Phage-Enabled Nanotechnology and Novel Sensing Architectures in Ribodevices

by

Griffin Cooper McCutcheon

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2022 by the
Graduate Supervisory Committee:

Alexander Green, Chair
Rizal Hariadi
Nicholas Stephanopoulos
Xiao Wang

ARIZONA STATE UNIVERSITY

August 2022

ABSTRACT

Two distinct aspects of synthetic biology were investigated: the development of viral structures for new methods of studying self-assembly and nanomanufacturing, and the designs of genetic controls systems based on controlling the secondary structure of nucleic acids. Viral structures have been demonstrated as building blocks for molecular self-assembly of diverse structures, but the ease with which viral genomes can be modified to create specific structures depends on the mechanisms by which the viral coat proteins self-assemble. The experiments conducted demonstrate how the mechanisms that guide bacteriophage lambda's self-assembly make it a useful and flexible platform for further research into biologically enabled self-assembly. While the viral platform investigations focus on the creation of new structures, the genetic control systems research focuses on new methods for signal interpretation in biological systems. Regulators of genetic activity that operate based on the secondary structure formation of ribonucleic acid (RNA), also known as riboswitches, are genetically compact devices for controlling protein translation. The toehold switch ribodevice can be modified to enable multiplexed logical operations with RNA inputs, requiring no additional protein transcription factors to regulate activity, but they cannot receive chemical inputs. RNA sequences generated to bind to specific chemicals, known as aptamers, can be used in riboswitches to confer genetic activity upon binding their target chemical. But attempts to use aptamers for logical operations and genetic circuits are difficult to generalize due to differences in sequence and binding strength. The experiments conducted demonstrate a ribodevice structure in which aptamers can be used semi-interchangeably to translate chemical inputs into the toehold switch paradigm, marrying the programmability and

orthogonality of toehold switches with the broad sensing potential of aptamer-based

ribodevices.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Viral Particles in Synthetic Biology

Bacteriophages and their constituent parts have a long history in pioneering the development of genetics[1]. With phages as model organisms, early tools and fundamental concepts were established such as DNA ligases, DNA/RNA hybridization, DNA transduction, cell free protein production, and of course phage display[2,3]. More recently, phage-related research led to the discovery and development of the CRISPR-Cas9 system[4]. While many of these research advances are the result of interrogating a natural function of phages, other studies have looked at how phage systems can be manipulated to serve other needs, such as nanoparticle templating[5–8], novel nanostructures and scaffolding[9–13], and biomaterial fabrication[14,15], many of which use the phage display technique.

Phage display is the modification of a viral capsid's major coat proteins to display a foreign protein or peptide of interest on the surface of the phage capsid[3]. The gene of the foreign protein is inserted into that of the major coat protein at either the C- or N-terminus depending on the species of phage and the mechanism of viral coat assembly. This forms a fusion protein with the foreign portion displayed on the outside of the protein coat. This technique can be used to attach proteins for a variety of applications, from characterizing protein binding sites[16,17], to screening cloned antibodies[18] and nucleating nanoparticle formation[5,7,19–21]. In the case of phage display with bacteriophage λ, the major coat protein gene is gpV and the location of fusion modifications is the gpV C-terminus[22,23].

1.2 Tail Assembly Mechanisms of Bacteriophage λ

Bacteriophage λ (lambda phage) is a model phage whose capsid's major structural sections can be distinguished as "head" and "tail" portions. The proteins responsible for these sections can be produced separately to form only a head or tail portion of the capsid[22–25]. The genes required for assembly of the tail section are gpZ, gpU, gpV, gpG, gpT, gpH, gpM, gpL, gpK, gpI, and gpJ. The current understanding of lambda phage biology indicates that the gpL, gpK, gpI, and gpJ proteins form an initiator complex, into which a subassembly of the "tape measure" protein gpH, with gpG and gpGT proteins, bind. From there, the major coat protein gpV polymerizes around the gpH subassembly with assistance from gpM, forming the bulk of the tail fiber. The tail fiber length is determined by the length of the gpH protein, and gpV polymerization is stopped by a hexamer of the gpU protein, then the assembly is capped by gpZ[24–27]. While the precise mechanism is unknown, some models assume that the N terminus of gpH inhibits gpU from binding, preventing capping until the gpV polymerization has reached the length of gpH[24].
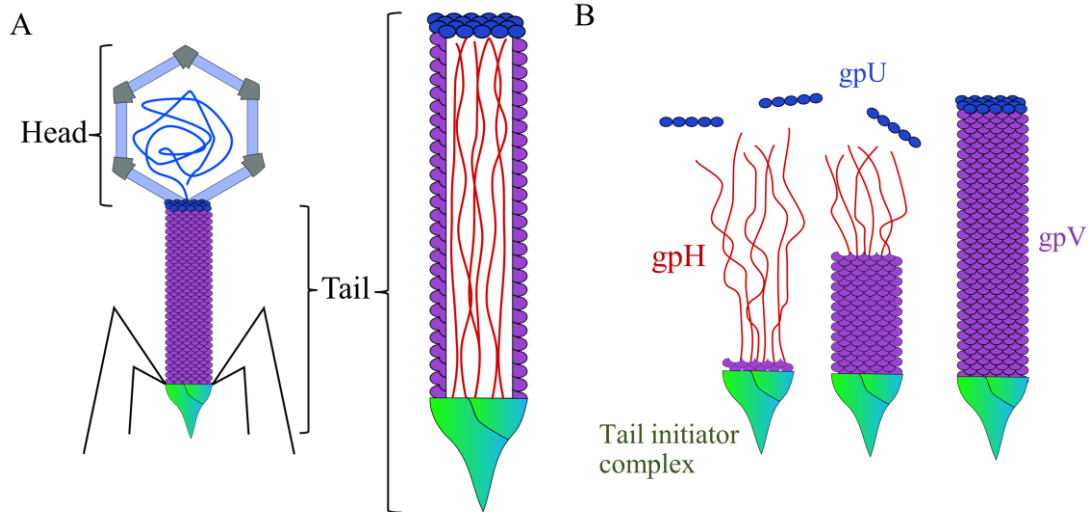
Figure 1.1 Overview of Lambda Phage Capsid Structure. A) General lambda phage structure. The bulk lambda phage capsid can be divided into the "head" and "tail" portions. The tail consists of a conical tip, the tubular main body, and an end capping assembly. B) Simple tail formation schematic. Current models of the tail formation process suggest the following core steps in assembly. First, an initiator complex consisting of at least the gpL, gpK, gpI, and gpJ proteins forms. Concurrently, three gpH form a fibrous subassembly with the help of chaperone proteins (gpG and gpGT) and the C terminus of the gpH bind the initiator complex subassembly. Following subassembly binding, the major coat protein gpV polymerizes onto the gpH starting at the initiator complex, displacing gpG and gpGT in the process. Once it reaches the N terminus of the gpH assembly, polymerization temporarily stops and gpU attaches to the end of the gpV, finalizing the tail's extension.

1.3 Mutation-sensitive RNA Ribodevices for Translational Control

One mechanism of controlling cellular outputs, ribodevices, revolves around translational control by using the secondary structure of RNA to interfere with the translational process. To properly explain the designs and principles described later, it is useful to examine a foundational ribodevice, the toehold switch[28], and some of its derivations[29–31]. In the toehold switch (Fig 1.2A), the region surrounding a gene's RBS is modified to hide the RBS in a hairpin loop, leading to repression of the gene. This repression can then be disabled by introducing a corresponding "trigger" that binds

3

upstream of the hairpin and sufficiently unwinds the hairpin such that the ribosome can access the RBS. Hairpin and trigger sequences are designed *in silico* using algorithmic design software such as NUPACK or MoiRNAiFold[32–37]. Experimental screening of generated designs has resulted in devices capable of shifting gene expression by over 400-fold *in vivo*[28].
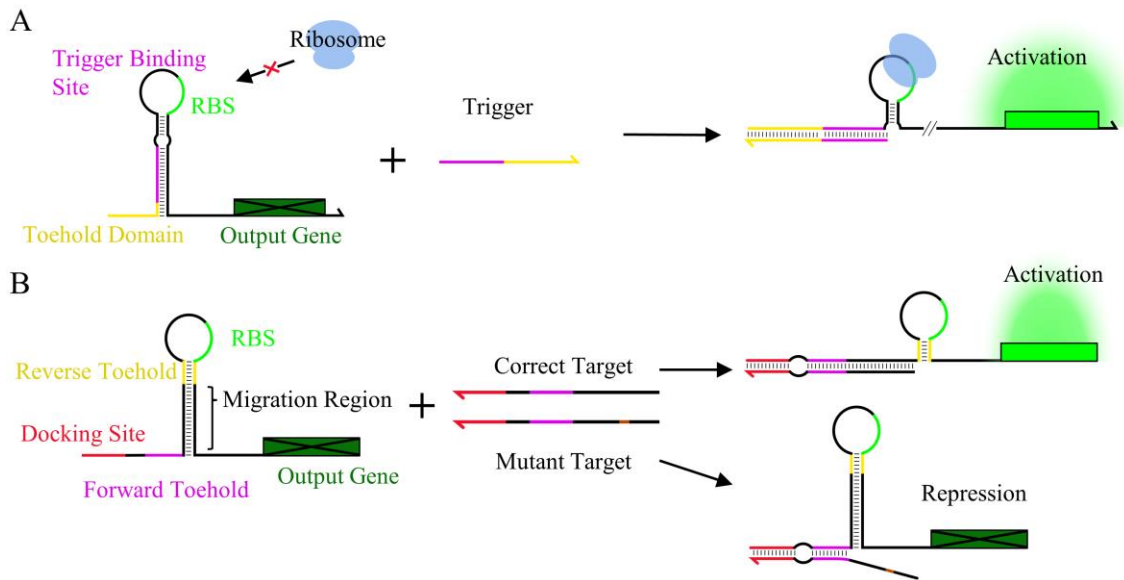
Figure 1.2 Toehold Switch and SNIPR Mechanisms. A) The toehold switch is a ribodevice that accepts other RNA strands as inputs. The RNA hairpin obscures the ribosome binding site from being accessed until it has been unwound by a target trigger RNA. B) With a similar structure as the toehold switch, the free energy of their trigger-bound and unbound hairpins is very similar, such that single point mutations in the trigger sequence will not activate the device.

Numerous logical architectures have been derived from the toehold switch, including multiplexed AND, OR[38], and NOR gates[30], while the core functionality of the device has been applied extensively to diagnostic applications[39–41]. The SNIPR system[42] (Fig 1.2B) is a particularly relevant diagnostic derivative of the toehold switch. Designed to detect single base pair mutations, the single-nucleotide-specific programmable riboregulator (SNIPR) ribodevice has a "docking" site that enables target diagnostic RNA

to partially hybridize. After hybridizing the remainder of the strand can invade the repressing hairpin like a standard toehold switch trigger, outcompeting with the complementary region due to the slight thermodynamic favorability conveyed by the toehold region. However, the toeholds in these designs are selected such that the free energy difference between the ON and OFF states is very small, approximately -1 to -3 kcal/mol. Because of this, any shift in the invading sequence shifts the thermodynamic favorability and prevents the ON state from dominating. This allows over 100-fold signal differentiation between single-nucleotide mismatch mRNA targets[42].

1.4 RNA Aptamers

RNA aptamers are short, single stranded oligomers that bind to specific ligands with high affinity and specificity. Novel aptamers can be created through the process known as Systematic Evolution of Ligands by Exponential Enrichment (SELEX)[43]. In this process, random libraries of DNA with identical flanking sequences are created, then their RNA counterparts are expressed and incubated in a column containing a desired binding ligand. The weak or non-binding oligomers are eluted with high-salt washes while the remaining bound oligomers are washed with water. The binding oligomers are treated with reverse transcriptase to convert them into DNA, then amplified with PCR. This process is repeated to purify and enrich the oligo pool until only tightly binding sequences remain, creating aptamers. Since the development of SELEX, aptamers have been generated for a wide variety of targets with applications in biological research[44–46], small molecule sensing[47–49], and medical diagnostics[50–52]. Some RNA aptamers undergo significant conformational change upon target ligand binding and this shift can be used to create various types of biomolecular sensors. For example, target binding can drive a

conformational change that interferes with RNA translation as with the Gallivan family

of theophylline aptamer-based devices[49].

1.5 Light-Sensitive RNA Binding Protein PAL

The PAL protein identified by Weber et al.[53] is an RNA-binding protein dimer

that binds specific RNA stem loop motifs under exposure to blue light and that can

modulate gene activity by binding mRNA. PAL contains three functional subunits, a

Light-Oxygen-Voltage (LOV) photoreceptor domain, an AmiR and NasR transcription

antitermination regulator (ANTAR) RNA binding domain, a Per-ARNT-Sim (PAS), and

a linker domain that connects the LOV and and ANTAR regions. Functionally, the LOV

domain blocks RNA access to the ANTAR domain until exposed to blue light, which

causes conformational changes that reverse the autoinhibitory effect of the LOV domain,

allowing RNA binding. The binding of the PAL protein to RNA is sufficient to interfere

with ribosomal access when placed immediately upstream of an RBS. Weber et al

reported an ON:OFF ratio of around 10-fold in *E. coli* when comparing exposure with

and without blue-light. Further experimentation in HeLa cells showed ON:OFF ratios of

around 4 to 6.5-fold when the RNA binding motif is placed near the 5' mRNA terminus.

1.6 References

1.   Casjens, S. R. & Hendrix, R. W. Bacteriophage lambda: Early pioneer and still relevant. *Virology* vols. 479–480 310–330 (2015).

2.   Keen, E. C. A century of phage research: Bacteriophages and the shaping of modern biology. *BioEssays* **37**, 6–9 (2015).

3.   Smith, G. P. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science (1979)* **228**, 1315–1317 (1985).

4.   Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* vol. 482 331–338 (2012).

5.   Atanasova, P. *et al.* Virus-templated synthesis of ZnO nanostructures and formation of field-effect transistors. *Advanced Materials* **23**, 4918–4922 (2011).

6.   Gerasopoulos, K., McCarthy, M., Royston, E., Culver, J. N. & Ghodssi, R. Microbatteries With Tobaco Mosaic Virus Templated Electrodes. in *2008 IEEE 21st International Conference on Micro Electro Mechanical Systems*.

7.   Nam, K. T. *et al.* Virus-Enabled Synthesis and Assembly of Nanowires for Lithium Ion Battery Electrodes. *Science* **312**, 885-888 (2008).

