing genetic circuits, and a cell-free platform provides researchers with higher control and flexibility of timing over every step of circuits' execution.

The performance of the circuits was evaluated by a vector proximity (VP) metric measuring the discrepancy between the experimental output and the expected logic table. The logic table and the experimental measurements were represented by a "truth table vector" and a "signal vector," respectively, in a four-dimensional space[30] (Fig. 9). The angular difference between the two vectors was measured: a 0° angle indicates that the experimental results represent the intended logic table perfectly, and a 90° angle indicates that the experimental output demonstrates completely wrong output. Within 11 circuits tested, only 5 circuits (36%) showed a VP angle that was smaller than 15° using the original experimental procedure (Table 1). With the modified procedure, the circuits' performance improved, and the VP angle of all circuits systematically decreased, resulting in 6 of 8 circuits (80%) with a VP angle smaller than 15°.

Figure 7: Results of 2-input-1-output genetic circuits with original experimental protocol

Twelve 2-input-1-output excision circuits with Cre and PhiC recombinases are tested in a cell-free environment. To set up the experiment, reporter plasmids, recombinase plasmids, and water were pre-mixed. The mixture of plasmids was added to the PURE reaction buffer for 5hr of incubation at 37°C.

Data shown are the means of technical triplicate samples with error bars indicating +1 standard deviation. P-values were calculated as described in the methods.

26
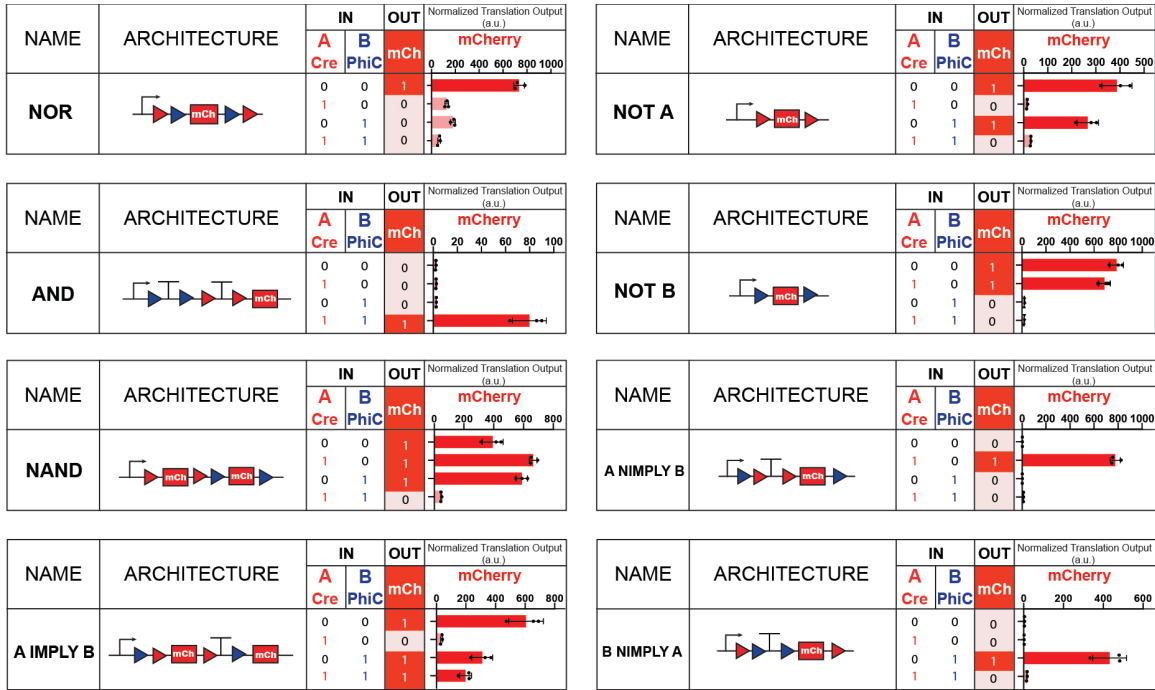
Figure 8: Results of 2-input-1-output genetic circuits with optimized experimental protocol

Eight 2-input-1-output excision circuits with poor performance were tested with optimized experimental procedures. To set up the experiment, recombinase plasmids and water were pre-mixed and incubated for 30min at 37°C, and then reporter plasmids were added to the reaction and incubated for another 5hr.

Data shown are the means of technical triplicate samples with error bars indicating +1 standard deviation. P-values were calculated as described in the methods.
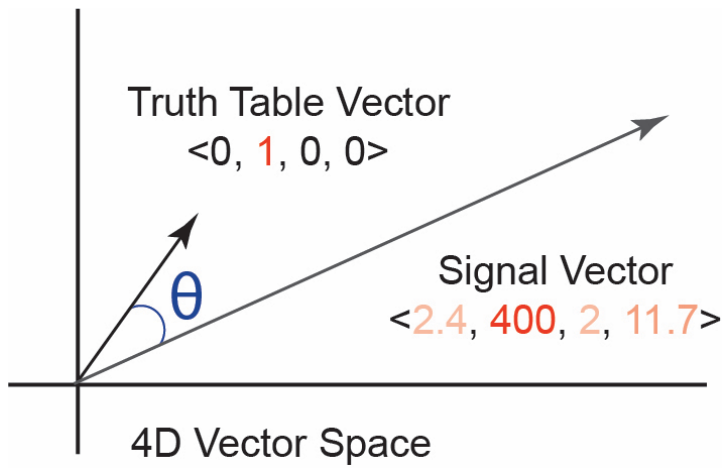
Figure 9: Vector proximity matric analysis

(A) Each genetic circuit's truth table is mapped into a four-dimensional vector called a "Truth Table Vector." The experimental measurement of each genetic circuit is also mapped into a four-dimensional vector called a "Signal Vector." The angle θ between the Truth Table Vector and Signal Vector is calculated for vector proximity (VP) metric analysis to evaluate the genetic circuit's correctness.

Table 1. Vector proximity (VP) angles for 2-input-1-output circuits with original experimental procedure.

2 Input 1 Output Circuits without Delay			
Circuit Type	VP Angle	Circuit Type	VP Angle
NOR	51.7	NOT A	35.7
AND	3.2	NOT B	45.0
NAND	30.0	A NIMPLY B	45.0
A IMPLY B	22.7	B NIMPLY A	2.5
A	2.9	TRUE	0
B	13.1		

Table 2. Vector proximity (VP) angles for 2-input-1-output circuits with optimized experimental procedure.

2 Input 1 Output Circuits with Delay			
Circuit Type	VP Angle	Circuit Type	VP Angle
NOR	30.0	NOT A	10.1
AND	2.9	NOT B	2.9
NAND	4.4	A NIMPLY B	1.7
A IMPLY B	15.8	B NIMPLY A	2.6

1.6 Construction and characterization of 2-input-4-output logic gates

Although 2-input-1-output circuits have successfully developed in cell-free system systems[38], a general strategy for creating N-input-M-output has yet to be developed. The Wong lab has previously developed a 2-input-4-ouptut 'Boolean logic and arithmetic through DNA excision' (BLADE) in mammalian cells to exploit the features of site-specific recombinases to enable N-input-M-output combinatorial computation[27]. With two inputs, the circuit requires four different outputs, $Z = Z_{AB} = Z_{00}$, Z_{10}, Z_{01}, and Z_{11}, to represent 2^N $= 2^2 = 4$ (N indicates input number) different states of inputs A and B (Fig. 10). To investigate the potential of building N-input-M-output in a cell-free environment, we implemented the 2-input BLADE circuits with terminators, fluorescent, and luminescent genes. One key element in the BLADE circuit is the heterospecific site, such as the Cre recombinase's lox66/72 and lox2272. Although they are only a few base pairs different, lox66/72 and lox2272 can only retain Cre excision capability when paired with their corresponding recognition sites, which means lox2272 can only be paired with lox2272 to trigger Cre recombination activity while lox66 can only be paired with lox72.

In the pilot test of the 2-input-4-output reporter, a low fluorescence signal was observed for GFP and mCherry (Fig. 11A). Based on the previous experience of reporter characterization, I hypothesized that this is due to the insert sequences positioning in front of the Z_{00} addresses. In contrast to the simple excision circuit, the BLADE reporter features increased complexity and more diverse recombinase recognition sites, resulting in a longer insert sequence between the promoter and the expressing gene. It's worth noting that, in addition to the PhiC attB site and loxP2272 site in the reporter design, an extra 40bp of U1
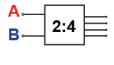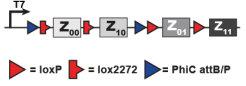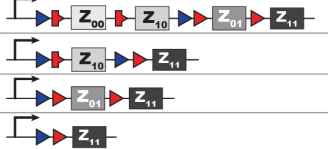
sequence is present to guide the Gibson reaction during molecular cloning, further elongating the insert in front of the expressing gene. Consequently, I postulated that such an extended sequence in front of the expressing cassette might adversely impact the expression efficiency of the output signals.

To test this hypothesis, three versions of inserts were designed and positioned in front of a constitutive mCherry gene (Fig. 11B). Design 1 is the insert sequence in the original reporter, design 2 is made by removing the U1 sequence from design 1 to investigate the effects of U1 on gene expression, and design 3 is the insert sequence taken from the excision reporter, which was proved to show the high fluorescence output signal in the excision circuit (Fig. 11B). Interestingly, design 2 successfully leads to a significant 18-fold increase in output expression by eliminating the 40bp U1 and substantially shortening the distance between the promoter and the expressing gene. Moreover, design 3 also demonstrated a 2-fold higher output expression than design 2, even with only a 1bp difference in their length (Fig. 11B). This suggests that, in addition to insert length, the sequence design itself may also influence gene expression. One hypothesis of such phenomenon is that mRNA secondary structure might play a vital role in determining the expression efficiency of the reporter gene.

With the promising data shown, we replaced the insert in the BLADE reporter with the design 3 insert to develop a BLADE reporter version 2 (Fig. 11C). The experimental data showed that the fluorescence signal largely increased with the optimized reporter (Fig. 11C-D). However, leaky expression is observed for all three addresses, indicating that the terminator sequence at the end of each address is too weak. Both the original and version

2 BLADE reporter designs employed the T7 terminator at the end of each address to simplify the molecular cloning process, as the synTerm posed challenges for PCR and Gibson Assembly due to its repetitive sequences. Unfortunately, the T7 terminator proved inadequate for the BLADE circuit, leading to leaky expression.
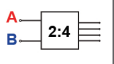
To address this issue and entirely suppress the leaky expression, I replaced all T7 terminators with synTerm, creating BLADE reporter version 3. This modification proved successful, providing the optimal performance for the BLADE circuit as anticipated (Fig. 11E). The VP angle metric for the BLADE circuits was evaluated to be less than 25% for all fluorescent and luminescent outputs (Table 3).

LOGIC GATE			AVERAGE SINGLE WELL RESULTS					
NAME	SYMBOL	ARCHITECTURE	INPUT		OUTPUT			
			A Cre	B PhiC	Z_{10}	Z_{01}	Z_{11}	Recombined Architecture
2-input DECODER			0	0	0	0	0	
			1	0	1	0	0	
			0	1	0	1	0	
			1	1	0	0	1	

Figure 10: Schematic for the mechanism of 2-input-4-output BLADE circuit

2-input-4-output BLADE reporter comprises 4 different addresses, Z_{00}, Z_{01}, Z_{10}, and Z_{11}, each indicating one state of the absence or presence of the two inputs. When both inputs are absent, the gene in address Z_{00} will be expressed. When only one of the inputs is present, the gene in address Z_{10} or Z_{01} will express. When both inputs are present, the gene in the address Z_{11} will be expressed.
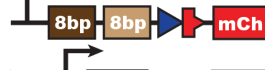
A

B

Design 1: original design

Design 2: remove U1
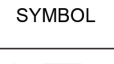
Design 3: remove U1 and use 15bp insert

C

= loxP   = lox2272   = PhiC attB/P

D

E

Figure 11: Optimization of 2-input-4-output BLADE reporter.

(A) BLADE reporter version 1 has low fluorescent expression for GFP and mCherry.
(B) There are three designs of insert sequence for the BALDE reporter. Design 1 is the original insert sequence from the BLADE pilot study, containing 2 random 8bp sequences and a U1 sequence. Design 2 cuts out the U1 sequence from design 1. Design 3 is the insert sequence taken from the excision circuits. The results show that Design 3 has the highest NTO for all output signals.
(C) Comparison of version 1 and 2 BLADE reporters. BLADE reporter version 2 is modified from version 1 by replacing the original insert sequence with insert design 3.
(D) BLADE reporter version 2 successfully resolves issues of low reporter expression efficiency but has high basal expression.
(E) BLADE reporter revision 3, modified from version 2 using synTerm to minimize basal expression, presenting 4 output states as expected in the logic table.

Data shown are the means of technical triplicate samples with error bars indicating +1 standard deviation. P-values were calculated as described in the methods.

Table 3. Vector proximity (VP) angles for 2-input-4-output decoders.

2 Input 4 Output Decoder					
Output	VP Angle	Output	VP Angle	Output	VP Angle
GFP	15.1	mCherry	8.1	NLuc	9.0

1.7 Conclusions and future perspective

In the present study, I introduce CRIBOS, a recombinase-based multiplex genetic circuit platform in a cell-free environment. I used CRIBOS to develop a 2-input-4-output Boolean logic gate that has never been documented in any cell-free system. Our results show that recombinases can function in cell-free conditions with high activity and orthogonality, enabling constructions of multiplex genetic logic gates in a cell-free system.

Through careful characterization, we discovered multiple design rules for recombinase-based Boolean logic gates in the cell-free environment, which are translatable throughout various recombinase logic gate designs. The challenges encountered during the optimization and troubleshooting process illustrate that applying genetic circuits developed in cells to cell-free is not a simple conversion. Although cell-free is a natural environment for circuit development and can provide benefits and convenience that cellular platforms and animal models cannot provide, it poses challenges and issues that are not present in any other platforms. For instance, while living organisms can offer genetic circuits with continuous and durable energy sources and building blocks, a cell-free environment has limited means to regenerate high-energy compounds and starting materials needed for protein synthesis and circuit components. This illustrates that proper resource allocation is critical for cell-free operations. The different spatial density in cell-free poses another challenge for converting cellular genetic circuits to cell-free. Cell-free systems are much more dilute than the cell lysate, not just in the concentration of the transcription and translation machinery but also in the degree of macromolecular crowding[39–42]. Such differences in spatial density can influence the biochemical reaction rates and equilibria,

changing the efficiency and competitive effect of protein-protein and protein-DNA binding. The biochemical characteristics of recombinase-recognition sites binding might also be affected in the cell-free reaction, indicated by the complete shutoff of the loxP reporter signal by Cre recombinase due to the high Cre-loxP binding affinity post recombination. While lox66/72 sites were previously used for fixing the directionality of the recombination, few studies have reported its functions in improving circuit output signals in cells or animal models because the native Cre-loxP system can function properly and efficiently in those platforms, and there is no need to replace the loxP sites into lox66/72 sites. These discrepancies between cell and cell-free systems will be critical considerations for building complex genetic circuits and expanding the cell-free bio-computation toolkits.

The capability of cell-free recombinases to construct complex multi-input-multi-output circuits was also demonstrated with CRIBOS. Using two of the six orthogonal recombinases, the BLADE circuit that was previously developed in mammalian cells was also designed and optimized in a cell-free environment. With the BLADE circuit as the demonstration, more complex genetic circuits can also be generated with CRIBOS. Previously, we developed a framework for building multi-input-multi-output genetic circuits in mammalian cells. Using the 2-input BLADE design as a basic structure, genetic circuits with three or more inputs can be efficiently designed and implemented. Besides engineering circuits to detect the concurrence of multiple events, recombinase-based temporal circuits can also be designed to detect sequences of events, providing an extra layer of information (Fig. 12). Combined with CRIBOS, such complex systems can be expanded to a cell-free environment to broaden the toolkits further.
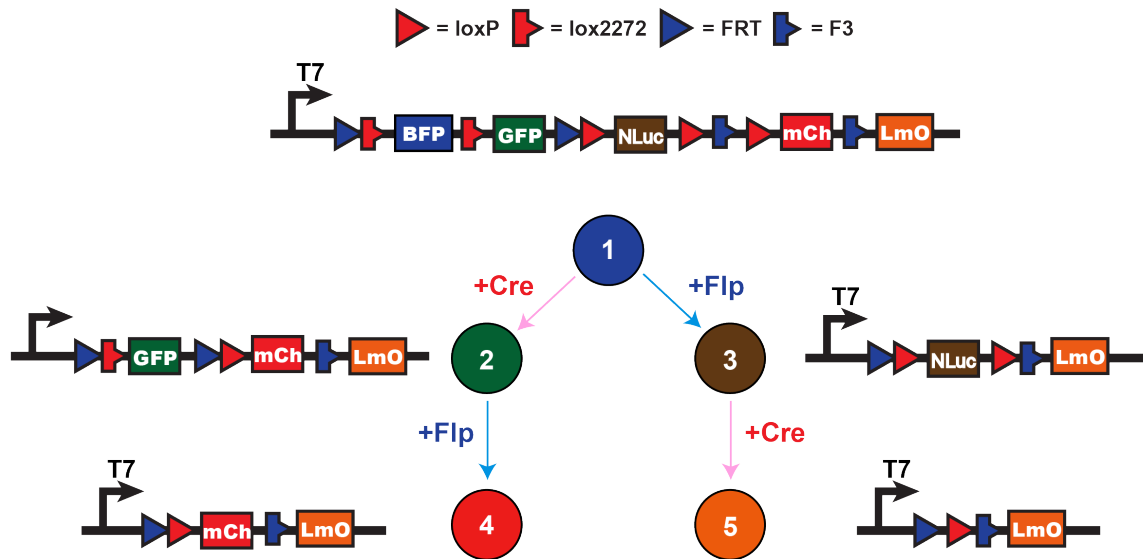
Figure 12: Schematic of mechanism for a temporal circuit reporter

The temporal circuit reporter involves 5 different stages to record the order of occurrences of 2 different inputs. Two pairs of heterospecific recognition sites for Cre (loxP and lox2272) and Flp (FRT and F3) recombinases will be used. The presence of each recombinase will make excisions on their corresponding recognition sites.

1.8 Materials and methods

Strains and growth medium

NEB 5-alpha Competent E. coli (New England Biolabs, C2987U) was used for routine molecular cloning. E. coli strain Rosetta 2 (DE3) pLysS (Novagen, 71401) was used for recombinant protein expression. BL21(DE3) Competent E. coli (New England Biolabs, C2527H) was used for recombinase circuit testing. Plasmids were transformed into corresponding bacterial strains based on the manufacturer's procedures. LB broth (Thermo Fisher Scientific, BP1426-2) supplemented with appropriate antibiotic(s) (100 μg/ml carbenicillin (Thermo Fisher Scientific, 10177012), 50 μg/ml kanamycin (Bio Basic, KB0286-25) and/or 25 μg/ml chloramphenicol (Sigma-Aldrich, C0378-25G)) was used as a bacterial growth medium.

Molecular cloning

DNA oligonucleotides for cloning and sequencing were synthesized by ThermoFisher Scientific. Genes encoding fluorescent, luminescent proteins and 2X terminators were synthesized as gBlocks (Integrated DNA Technologies). All constructs were created using standard restriction enzyme digestion, Gibson isothermal assembly, and Unique Nucleotide Sequence (UNS) Guided assembly procedures into a pET15b plasmid. UNS Guided assembly utilizes a mechanism similar to Gibson's isothermal assembly but with standard homology sequences that were computationally optimized. Constructed plasmids were transformed and maintained in NEB 5-alpha (New England Biolabs (NEB) C2987H) Escherichia coli competent cells at 37°C overnight before miniprep (Epoch Life Sciences,

2160250).

In vitro protein synthesis

NEB PURExpress *In Vitro* Protein Synthesis Kit (New England Biolabs, E6800S) supplemented with RNase Inhibitor (Sigma-Aldrich, 03335402001) was used to generate data for cell-free experiments based on the manufacturer's protocol. The plasmids, DNAs, aTFs, and small molecules are first premixed before combining with the cell-free reaction mix. Then, 5µL of cell-free circuits were incubated in each well of a 384-well clear bottom black plate (Corning, 3542) at 37°C for 5hr. An endpoint measurement, or dynamic measurement, was performed on a plate reader (SpectraMax M5, Molecular Devices) with an excitation wavelength of 580nm and emission wavelength of 611nm for mCherry, an excitation wavelength of 485nm and emission wavelength of 525nm for GFP, and an excitation wavelength of 381nm and emission wavelength of 445nm for BFP. Nano-Glo Luciferase Assay (Promega, N1110) was used for NLuc luminescence detection. Nano-Glo Luciferase Assay Substrate was diluted 1:50 by Nano-Glo Luciferase Assay Buffer. The luminescence level was measured immediately after a volume of the diluted substrate equal to 6X the volume of the cell-free reaction was added to each well.

Plate reader quantification and MEF standardization

A Texas Red Dye (Sigma-Aldrich, F6377-100G) was used to convert arbitrary mCherry fluorescence intensity into nanomolar equivalent Texas Red Dye (labeled as normalized translational expression). A serial dilution was prepared with PBS from a 500uM stock solution. The samples were prepared with triplicates, and fluorescence values were

measured at an excitation wavelength of 580nm and emission wavelength of 611nm on a plate reader (SpectraMax M5, Molecular Devices). Fluorescence for a concentration in which a single replicate saturated at the plate reader was excluded from the analysis. The rest of the measurements were averaged and formed a linear regression line, which is used to convert the measured arbitrary fluorescence value to the concentration of the Texas Red Dye (Fig. 13). Standard curves were created with the same process for each plate reader for data collection to normalize the results. The exact process was used to convert GFP fluorescence intensity into nanomolar equivalent Fluorescein (Sigma-Aldrich, F6377) and NanoLuc luminescence intensity into nanomolar equivalent NanoLuc purified protein (Promega, E499A) (Fig. 13). All fluorescence and luminescence output are normalized to the nanomolar equivalent of corresponding dye or proteins, and then present on the figure as "Normalized Translational Output (NTR)."

Vector proximity computational analysis

Each 2-input-1-output or 2-input-4-output Boolean function corresponds to a logic table with four rows and one output column or four rows and four output columns, respectively; the output column in each Boolean function was mapped to a 4-dimensional vector a, called the "truth table vector." The fluorescent reporter output signal of each genetic circuit measured from experimental implementation was also mapped to a 4-dimensional vector b, called the "signal vector." We evaluated the accuracy of a genetic circuit by computing the angle θ between the vectors a and b using the formula:

$$\theta = cos^{-1}(\frac{a \cdot b}{|a||b|})$$

When computing $\boldsymbol{\theta}$, we capped the signal values in $\boldsymbol{b}$ to a maximum of 400nM Texas Red, 60nM Fluorescein, and 400nM NanoLuc purified protein. The angular difference ranges from 0° (best) to 90° (worst).

Statistical analysis

All cell-free *in vitro* experiments involved setting up PURE *in vitro* protein synthesis expression of DNA into n = 2 or 3 reactions. Fluorescence and luminescence intensities for each reaction were averaged, and the standard deviation was taken. Data between groups were evaluated with unpaired two-tailed t-tests to evaluate statistical significance. P values are reported (not significant = $p > 0.05$, * = $0.01 < p < 0.05$, ** = $0.001 < p < 0.01$, *** = $0.01 < p < 0.001$, **** = $0.001 < p < 0.0001$).