8.   Nam, K. T., Reelle, B. R., Lee, S. W. & Belcher, A. M. Genetically Driven Assembly of Nanorings Based on the M13 Virus. *Nano Letters* **4**, 23–27 (2004).

9.   Luo, Q., Hou, C., Bai, Y., Wang, R. & Liu, J. Protein Assembly: Versatile Approaches to Construct Highly Ordered Nanostructures. *Chemical Reviews* **116** 13571–13632 (2016).

10.   Sugimoto, K., Kanamaru, S., Iwasaki, K., Arisaka, F. & Yamashita, I. Construction of a Ball-and-Spike Protein Supramolecule. *Angewandte Chemie International Edition* **45**, 2725–2728 (2006).

11.   Huang, Y. *et al.* Programmable assembly of nanoarchitectures using genetically engineered viruses. *Nano Letters* **5**, 1429–1434 (2005).

12.   Soto, C. M. & Ratna, B. R. Virus hybrids as nanomaterials for biotechnology. *Current Opinion in Biotechnology* vol. 21 426–438 (2010).

13.   Holder, P. G. *et al.* Dramatic thermal stability of virus-polymer conjugates in hydrophobic solvents. *Langmuir* **26**, 17383–17388 (2010).

14.   Jackson, K., Peivandi, A., Fogal, M., Tian, L. & Hosseinidoust, Z. Filamentous Phages as Building Blocks for Bioactive Hydrogels. *ACS Applied Bio Materials* **4**, 2262–2273 (2021).

15.   Cao, B. *et al.* Bacteriophage-based biomaterials for tissue regeneration. *Adv Drug Deliv Rev* **145**, 73 (2019).

16.   Hoffmann, S. *et al.* Competitively selected protein ligands pay their increase in specificity by a decrease in affinity. *Molecular BioSystems* **6**, 126–133 (2009).

17.   Zhang, J. *et al.* Phage-Derived Fully Human Antibody scFv Fragment Directed Against Human Vascular Endothelial Growth Factor Receptor 2 Blocked Its Interaction with VEGF. (2012).

18.   Wiebe, J., Kramer, K. & Skerra, A. A System for Repertoire Cloning and Phage Display of Murine and Leporid Antibody Fragments Drug Development of PASylated Interleukin-1 Receptor Antagonist View project Enzyme engineering View project. *Article in Journal of AOAC International* (2010.

19.   Pires, D. P., Cleto, S., Sillankorva, S., Azeredo, J. & Lu, T. K. Genetically Engineered Phages: a Review of Advances over the Last Decade. *Microbiology and Molecular Biology Reviews* **80**, 523–543 (2016).

20.   Briggs, B. D. & Knecht, M. R. Nanotechnology meets biology: Peptide-based methods for the fabrication of functional materials. *Journal of Physical Chemistry Letters* vol. 3 405–418 (2012).

21.   Love, A. J. *et al.* A genetically modified tobacco mosaic virus that can produce gold nanoparticles from a metal salt precursor. *Frontiers in Plant Science* **6**, (2015).

22.   Rajagopala, S. v., Casjens, S. & Uetz, P. The protein interaction map of bacteriophage lambda. *BMC Microbiology* **11**, (2011).

23.   Katsura, I. & Hendrixt, R. W. Length Determination in Bacteriophage Lambda Tails. *Cell* vol. 39 (1984).

24.   Katsura, I. Mechanism of Length Determination In Bacteriophage Lambda Tails. *Adv. Biophys* vol. 26 (1990).

25.   Xu, J., Hendrix, R. W. & Duda, R. L. Chaperone-protein interactions that mediate assembly of the bacteriophage lambda tail to the correct length. *Journal of Molecular Biology* **426**, 1004–1018 (2014).

26.   Xu, J., Hendrix, R. W. & Duda, R. L. A balanced ratio of proteins from gene gand frameshift-extended gene GT is required for phage lambda tail assembly. *Journal of Molecular Biology* **425**, 3476–3487 (2013).

27.   Pell, L. G. *et al.* The X-Ray Crystal Structure of the Phage λ Tail Terminator Protein Reveals the Biologically Relevant Hexameric Ring Structure and Demonstrates a Conserved Mechanism of Tail Termination among Diverse Long-Tailed Phages. *Journal of Molecular Biology* **389**, 938–951 (2009).

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

29.   Green, A. A. *et al.* Complex cellular logic computation using ribocomputing devices. *Nature* **548**, 117–121 (2017).

30.   Kim, J. *et al.* De novo-designed translation-repressing riboregulators for multi-input cellular logic. *Nature Chemical Biology* **15**, 1173–1182 (2019).

31.   Hong, F. *et al.* Precise and Programmable Detection of Mutations Using Ultraspecific Riboregulators. *Cell* **180**, 1018-1032.e16 (2020).

32.   Minuesa, G., Alsina, C., Garcia-Martin, J. A., Oliveros, J. C. & Dotu, I. MoiRNAiFold: A novel tool for complex in silico RNA design. *Nucleic Acids Research* **49**, 4934–4943 (2021).

33.   Wolfe, B. R. & Pierce, N. A. Sequence Design for a Test Tube of Interacting Nucleic Acid Strands. *ACS Synthetic Biology* **4**, 1086–1100 (2015).

34.   Wolfe, B. R., Porubsky, N. J., Zadeh, J. N., Dirks, R. M. & Pierce, N. A. Constrained Multistate Sequence Design for Nucleic Acid Reaction Pathway Engineering. *J Am Chem Soc* **139**, 3134–3144 (2017).

35.   Zadeh, J. N., Wolfe, B. R. & Pierce, N. A. Nucleic acid sequence design via efficient ensemble defect optimization. *Journal of Computational Chemistry* **32**, 439–452 (2011).

36.   Dirks, R. M., Lin, M., Winfree, E. & Pierce, N. A. Paradigms for computational nucleic acid design. *Nucleic Acids Research* **32**, 1392–1403 (2004).

37.   Dotu, I. *et al.* Complete RNA inverse folding: computational design of functional hammerhead ribozymes. *Nucleic Acids Research* **42**, 11752–11762 (2014).

38.   Green, A. A. *et al.* Complex cellular logic computation using ribocomputing devices. *Nature* **548**, 117–121 (2017).

39.   Ma, D., Shen, L., Wu, K., Diehnelt, C. W. & Green, A. A. Low-cost detection of norovirus using paper-based cell-free systems and synbody-based viral enrichment. *Synthetic Biology* **3**, (2018).

40.   Pardee, K. *et al.* Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell* **165**, 1255–1266 (2016).

41.   Sadat Mousavi, P. *et al.* A multiplexed, electrochemical interface for gene-circuit-based sensors. *Nature Chemistry* **12**, 48–55 (2020).

42.   Hong, F. *et al.* Precise and Programmable Detection of Mutations Using Ultraspecific Riboregulators. *Cell* **180**, 1018-1032.e16 (2020).

43.   Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature 1990 346:6287* **346**, 818–822 (1990).

44.   Paige, J. S., Nguyen-Duc, T., Song, W. & Jaffrey, S. R. Fluorescence imaging of cellular metabolites with RNA. *Science* **335**, 1194 (2012).

45.   Wu, J. *et al.* Live imaging of mRNA using RNA-stabilized fluorogenic proteins. *Nat Methods* **16**, 862–865 (2019).

46.   Dolgosheina, E. v. *et al.* RNA Mango aptamer-fluorophore: A bright, high-affinity complex for RNA labeling and tracking. *ACS Chemical Biology* **9**, 2412–2420 (2014).

47.   Nakatsuka, N. *et al.* Aptamer–field-effect transistors overcome Debye length limitations for small-molecule sensing. *Science (1979)* **362**, 319–324 (2018).

48.   Baker, B. R. *et al.* An electronic, aptamer-based small-molecule sensor for the rapid, label-free detection of cocaine in adulterated samples and biological fluids. *J Am Chem Soc* **128**, 3138–3139 (2006).

49.   Wrist, A., Sun, W. & Summers, R. M. The Theophylline Aptamer: 25 Years as an Important Tool in Cellular Engineering Research. *ACS Synthetic Biology* vol. 9 682–697 (2020).

50.   Liu, S., Xu, Y., Jiang, X., Tan, H. & Ying, B. Translation of aptamers toward clinical diagnosis and commercialization. *Biosensors and Bioelectronics* **208**, 114168 (2022).

51.   Osborne, S. E., Matsumura, I. & Ellington, A. D. Aptamers as therapeutic and diagnostic reagents: problems and prospects. *Current Opinion in Chemical Biology* **1**, 5–9 (1997).

52.   Marnissi, B., Kamali-Moghaddam, M., Ghram, A. & Hmila, I. Generation of ssDNA aptamers as diagnostic tool for Newcastle avian virus. *PLOS ONE* **15**, e0237253 (2020).

53.   Weber, A. M. *et al.* A blue light receptor that mediates RNA binding and translational regulation. *Nature Chemical Biology* **15**, 1085–1092 (2019).

CHAPTER 2

RE-IMAGINING LEGACY PHAGES FOR VIRAL SELF ASSEMBLY

2.1 Introduction and Literature Review

Viral particles possess useful traits for a variety of self-assembling applications owing to their capacity to be tuned through genetic manipulation for different morphological[1,2], reactive [3,4], and electrochemical properties[5]. Viruses have been applied towards nanoparticle fabrication [5–8], self-assembling materials [9,10], nano-factories [11–15], and structural biomaterials [9,10], but is has often been the same well characterized phages used in each study. We ask if other phages could be modified into efficient and flexible platforms to serve novel bioengineering applications. To that end this study explored the well characterized bacteriophage λ's (lambda phage) capacity as a modular system for fabricating nanoparticles and serving as a scaffold for chemical and enzymatic reactions. Lambda phage's capsid structure consists of a "head" portion attached to a non-contractile "tail", which may or may not have fibers (Fig 1.1A). The proteins responsible for the capsid tail portion can be expressed from a plasmid vector to create viral particles with only the tail portion of lambda phage capsid rather than the full structure. The current model of lambda phage tail assembly suggests that the "tape measure" protein gpH determines the length of the tail by providing a subassembly complex upon which chaperone proteins guide the main coat protein gpV to assemble, simultaneously preventing "capping" proteins from binding until the tail has reached its full length[2,16]. The major coat protein, gpV, can be modified via C-terminal protein fusions to display foreign proteins along the length of the viral tail particle[17,18]. Such fusions can be used as binding or recognition sites for other proteins.
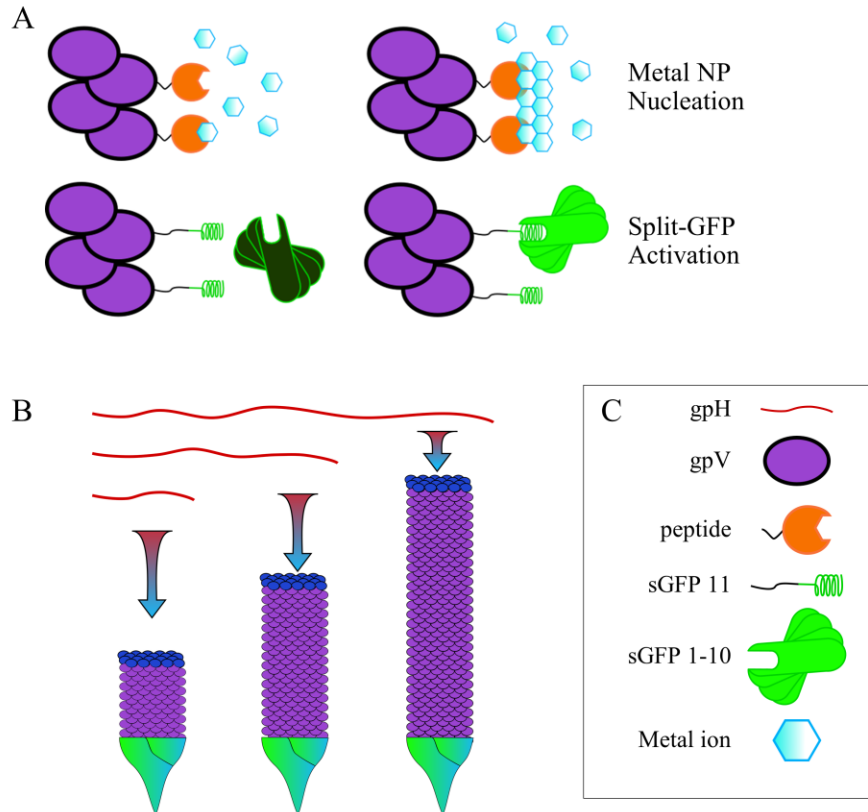
Figure 2.1 Mechanisms of Lambda Phage Modification. A) Modifications of the major coat protein gpV. Through the phage display technique, functional groups such as metal nucleating peptides can be coated along the tail surface. B) Tape Measure protein impact on lambda phage tail length. Broadly speaking, the length of the tape measure protein gpH directly impacts lambda phage tail length. As past studies have demonstrated that truncating gpH shortens the phage tail, it is also possible that extending gpH artificially will elongate the tail. C) Reference key.

To test how we might optimize the lambda phage for different purposes, we tested three primary modification pathways to produce viral tail particles. First, we expressed the split-fluorescent protein peptide tags sfGFP11 and sfCherry11 along the tail surface. These split protein units fluoresce when bound [19–21], allowing us to test for functionality in co-localizing other proteins as one might expect of an enzymatic scaffolding application. Second, we attempted to use the tails as templates for nanoparticle formation with metal nucleating peptides, testing how lambda phage might be modified towards nanofabrication or bioremediation applications. Finally, we tested control of the phage

12

tail length by editing the gpH sequence, either artificially reducing or extending the length. In this way we created versions of the lambda phage's tail geared towards different purposes such as nanoparticle templating or scaffolding within living cells.

2.2 Results

2.2.1 Viral coat Knockout and Co-expression

The portion of lambda phage genome responsible for tail-protein production was placed on a plasmid (pλ) under Isopropyl ß-D-1-thiogalactopyranoside (IPTG) inducible expression. The large size of the lambda phage tail plasmid, over 15 kilobases, made sequence modification difficult for Gibson assembly or Overlap-Extension PCR mutagenesis methods. To increase testing throughput for this and later experiments, a gpV knockout version of the lambda tail plasmid (pλΔgpV) was created as well as a plasmid expressing only gpV (pλ-gpVonly). The gpV-only plasmid was assembled from a pCDF backbone with gpV under IPTG-inducible expression. The pλΔgpV and pλ-gpVonly plasmids were co-transformed into DE3 cells and tails were extracted using PEG precipitation, then imaged with TEM and AFM. Imaging confirmed the formation of tails in both cases (Appendix A, Fig S2.1), although a longer average tail length was observed in the case on the double-transformed knockouts. This was likely caused by an excess of gpV relative to the gpU protein, which can lead to the formation of a polytail, an aberrant tail structure of indeterminant length[2,22]. Due to the formation of these structures, tail formation via pλΔgpV and pλ-gpVonly co-expression was not used for later investigations into length control.