Standard Curve
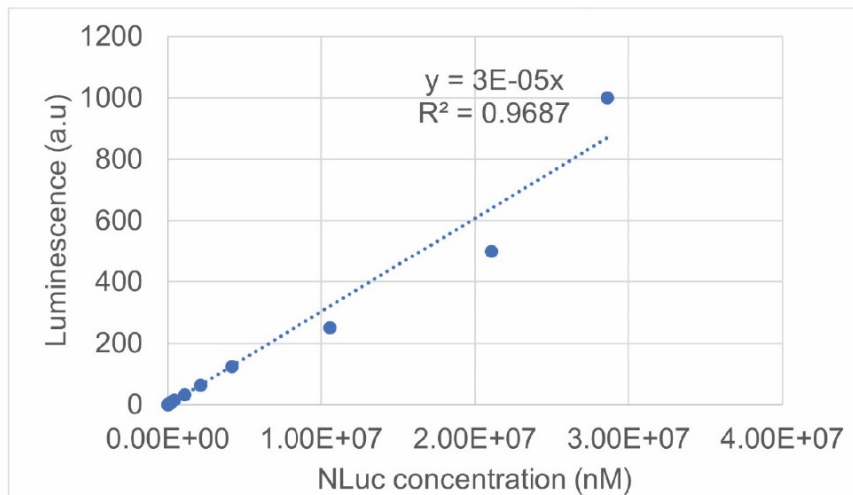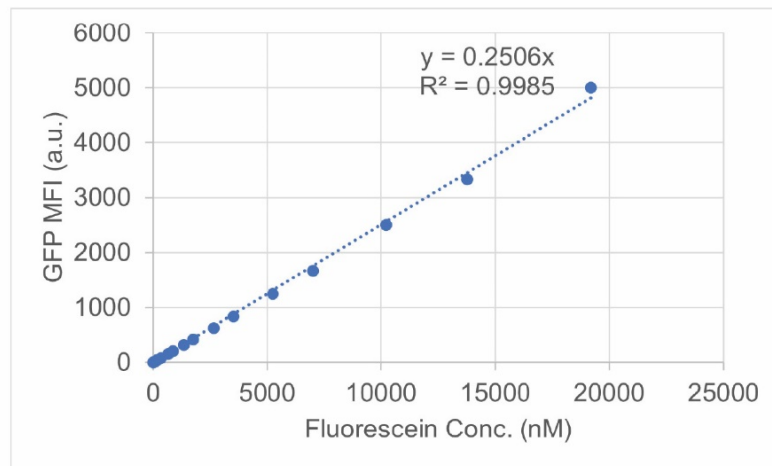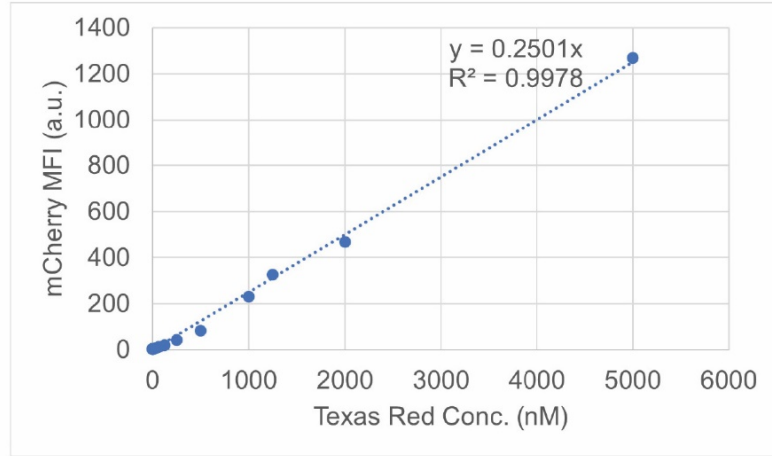
Figure 13: Standard curves for fluorescence and luminescence expression.

Arbitrary units of fluorescence and luminescence were standardized to nM concentration of Texas Red, Fluorescein (FITC), and NanoLuc proteins. In the three representative examples shown here, the Texas Red and FITC dilution series was prepared with PBS, and NanoLuc was prepared in Promega Nano-Glo Luciferase Assay Buffer (Promega, E499A). The dilution series was measured on a plate reader using the same settings for measuring mCherry, GFP, and NanoLuc from PURE reactions. Data shown are for n=3 replicates.

CHAPTER TWO: Implementing an aTF-based sensing platform to CRIBOS

2.1 Introduction

Having established site-specific recombinases as DNA executors in a cell-free system, the subsequent research aimed at developing a sensing platform compatible with recombinase-based Boolean circuits to broaden the potential applications of the CRIBOS system.

Allosteric transcription factors (aTFs) are a group of proteins involved in the process of DNA transcription regulation. By binding to target DNA sequences, aTFs can either activate or suppress DNA transcription, thereby controlling protein expression profiles within cells across different temporal and spatial dimensions. Natural aTFs, inherent in bacterial cells, act as an intrinsic system responding to chemicals in their surroundings and modulating the behavior of bacterial cells[43]. A diverse pool of aTFs was discovered and engineered in bacterial cells to sense a wide range of small molecules and environmental signals, including antibiotics[44,45], heavy metals[46–50], metabolites[51–54], aromatic compounds[55,56], and changes in pH and osmolarity[57]. They play a crucial role in modifying the protein expression profile during different phases of growth and replication, regulating metabolic pathways, cell division, and defense mechanisms against toxins. In addition to natural environmental stimuli, efforts have been invested in engineering aTFs to respond to novel ligands, broadening their chemical sensing spectrum and facilitating more diverse applications[51,58,59].

Driven by high modularity, robustness, and a remarkable capacity for modification, aTFs have been engineered into exceptional tools for gene regulation. Their extensive

range of potential sensing chemicals has positioned them as invaluable assets in fundamental research, protein biomanufacturing, and environmental sensing across diverse cellular contexts[46-48].

The versatility of aTFs extends from cellular to cell-free systems. Notably, their robust and sensitive application was demonstrated in water contaminant detection[18]. Building upon these advantageous features and leveraging insights from previous studies on aTFs, we integrated aTFs into the CRIBOS system, unlocking new dimensions in genetic circuit engineering within cell-free environments.

2.2 Characterization for aTFs-based biosensors in the PURE system

To validate the compatibility of aTFs-based sensors with the PURE cell-free system, an initial assessment for aTFs biosensors was performed. TetR, a transcription repressor responsive to tetracycline, was chosen as the first aTF for pilot synthetic circuit testing due to its outstanding sensitivity, robust functionality, and rapid response demonstrated in the previous research. In the pilot experiment, the tetO operator, specifically recognized by the tetR biosensor, was cloned upstream of a constitutively expressing mCherry gene (Fig. 14A). In the absence of tetracycline, tetR will tightly bound to tetO, suppressing the progress of T7RNAP and transcription of the output mCherry gene. When tetracycline is introduced to the system, it will bind to tetR and result in a conformational change in the protein structure, prompting the dissociation of tetR from tetO and facilitating the expression of the output gene (Fig. 14A).

Prior to the pilot test, tetR protein was produced and extracted from Escherichia coli (E. coli) and was subsequently purified via histone columns and then heparin columns through fast protein liquid chromatography (FPLC). To set up the experiment, tetR protein and reporter plasmid were co-incubated within the cell-free system, both in the presence and absence of tetracyclines. The results demonstrate remarkable performance of the tetR biosensors, evidenced by a 13-fold change in mCherry expression between conditions with and without small molecules (Fig. 14B). Comparison between the Reporter alone and Reporter+tetR protein conditions revealed that the tetR protein effectively suppressed the transcription of the mCherry gene, yielding a fluorescence level that is close to the baseline.

Following the same procedure as the tetR pilot experiment, an array of diverse aTFs was also tested in the PURE system. Among these candidates, tetR and smtB showed the best performance and highest protein production yield and thus were chosen for future circuit construction (Fig. 14B–C).

50
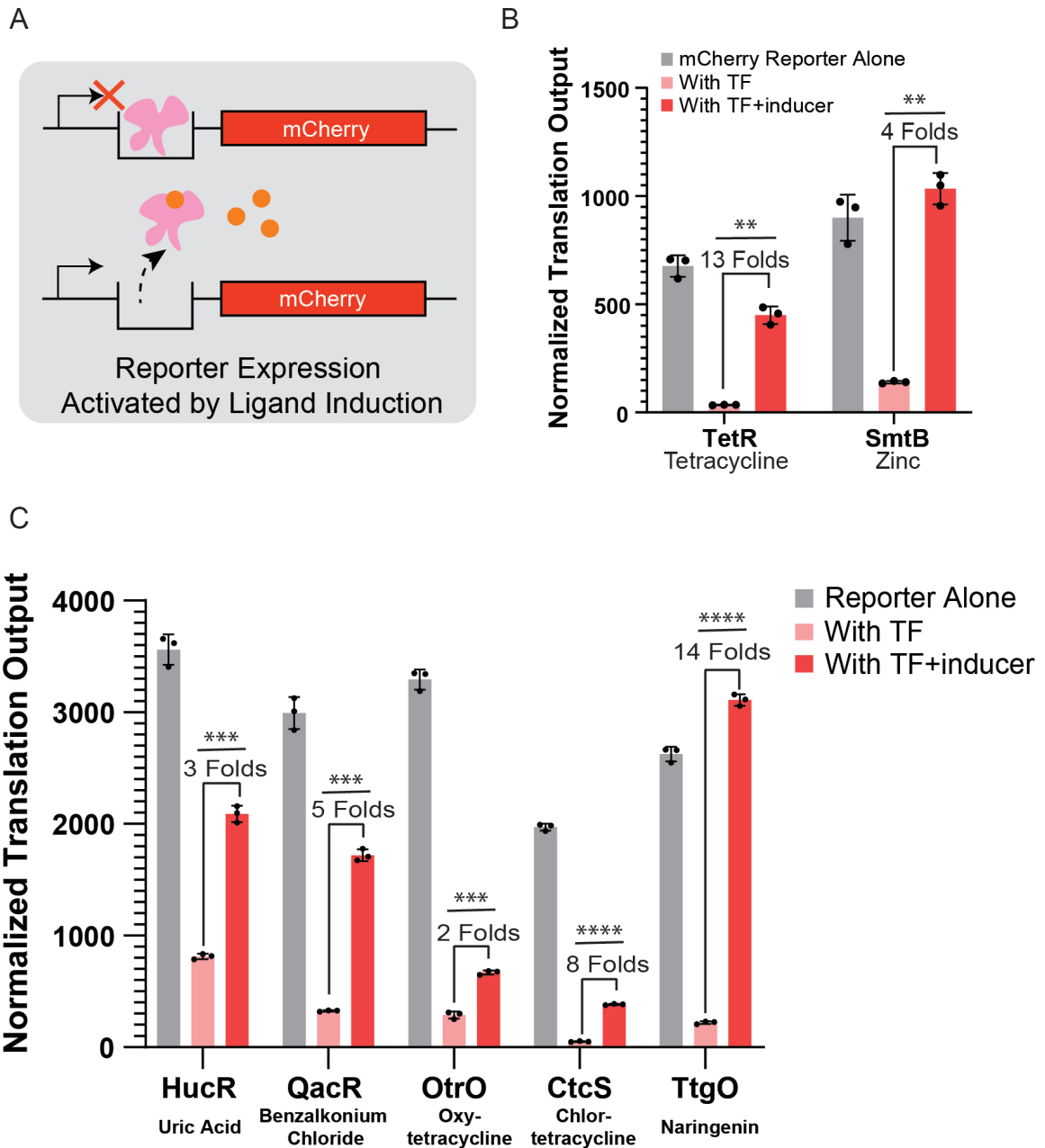
Figure 14: Characterization for aTFs-based biosensors in the PURE system

(A) Schematic of mechanism for aTFs pilot test. A cognate operator was inserted upstream of a mCherry reporter gene. When small molecules are absent, purified aTFs will bind to the operator to suppress mCherry reporter transcription, while when present, small molecules will bind to aTFs and then release it from the operator, leading to transcription and translation of reporter gene.
(B) Constitutive mCherry expression implemented with tetR and smtB biosensors. tetR and smtB proteins were produced, extracted and purified before being added to the PURE reaction mix. Both

tetR and smtB biosensors can successfully suppress mCherry expression and activate upon inductions of corresponding inducers.

(C) Potential aTFs systems to control constitutive mCherry expression in the cell-free system.

Data shown are the means of technical triplicate samples with error bars indicating +1 standard deviation. P-values were calculated as described in the methods.

2.3 Recombinase excision circuits for tetracycline and zinc sensing

With a list of potential aTFs that can function with high specificity, sensitivity, and robustness in the PURE system, next, I interfaced the cell-free recombinase genetic circuits with the best-performing aTFs, tetR and smtB, to regulate recombinase transcription. To control the transcription of recombinase, the cognate operator sequence is placed after the promoter of recombinase plasmids. We reason that the binding of aTFs to the operator will prevent T7RNAP from accessing the promoter while ligand binding will result in a conformation change of the aTFs, causing its release from the operator and thus triggering transcription of recombinase and the downstream excision circuit (Fig. 15A).

Dose-response curves for aTFs were performed to determine optimal aTF concentrations (Fig. 15B-D). The results demonstrate that different recombinases might require different optimal aTF concentrations. The reason for this is the difference in recombinase activity. Compared with PhiC and Bxb1, Cre recombinase is much more active and efficient. Thus, even a tiny amount of leaky Cre expression will result in a strong fluorescent reporter signal, making robust repression necessary to shut down the leaky Cre expression entirely. At the optimal concentrations of aTFs, the introduction of tetracycline and zinc effectively triggers the activation of Cre, PhiC, and Bxb1 excision circuits, yielding a robust greater-than-6-fold change in mCherry expression levels between the absence and presence of the corresponding inducers (Fig. 15E-F).