13

2.2.2 Testing Scaffolding Function

We sought to test the use of scaffolding with a simple and easily visualized test case. To that end we tested the split-GFP system's ability to function as a fusion protein with the major coat protein gpV[20] (Fig 2.1B). The outward-facing N-terminus of gpV was fused to 3-, 5- and 7-fold repeats of the GFP-11 and sfCherry-11 peptides with an amino acid linker (GG) between the gpV sequence and the peptides, and spacer amino acid sequences (GGSGG) between each peptide repeat. Functionality of the system was first tested by co-expressing the sfCherry-(1-10) split proteins alongside their respective gpV fusion proteins and the knockout pλΔgpV plasmid in *E. coli* DE3 cells and measuring fluorescence in 96 well plates (Fig 2.2A). This confirmed that neither the gpV protein nor the tail-formation process interfered with split-protein binding. However, we did not observe an increase in fluorescence with greater peptide repeats as previous studies have[19].

Having confirmed that the gpV protein did not prevent split-protein binding and fluorescence, the full tail extraction protocol was performed on DE3 cells expressing the knockout pλΔgpV plasmid and the gpV-fusion proteins. Visual observation of extract showed obvious shifts in color intensity that corresponded with prior fluorescence measurements (Fig2.2C), while atomic force microscopy (AFM) confirmed that the gpV fusions formed tail particles (Fig. 2.2B). Between the *in vivo* studies and AFM visualization of tail formation, we demonstrated how lambda phage tail particles might be used as a binding zone for proteins to assemble.
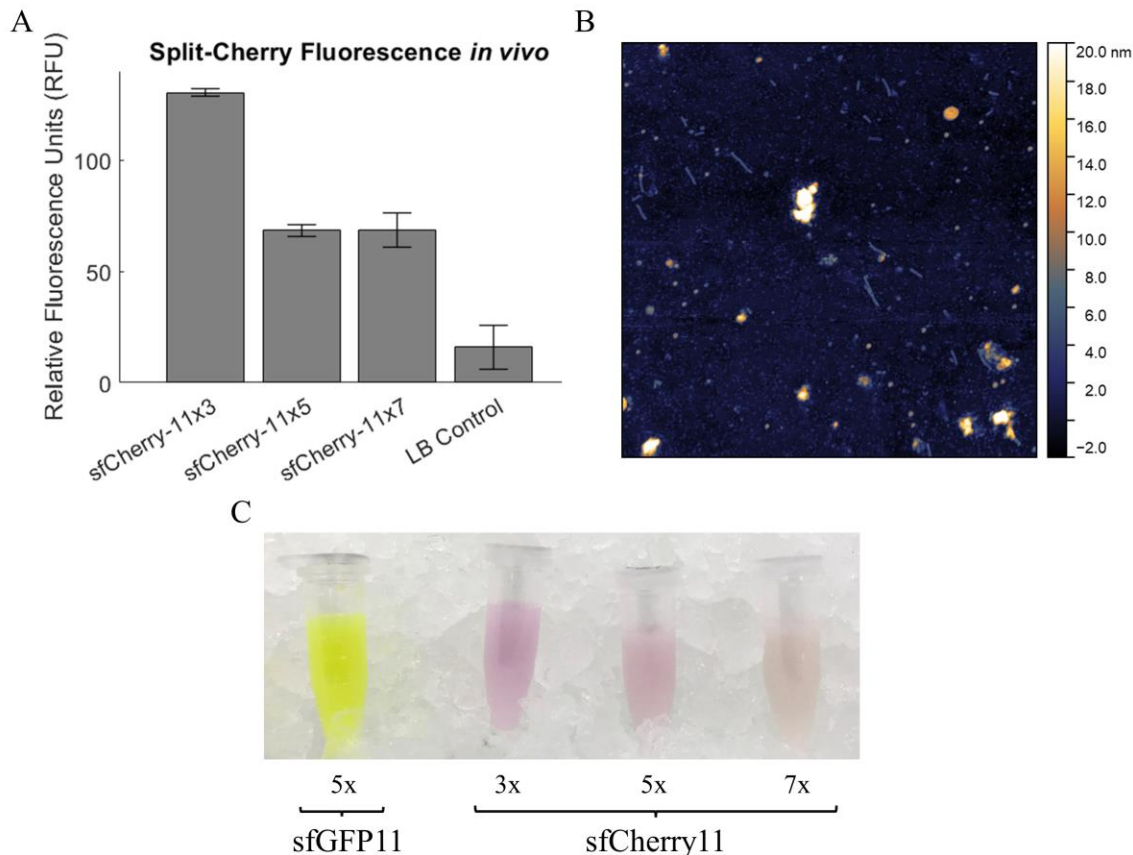
Figure 2.2 Split Protein Function in Lambda Phage Tails. A) Fluorescence of the sfCherry repeat peptides on lambda phage tails. The sfCherry peptide was repeated three, five, or seven times as indicated. B) Tail formation with sfGFP11x5 peptide fusion. Image taken from AFM scan of purified tail product with the sfGFP11x5 peptide fusion. C) Visual observation of split-protein binding. Image taken from midpoint of tail purification process. Samples are lysate from cells co-expressing the pλΔgpV plasmid and a pλΔ-gpVonly plasmid with a split protein fusion peptide. Upper labels indicate the number of peptide repeats in the fusion, while lower labels indicate the split protein peptide associated with the sample. Samples retained visual indicators of binding during tail precipitation.

2.2.3 Lambda Phage as a Nanoparticle Template Platform

A variety of modified pλ-gpVonly plasmids were generated in which known metal-nucleating peptides or proteins (Table 2.1) were fused to the N-terminus of gpV along with a flexible amino acid linker (GGGSGGGS). Tails were produced in bacteria through co-expression of pλΔgpV and modified pλ-gpVonly, then purified using PEG extraction and placed under reducing conditions with different metal ions. Some gpV

modifications inhibited proper tail formation, in which case tails were produced vial co-expression of the standard pλ plasmid rather than pλΔgpV. The goal in these cases was to incorporate the peptide additions at a concentration low enough to avoid tail mis-assembly. After purification, tails were visualized under AFM to confirm assembly and the presence of modified gpV protein in tails produced by co-expression with pλ was confirmed by PAGE-SDS gel analysis (Appendix A, Fig S2.2). While the focus for these actions was to develop nanoparticle-forming tails, such results also demonstrate how proteins that exceed the typical capacity for phage display could be incorporated into biomolecular scaffolding.

Table 2.1 Peptides Tested for Display on Lambda Phage Tail. The peptides listed here are those we attempted to display and form nanoparticles with using the lambda phage tail as a nucleating body and templating agent.

Peptide	AA sequence	Known metal nucleation	Integrated into tail?	AFM-indicated ion nucleation
Ge34[23]	TGHQSPGAYAAH	Ge	Yes	Co, Cu, Fe, Ni
Tetraglutamate[5]	EEEE	Co, Fe	Yes	Co, Fe
dTiRRK[24]	RKKRKKRKKRKKGGGW	Ti	Yes	Co, Cu, Cd, Fe, Mg, Zn
Ge8[23]	SLKMPHWPHLLP	Ge, Ag	Yes	-
KDTK[25]	KTEYVDERSKSLTVDLTK	Zn	Yes	-
EC7G[26]	ECECECECECECECG	Cd	Yes	-
MT_AGABI[27]	GDCGCSGASSCTCASGQCTCSG CGK	Cu, Cd	Yes	-
Ferritin[28]	*	Fe	Yes†	-
OPN_S1[29]	GLRSKSKKFRRPDIQYPDATDED ITSHM	Ca	Yes†	-
MT3_Mouse[27]	*	Cu, Cd	Yes†	-
MT4B_ARATH[27]	*	Cu, Cd	Yes†	-
HHTC[25]	HNLGMNHDLQGERPYVTEGC	Cu	N	-
DC7G[26]	DCDCDCDCDCDCDCG	Cd	N	-
PCS1_ARATH[27]	*	Cu, Cd	N	-
Gold Binding Motif[5]	LKAHLPPSRLPS	Au	N	-

* Sequence too long for display constraints, see Appendix A, Table S2.1 for full sequence
† Tail integration possible with co-expression of unmodified gpV

After confirming tail formation via AFM, the purified products were combined with metal ions in the presence of a reducing agent to initiate nucleation onto the tail surface. Initial tests showed the formation of visible precipitate with tetraglutamate-

displaying tails when reacted with $CoCl_2$. AFM imaging of the precipitate suggested these to be networked crystals of virus tails (Appendix A, Fig S2.3), requiring us to dilute the purified tails until we could visualize individual tails after metal ion reduction. Once individual tails could be discerned in AFM post-reduction, we compared the impact of metal salt exposure on standard tails versus peptide modified tails under reducing conditions (Fig 2.3A). Profilometry of the observed tail particles showed increasing tail diameter, indicating aggregation of nucleated metal ions along the tail surface. The diameter of unmodified tails averaged 5.5 nm, with that number shrinking slightly after when exposure to metal ions under reducing conditions. By comparison, peptide-modified tails increased from an average of 6.7 nm to 10 nm, with a significant portion reaching up to 14 nm (Fig 2.3B).
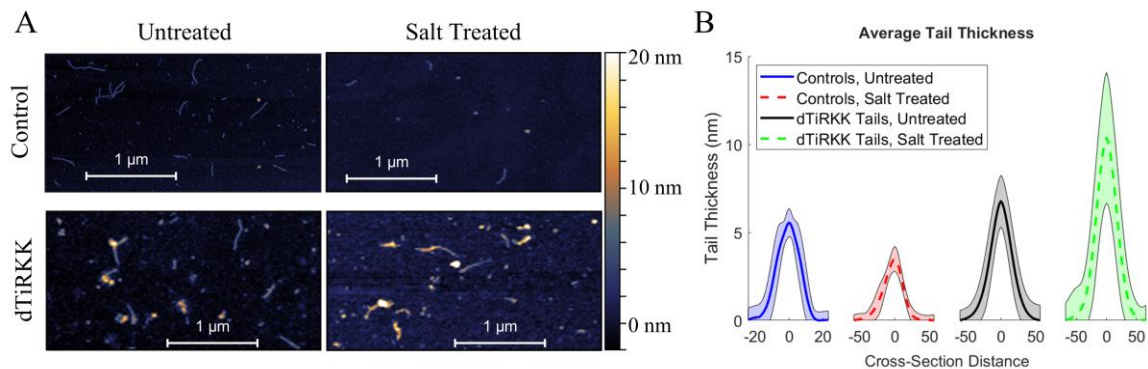
Figure 2.3 Nanoparticle Nucleation on Tails. A) AFM images of Metal Salt-Treated Tails. Upper images are unmodified lambda phage tail particle controls while bottom images are tails modified to display the dTiRRK peptide through major coat protein fusion. Left side images are of tails that have not been exposed to magnesium metal salts under reducing conditions, right are tails after such exposure. B) Tail Heights: Representative Cross Sections. Average heights of the tails as determined by AFM profilometry. Tail profiles were averaged and plotted with standard deviation represented by the shaded area in corresponding colors. Sample sizes are as follows: Untreated Controls = 24, Treated Controls = 10, Untreated dTiRRK = 71, and treated dTiRRK = 64.

After tail-profilometry analysis suggested that the peptide-fusion tails were thickening in due to metal ion reduction, we used SEM-based Energy-Dispersive X-ray Spectroscopy (EDS) analysis to confirm the case of cobalt crystallization in the case of the gpV-dTiRRK peptide fusion (Appendix A, Fig S2.4). Based on these results, we used tail profile measurement with AFM to screen for nanoparticle formation. Peptide-coated tails were screened against iron, cobalt, cadmium, nickel, zinc, magnesium, and copper ions. While our literature review often indicated a specific target or affinity ion for a given peptide, our screening showed that peptides often served as nucleating agents for multiple ions (Table 2.1). Altogether, this provided a compelling case for the use of lambda phage tail particles as templates for nanoparticle formation.