After the optimal aTFs concentration was determined, a dose-response curve for small molecule inducers was performed to determine the detection window of the recombinase-based sensing circuits (Fig. 16A-B). It is essential to consider that the

detection window for inducers varies among recombinase genetic circuits due to differences in their activities. For instance, stronger recombinases such as Cre necessitate higher concentrations of aTFs to fully deactivate circuit activity. As a result, these circuits exhibit a detection window that is shifted to the right when compared to weaker recombinases like PhiC and Bxb1, detecting higher dosages of inducers. This diversity in recombinase activities allows us to broaden the detection window for inducers, as illustrated in Fig. 16A and 16B. Furthermore, we assessed the limit of detection (LoD) for both zinc and tetracycline-induced sensors to determine the input dose that achieves detection above negative controls (Table 4). The fold change was calculated to determine the maximum change in reporter output. We found the limit of detection (LoD) for zinc-induced recombinase sensing was 1.17 uM for Cre, 2.69 uM for PhiC, and 4.45 uM for Bxb1, with fold changes of 20.7x, 8.14x, and 12.8x respectively. These values indicate that for zinc-induced sensing, Cre achieves both the highest fold change in reporter response and detection at the lowest dose of zinc input. For tetracycline-induced sensing, the detection window was between 1.08 – 59.0 uM for Cre, 0.254 – 19.6 uM for PhiC, and 0.366 – 22.5 uM for Bxb1, with fold changes of 14.0x, 14.6x, and 5.76x respectively. These results show that while PhiC has an initial lower limit of detection, Cre has the broadest range in detectable tetracycline. Moreover, tetracycline at high dosages (>12.5uM) was also observed to negatively affect the reporter gene expression. This is most likely due to tetracycline's inhibition of bacterial protein synthesis by preventing the association of aminoacyl-tRNA with the bacterial ribosome, as reported in the previous study[61] (Fig. 16A). Additionally, we plotted the ROC curves and calculated AUC to quantify assay

performance across sensitivity and specificity classification thresholds (Fig. 17). For tetracycline-induced sensing, the ROC curve analysis showed AUC values of 0.857 for Cre, 0.730 for PhiC, and 0.841 for Bxb1. These results demonstrate the promising ability of the assay to distinguish both true positive and true negative results.

In addition to the tetracycline and zinc biosensors, uric acid-sensing aTF has also proved to be compatible with Cre, PhiC, and Bxb1 recombinase excision circuits and thus can be considered as a potential candidate to expand the repertoire of applications for CRIBOS biosensing (Fig. 16C).
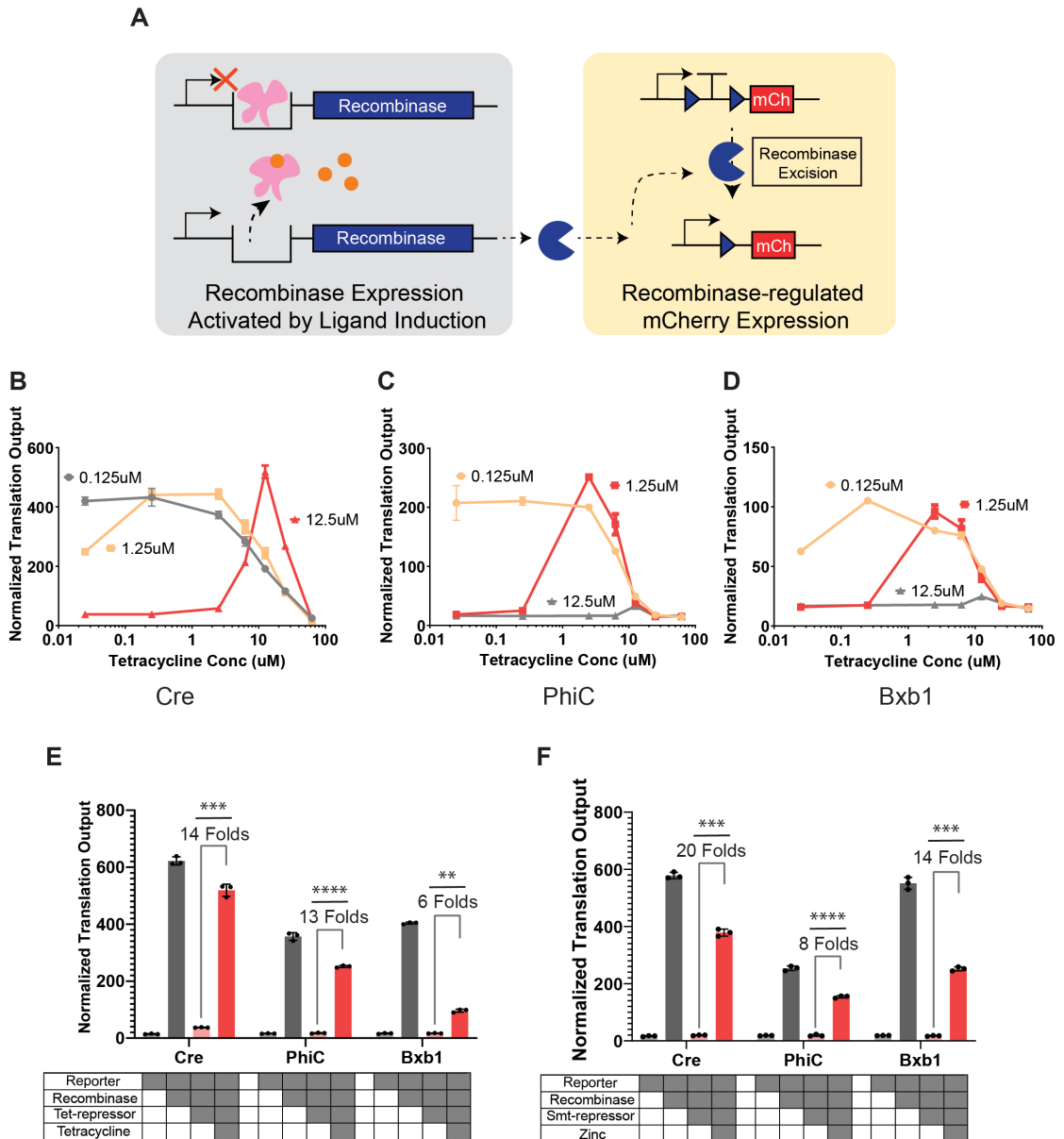
Figure 15: Recombinase excision circuits for tetracycline and zinc Sensing

(A) Schematic of mechanism for small molecule-sensing recombinase excision circuits. Operator sequences are cloned upstream of recombinase-expressing genes. The presence of inducers can trigger recombinase transcription and thus activate the downstream excision circuits.

(B-D) Tetracycline dosage curves for (B) Cre, (C) PhiC, and (D) Bxb1 under different dosages of tetR proteins. The results indicate that the optimal tetR protein dosage is 12.5uM for the Cre excision circuit and 1.25uM for the PhiC and Bxb1 excision circuits.

(E-F) Cre, PhiC, and Bxb1 excision circuits can successfully detect (E) tetracycline and (F) zinc ions when implemented with aTFs-based biosensors.

Data shown are the means of technical triplicate samples with error bars indicating +1 standard deviation. P-values were calculated as described in the methods.
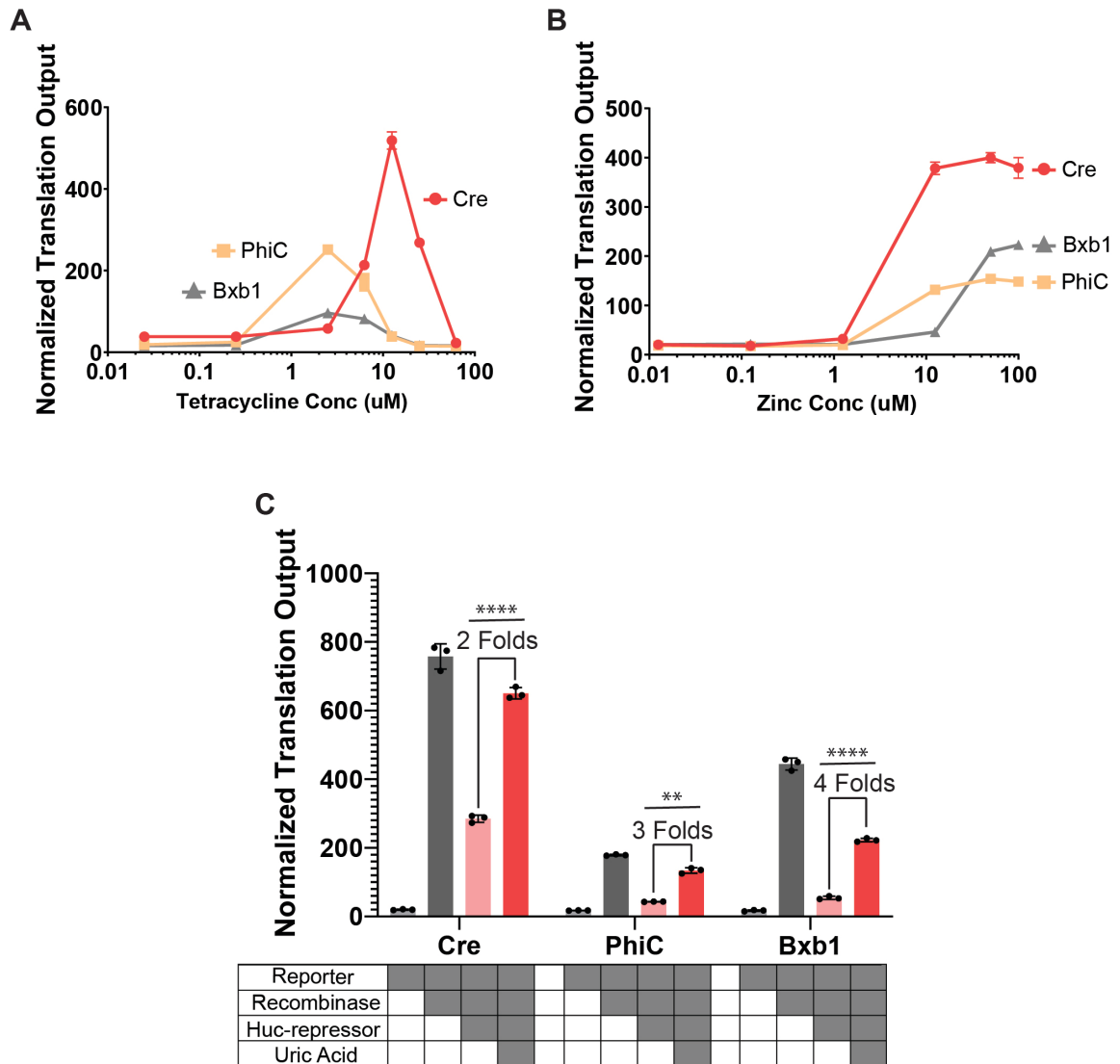
Figure 16: Dynamic range of tetracycline and zinc ion sensing by recombinase excision circuits

(A-B) Detection windows of (A) tetracycline and (B) zinc ions by Cre, PhiC, and Bxb1 recombinase excision circuits are determined.
(C) HucR-implemented recombinase excision circuits can efficiently detect uric acid.

Data shown are the means of technical triplicate samples with error bars indicating +1 standard deviation. P-values were calculated as described in the methods.

Table 4. Detection window, dynamic range and limit of detection for aTFs-based recombinase excision circuits.

tetR-tetracycline Biosensor			smtB-Zinc Biosensor		
Recombinase	Detection Window (uM)	Dynamic Range of Sensing (uM)	Recombinase	Limit of Detection (uM)	Dynamic Range of Sensing (uM)
Cre	1.08 – 59.0	14.0	Cre	1.17	20.7
PhiC	0.254 – 19.6	14.6	PhiC	2.69	8.14
Bxb1	0.366 – 22.5	5.76	Bxb1	4.45	12.8
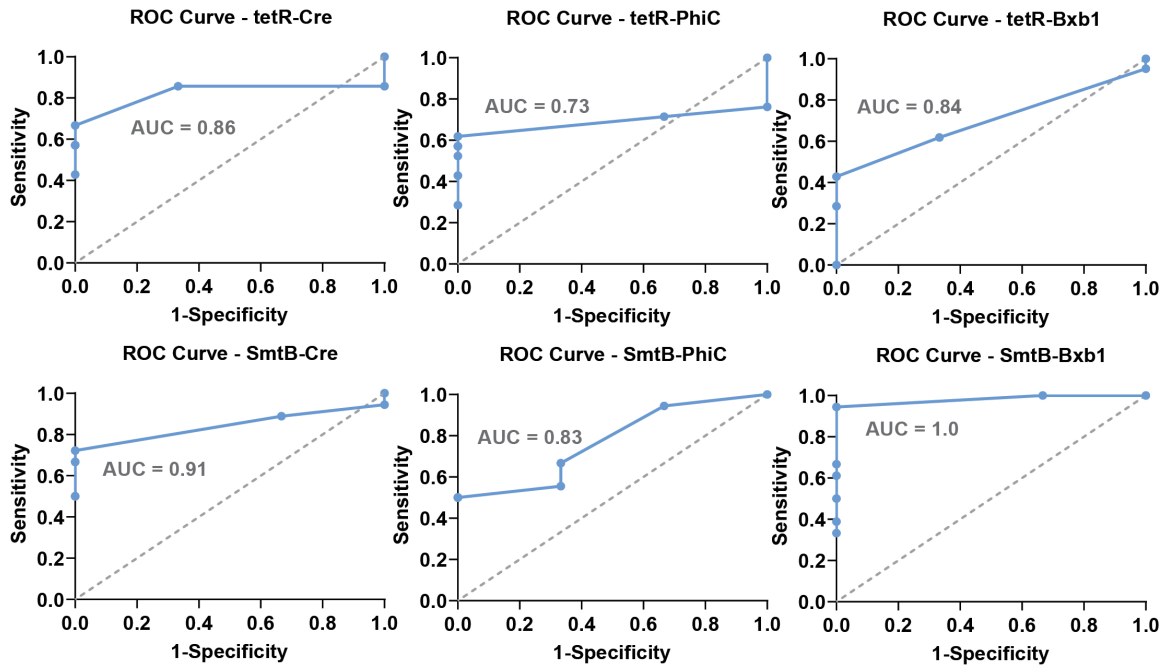
Figure 17: ROC curves (receiver operating characteristic curve) and AUC (Area Under the ROC Curve) for aTFs-controlled recombinase excision circuits

2.4 Tetracycline and zinc sensing multi-input-multi-output Boolean logic gates

Having successfully constructed small molecule sensing recombinase excision circuits within the PURE system, we next delved into establishing more intricate multi-input-multi-output Boolean logic gates for environmental sensing in the cell-free context. TetR and smtB, both extensively studied, were chosen for constructing multiplex sensing circuits due to their highest protein production yield and best performance demonstrated in the previous experiments.

In the construction of genetic circuits, the respective cognate operator sequences, tetO for tetR and smtO for smtB, were cloned upstream of Cre and PhiC recombinases, respectively (Fig. 18A). Subsequently, tetR-controlled Cre and smtB-controlled PhiC were integrated into 2-input-1-output circuits and a 2-input-4-output BLADE decoder. Leveraging the previously optimized experimental procedure, eight 2-input-1-output circuits and the BLADE decoder were successfully assembled (Fig. 18B-C). The 2-input 1-output circuits with biosensors yield 7 over 8 circuits (88%) with a VP angle smaller than 15° and the VP angle metric for the BLADE circuits was evaluated to be less than 25°, for all fluorescent and luminescent outputs (Table 5 and 6). They allowed for the parallel detection of tetracycline and zinc ions, proving the versatility and applicability of CRIBOS and aTFs-based biosensing platforms.

A

Cre Expression Activated
by Tetracycline Induction

Phi Expression Activated
by Zinc Induction

Recombinase
Excision

mCherry Expression of
2-input 1-output Reporter

B

C

Figure 18: Multi-input-multi-output circuits for small molecule sensing

(A) Schematics of implementing biosensors with multi-input-multi-output Boolean logic gates. TetR biosensor was implemented to control Cre transcription, while the smtB biosensor was implemented to control PhiC transcription.
(B) Eight 2-input-1-output excision circuits were combined with tetR and smtB sensors to detect tetracycline and zinc ions in parallel. To set up the experiment, recombinase plasmids and water were pre-mixed and incubated in the cell-free reaction for 30 minutes at 37°C, and then reporter plasmids were added to the reaction and incubated for another 5hr.
(C) The 2-input-4-output BLADE circuit was combined with tetR and smtB sensors to detect tetracycline and zinc ions in parallel. To set up the experiment, recombinase plasmids and water were pre-mixed and incubated in the cell-free reaction for 30 minutes at 37°C, and then reporter plasmids were added to the reaction and incubated for another 5hr.

Data shown are the means of technical triplicate samples with error bars indicating +1 standard deviation. P-values were calculated as described in the methods.

Table 5. Vector proximity (VP) angles for 2-input-1-output circuits with aTFs-based biosensors.

2 Input 1 Output Circuits With aTFs-based Biosensors			
Circuit Type	VP Angle	Circuit Type	VP Angle
NOR	24.3	NOT A	9.3
AND	1.3	NOT B	5.7
NAND	4.1	A NIMPLY B	1.7
A IMPLY B	9.2	B NIMPLY A	13.3

Table 6. Vector proximity (VP) angles for 2-input-4-output decoders with aTFs-based biosensors.

2 Input 4 Output Decoder with aTFs-based Biosensors					
Output	VP Angle	Output	VP Angle	Output	VP Angle
GFP	10.5	mCherry	7.8	NLuc	23.4

2.5 Conclusions and future perspective

In this chapter, we showcase a practical application of the CRIBOS platform by integrating it with aTFs-based water contaminant sensors. This successful implementation validates the feasibility of complex bio-computation with recombinase genetic circuits in cell-free conditions and opens avenues for real-world applications such as environmental monitoring and water contamination detection. Incorporating aTFs enhances the versatility of CRIBOS, making it a valuable tool from both academic research and industry perspectives.

While our demonstration focused on tetracycline, zinc ions, and uric acid biosensors, it is crucial to recognize that a vast array of aTFs designed for different chemicals and environmental stimuli has been discovered or engineered. This diversity enables the potential integration of various aTFs into the CRIBOS platform for more comprehensive and multiplex sensing applications in future research. The robust and efficient nature of cell-free recombinases establishes a dependable foundation, allowing for flexible modifications and integrations of diverse circuit modules within CRIBOS.

Expanding beyond the scope of aTFs-based sensors, CRIBOS can potentially be adaptable to accommodate other sensor types. For instance, in mammalian cells, recombinase genetic circuits have demonstrated compatibility with chemically induced dimerization (CID)-based sensors (Fig. 19). The successful integration of CID-based sensors with recombinase genetic circuits introduces the potential to detect changes in temperature, light, and a broader range of chemicals. With its high modularity and adaptability, exploration can be done to implement CRIBOS with the CID-based platform,

broadening its potential sensing capabilities and making it a versatile platform for diverse applications.

By incorporating aTFs-based small molecule sensors, the CRIBOS platform simultaneously responds to multiple chemical inducers. This capability is particularly relevant for multiplex environmental sensing and water contaminant detection, further emphasizing the practical utility and significance of the CRIBOS system in various domains.

65
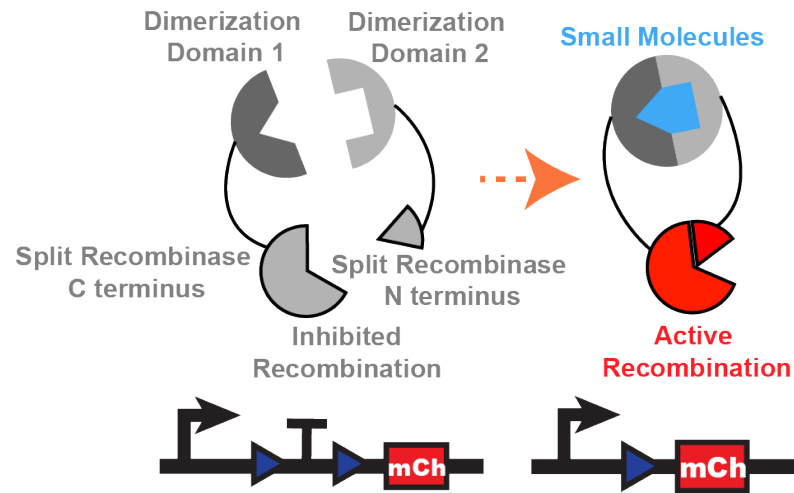
Figure 19: Schematic of mechanism for CID-based biosensor

CID contains two dimerization domains (indicated as 1 and 2 in the figure), which can be recruited to form a dimer with a specific target environmental stimulus present. Recombinases are split into two catalytically inactive halves, and each half is attached to one of the CID domains. Two inactive halves of the recombinases are separated when the target input is absent. In contrast, when input is present, two CID domains will be recruited and form a dimer, reconstituting two halves of recombinases to form an active recombinase, thus triggering the downstream genetic circuit execution.

2.6 Materials and methods

Protein production and purification

All aTFs expression plasmids were purchased from Addgene. For aTFs production, plasmids were transformed into Rosetta 2 (DE3) pLysS E. coli strain. A single colony of transformed bacteria was cultured in liquid LB media overnight. The next day, overnight cultures were then diluted 1:100 into 200mL-500mL culture media. OD of culture media was measured by a plate reader every 30 minutes. When OD reached 0.5-0.8, culture was induced with 100uM Isopropyl β-D-thiogalactoside (IPTG) (Sigma-Aldrich, I6758-5G) overnight at 30°C. After overnight IPTG induction, bacterial culture was pelleted through a centrifuge at 3500Xg for 20 minutes. Every 1g of bacterial pellet is resuspended in 10ml SoluLyse lysis buffer (Amsbio, L200500) supplemented with 1 tablet of protease inhibitor cocktail (Sigma-Aldrich, 11836170001), 40mM imidazole (Fisher Scientific, O3196-500), 200uL Lysozyme (Sigma-Aldrich, L4919-500MG) and 0.5uL Benzonase (EMD Millipore, 71205-3). Lysozyme was pre-resuspended to 10mg/mL in water and stored at -80°C upon arrival. The resuspended cell pellet was then placed on a mixer for 20 minutes and centrifuged at 4000g for 10 minutes at 4°C to remove bacterial debris post-incubation. The bacterial supernatant was then sterilely filtered with a 0.22μm filter (EMD Millipore, SCGP00525) to remove bacterial precipitate and avoid clogging in FPLC columns.

Clarified supernatants were purified using His-tag affinity chromatography with a HisTrap column (Cytiva, 17-0406-01), followed by Heparin chromatography with a heparin column (Cytiva, 17-5247-01) using an AKTAxpress fast protein liquid chromatography system. The fractions collected from the FPLC were buffer-exchanged and concentrated utilizing

a protein concentrator (Thermo Scientific, 88515). Protein concentrations were evaluated using Nanodrop. The purity and size of proteins were validated with SDS-PAGE and western blot. Purified proteins were stored in PBS at 4°C.

Data analysis

For each inducer and recombinase combination, the limit of blank (LoB) was calculated as the mean of the blank added to three times the standard deviation of the blank. The limit of detection (LoD) was calculated as the LoB plus three times the standard deviation of low concentration sample. The dynamic range was computed as the mean expression of highest output for tetracycline and zinc positive samples over the mean expression of negative samples. False positive rate and true positive rate were used to construct the receiver operating characteristic (ROC) curve and compute the area under the curve (AUC).

CHAPTER THREE: Paper-based CRIBOS for biological memory storage

3.1 Introduction

The ability to record events is a defining characteristic of complex systems. Human history has been recorded using the written word to remember facts, dates, and historical figures. Computers keep "memory" in silicon chips to record data, communicate, and save states. Human brains memorize knowledge and information for individual adaptation, while organisms memorize genetic information for evolution. This capability to record crucial biological events not only shapes research but also extends to diverse applications in disease diagnostics, medical therapeutics[62], biofuel manufacturing[63], and agricultural production[6]. For example, immunological memory enables the development of innovative therapeutics such as vaccines and immunotherapy[64,65]. Additionally, DNA barcoding plays a pivotal role in detecting invasive organisms and uncovering novel species[66–68]. These applications underscore the far-reaching impact of biological memory storage in advancing various fields of science and technology.