2.2.4 Viral Tail Length Control

Based on the model of Katsura and Hendrix[2], changing the number of amino acids in the gpH protein has a proportional impact on the overall length of the tail assembly that corresponds with the function $L = 0.15 \cdot n \cdot \cos(\theta)$, in which L is the length in nanometers (nm), n is the number of amino acids removed, and θ is the is the angle between the axis of the pH protein's a-helix structure and that of the tail tube[2]. With this model as a guide, we attempted to shorten and extend the tail assembly length by changing the number of amino acids in the primarily alpha-helical mid region of gpH. To shorten the tail, we removed portions of the gpH protein sequence from the viral tail plasmid corresponding to roughly 50 nm and 20 nm of overall length. While studies into unstructured tail elongation and polymerization have been conducted[2,22], we were unable to find a study that attempted to artificially extend a lambda phage tail in tightly controlled manner.
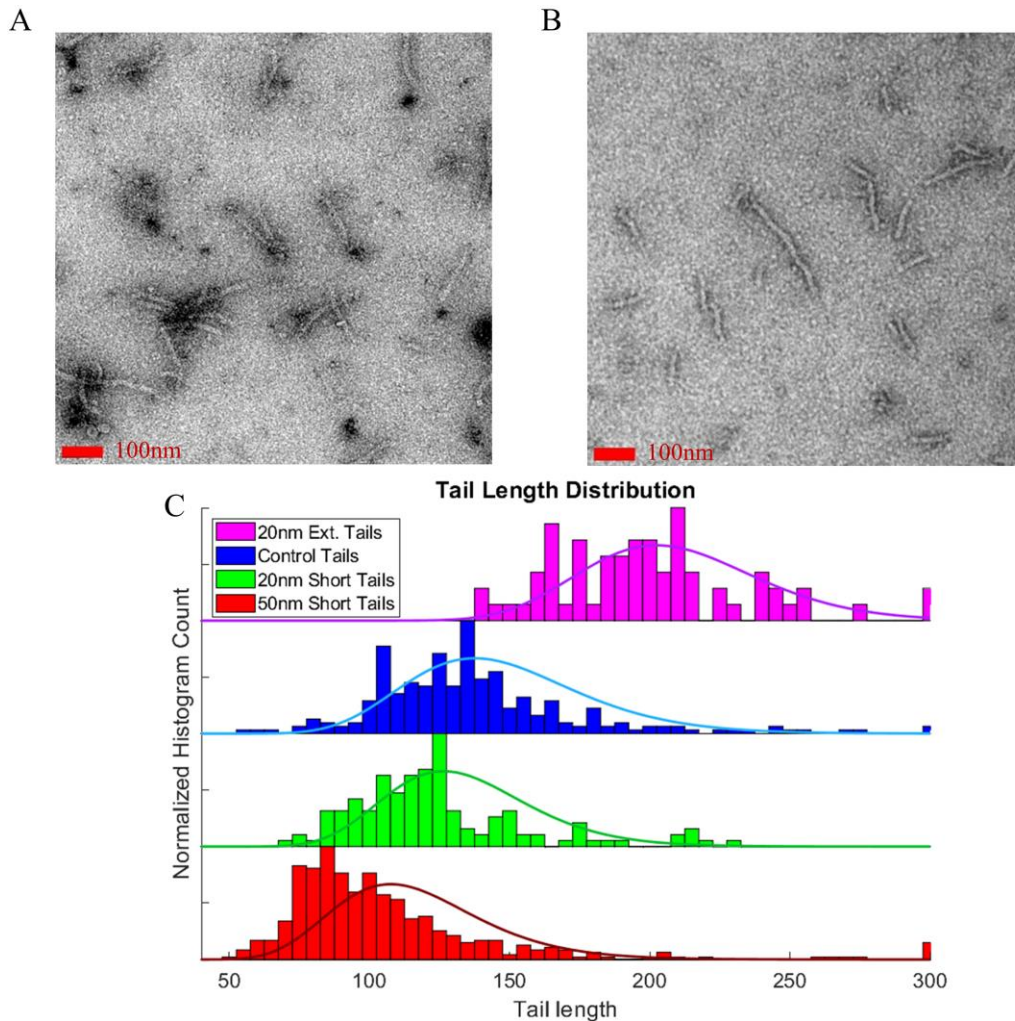
Figure 2.4 Tail Length Modification. A) TEM images of unmodified tails. B) Tails shortened by 50nm. By truncating the "tape measure" protein gpH, the average tail length of the lambda phage can be tuned. C) Histogram of lambda tail modification effects. Histograms display normalized sample counts of measured tail lengths with lognormal distribution overlays. Distribution peaks and sample counts are as follows: 20nm Extended Tails peak = 202, n = 66. Control Tails peak = 138 nm, n = 241. 20nm Short Tails peak = 126 nm, n = 129. 50nm Short Tails peak = 108 nm, n = 411.

We attempted to elongate the lambda phage tail in a controlled, tunable fashion by inserting synthetic repeats of the gpH amino acid sequence along its sequence. While in many cases this appeared to disrupt tail assembly, one functioning insert arose from an 18 nm extension sequence placed after the 150th amino acid. The formation of tails and their resulting length change was initially visualized under TEM, with large scale data

19

collection conducted through AFM profiles. As Figure 2.4C shows, the viral particles were not uniform in size but adjustments to gpH length had a clear impact on the overall length distribution. The elongated tails exhibited a much greater shift in their distribution compared to their truncated and control counterparts. The lower observed tail density for these samples and failure of other elongation inserts suggests that the inserted sequences disrupted gpV polymerization. If the disruption was only to local polymerization, then the large distribution shift may be explained by a survivor bias of sorts. We would only observe the longer tails that managed to successfully polymerize past the 150th amino acid insertion point.

2.3 Discussion and Conclusion

By placing the responsible proteins onto a single bacterial expression plasmid, we avoided the complications that arise in phagemid-based production of viral structures. Such a production method is not universally possible amongst tail-forming phages. Other viruses used in nanostructure templating, such as Tobacco Mosaic Virus (TMV) and M13 phage[5,30], require phagemid co-production as the packaged DNA determines the tail structure's length. The lambda phage's protein-based length determination means that the production system can be simplified from a multi-step infection and harvesting process to a straightforward bacterial-expression. This has also allowed us to more easily pursue different protein modifications and co-expression applications, despite the larger overall size of the lambda phage genome. While other researchers may find phagemid production to function better for their particular research goals, the simplicity of plasmid-based production and modification enabled a quick prototyping pathway that was advantageous for testing a wide variety of modifications. Another significant benefit to the plasmid

20

expression format is that expression of phage particles can be tuned to allow for bacterial growth alongside tail production. This aspect may be of greater interest to bioremediation applications in which bacterial growth is favorable, or to scaffolding applications for fermentation reactions in which viral particles might function as mini organelles for particular reactions.

Although we were able to influence the length distribution of the tails with some degree of accuracy, the viral structures are not mono-disperse as hoped. Past studies in which wild-type lambda tails were produced via plasmid and phagemid production suggested near uniform size[31]. Studies that measured length distribution indicated a 4-8 nm standard distribution of length[2], whereas our wild-type tails had a normal standard deviation of 30 nm. This may be attributed to viral particle purification methods, as prior studies utilized ultracentrifugation along a glycerol gradient after PEG precipitation, while we imaged directly from resuspended precipitate. Alternatively, our production mechanism excludes a significant portion of the viral genome and it is possible that we excluded the assistant regulatory elements that ensure even production of tail assembly proteins elements and that lead to our broader distribution. The requirement of balanced protein production is further supported by observations of tails formed by double transformation of standard tail production plasmid and the gpV production plasmid during nanoparticle formation investigation. These tails often exceeded the standard length by hundreds of nanometers, occasionally forming polytails[22] that extended several micrometers. Based on our observations and literature review, additional purification steps and/or phagemid production of viral particles may resolve length distribution issues. If bacterial expression is still desirable, the incorporation of additional regulatory

elements on the production plasmid may improve results at the cost of a more unwieldly prototyping system.

In our analysis of the metal-nucleation potential for lambda phage particles we focused on the breadth of potential particles and ease of modification and testing rather than in-depth analysis of specific material fabrication, properties, or applications. AFM profilometry indicated numerous metal/mineral species forming on viral tails, and SEM EDS confirmed the presence of expected elemental components. However further analysis of how lambda phage's coat morphology impacts the nucleation process and grain structure may be desirable for researchers seeking to use it as a templating mechanism.

2.4 Methods

2.4.1 Lambda Phage Tail Particle, Plasmid Assembly

The lambda phage genome was provided by the Duda lab. Gibson assembly was used to PCR amplify the region containing the tail assembly proteins and to insert them into a backbone containing the ColA origin of replication and Kanamycin resistance. Tail assembly sequences were inserted into a backbone containing a CloDF13 origin of replication and placed downstream of a T7 promoter with LacO repression.

2.4.2 Viral Tail Particle Production

The tail assembly plasmids were transformed into DE3 cells. Colonies were picked and overnight cultures grown in 2ml of LB with 50 µg/ml Kanamycin. Cultures were then diluted 1:100 into 30 ml of fresh LB with Kanamycin and grown until the absorbance at 600 nm was between 0.40-0. Then IPTG was added to a concentration of 1

µM and the culture left to grow at room temperature (23 °C) for 18 hours. Cultures were then placed on ice and centrifuged at 4200 RPM for 10 minutes at 4 °C. Cell pellets were resuspended in 1 ml of Tris-EDTA buffer (50 mM Tris-HCl, 5 mM EDTA, pH 8.0) and lysozyme added to a concentration of 1 mg/ml. Cultures were incubated on ice for one hour then moved to a -80 °C freezer for one hour. Cultures were repeatedly thawed on ice, then moved back to -80 °C conditions for two cycles, after which they were spun down at 13000 RPM for 30 minutes at 4 °C. The supernatant was drawn off and added to an equal volume of Polyethylene-glycol (PEG) precipitation buffer (10% PEG-8000, 1M NaCl), gently mixed, and stored at 4 °C for 16-24 hours, during which viral precipitates formed. After precipitation, the tubes of supernatant were spun down at 13000 RPM for 30 minutes at 4 °C and the supernatant discarded. The remaining pellet of viral precipitate was then resuspended in de-ionized H_2O.

2.4.3 Length Modification with gpH

The lambda protein pH is theorized to control bulk tail-region length by preventing the capping of the tail during its self-assembly[2,16,22]. As the protein is broadly alpha-helical in structure, reductions in protein length lead to approximate tail length reductions following the function $L = 0.15 \cdot n \cdot \cos(\theta)$, in which L is the length in nanometers (nm), n is the number of amino acids removed, and θ is the is the angle between the axis of the pH protein's a-helix structure and that of the tail tube, estimated to be between 0.88 – 0.95 radians[2]. Based on this formula, amino acid truncations corresponding to 10, 20, 50, and 100 nm were determined. These sections were excised from the middle region of the tail region that prior studies have shown to allow for

excision via whole-plasmid PCR amplification followed by assembly with the Kinase-Ligase-DpnI mix (New England Biolabs, M0554).

Tail extension was accomplished by inserting synthetic copies of its sections into various locations along the gpH coding region. These copies were created by randomly selecting portions of the gpH protein and changing the codon distribution of the sequence to reduce the sequence similarity to the original copy, ideally to less than 50%. Insertion locations were chosen at random within the same alpha-helical region known to permit excision. Synthetic sequences were ordered from IDT as ultramers and assembly PCR used to join them. Gibson assembly was used to inserting the synthetic portions of gpH.

2.4.4 Viral Coat Modification

Following a literature review of peptides that exhibit metal nucleation and/or precipitation properties, we identified several peptides (Table 2.1) as potential nanoparticle synthesis tools. Following the principles of phage display systems, we inserted the peptide sequences at the C-terminus of the lambda phage tail's major coat protein (pV) separated by a linker peptide (GGGSGGGS). Modification of the primary tail assembly plasmid proved difficult owing to its size, so a second plasmid was fabricated containing only the pV protein and expressed alongside a tail assembly plasmid with a pV knockout. In some cases this appeared to prevent proper viral tail formation, possibly due to either the large size of some inserts or because the charged peptide structures interfering with tail assembly. In these cases, a non-knockout version of the viral tail assembly plasmid was used. Tails produced in this manner were tested for peptide inclusion by SDS-PAGE gel analysis (Appendix A, Fig S2.2), and all samples

24

were screened for tail formation via imaging with Atomic Force Microscopy (AFM). The same steps were used for the insertion of the split fluorescent peptides sfCherry11 and sfGFP11.

2.4.5 Observation of Tail Modifications with AFM

Modified and unmodified tails were viewed under AFM and TEM. AFM samples were deposited on freshly cleaved mica sample plate. Samples were deposited directly onto plate after diluting 1:20 in TAE-Mg buffer (TAE-Mg buffer: 40 mM Tris, 20 mM acetic acid, 2.5 mM EDTA, 12.5 mM $MgCl_2$, pH 7.5), incubated for 5 minutes, then dried with pressurized nitrogen before air imaging in contact mode.

2.4.6 Observation of Tails with TEM

For TEM imaging, samples were placed on glow discharge-treated copper TEM grids and negatively stained using uranyl formate before imaging on a Philips CM200-FEG high resolution TEM microscope. Uranyl formate stain was prepared by adding 37 mg of uranyl formate to a glass vial along with 5 mL of boiling water to it. The solution was stirred with a magnetic stirrer for 5 minutes in the dark. After 5 minutes, 50 µL of 2M NaOH solution was added and stirred for an additional 5 minutes in the dark. Then the solution was passed through a 0.22 micron syringe filter into opaque Eppendorf tubes. To stain the samples, 2uL of purified tail samples were placed on freshly negatively glow discharged grids. Samples adsorbed for 30 seconds on the grid then excess was removed with filter paper. Grids were then touched to 10 µl drops of water and the water immediately blotted with filter paper twice. After these washing steps the grid was touched to a 10 µl uranyl formate droplet and the stain was quickly blotted with filter

paper twice. Grids were then placed face up on clean filter paper, covered with a petri

dish and allowed to dry for 5 minutes before imagining.

2.4.7 Nanoparticle Formation and Screening

Following the protocol of Nam et al.[5], successfully modified tails were placed in

solution with metal ions of varying concentrations corresponding to their affinity in the

literature and incubated on ice for 30 minutes. A reducing agent (1 mM NaBH$_4$) was then

introduced and the samples allowed to incubate on ice for an additional 30 minutes.

Unmodified tails and Phosphate Buffered Saline (PBS) solution were used as negative

controls. After reduction, samples were imaged using contact-mode AFM and height

profiles of the tails were analyzed for signs of increasing thickness. Focusing on the

tetraglutamate peptide's reaction with cobalt, we conducted elemental analysis via

Scanning Electron Microscopy – Energy-Dispersive X-ray Spectroscopy (SEM-EDS)

analysis, the results of which showed cobalt co-localization with the viral tail aggregates,

confirming through both morphological and elemental analysis that the tails are forming

nanoparticles along their surface.

2.4.8 Split-Fluorescent System, Plate Reader Measurements

Split-fluorescent protein binding was tested by co-expressing the sfGFP-(1-10) or

sfCherry-(1-10) split proteins alongside their respective gpV fusion proteins in a 96-well

plate. *E. coli* DE3 cells were double transformed with pλ-gpVonly-sfCherry11 or pλ-

gpVonly-sfGFP11 fusion proteins and the knockout pλΔgpV plasmid and plated with

dual-antibiotic selection. Colonies were cultured overnight in Luria Broth with

kanamycin (50 mg/mL) and spectinomycin (50 mg/mL), then diluted to an absorbance of

0.2 OD_{600} and incubated at 180 RPM at 37 °C for two hours before plating into 96-well

plates and inducing with 1 mM IPTG. The plate was then placed in a Biotek Synergy

Neo2 Hybrid Multi-Mode Reader and held at 37 °C while fluorescence measurements

were taken. Fluorescence readings were taken with excitation/emission pairs of 480/520

nm when measuring sfGFP binding, and 580/610 nm when measuring sfCherry binding.

2.5 References

1. Iannolo, G., Minenkova, O., Petruzzelli, R. & Cesareni, G. *Modifying Filamentous Phage Capsid: Limits in the Size of the Major Capsid Protein*. *J. Mol. Biol* vol. 248 (1995).

2. Katsura, I. & Hendrixt, R. W. Length Determination in Bacteriophage Lambda Tails. *Cell* vol. 39 (1984).

3. Hoffmann, S. *et al.* Competitively selected protein ligands pay their increase in specificity by a decrease in affinity. *Molecular BioSystems* **6**, 126–133 (2009).

4. Zhang, J. *et al.* Phage-Derived Fully Human Antibody scFv Fragment Directed Against Human Vascular Endothelial Growth Factor Receptor 2 Blocked Its Interaction with VEGF. (2012) doi:10.1002/btpr.1559.

5. Nam, K. T. *et al.* Virus-Enabled Synthesis and Assembly of Nanowires for Lithium Ion Battery Electrodes. *Science* **312**, 885-888 (2008).

6. Atanasova, P. *et al.* Virus-templated synthesis of ZnO nanostructures and formation of field-effect transistors. *Advanced Materials* **23**, 4918–4922 (2011).

7. Gerasopoulos, K., McCarthy, M., Royston, E., Culver, J. N. & Ghodssi, R. Microbatteries With Tobaco Mosaic Virus Templated Electrodes. in *2008 IEEE 21st International Conference on Micro Electro Mechanical Systems*.

8. Nam, K. T., Reelle, B. R., Lee, S. W. & Belcher, A. M. Genetically Driven Assembly of Nanorings Based on the M13 Virus. *Nano Letters* **4**, 23–27 (2004).

9. Jackson, K., Peivandi, A., Fogal, M., Tian, L. & Hosseinidoust, Z. Filamentous Phages as Building Blocks for Bioactive Hydrogels. *ACS Applied Bio Materials* **4**, 2262–2273 (2021).

10. Cao, B. *et al.* Bacteriophage-based biomaterials for tissue regeneration. *Adv Drug Deliv Rev* **145**, 73 (2019).

11.   Luo, Q., Hou, C., Bai, Y., Wang, R. & Liu, J. Protein Assembly: Versatile Approaches to Construct Highly Ordered Nanostructures. *Chemical Reviews* vol. 116 13571–13632 (2016).

12.   Sugimoto, K., Kanamaru, S., Iwasaki, K., Arisaka, F. & Yamashita, I. Construction of a Ball-and-Spike Protein Supramolecule. *Angewandte Chemie International Edition* **45**, 2725–2728 (2006).

13.   Huang, Y. *et al.* Programmable assembly of nanoarchitectures using genetically engineered viruses. *Nano Letters* **5**, 1429–1434 (2005).

14.   Soto, C. M. & Ratna, B. R. Virus hybrids as nanomaterials for biotechnology. *Current Opinion in Biotechnology* vol. 21 426–438 (2010).

15.   Holder, P. G. *et al.* Dramatic thermal stability of virus-polymer conjugates in hydrophobic solvents. *Langmuir* **26**, 17383–17388 (2010).

16.   Xu, J., Hendrix, R. W. & Duda, R. L. Chaperone-protein interactions that mediate assembly of the bacteriophage lambda tail to the correct length. *Journal of Molecular Biology* **426**, 1004–1018 (2014).

17.   Smith, G. P. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science (1979)* **228**, 1315–1317 (1985).

18.   Keen, E. C. A century of phage research: Bacteriophages and the shaping of modern biology. *BioEssays* **37**, 6–9 (2015).

19.   Kamiyama, D. *et al.* Versatile protein tagging in cells with split fluorescent protein. *Nature Communications* **7**, (2016).

20.   Kakimoto, Y. *et al.* Visualizing multiple inter-organelle contact sites using the organelle-Targeted split-GFP system. *Scientific Reports* **8**, (2018).

21.   Feng, S. *et al.* Improved split fluorescent proteins for endogenous protein labeling. *Nature Communications* **8**, (2017).

22.   Katsura, I. Mechanism of Length Determination In Bacteriophage Lambda Tails. *Adv. Biophys* vol. 26 (1990).

23.   Dickerson, M. B., Sandhage, K. H. & Naik, R. R. Protein- and peptide-directed syntheses of inorganic materials. *Chemical Reviews* **108**, 4935–4978 (2008).

24.   Briggs, B. D. & Knecht, M. R. Nanotechnology meets biology: Peptide-based methods for the fabrication of functional materials. *Journal of Physical Chemistry Letters* vol. 3 405–418 (2012).

25.   Kožíšek, M. *et al.* Molecular design of specific metal-binding peptide sequences from protein fragments: Theory and experiment. *Chemistry - A European Journal* **14**, 7836–7846 (2008).

26.   Satofuka, H., Fukui, T., Takagi, M., Atomi, H. & Imanaka, T. *Metal-binding properties of phytochelatin-related peptides*. *Journal of Inorganic Biochemistry* vol. 86 www.elsevier.com/locate/jinorgbio (2001).

27.   Park, T. J., Lee, S. Y., Heo, N. S. & Seo, T. S. In vivo synthesis of diverse metal nanoparticles by recombinant escherichia coli. *Angewandte Chemie - International Edition* **49**, 7019–7024 (2010).

28.   Huang, J. *et al.* Bio-inspired synthesis of metal nanomaterials and applications. *Chemical Society Reviews* **44**, 6330–6374 (2015).

29.   Lee, J. Y. *et al.* Assembly of collagen-binding peptide with collagen as a bioactive scaffold for osteogenesis in vitro and in vivo. *Biomaterials* **28**, 4257–4267 (2007).

30.   Love, A. J. *et al.* A genetically modified tobacco mosaic virus that can produce gold nanoparticles from a metal salt precursor. *Frontiers in Plant Science* **6**, (2015).

31.   Xu, J., Hendrix, R. W. & Duda, R. L. A balanced ratio of proteins from gene g and frameshift-extended gene GT is required for phage lambda tail assembly. *Journal of Molecular Biology* **425**, 3476–3487 (2013).

CHAPTER 3

INTEGRATION OF CHEMICAL AND OPTICAL SENSING INTO TOEHOLD

SWITCH-LIKE RIBODEVICES

3.1 Abstract

As mRNA-based methods of controlling genetic output and logical computation continue to develop, one aspect to be enhanced is the range and complexity of input sensing. Current ribodevices of the toehold switch variety benefit from orthogonality and programmability, but only directly detect other nucleic acids as inputs. Current sensing mechanisms would require the use of proteins tied to transcription factors or other non-RNA sensing mechanisms to convert inputs such as small molecules, proteins, or light into an RNA signal that toehold switches can interpret. As the benefits of toehold switch devices include their lower genetic footprint relative to protein-based sensors and stability for point-of-care use when freeze-dried as DNA, a switch that directly detects non-RNA inputs would greatly increase the utility of toehold switch-like ribodevices. We first sought to create such devices by integrating light-sensitive RNA binding proteins (RBPs) into toehold switch design, but this system was not as robust as necessary for continued use. However, the design insights gained from the attempt led us to apply these principles towards aptamer-based devices, designing switchable ribodevices with interchangeable aptamer binding sites for activation or deactivation of the output gene upon target binding in a manner analogous to the toehold-switch derived SNIPR system. Here we describe the testing and optimization of these devices, dubbed SNIPR Aptamer Seesaw (SAS) and Peripheral Action Seesaw (PAS) switches, for use with a S-adenosyl-

methionine (SAM) aptamer and the DFHBI-1T-binding Broccoli aptamer, as well as the role of our light-sensitive RBP integration investigation that led to their creation.

3.2 Introduction and Literature Review

Many classes of ribonucleic acid-based regulatory mechanisms and designs have been either created or discovered in bioengineering, generally referred to here as ribodevices. Two common categories of ribodevice designs are aptamer-based chemical sensors and RNA-sensing logical ribodevices, with more recent ribodevices attempting to blend logical control with chemical sensing. For example, some early synthetic ribodevices were the many theophylline chemical sensors[1]. These were created by inserting the SELEX-derived theophylline aptamer upstream of output genes such that the aptamer binds and hides the ribosome binding site (RBS), or by combining the aptamer with the self-cleaving hammerhead ribozyme to disrupt translation[1,2]. While highly effective and capable of clear switch-like behavior, genetic logic circuits using such devices still requires multiple protein inputs, increasing the number of moving parts, genetic data, and energy required to operate the circuit[3]. By comparison, RNA-sensing ribodevices such as the toehold switch devices and their derivations use only RNA oligomers to create logical circuits. They require less resources and take up less genetic space, important considerations for point of care diagnostics[4–7] use and many methods of inserting logical circuits into the genome of living systems[8,9].

Toehold switch devices typically obscure the RBS in a hairpin loop, preventing translation until the proper RNA trigger sequence binds and unveils the RBS[10]. These devices have few constraints besides the RBS and start codon sites, enabling a wide

variety of orthogonally operating devices to be created and arranged for multiplexed logical operations[11], as well as reversed to repress translation upon binding of trigger RNA[12]. However, toehold switch devices only accept RNAs as sensor inputs. A ribodevice that enables translational control through arbitrary sensing without the need for additional protein interaction would advance the field by combining the vast range of potential chemical inputs enabled by SELEX-derived aptamers with the genetic compactness and orthogonality of toehold switch-like ribodevices. To create such devices, we first attempted the integration of light-responsive control into toehold switch devices, as there is a need for compact multi-logic interpretation in complex environments like human cells[9], as well as more-easily programmed logic systems in light-based fabrication techniques used for biomaterial fabrication[13]. Using the PAL protein described by Weber et al.[14], a RBP that binds specific RNA motifs upon exposure to 440 nm light, we developed some methods for integrating the light-sensing capabilities of PAL into toehold switch systems. The PAL system was not sufficiently robust for us to continue research, as described later, but it led to us developing a new type of device based on the toehold switch-derived ribodevices known as the single-nucleotide-specific programmable riboregulator (SNIPR)[7].
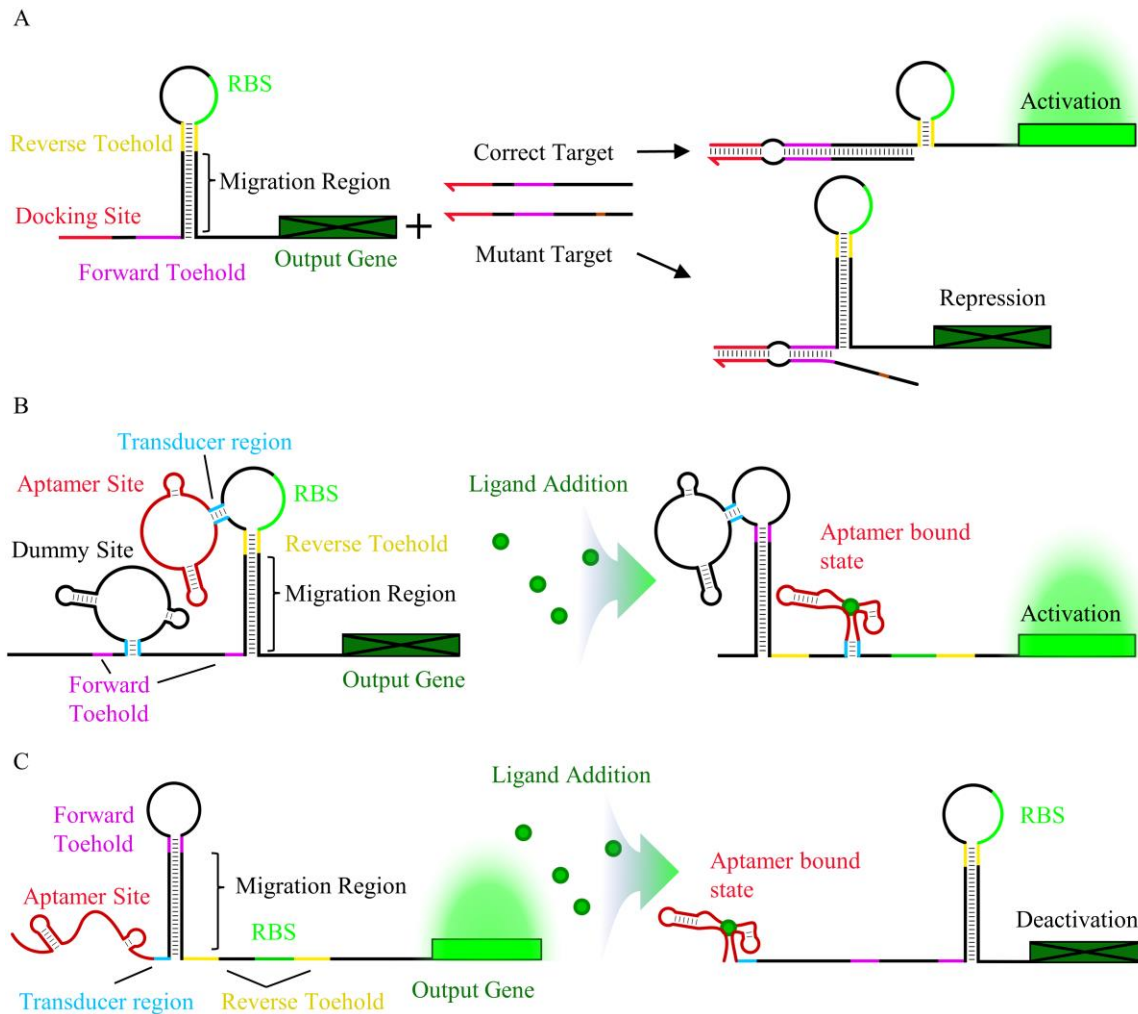
Figure 3.1 Schematics of SNIPR, SAS, and PAS Switches. A) SNIPR sensor schematic. The SNIPR system is similar to a standard toehold switch, but contains competing forward and reverse toeholds to drive a narrow thermodynamic preference for the bound state when a correct targeting strand binds the switch. However, a single nucleotide difference is sufficient to make the activated state thermodynamically unfavorable. B) SAS Switch schematic. Unlike the SNIPR system, the SAS is a single-stranded device with the two potential binding conformations for the migration region, one that forms the forward toehold hairpin (ON) and one that forms the reverse toehold hairpin containing the RBS (OFF). Like SNIPR, the OFF state is favored until the target ligand binds the aptamer and slightly disrupts the stability of the reverse toehold, pushing the thermodynamic balance to favor the ON configuration. The length of the transducer region assists in tuning the aptamer's impact on the conformation change. C) The PAS system is built off the same assumptions as the SAS switch, except that the conformation change primarily disrupts the 3' end of the reverse hairpin region in the OFF configuration, or 5' end of the forward hairpin in the ON configuration.

In the SNIPR system (Fig 3.1A), a hairpin blocks expression of an output gene such as GFP, like a toehold switch. The 5' sequence upstream of the hairpin contains a docking site to which the target RNA will bind. Unlike a typical toehold switch, the wound (OFF) and the unwound (ON) hairpin states with bound target RNA have very similar thermodynamic stabilities, typically favoring unwinding by about -1 to -3 kcal/mol. The target RNA will fully unwind to the ON state if it matches the hairpin exactly, activating translation of the output gene. However, any mutations in the hairpin-binding region will disrupt the thermodynamic stability of binding such that the OFF state dominates and the hairpin remains intact. This the hairpin-trigger binding section is the "migration region" that shifts its binding state to between the ON and OFF device states.