DNA has been used to transmit genetic information to the following generations because of its high stability, storage capacity, small volume, and inheritance. Moreover, employing DNA as the substrate for memory storage facilitates direct interfacing with sensors capable of detecting significant biological and environmental signals. This integration allows for the activation of biological programs in response to various environmental changes and bridges the gap between sensing and biological responses. Site-specific recombinase permanently modifies DNA sequence with high fidelity and efficiency, serving as a powerful writer to encode biological information into the DNA-

based memory storage medium. Thus, with the success of building recombinase-based Boolean logic gates in a cell-free environment, we created a portable, robust, and stable biological memory storage device with the CRIBOS platform (Fig. 20).
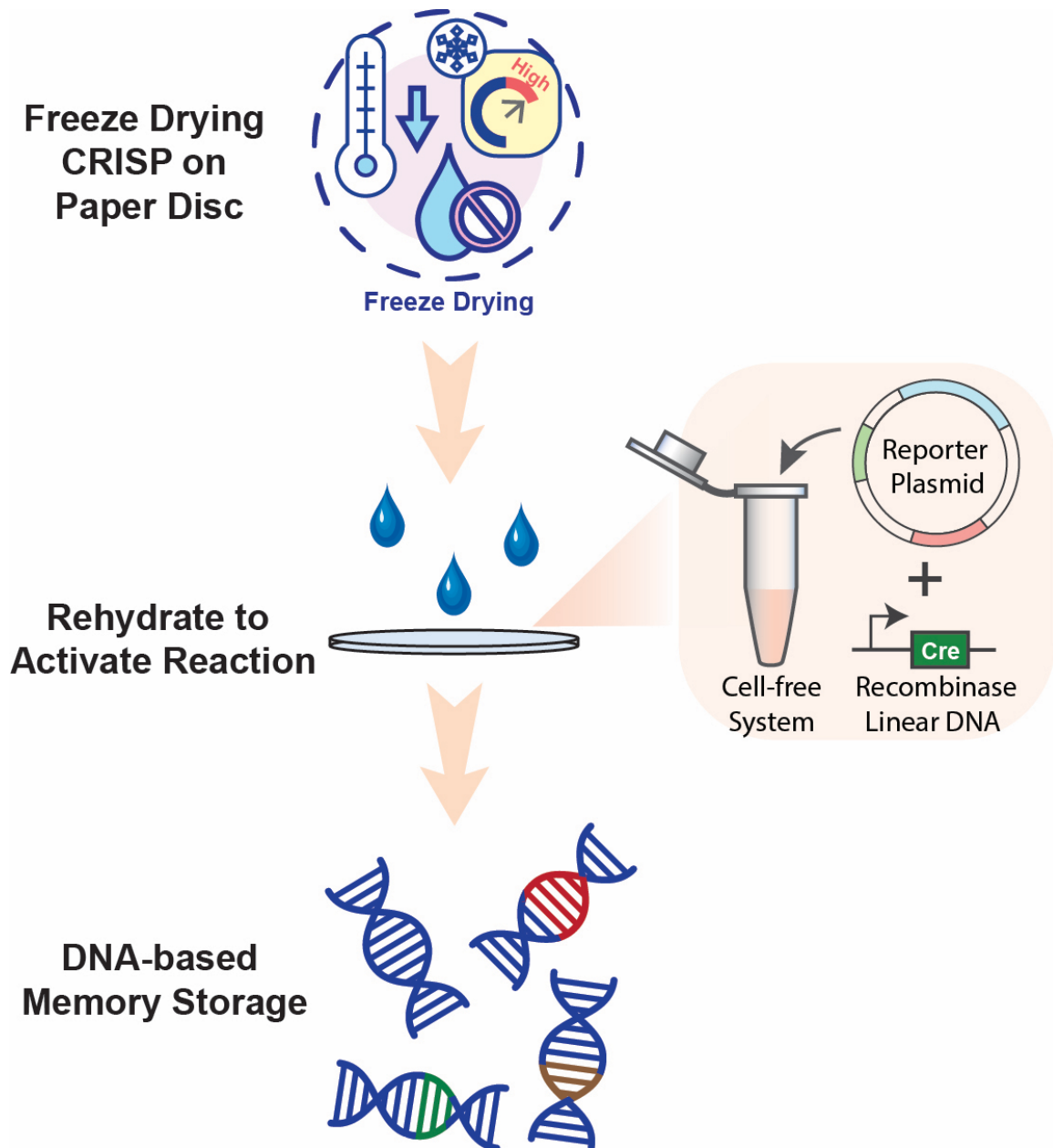
Figure 20: DNA-based memory storage with cell-free recombinase genetic circuits

A DNA-based memory storage device can be developed with cell-free recombinase genetic circuits. The cell-free reaction mix with recombinase genetic circuits will first be lyophilized on a paper disc. The dried reaction will be activated through water rehydration. When the recombinase transcription and expression are triggered by environmental stimuli, recombinase will execute and rearrange the DNA sequence on the reporter plasmid, where biological memory will be stored. Such information can be collected from the paper disc and analyzed through sequencing or PCR upon retrieval.

3.2 Paper-based CRIBOS

To evaluate the portability potential, we apply CRIBOS in a paper-based setting. One benefit of the PURE system is that it can be lyophilized on a sterile and abiotic paper-based platform. The lyophilized paper-based cell-free reaction is stable at room temperature for over a year and can be activated through rehydration[69,70]. In contrast, liquid-based reactions must be stored at -80°C to maintain enzyme activity.

In addition to device portability, memory stability is critical for a memory storage device. DNA memory can not only be effectively modified on paper through recombinase circuits but can also be stably preserved under room temperature for over four months upon activation (Fig. 21B). Paper-based CRIBOS with a plasmid reporter DNA and linear recombinase expressing gene was activated on day 0 through rehydration, and information was stored on the paper. At different time points post-activation, memory stored in the paper-based device was retrieved and analyzed. The reporter DNA with biological memory was retrieved through elution, and then collected DNA was used to transform bacteria to express mCherry reporter output (Fig. 21A). The recombinase excision rate was evaluated based on the percentage of mCherry-expressing bacteria colonies. High excision efficiency in the Reporter+Cre condition indicates that the information was well-preserved at room temperature, and no DNA mutation is observed based on the sequencing results (Fig.22).

A

B
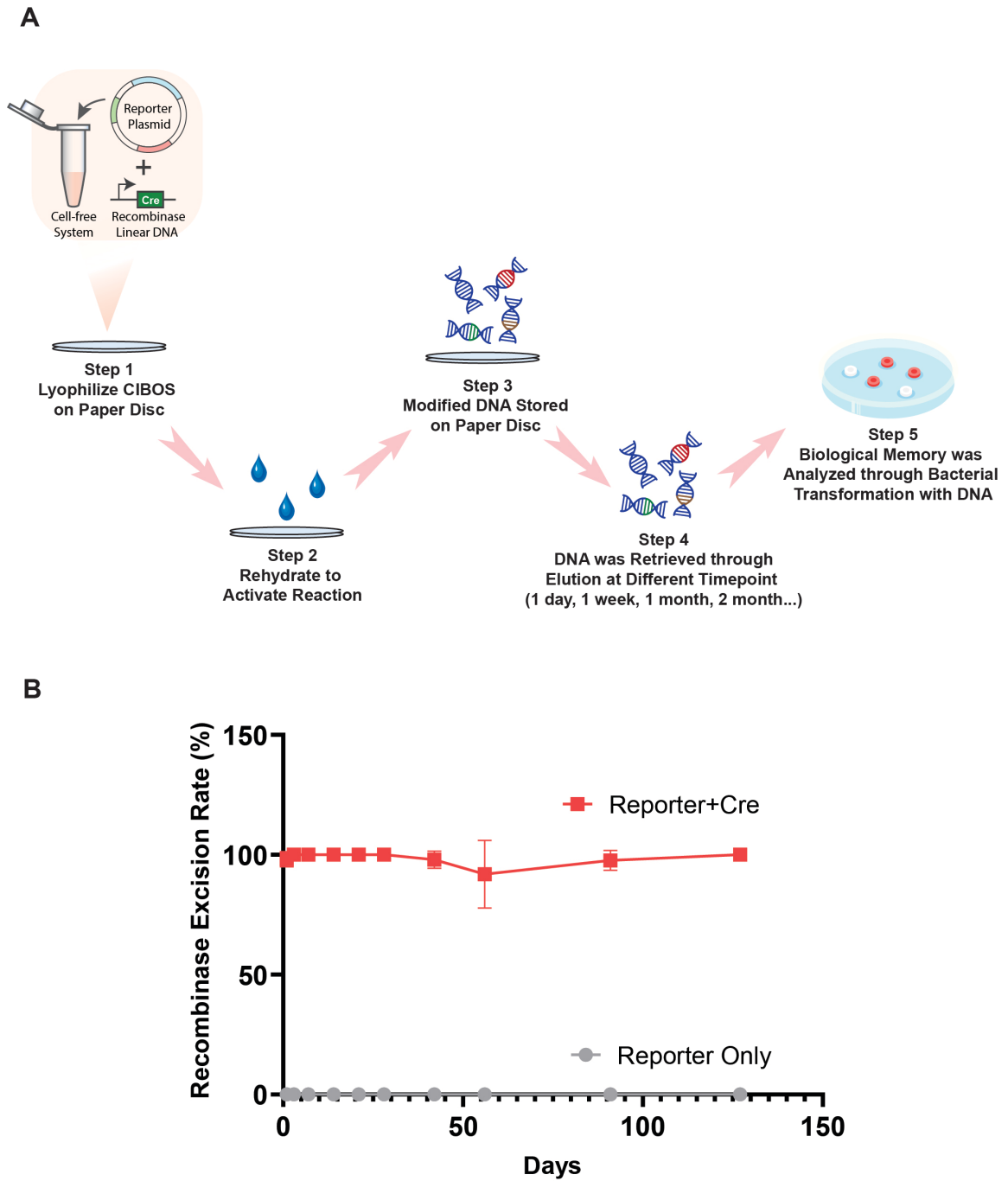
Figure 21: DNA-based memory storage device

(A) Schematic of the experimental process for paper-based cell-free recombinase setup. Reporter plasmid and linear recombinase DNA are mixed with PURE reaction mix and then freeze-dried onto a paper disc for stable storage under room temperature. Then, the reactions are activated through rehydration. Modified DNA-based memory was stored on the paper disc for different

lengths of duration, ranging from 1 day to over 4 months. On the day of analysis, DNA-based memory is retrieved from a paper disc and transformed into bacteria for signal output expression. (B) Memory stability evaluation for paper-based memory storage devices. Paper-based CRIBOS was activated through rehydration on day 0. DNA-based memory was collected on different days up until six months. Results showed that the DNA-based memory can be stored on a paper disc at room temperature with high stability with minimal sequence mutation.

Data shown are the means of technical triplicate samples with error bars indicating +1 standard deviation. P-values were calculated as described in the methods.

\>> Reporter Alone sequence:
TAATACGACTCACTATAGGAGCGGCCGCACTAGTTACCGTTCGTATAGCATACATT
ATACGAAGTTATCCAGGCATCAAATAAGGATCCAAACTCGAGTAAGGATCTCCTAG
CATAACCCCGCGGGGCCTCTTCGGGGGTCTCGCGGGGTTTTTTGCTGAAAGAA
GCTTCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGT
TGTTTGTCGCTGCGCGGCCGCCCTAGCATAACCCCTTGGGGCCTCTAAACGGGT
CTTGAGGGGTTTTTTGGTCGACCTAGCATAACCCCGCGGGGCCTCTTCGGGGGT
CTCGCGGGGTTTTTTGCTGAAAGAAGCTTCAAATAAAACGAAAGGCTCAGTCGAA
AGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGCTGCGCGGCCGCCCTAGCATA
ACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCCATAACTTCGTATA
GCATACATTATACGAACGGTACCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGA

\>> Reporter+Cre White Colony sequence:
TAATACGACTCACTATAGGAGCGGCCGCACTAGTTACCGTTCGTATAGCATACATT
ATACGAAGTTATCCAGGCATCAAATAAGGATCCAAACTCGAGTAAGGATCTCCTAG
CATAACCCCGCGGGGCCTCTTCGGGGGTCTCGCGGGGTTTTTTGCTGAAAGAA
GCTTCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGT
TGTTTGTCGCTGCGCGGCCGCCCTAGCATAACCCCTTGGGGCCTCTAAACGGGT
CTTGAGGGGTTTTTTGGTCGACCTAGCATAACCCCGCGGGGCCTCTTCGGGGGT
CTCGCGGGGTTTTTTGCTGAAAGAAGCTTCAAATAAAACGAAAGGCTCAGTCGAA
AGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGCTGCGCGGCCGCCCTAGCATA
ACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCCATAACTTCGTATA
GCATACATTATACGAACGGTACCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGA

\>> Reporter+Cre Red Colony sequence:
TAATACGACTCACTATAGGAGCGGCCGCACTAGTTACCGTTCGTATAGCATACATTA
TACGAACGGTACCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGA

Figure 22: Sequencing validation of DNA-based memory storage device

DNA-based memory validation through sequencing. DNA eluted from paper discs was used to transform bacteria, and colonies were picked. The three representative examples shown here are from a Reporter (white colony) only condition and a Reporter+Cre (white and red) condition. The T7 promoter is indicated by green, Cre recognition sites are blue, terminators are red, and ribosome binding site is yellow. Sequencing data indicates that for mCherry positive colony, the terminator sequence between the recognition sites was successfully excised, while for white colony, the terminator still locates in front of the reporter gene.

3.3 Conclusions and future perspective

In conclusion, the successful application of CRIBOS in a paper-based setting represents a significant advancement in the portability and stability of genetic circuit platforms. Notably, the paper-based CRIBOS system stands as the pioneering genetic circuit with memory capabilities, showcasing an unprecedented longevity of over four months with minimal resource, energy costs, and maintenance requirements. The exploration and development of methods for encoding, storing, and retrieving memory from the paper-based CRIBOS were conducted to validate the device's functionality and procedure's versatility.