While the SNIPR ribodevice was designed to detect mutations to target sequences, the shift in thermodynamic balanced could come from any input. To this end, we redesigned the SNIPR concept as a single-component device with aptamer binding as the thermodynamic lever to drive ON:OFF state switching. Here we describe how these redesigns led to two types of functioning designs, the SNIPR Aptamer Seesaw (SAS), and the Peripheral Action Seesaw (PAS) switches (Fig 3.1B,C). We tested the SAS designs across the S-adenosyl-methionine (SAM) and Adenosine 5'-diphosphate (ADP) metabolite aptamers[15,16], as well as the fluorogenic Broccoli aptamer[17]. Testing of the PAS switches was conducted primarily with the SAM aptamer. With these devices we were able to demonstrate activating and deactivating switches under the SAS architecture, and deactivating switches under the PAS architecture.

34

3.3 Results and Discussion

3.3.1 PAL Binding in Toehold Switch Systems

We started by interrogating the PAL protein's ability to bind and prevent translation in proven switch designs. Using the proven ACTS type II AND toehold switches designed by Green et al.[11], we modified the existing toehold switches' AND trigger pairs to bind PAL. We theorized that PAL could bind the triggers and prevent proper switch unwinding. To enable PAL-trigger binding we removed the trigger-assembly domain that binds together a typical multi-input AND trigger, placing the PAL-binding motif at two locations in the triggers' sequence (Fig 3.2A,B). The switch sensor with a GFP reporter was placed on a medium copy number plasmid under IPTG induction. The PAL protein and the triggers sequences were placed in a plasmid with a high copy number origin of replication with PAL expression under arabinose control and trigger expression under IPTG control. The switches were tested by co-expressing the PAL-trigger and sensor plasmids under varying IPTG and arabinose concentrations and different light exposure. Cells exposed to 440 nm light showed a maximum 3.5-fold reduction in background- and absorbance-corrected fluorescence versus those grown in darkness with 1 mM IPTG and 4mM arabinose induction of the switch, trigger, and PAL (Fig 3.2C). Triggers with PAL binding motifs placed at the 3' end of the trigger binding site showed a clear performance advantage over those with their motifs located in between the toehold domain and trigger binding site.
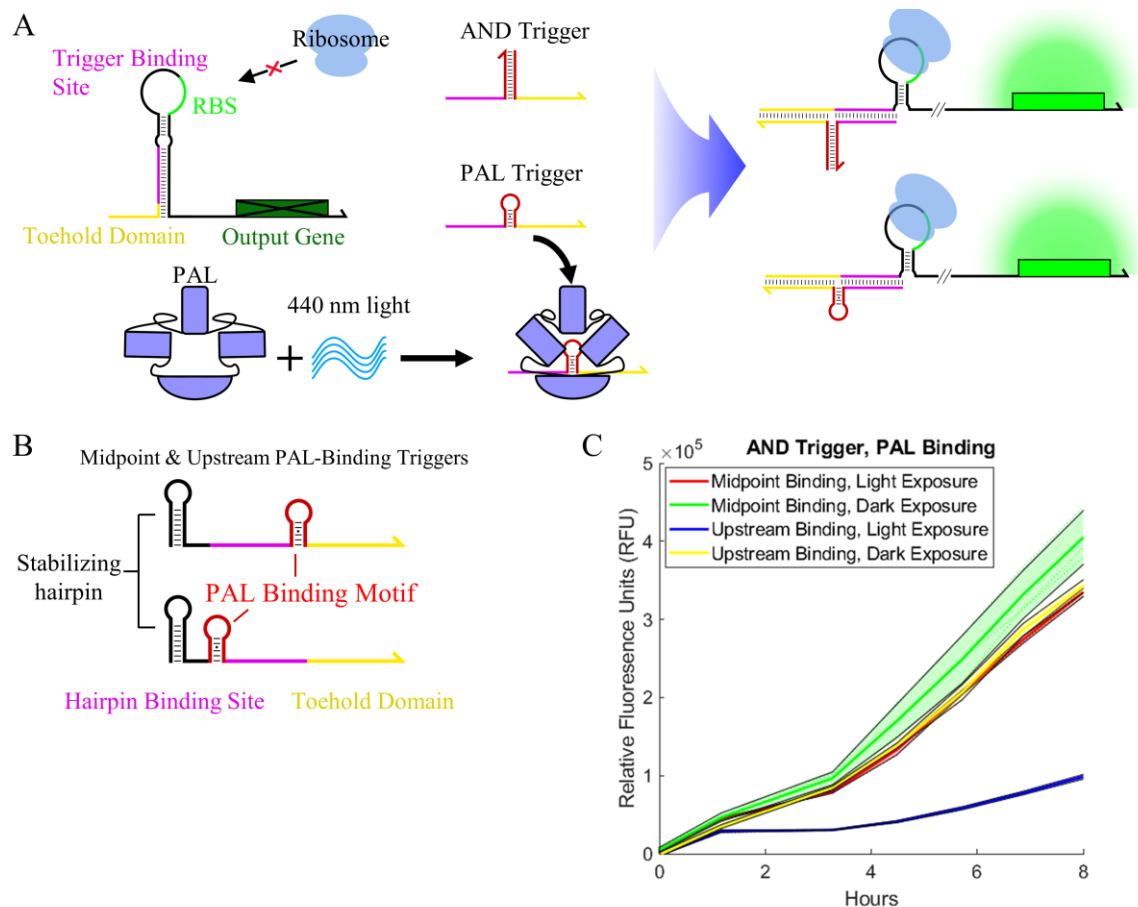
Figure 3.2 Light-sensitive Toehold Switch Performance. A) Schematic of toehold switches with PAL-sequestered triggers. The AND toehold switch design uses two trigger strands that assemble to unwind the hairpin. The modified PAL-sequestered triggers place the PAL-binding motif within a single stranded version of this trigger. Exposure to 440 nm light activates PAL binding, sequestering modified triggers. B) Schematics of trigger modifications tested. The PAL binding motif was placed at a "midpoint" location (top) and an "upstream" location (bottom). C) Testing PAL-sequestration of triggers *in vivo*. DE3 cells were grown at 27 °C and induced with 1mM IPTG and 4mM Arabinose to produce AND Toehold switches, modified triggers, and PAL protein. Samples were grown in 96-well plates with or without 440 nm light exposure and measured approximately once per hour. Plotted data is absorbance corrected with background fluorescence subtracted. The upstream binding site showed a 3.5-fold reduction in GFP fluorescence with 440nm light exposure. Error for plotted samples in (C) are from the standard deviation of n = 3 replicates and represented by the corresponding colored area.

While the dynamic range was not an improvement over that described in the

original Weber et al paper, these devices marked a successful demonstration of

optogenetic integration into the toehold switch framework. When considering how to

improve the performance of these devices, one of the first considerations was the balance of PAL proteins to the triggers. As toehold switches typically use high ratios of trigger to sensor to generate high dynamic ranges, using the same relative amounts of trigger would likely overwhelm the binding capacity of the PAL protein. This led us to consider designs that would shift the stoichiometry in favor of PAL binding, such as placing the trigger and sensor sequence on a single RNA strand.

3.3.2 PAL binding of Thermodynamically Balanced Switches

The binding of PAL to its RNA motif must place steric strain on the surrounding RNA, disrupting any attached structure. One such device that is very sensitive to changes in free energy and structure is the SNIPR system; a free energy difference of less than 3kcal/mol between binding states is sufficient to enable 100-fold shifts in device output[7]. We hypothesized that PAL binding would be sufficient to perturb a single stranded RNA system that was thermodynamically balanced between an active and inactive reporter state. As Figure 3.3 illustrates, we designed a single-stranded ribodevice consisting of upstream "forward" and downstream "reverse" hairpin loops with a migration region between them and an RBS located in the reverse loop. Each loop has either a PAL-binding motif or a non-functional dummy sequence (Fig 3.3A). Switches are designed such that the forward (ON) or reverse (OFF) hairpin will dominate as the migration region attempts to alternate states, with the dominant hairpin containing the PAL motif. Much like the SNIPR designs, the difference in thermodynamic stability between the ON and OFF states is small. We theorized that PAL binding would disrupting the dominant hairpin and pushing the migration region to favor binding the opposite hairpin state. The hairpin gap between the PAL motif and the dominant hairpin was varied to modulate the

impact of PAL binding. The relative stability of the alternate hairpins was tuned by modifications to the forward and reverse toehold binding regions that help initiate formation of the OFF and ON states, respectively. Using the NUPACK software suite[18–21], a library of designs was created with different forward toehold, reverse toehold, and hairpin gap lengths. The PAL motifs were placed on either the forward or reverse hairpin to enable, respectively, deactivation or activation upon binding. Designs were scored and selected based on their structural defect and the difference in free energy between ON and OFF states.
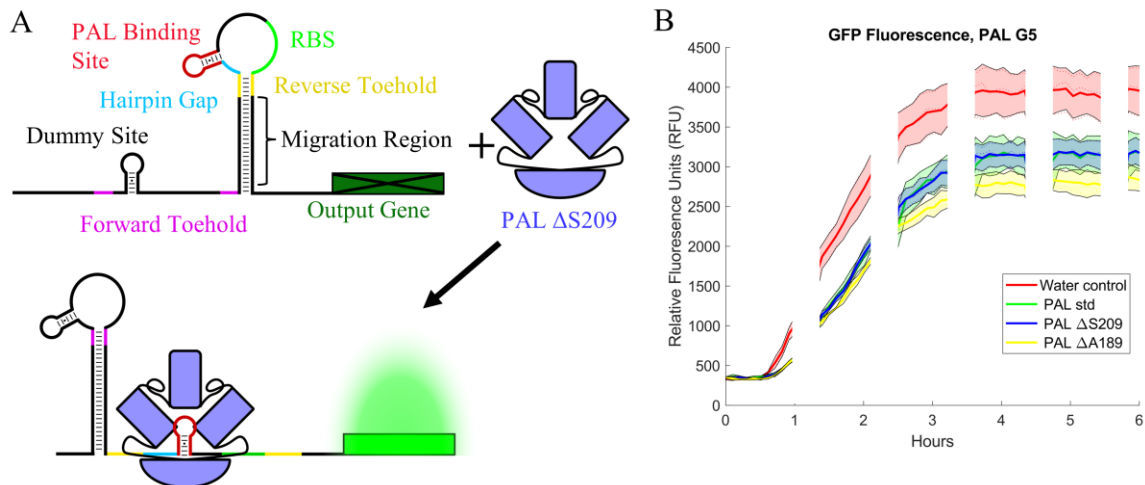
Figure 3.3 Thermodynamically Balanced PAL Switch. A) Schematic of activating PAL-SNIPR switch. The PAL-SNIPR switches are similar overall to the later SAS switches. The primary differences are the use of a protein-binding motif rather than an aptamer, and the choice to vary the hairpin-to-binding site gap rather than use a transducer hairpin when designing switches. The PALΔS209 mutant binds the RNA motif, disrupting the nearby hairpin and allowing the alternate hairpin state to dominate. B) Output of activating PAL-SAS switch designated PAL_G5. The best performing device elicited a 1.16-fold shift in GFP fluorescence, lower than the Broccoli SAS switches. Data gaps are periods of blue-light exposure where no measurements were conducted. Error for plotted samples in (B) are from the standard deviation of n = 3 replicates and represented by the corresponding colored area.

Selected devices were inserted into the pSNIPR plasmid used by Hong et al.[7] with a T7 promoter and GFPmut3b reporter containing a C-terminal ASV degradation tag[22].

38

We ran the testing in the PURExpress cell free expression system (New England Biolabs) to enable faster screening than *in vivo* testing. In addition to standard PAL, we utilized its mutant derivatives as controls. The PALΔS209 mutant always binds to its corresponding RNA motif, regardless of light conditions, while the PALΔA189 mutant possesses no binding activity and serves as a negative control. To produce the PAL protein in the PURExpress system, PAL DNA under the control of a T7 promoter was PCR amplified and purified with spin-column filtration (Qiagen, 28106) to be added at controlled PAL-DNA to sensor ratios. Screening the switch behavior with both standard PAL, PAL mutants, and non-PAL controls enabled us to account for the decreased GFP fluorescence caused by PAL production (Fig 3.3B). Samples were periodically exposed to light during measurement in the cell-free system to evaluate the ability of standard PAL to activate the devices upon light exposure.

While no deactivating switches were produced, some results suggested designs to be capable of binding and increasing GFP output relative to the non-binding ΔA189 mutant control (Fig 3.3B). However, the dynamic range was much worse than our prior trigger-binding system, with only a sixteen percent increase in GFP output between switches transcribed with PALΔS209 and PALΔA189 mutants expressed. Given the binding site proximity to the RBS, it is possible that whatever switching behavior incurred by the PAL binding may have been summarily inhibited by the PAL partially obscuring the RBS, or that the size of the bound PAL was sufficient to disrupt both the ON and OFF states when bound, preventing strong switch behavior. In any case, the poor output and complexity of testing this optogenetic system discouraged us from further testing, as our results indicated that the PAL binding system was not sufficiently robust to

be utilized in this manner. However, the underlying principles of this device were sound, leading us to pursue other mechanisms enable ON:OFF state switching.

3.3.3 SAS Switch Design and Testing for SAM, ADP metabolites

We modified the thermodynamically balanced device from the PAL investigations into what we refer to as the SNIPR Aptamer Seesaw, or SAS, switch (Fig 3.1B). The SAS switches are nearly identical to those designed for PAL-based switching with a few key differences described here. First, the dominant hairpin loops contain an aptamer sequence rather than the PAL-binding motif and second, we used the length of the hairpin connecting the aptamer to the hairpin loop to attempt to modulate the binding impact rather than the spacer distance between the aptamer and the dominant hairpin. Switches are designed such that the forward (ON) or reverse (OFF) hairpin will dominate as the migration region attempts to alternate states, with the dominant hairpin containing the aptamer sequence. In our first round of designs, SAS switches used aptamer sites that bound S-adenosyl-methionine (SAM) and adenosine 5'-diphosphate (ADP) with GFP as the reporter protein. We sourced SAM and ADP aptamer sequences from the literature[15,16,23,24] for their hairpin sites and we placed non-functional "dummy" sequences with similar structures on their counterpart loops to encourage the formation of ON and OFF states of similar thermodynamic favorability. The switches were designed in the same manner as prior designs, using the NUPACK software suite to generate designs, then[33–36] scoring and selecting devices to test based on their defect and the free energy difference between ON and OFF states.
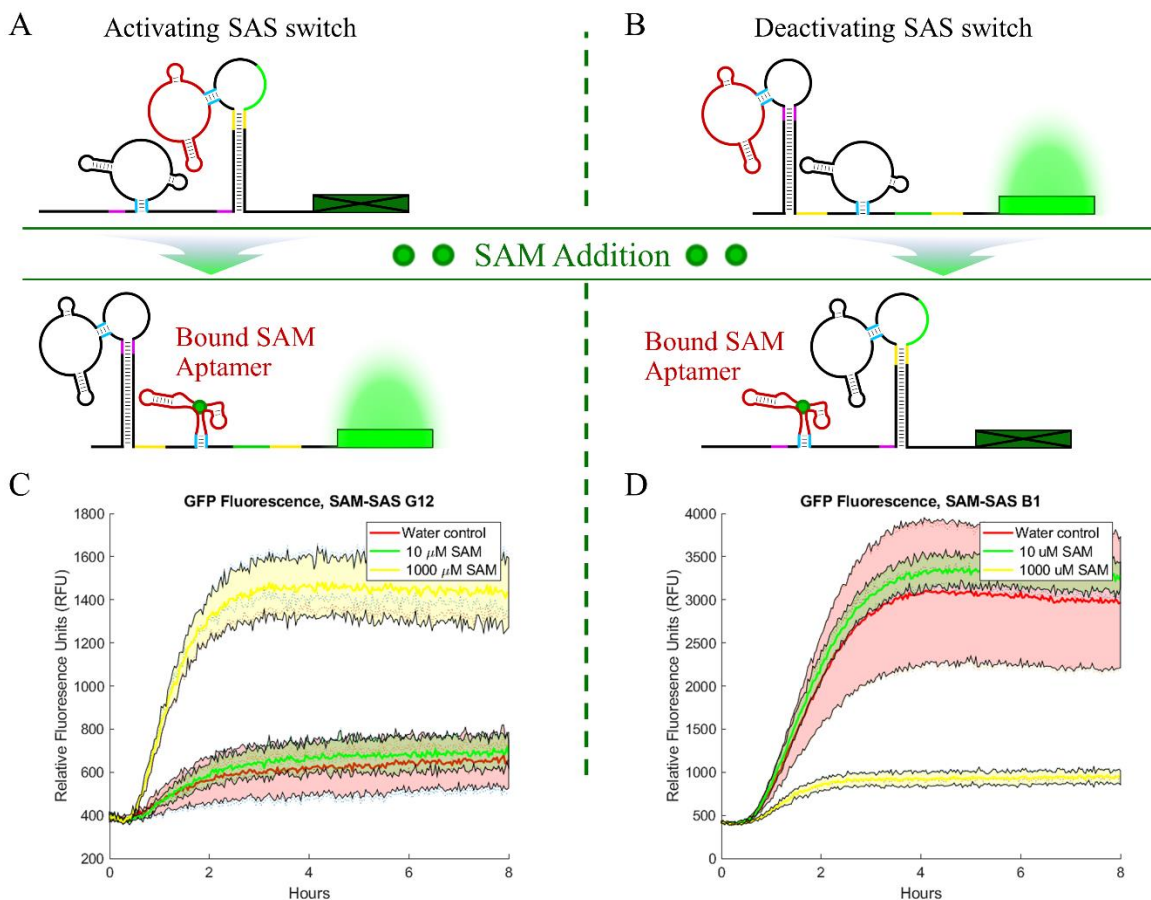
Figure 3.4 SAM-responsive SAS switches. A) Schematic of activating SAM-SAS switch. The aptamer site of activating switches is located in the reverse hairpin loop. Upon the addition of the SAM metabolite, the free energy shift of the aptamer-ligand interaction drives a conformation change from the OFF state (above) to the ON state (below). B) Schematic of deactivating SAM-SAS switch. The aptamer site of activating switches is located in the forward hairpin loop. Like the activating SAS devices, the aptamer-ligand interaction drives a conformational change to switches from the ON state (above) to the OFF state (below). C) GFP expression increased 2.3-fold with the addition of 1000 µM SAM metabolite. D) GFP expression decreased 3.6-fold with addition of 1000 µM SAM metabolite. Error for plots in (C) and (D) are from the standard deviation of n = 3 replicates and represented by the corresponding colored area.

Following plasmid assembly into the same backbone as prior PAL tests, switches were tested both *in vivo* and *in vitro*. *In vivo* testing showed little bulk change in fluorescence, possibly due to the more complex nature of the cytoplasm preventing a single state from dominating, and possibly due to incomplete control over the presence metabolites in living environments. To remove any confounding factors and test the

41

device concept, all further evaluation was conducted using cell free production in the PURExpress system (New England Biolabs). In this controlled environment, we tested the impact of SAM and ADP addition in concentrations ranging from 1 µM to 1000 µM. We successfully designed several devices that activated or deactivated in the presence of SAM (Fig 3.4). These devices exhibited a 2.3-fold ON:OFF shift in the case of SAM-activating devices, and 3.6-fold ON:OFF shift in the case of SAM-deactivating devices. Unfortunately, full evaluation of the ADP sensors was prevented by the metabolite's non-specific inhibitory effect on the PURExpress system (Appendix B Figure S3.2). While we successfully demonstrated the core SAS switch functionality, our inability to test ADP sensing led us to we design additional switches with the Broccoli aptamer to test if the SAS architecture could serve as a generalized design for interchangeable aptamer use.

3.3.4 SAS Switch Testing with the Broccoli Aptamer

The Broccoli aptamer binds to and induces fluorescence in (Z)-4-(3,5-difluoro-4-hydroxybenzylidene)-2-methyl-1-(2,2,2-trifluoroethyl)-1H-imidazol-5(4H)-one (DFHBI-1T)[17], which is a structural mimic of the GFP fluorophore, 4-hydroxybenzylidene-imidazolinone (HBI). The Broccoli aptamer was chosen so that the fluorescent properties of the ligand would help us confirm if the generated SAS switches were allowing proper aptamer-ligand binding complexes to form, as well as avoid interacting with any metabolic pathways as we may have seen when testing the ADP-sensitive switches. The GFPmut3 reporter was removed and an mCherry reporter used in its place for independent signal measurement. Additionally, the structure of the dummy aptamer site was modified to simplify device design within the NUPACK software suite. While the SAM SAS switches used structural mimics of the aptamers to balance the relative

stability of each hairpin state, the Broccoli SAS switches used a simple 8-basepair dummy hairpin to create similar free energy states while minimizing the probability of base pair mismatches in the design generation phase (Fig 3.5A).

When testing the first set of broccoli designs, we started with cell-free screening. A minor complicating factor was the Broccoli-DFHBI-1T complex's dependence on magnesium ions present for full fluorescence. Since magnesium shifts the stability of RNA devices and interferes with the PURExpress cell free system[25] (Appendix C, Fig 3.3), we used a 1:10 dilution of the standard DFHBI-1T buffer to minimize its impact on plasmid expression. In testing the impact of the buffer composition on our devices, we observed the buffer content enabling switch behavior as well. While excess amounts of the DFHBI-1T buffer could nonspecifically inhibit reporter expression, the overall shift in RNA stability that the buffer ions provided was enough to demonstrate some state-switching (Appendix B, Fig S3.4). Cell free screening of the Broccoli sensors was conducted with DFHBI-1T concentrations ranging from 0.4 µM to 40 µM, revealing candidates for both activation and deactivation. However, these had very low dynamic ranges, with the best activating switch demonstrating only a 1.8-fold ON:OFF ration, and only a 1.3-fold ON:OFF ration in the case of de-activating designs (Fig 3.5).
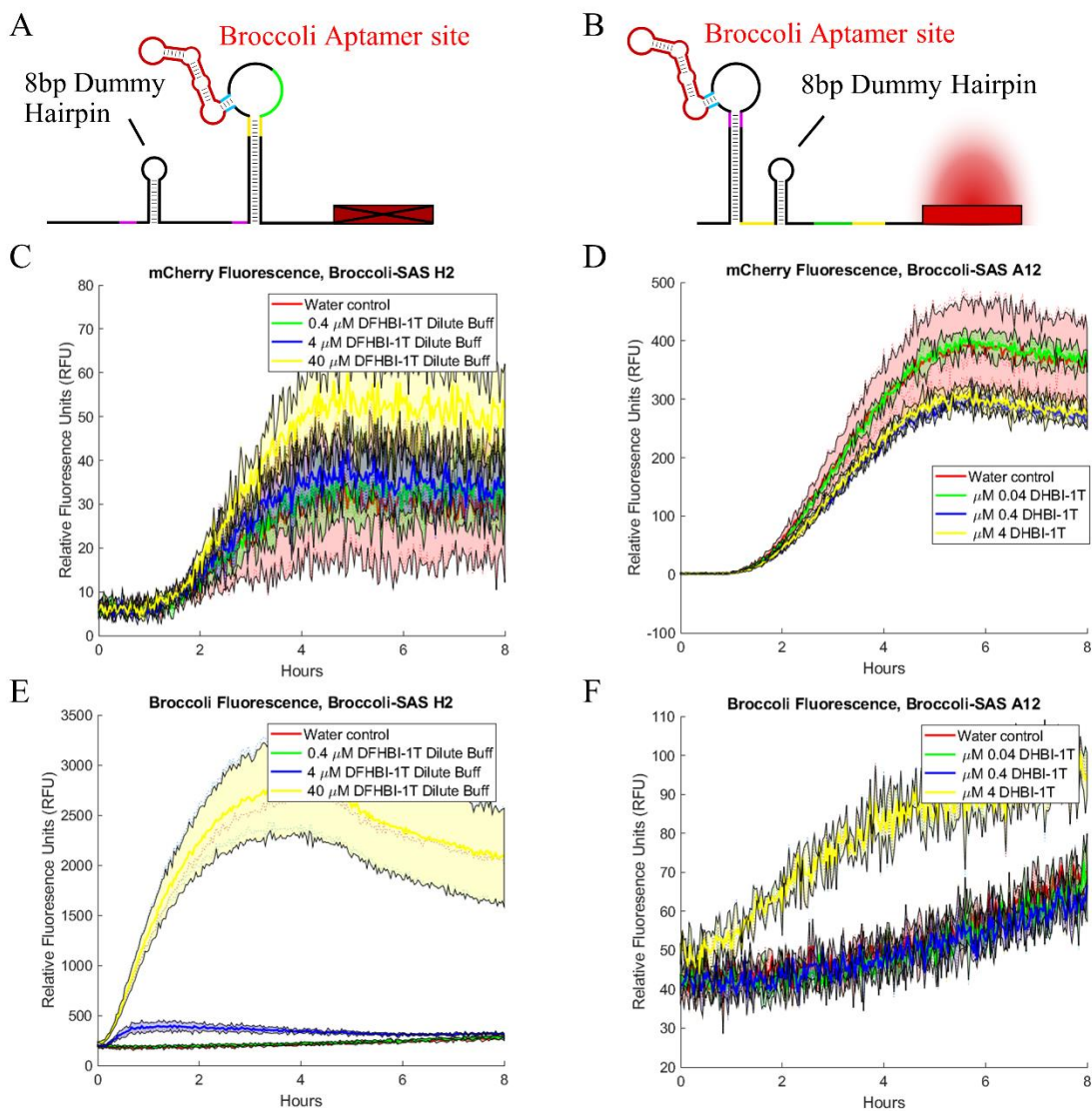
Figure 3.5 Broccoli-responsive SAS switches. The SAS switches with Broccoli differ from the devices created for SAM- and ADP-based switching in both activating (A) and deactivating (B) switches. Broccoli-SAS switches use an 8bp dummy hairpin rather than a dummy site of identical structure and use the mCherry reporter gene rather than GFP. The Broccoli-SAS switches showed less switching behavior than the SAM-SAS switches. (C) The highest activating switch, designated Broc H2, showed a 1.8-fold increase in fluorescence. while the greatest deactivating switch, designated Broc A12, showed a 1.3-fold reduction in fluorescence (D). Despite lower net fluorescence and dynamic range, the use of the Broccoli aptamer site allowed us to confirm that the aptamer sites form and bind to targets as expected (E,F). Error for plots in (C, D, E, F) are from the standard deviation of n = 3 replicates and represented by the corresponding colored area.

Analysis of predicted minimum free energy (MFE) structures for our different

aptamers indicated that the SAM aptamer possess a less strictly defined secondary

structure compared to the Broccoli aptamer (Appendix B, Fig S3.5). Past studies have

noted that while the full Broccoli aptamer structure is reasonably stable, the presence of

the DFHBI-1T ligand alone may not be sufficient to induce structural shift. Filonov et al

note that Broccoli-DFHBI-1T activation was overruled by the strength of a cyclic di-

GMP aptamer[17], and the SAM aptamer used in our study has been shown to induce

conformational changes in other RNA-fluorophore complexes[24]. Crystallographic

investigations into the bound and unbound structure of similar fluorophore complexes

such as the Spinach aptamer noted that DFHBI binding caused only local structural

perturbations rather than a global conformational change on the aptamer structure[26]. This

suggests that the binding of DFHBI-1T to the Broccoli aptamer does not cause a

significant structural shift, leading to a reduced shift in reporter output compared to the

SAM binding SAS switches.

3.3.5 Piecewise Modification of SAM SAS Devices and Parameter Analysis

To better understand the limits of the SAS system, we modified the structural

parameters of successful switches to test if an individual parameter drives general device

performance. Using three successful SAM switches to compare baseline performance, we

tested if such modifications would the switch dynamic range or increase its sensitivity to

the SAM target. The devices tested were two deactivating SAS devices, designated SAM

B1 and SAM D2, and the activating SAM G12 switch. The structural parameters

modified were forward toehold length, reverse toehold length, aptamer site loop length,

and transducer length. Length modifications were additions and subtractions of one to

two nucleotides for all parameters except the aptamer site loop, whose modifications added or subtracted two or four nucleotides.