While the current memory retrieval method is proved to be cost- and resource-effective, future studies will explore more sensitive and rapid detection methods for field applications. Moreover, the CRIBOS memory devices can be tested and characterized under an array of reaction conditions, fostering adaptability to diverse environments. Although the current application of CRIBOS is limited to a paper-based setting, the proven functionality of cell-free reactions in various environments, such as microfluidics and flow strips, opens avenues for future research. Subsequent studies may focus on integrating CRIBOS with microfluidic devices to further minimize the resources required for individual applications.

In summary, the developed CRIBOS system not only establishes a portable, robust, and stable biological memory storage device but also paves the way for continued innovation and expansion of its utility across different settings and applications (Fig. 23).
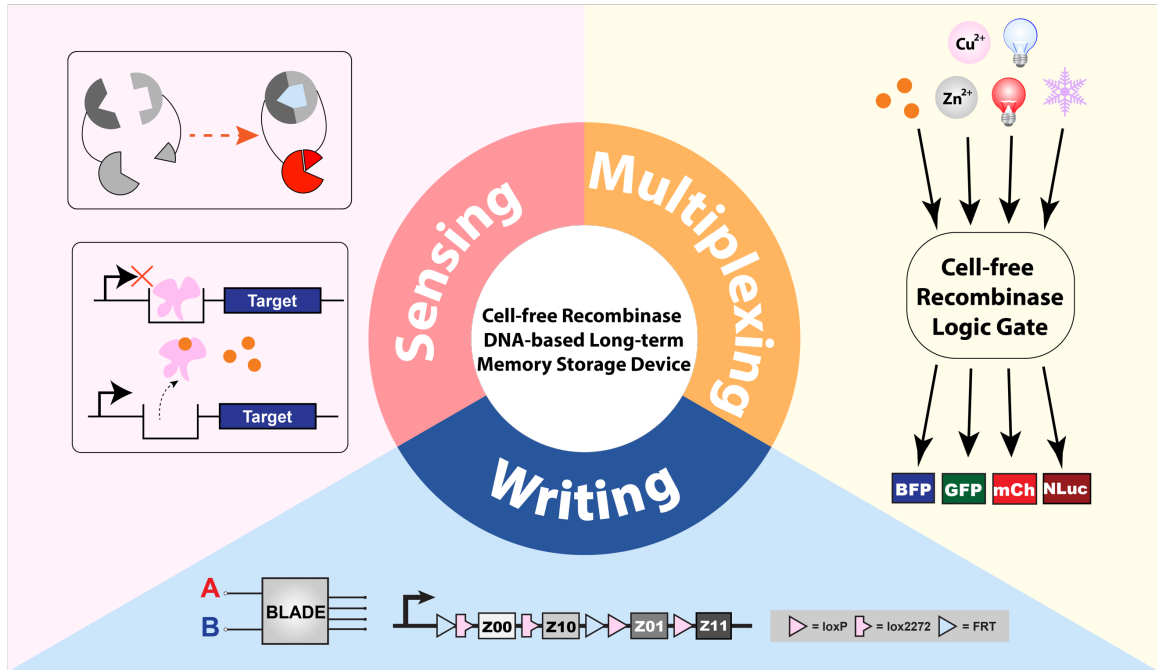
Figure 23: Overview of a DNA-based long-term memory storage device

Combining Chapters 1, 2, and 3, in this project, we developed a DNA-based long-term memory storage device with CRIBOS, a cell-free recombinase Boolean logic platform. By implementing aTF-based sensors with cell-free Boolean logic gates, such a device allows for multiplex environmental sensing and long-term recording, achieving the first biological memory recording system that can function under low-cost and low-resource settings.

3.4 Materials and methods

Freeze-drying

Filter paper (Cytiva, 1442-070) for depositing and freeze-drying cell-free reactions was first treated with 5% bovine serum albumin (Sigma-Aldrich, A3059-100G) overnight at 37°C. The paper was then cut into 2-mm-diameter paper discs with a biopsy punch and placed at the bottom of a 96-well PCR plate. Before lyophilization, 2μL of cell-free reaction mix was applied directly onto the paper disc, and the PCR plate was sealed by porous AeraSeal Sealing Films (Excel Scientific, B-100). Post-freezing by liquid nitrogen, the PCR plate was immediately sent for lyophilization overnight. After lyophilization, the paper discs were then inserted into the bottom of the 384-well plate and activated by adding 2μL of autoclaved DI water.

Molecular cloning for linear DNA

Linear DNA template with T7 promoter, followed by recombinase expressing gene and a T7 terminator, was PCR (CoWin Biosciences CW2965F) amplified from the corresponding recombinase plasmids. Amplified DNA was purified with a PCR cleanup kit (Epoch Life Science 2360250), and band size and sequence were verified through gel electrophoresis and sequencing.

CHAPTER FOUR: Conclusion and discussion

As we delve into the world of synthetic biology, we aim to provide useful tools for engineering complex, logical control over gene expression. Synthetic genetic circuits, in particular, have demonstrated incredible promise in this field. Advances in synthetic circuit design and fabrication have enabled biological systems to be programmed in a similar manner as computers, with user-defined inputs and outputs in the form of proteins and regulatory RNAs.

In this project, in order to enrich the limited toolkit and implementation of bio-computing devices for the cell-free system, we engineered CRIBOS, the first and largest comprehensive Boolean logic platform of site-specific recombinase-based gene circuits that can execute complex multi-input-multi-output biocomputation in the cell-free environment, upon induction by various environmental stimuli, including antibiotics such as tetracycline and heavy metal zinc. Leveraging CRIBOS, we developed 3 categories of more than 20 different logic gates of 6 recombinases to function with high orthogonality and high efficiency in the cell-free system. The construction of a 2-input-4-output Boolean decoder demonstrates the potential of building multi-input-multi-output Boolean logic gates in the cell-free environment through an array of orthogonal recombinases, further enhancing the complexity of Boolean logics to be constructed.

In addition, through characterization of the simplest excision circuits, we identified the essential design rules for cell-free recombinase-based logic gates which were proved to be widely applicable to different types of recombinase-based logic in the cell-free environment. The streamlined reporter optimization process can significantly facilitate the

potential expansion of the CRIBOS to generate highly complex Boolean logic gates in the cell-free system.

Moreover, we extend CRIBOS's applicability to a paper-based setting, offering improved portability and stability for genetic circuit platforms. The paper-based CRIBOS represents a memory device capable of sustaining functionality for over 4 months with minimal resource and energy costs. We explore methods for encoding, storing, and retrieving memory from paper-based CRIBOS, laying the groundwork for future advancements in field-deployable memory storage devices. While our focus remains on paper-based applications, the adaptability of cell-free reactions enables potential integration with diverse platforms such as microfluidics and flow strips. Future endeavors may explore combining CRIBOS with microfluidic devices to create portable memory storage devices for continuous and multiplex environmental sensing under resource-constrained settings.

In summary, with the revolutionary properties of the cell-free genetic circuits, such as high circuit robustness, low pathway interference, high composition control, and amenability for modularity, CRIBOS will accelerate the toolkit expansion for Boolean logic design in a cell-free environment, with enormous potential in basic academic research, industrial manufacturing, and environmental monitoring. It is the most advanced toolkit for Boolean genetic logic gate in the cell-free system, which we believe will ultimately lead to a more thorough understanding of the complex biological genetic circuit design and more diverse applications in the field of detection, diagnostics, biomanufacturing, and therapeutics.

BIBLIOGRAPHY

1. Beitz AM, Oakes CG, Galloway KE. Synthetic gene circuits as tools for drug discovery. *Trends in Biotechnology*. 2022;40(2):210–225. doi:10.1016/J.TIBTECH.2021.06.007

2. Riglar DT, Silver PA. Engineering bacteria for diagnostic and therapeutic applications. *Nature Reviews. Microbiology.* 2018;16(4):214–225. doi:10.1038/nrmicro.2017.172

3. Jung C, Ellington AD. Diagnostic applications of nucleic acid circuits. *Accounts of Chemical Research*. 2014;47(6):1825–1835. doi:10.1021/AR500059C/ASSET/IMAGES/LARGE/AR-2014-00059C_0009.JPEG

4. Slomovic S, Pardee K, Collins JJ. Synthetic biology devices for in vitro and in vivo diagnostics. *Proceedings of the National Academy of Sciences of the United States of America*. 2015;112(47):14429–14435. doi:10.1073/PNAS.1508521112/ASSET/4DBE653E-8604-42A8-8C71-CA6E5838A898/ASSETS/GRAPHIC/PNAS.1508521112FIG02.JPEG

5. Robinson CM, Short NE, Riglar DT. Achieving spatially precise diagnosis and therapy in the mammalian gut using synthetic microbial gene circuits. *Frontiers in Bioengineering and Biotechnology*. 2022;10:959441. doi:10.3389/FBIOE.2022.959441/BIBTEX

6. de Lange O, Klavins E, Nemhauser J. Synthetic genetic circuits in crop plants. *Current Opinion in Biotechnology*. 2018;49:16–22. doi:10.1016/J.COPBIO.2017.07.003

7. Goold HD, Wright P, Hailstones D. Emerging Opportunities for Synthetic Biology in Agriculture. *Genes*. 2018;9(7):341. doi:10.3390/GENES9070341

8. Saltepe B, Kehribar EŞ, Su Yirmibeşoğlu SS, Şafak Şeker UÖ. Cellular Biosensors with Engineered Genetic Circuits. *ACS Sensors*. 2018;3(1):13–26. doi:10.1021/ACSSENSORS.7B00728/ASSET/IMAGES/LARGE/SE-2017-00728E_0005.JPEG

9. Bernard E, Wang B. Synthetic cell-based sensors with programmed selectivity and sensitivity. *Methods in Molecular Biology*. 2017;1572:349–363. doi:10.1007/978-1-4939-6911-1_23/COVER

10. Xu P. Production of chemicals using dynamic control of metabolic fluxes. *Current Opinion in Biotechnology*. 2018;53:12–19. doi:10.1016/J.COPBIO.2017.10.009

11. Brophy JAN, Voigt CA. Principles of Genetic Circuit Design. *Nature Methods*. 2014;11(5):508. doi:10.1038/NMETH.2926

12. Bundy BC, Hunt JP, Jewett MC, et al. Cell-free biomanufacturing. *Current Opinion in Chemical Engineering*. 2018;22:177–183. doi:10.1016/J.COCHE.2018.10.003

13. Hodgman CE, Jewett MC. Cell-free synthetic biology: Thinking outside the cell. *Metabolic Engineering*. 2012;14(3):261–269. doi:10.1016/J.YMBEN.2011.09.002

14. Swartz J. Developing cell-free biology for industrial applications. *Journal of Industrial Microbiology & Biotechnology*. 2006;33(7):476–485. doi:10.1007/S10295-006-0127-Y

15. Brookwell A, Oza JP, Caschera F. Biotechnology Applications of Cell-Free Expression Systems. *Life (Basel, Switzerland)*. 2021;11(12). doi:10.3390/LIFE11121367

16. Laohakunakorn N. Cell-Free Systems: A Proving Ground for Rational Biodesign. *Frontiers in Bioengineering and Biotechnology*. 2020;8:557668. doi:10.3389/FBIOE.2020.00788/BIBTEX

17. Tinafar A, Jaenes K, Pardee K. Synthetic Biology Goes Cell-Free. *BMC Biology*. 2019;17(1):1–14. doi:10.1186/S12915-019-0685-X

18. Jung JK, Alam KK, Verosloff MS, et al. Cell-free biosensors for rapid detection of water contaminants. *Nature Biotechnology 2020 3812*. 2020;38(12):1451–1459. doi:10.1038/s41587-020-0571-7

19. Mustafa MI, Makhawi AM. Sherlock and detectr: CRISPR-cas systems as potential rapid diagnostic tools for emerging infectious diseases. *Journal of Clinical Microbiology*. 2021;59(3). doi:10.1128/JCM.00745-20/ASSET/F83C19DF-665D-4335-9723-264021982D57/ASSETS/IMAGES/LARGE/JCM.00745-20-F0002.JPG

20. Sherlock ME, Sudarsan N, Stav S, Breaker RR. Tandem riboswitches form a natural Boolean logic gate to control purine metabolism in bacteria. *eLife*. 2018;7. doi:10.7554/ELIFE.33908

21. Marchisio MA. In silico design and in vivo implementation of yeast gene Boolean gates. *Journal of Biological Engineering*. 2014;8(1):1–15. doi:10.1186/1754-1611-8-6/FIGURES/7

22. Lohmueller JJ, Armel TZ, Silver PA. A tunable zinc finger-based framework for Boolean logic computation in mammalian cells. *Nucleic Acids Research*. 2012;

40(11):5180. doi:10.1093/NAR/GKS142

23. Bressler EM, Adams S, Liu R, et al. Boolean logic in synthetic biology and biomaterials: Towards living materials in mammalian cell therapeutics. *Clinical and Translational Medicine*. 2023;13(7):e1244. doi:10.1002/CTM2.1244

24. Hou B, Zhou L, Wang H, et al. Engineering Stimuli-Activatable Boolean Logic Prodrug Nanoparticles for Combination Cancer Immunotherapy. *Advanced Materials*. 2020;32(12):1907210. doi:10.1002/ADMA.201907210

25. Moore SJ, Lai HE, Chee SM, et al. A Streptomyces venezuelae Cell-Free Toolkit for Synthetic Biology. *ACS Synthetic Biology*. 2021;10(2):402–411. doi:10.1021/ACSSYNBIO.0C00581/ASSET/IMAGES/LARGE/SB0C00581_000 4.JPEG

26. Roquet N, Soleimany AP, Ferris AC, Aaronson S, Lu TK. Synthetic recombinase-based state machines in living cells. *Science*. 2016;353(6297). doi:10.1126/SCIENCE.AAD8559

27. Weinberg BH, Pham NTH, Caraballo LD, et al. Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. *Nature Biotechnology*. 2017;35(5):453–462. doi:10.1038/nbt.3805

28. Razavi S, Su S, Inoue T. Cellular Signaling Circuits Interfaced with Synthetic, Post-Translational, Negating Boolean Logic Devices. *ACS Synthetic Biology*. 2014;3(9):676–685. doi:10.1021/SB500222Z/SUPPL_FILE/SB500222Z_SI_008.AVI

29. Ubina T, Vahedi-Hunter T, Agnew-Svoboda W, et al. ExBoX – A simple Boolean exclusion strategy to drive expression in neurons. *Journal of Cell Science*. 2021; 134(20). doi:10.1242/JCS.257212/272145/AM/EXBOX-A-SIMPLE-BOOLEAN-EXCLUSION-STRATEGY-TO-DRIVE

30. Weinberg BH, Pham NTH, Caraballo LD, et al. Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. *Nature Biotechnology*. 2017;35(5):453–462. doi:10.1038/nbt.3805

31. Sheets MB, Tague N, Dunlop MJ. An optogenetic toolkit for light-inducible antibiotic resistance. *Nature Communications*. 2023;14(1). doi:10.1038/S41467-023-36670-2

32. Mairhofer J, Wittwer A, Cserjan-Puschmann M, Striedner G. Preventing T7 RNA polymerase read-through transcription-A synthetic termination signal capable of improving bioprocess stability. *ACS Synthetic Biology*. 2015;4(3):265–273. doi:10.1021/SB5000115/SUPPL_FILE/SB5000115_SI_004.ZIP

33.   Zhang Z, Lutz B. Cre recombinase-mediated inversion using lox66 and lox71: method to introduce conditional point mutations into the CREB-binding protein. *Nucleic Acids Research*. 2002;30(17):e90. doi:10.1093/NAR/GNF089

34.   Grindley NDF, Whiteson KL, Rice PA. Mechanisms of site-specific recombination. *Annual Review of Biochemistry*. 2006;75:567–605. doi:10.1146/ANNUREV.BIOCHEM.73.011303.073908

35.   Merrick CA, Zhao J, Rosser SJ. Serine Integrases: Advancing Synthetic Biology. *ACS Synthetic Biology*. 2018;7(2):299–310. doi:10.1021/ACSSYNBIO.7B00308/ASSET/IMAGES/LARGE/SB-2017-003085_0005.JPEG

36.   Lehr FX, Hanst M, Vogel M, et al. Cell-Free Prototyping of AND-Logic Gates Based on Heterogeneous RNA Activators. *ACS Synthetic Biology.* 2019;8(9): 2163–2173. doi:10.1021/ACSSYNBIO.9B00238/ASSET/IMAGES/MEDIUM/SB9B00238_M009.GIF

37.   Ma D, Li Y, Wu K, et al. Multi-arm RNA junctions encoding molecular logic unconstrained by input sequence for versatile cell-free diagnostics. *Nature Biomedical Engineering*. 2022;6(3):298–309. doi:10.1038/s41551-022-00857-7

38.   Miyamoto T, Razavi S, Derose R, Inoue T. Synthesizing biomolecule-based boolean logic gates. *ACS Synthetic Biology*. 2013;2(2):72–82. doi:10.1021/SB3001112/ASSET/IMAGES/LARGE/SB-2012-001112_0006.JPEG

39.   Grasemann L, Lavickova B, Elizondo-Cantú MC, Maerkl SJ. Onepot pure cell-free system. *Journal of Visualized Experiments: JoVE*. 2021;2021(172). doi:10.3791/62625

40.   Lavickova B, Maerkl SJ. A Simple, Robust, and Low-Cost Method to Produce the PURE Cell-Free System. *ACS Synthetic Biology*. 2019;8(2):455–462. doi:10.1021/ACSSYNBIO.8B00427/SUPPL_FILE/SB8B00427_SI_002.XLSX

41.   Ge X, Luo D, Xu J. Cell-Free Protein Expression under Macromolecular Crowding Conditions. *PLoS One*. 2011;6(12). doi:10.1371/JOURNAL.PONE.0028707

42.   Li J, Gu L, Aach J, Church GM. Improved Cell-Free RNA and Protein Synthesis System. *PLoS One*. 2014;9(9). doi:10.1371/JOURNAL.PONE.0106232

43.   Li S, Li Z, Tan GY, Xin Z, Wang W. In vitro allosteric transcription factor-based biosensing. *Trends in Biotechnology*. 2023;41:1080–1095. doi:10.1016/j.tibtech.2023.03.001

44.   Colclough AL, Scadden J, Blair JMA. TetR-family transcription factors in Gram-negative bacteria: Conservation, variation and implications for efflux-mediated antimicrobial resistance. *BMC Genomics*. 2019;20(1):1–12. doi:10.1186/S12864-019-6075-5/FIGURES/6

45.   Aleksandrov A, Schuldt L, Hinrichs W, Simonson T. Tetracycline-Tet Repressor Binding Specificity: Insights from Experiments and Simulations. *Biophysical Journal*. 2009;97(10):2829. doi:10.1016/J.BPJ.2009.08.050

46.   Liu T, Ramesh A, Ma Z, et al. CsoR is a novel Mycobacterium tuberculosis copper-sensing transcriptional regulator. *Nature Chemical Biology*. 2007;3(1):60–68. doi:10.1038/NCHEMBIO844

47.   Antonucci I, Gallo G, Limauro D, et al. An ArsR/SmtB family member regulates arsenic resistance genes unusually arranged in Thermus thermophilus HB27. *Microbial Biotechnology*. 2017;10(6):1690. doi:10.1111/1751-7915.12761

48.   Endo G, Silver S. CadC, the transcriptional regulatory protein of the cadmium resistance system of Staphylococcus aureus plasmid pI258. *Journal of Bacteriology*. 1995;177(15):4437–4441. doi:10.1128/JB.177.15.4437-4441.1995

49.   Fang C, Zhang Y. Bacterial MerR family transcription regulators: activationby distortion: The mechanism of transcription regulation by MerR. *Acta Biochimica et Biophysica Sinica*. 2022;54(1):25. doi:10.3724/ABBS.2021003

50.   Leyn SA, Rodionov DA. Comparative Genomics of DtxR Family Regulons for Metal Homeostasis in Archaea. *Journal of Bacteriology*. 2015;197(3):451. doi:10.1128/JB.02386-14

51.   Taylor ND, Garruss AS, Moretti R, et al. Engineering an allosteric transcription factor to respond to new ligands. *Nature Methods*. 2016;13(2):177. doi:10.1038/NMETH.3696

52.   Yamamoto K, Ishihama A. Two different modes of transcription repression of the Escherichia coli acetate operon by IclR. *Molecular Microbiology*. 2003;47(1):183–194. doi:10.1046/J.1365-2958.2003.03287.X

53.   Anzai T, Kijima K, Fujimori M, Nakamoto S, Ishihama A, Shimada T. Expanded roles of lactate-sensing LldR in transcription regulation of the Escherichia coli K-12 genome: lactate utilisation and acid resistance. *Microbial Genomics*. 2023;9(5). doi:10.1099/MGEN.0.001015

54.   Pilalis E, Chatziioannou AA, Grigoroudis AI, Panagiotidis CA, Kolisis FN, Kyriakidis DA. Escherichia coli genome-wide promoter analysis: Identification of additional AtoC binding target elements. *BMC Genomics*. 2011;12:238.

doi:10.1186/1471-2164-12-238

55. Chen JX, Steel H, Wu YH, et al. Development of Aspirin-Inducible Biosensors in Escherichia coli and SimCells. *Applied and Environmetnal Microbiology*. 2019;85(6). doi:10.1128/AEM.02959-18

56. Park W, Padmanabhan P, Padmanabhan S, Zylstra GJ, Madsen EL. nahR, encoding a LysR-type transcriptional regulator, is highly conserved among naphthalene-degrading bacteria isolated from a coal tar waste-contaminated site and in extracted community DNA. *Microbiology*. 2002;148(Pt 8):2319–2329. doi:10.1099/00221287-148-8-2319

57. Fu D, Wu J, Gu Y, et al. The response regulator OmpR contributes to the pathogenicity of avian pathogenic Escherichia coli. *Poultry Science*. 2022;101(4): 101757. doi:10.1016/J.PSJ.2022.101757

58. F.M. L, Currin A, Dixon N. Directed evolution of the PcaV allosteric transcription factor to generate a biosensor for aromatic aldehydes. *Journal of Biological Engineering*. 2019;13(1):1–15. doi:10.1186/S13036-019-0214-Z/TABLES/2

59. Hersey AN, Kay VE, Lee S, Realff MJ, Wilson CJ. Engineering allosteric transcription factors guided by the LacI topology. *Cell Systems*. 2023;14(8):645–655. doi:10.1016/J.CELS.2023.04.008

60. Jung JK, Archuleta CM, Alam KK, Lucks JB. Programming cell-free biosensors with DNA strand displacement circuits. *Nature Chemical Biology*. 2022;18(4):385–393. doi:10.1038/s41589-021-00962-9

61. Chopra I, Roberts M. Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. *Microbiology and Molecular Biology Reviews*. 2001;65(2):232. doi:10.1128/MMBR.65.2.232-260.2001

62. Sedlmayer F, Aubel D, Fussenegger M. Synthetic gene circuits for the detection, elimination and prevention of disease. *Nature Biomedical Engineering*. 2018;2(6): 399–415. doi:10.1038/s41551-018-0215-0

63. Jagadevan S, Banerjee A, Banerjee C, et al. Recent developments in synthetic biology and metabolic engineering in microalgae towards biofuel production. *Biotechnology for Biofuels*. 2018;11(1). doi:10.1186/S13068-018-1181-1

64. Varadé J, Magadán S, González-Fernández Á. Human immunology and immunotherapy: main achievements and challenges. *Cellular & Molecular Immunology*. 2020;18(4):805–828. doi:10.1038/s41423-020-00530-6

65. Farber DL, Netea MG, Radbruch A, Rajewsky K, Zinkernagel RM. Immunological memory: lessons from the past and a look to the future. *Nature Reviews. Immunology*. 2016;16(2):124–128. doi:10.1038/nri.2016.13

66. Armstrong KF, Ball SL. DNA barcodes for biosecurity: invasive species identification. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. 2005;360(1462):1813–1823. doi:10.1098/RSTB.2005.1713

67. Fišer Pečnikar Ž, Buzan E V. 20 years since the introduction of DNA barcoding: From theory to application. *Journal of Applied Genetics*. 2014;55(1):43–52. doi:10.1007/S13353-013-0180-Y/TABLES/1

68. Lopez-Vaamonde C, Kirichenko N, Cama A, et al. Evaluating DNA Barcoding for Species Identification and Discovery in European Gracillariid Moths. *Frontiers in Ecology and Evolution*. 2021;9:626752. doi:10.3389/FEVO.2021.626752/BIBTEX

69. Smith MT, Berkheimer SD, Werner CJ, Bundy BC. Lyophilized Escherichia coli-based cell-free systems for robust, high-density, long-term storage. *Biotechniques*. 2014;56(4):186–193. doi:10.2144/000114158/ASSET/IMAGES/LARGE/FIGURE3.JPEG

70. Pardee K, Green AA, Ferrante T, et al. Paper-Based Synthetic Gene Networks. *Cell*. 2014;159(4):940–954. doi:10.1016/J.CELL.2014.10.004

CURRICULUM VITAE