Modifications to the sequence and structure of the tested switches allowed for improvements in the ON:OFF ratio of devices, increasing the ON:OFF fold-shift and sensitivity to the SAM metabolite at lower concentrations (Appendix B, Fig. S3.6B). However, the impacts of these modifications were not consistent. While changes to some areas more frequently caused a change in output, such as the aptamer site loop, the results of said changes were not consistent between samples, nor did they react in a linear manner. For example, increasing the SAM B1 switch's aptamer site loop length by four nucleotides improved the ON:OFF fold-shift by 29% and shortening it by four nucleotides increases the net fluorescent signal with minor change to the dynamic range, yet shortening this region by two nucleotides causes an 18.5-fold reduction in baseline fluorescence and a 20% reduction in fold-shift range. To better analyze the impact of these design parameters on the switch performance we conducted multivariate linear regression analysis on the designs using the ON:OFF ratio, as well as the ON and OFF fluorescent values themselves, as dependent variables. The independent variables included the structural design parameters of the forward toehold length, reverse toehold length, and transducer length, as well as calculated thermodynamic parameters such as state defect, minimum free energy, free energy shift between ON and OFF states ($\Delta\Delta G$), and several others.

Analysis of SAM SAS performance by multivariate linear regression contradicted our assumptions regarding which factors were most important for device selection and

design. The predictors of switch function with significant p-values ($p < 0.005$) were the ON state defect, OFF state defect, total structural defect, and Transducer region GC content (Appendix B, Table S3.1 and Table S3.2), with the latter two positively correlated with increasing ON:OFF ratios. Notably, the predicted free energy difference between ON and OFF states had no correlation with device performance while structural defect much more strongly predicted device function. Consideration of our design software's limitations aligns with this result. While we might be able to provide estimates for the free energy shift of aptamers-ligand binding, NUPACK and other openly available software packages for RNA design cannot readily incorporate such information into their calculations. The methods for running such calculations have been described[27,28] but there is not readily available design software that can determine such estimates. This leaves the structure defect predictions as the best predictive metric without extensive expansion and refinement of the switch design pipeline. This metric may still be confounded by aptamers that adopt conformations that NUPACK cannot properly represent, such as the Broccoli aptamer's G-quadruplex. Even so, the results of this analysis were used to better inform later design scoring and selection.

The significance of the transducer GC content presents several potentail implications for the SAS switch design. One interpretation is that the transducer functions as a kinetic trap. Consider that passive unwinding of an AA to UU basepair at 37 °C occurs 42 times faster than a CC to GG basepair despite a difference in free energy of only about 2.3 kcal/mol[29]. Given that the selection pipeline picks SAS devices with very similar ON and OFF states free energys, the kinetics of the aptamer site are a logical lever for controlling device state switching. This is somewhat reinforced by observations

47

of GC content in the branch migration region. GC content did not have significant correlation with switch ON:OFF performance but functioning switches had branch migration regions with a GC content of approximately 30-50%. If we assume that kinetics are the major determinant of hairpin unwinding, high GC content is unlikely to be conducive to fast switching behavior due to the slower pace of unwinding. For this reason GC content was evaluated during later design scoring and selection.

3.3.6 PAS Switch Design and Testing

With the core of our design concept demonstrated with the SAS switches, we tested other aptamer site placements to see if they might provide enhanced state-switching, leading to the development of what we call the Peripheral Action Site (PAS) switch. The PAS devices drew from our initial success with PAL-trigger sequestration. As the more effective PAL-binding triggers had the PAL binding location at their 3' end, we thought modifying the SAS design to place aptamer binding sites at the corresponding locations might prove successful. In these designs, a deactivating switch's aptamer site would be located upstream of the fully bound forward hairpin, with a single stranded transducer region separating it to modulate hairpin disruption (Fig3.6A). Conversely, the activating hairpins' aptamer site would be downstream of the fully bound reverse toehold hairpin, also with a single stranded transducer region between the hairpin and aptamer site to modulate hairpin disruption. As with prior testing, NUPACK was used to generate potential designs.
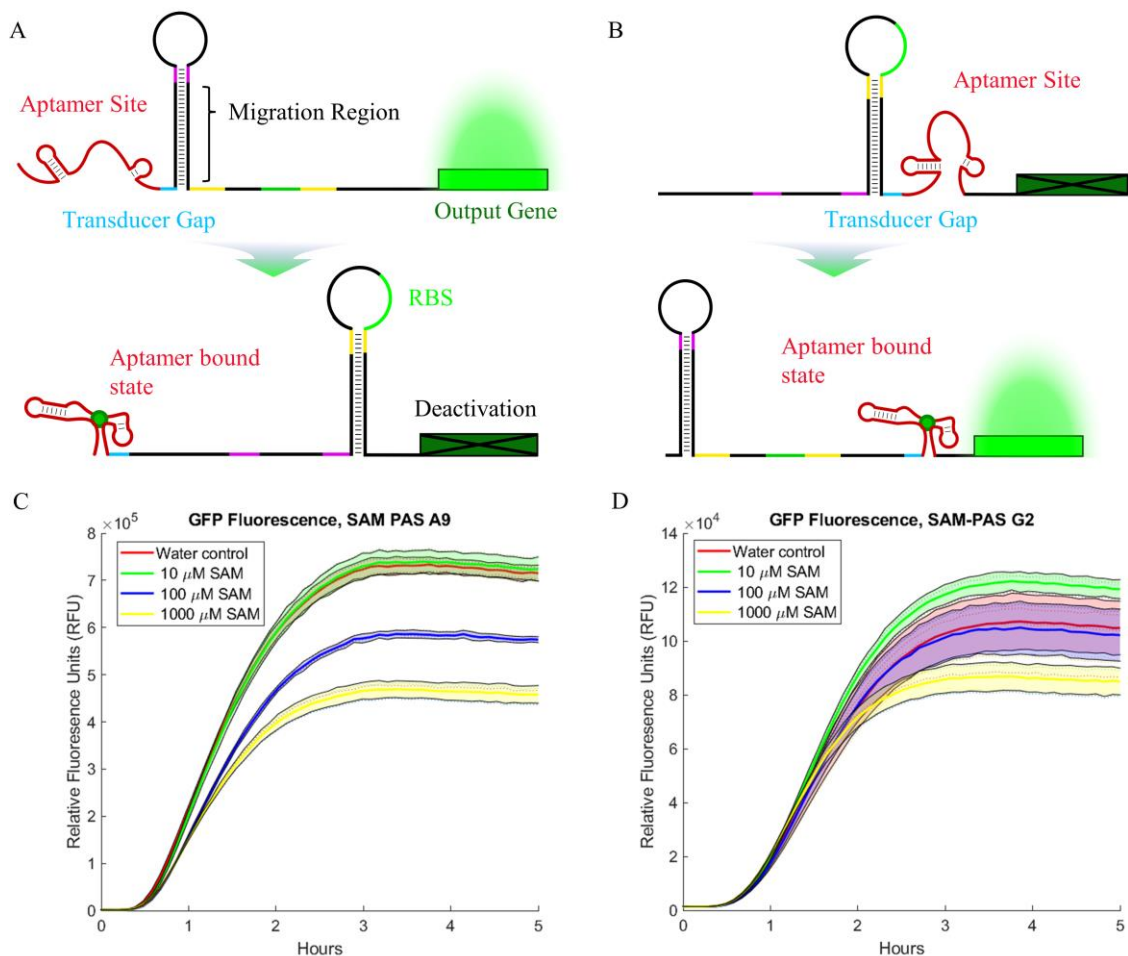
Figure 3.6 Peripheral Action Site Switch Using SAM Aptamer. A) Schematic of deactivating PAS switch operation. B) Schematic of activating PAS switch operation. Rather than the aptamer sites perturbing the dominant hairpin loop, the aptamer site disrupts the edge of the dominant loop's migration hairpin. C) Deactivating PAS switch function. Example Data from SAM-PAS switch A9. The PAS-SAM A9 device displayed a 1.56-fold decrease in fluorescence upon 1000 µM SAM exposure. D) Potential functioning of activating PAS switch. This was the only switch that showed potential function for the activating PAS designs, with a 15% increase at only 10 µM SAM exposure. However, higher concentrations showed a clear drop in activation. Error for (C) and (D) is from the standard deviation of n = 3 replicates and represented by the corresponding colored area.

Selected devices were assembled into the same plasmids as prior SAM devices

and testing was conducted in an identical manner using cell free expression. Several

deactivating devices were generated in this manner, but none of those tested showed

greater fold-shift designs successfully induced designs. A 1.56-fold shift from the SAM-

PAS A9 design was the maximum observed (Fig. 3.6C). Of the tested activating designs, none showed definitive activation. One of the designs, SAM-PAS G2, had some markers of activation with a 15% increase in GFP signal at only 10 µM SAM exposure. However, this effect dropped off at higher concentrations.

3.4 Conclusion

While not the primary focus of this text, the initial results from our PAL-binding studies indicate that protein-binding can be integrated into toehold switches with potentially minimal modification to existing trigger structures. This suggest that conferring the sensing capabilities of a given RNA binding protein to the toehold switch may prove a fruitful endeavor. With more effective light-sensitive RBP variants being developed in the wake of PAL protein discovery[30], new candidates exist for further integration of light-sensing capabilities into toehold switches. While determining the degree to which these capabilities could be extended into existing toehold switches is beyond the scope of this work, we might further venture that the OR logic demonstrated in prior toehold switch studies[11] could be maintained as the protein binding modifications we enacted only changed to the trigger portion of the toehold switch.

The SAS switch demonstrates a semi-generalizable approach for integrating different chemical sensors into the toehold switch design paradigm. The PAS architecture may function better with different aptamers or binding mechanisms, but the testing so far has demonstrated it to be less effective than the SAS designs. As differences in device performance between aptamers indicate, not all aptamers would function equally well in this design. Aptamers that exhibit conformational changes throughout their structure

50

upon binding target ligands are the most likely to prove strong switching behavior, as these can better drive non-local changes that disrupt forward or reverse hairpin stability to drive the SAS or PAS device state change. Screening designs via thermodynamic shift was complicated by the limitations of current RNA-interaction design software. While NUPACK is sufficient to evaluate the thermodynamics and design of RNA-only interactions, modeling the interactions of RNA with other chemicals in a design workflow is not something that readily available software can do. Expansion of the design pipeline used in prior toehold switch generating studies to include non-ribonucleic acid interactions would likely greatly improve our ability to generate functional designs. Defect estimates and heuristics based on the aptamer-ligand interaction are the remaining options with the current design pipeline. Nevertheless, the SAS switch provides a ribodevice framework for integration of an arbitrary chemical sensor into a toehold switch-like architecture and design pipeline.

3.5 Methods

3.5.1 Ribodevice Design Generation

Oligonucleotide sequences for the SASS devices were designed using version 3.2 of the NUPACK software suite[18–21]. The oligonucleotides were purchased from Integrated DNA Technologies.

3.5.2 SAS and PAS Switch Plasmid Assembly

SAS and PAS switch sequences were constructed through PCR and Gibson Assembly techniques, using the Phusion High-Fidelity PCR Master Mix (NEB, M0531L).and Gibson Assembly Master Mix (NEB, E2611L) products. Switch sequences

51

were inserted into a plasmid containing a pColA origin of replication, kanamycin resistance, and either GFPmut3 or mCherry fluorescent proteins under T7 expression. This plasmid was the pSNIPR plasmid, obtained from Addgene (Addgene, # 139463). The sequences were inserted immediately after the T7 promoter sequence, with a 21 nucleotide linker sequence between the 5' end of the migration region and the fluorescent protein start codon. In the case of activating PAS designs, the linker region immediately followed the aptamer site. Following amplification and assembly, plasmids were transformed into *E. coli* DH5α cells and selected for using LB agar plates with antibiotic. Colonies were sequenced via Sanger sequencing through Genewiz, Eton Bioscience Inc., and Quintara Biosciences. Successfully assembled plasmids were purified via miniprep purification (Qiagen, 27106X4).

3.5.3 PAL-Toehold Switch Design and Assembly

ACTS Type II AND Switch and trigger plasmids were obtained from the Green Lab. A pCDF backbone plasmid expressing the PAL protein under Arabinose-inducible control with spectinomycin resistance was provided by the Weber Lab. Trigger sequences of the AND switches were modified to replace the binding region with a fused PAL motif. Modification of these plasmids was completed using PCR amplification followed by KLD assembly. Amplification was verified using gel electrophoresis. Oligonucleotides used in the editing process were sourced from Integrated DNA Technologies. Phusion High-Fidelity PCR Master Mix (NEB, M0531L), and KLD Enzyme mix (NEB, M0554S) were used for PCR amplification and KLD assembly respectively. After KLD assembly, plasmids were transformed into *E. coli* DH5α cells and selected for using LB agar plates with antibiotic. Colony DNA was amplified using

the TempliPhi RCA kits (Cytiva Life Sciences, 25640050), then sequenced via Sanger sequencing through Genewiz and the Biodesign Sequencing Lab at ASU. Successfully assembled plasmids were extracted via QIAprep Spin Miniprep Kit purification (Qiagen, 27106X4).

3.5.4 Plate Reader Measurements for Cell-free Testing

Plate reader measurements were conducted using 384-well plates in a Biotek Synergy Neo2 Hybrid Multi-Mode Reader and a Biotek Cytation Hybrid Multi-Mode Reader. Fluorescence readings were taken with excitation/emission pairs of 480/520 nm when measuring GFP and Broccoli fluorophore fluorescence, and 580/610 nm when measuring mCherry fluorescence. Cell free reactions were run at 37 °C using the PURExpress® In Vitro Protein Synthesis Kit (NEB, E6800L) in 4 µl volumes. Reaction wells were surrounded by water-filled wells to prevent the reaction volume from evaporating. In the case of cell free testing of PAL designs, 5:1 ratios of PAL to sensor DNA were used for screening. This ratio resulted in sufficient PAL production to cause a small drop in GFP production alone, indicating that sufficient PAL would be produced under these conditions that could bind target sensor RNA and produce a noticeable shift in reporter output during cell-free screening.

3.5.5 Plate Reader Measurements for *in vivo* PAL-Toehold Switch Testing

Plate reader measurements were conducted using 96-well plates in a Biotek Synergy Neo2 Hybrid Multi-Mode Reader and a Biotek Cytation Hybrid Multi-Mode Reader. To measure PAL-based switching of toehold triggers, *E. coli* DE3 cells were triple transformed with the switch-GFP plasmid and PAL + trigger plasmid. Colonies

were cultured overnight in Luria Broth with kanamycin (50 mg/mL) and spectinomycin (50 mg/mL), then diluted to an absorbance of 0.2 OD_{600} and incubated at 180 RPM at 37 °C for two hours before plating into 96-well plates and inducing with either IPTG and/or arabinose at the concentrations described in the experiments. Plates were incubated at 25 °C under transilluminator exposure, with measurements taken every hour.

3.5.6 PAL Protein Purification

DE3 cells were transformed with plasmids expressing PAL or one of the mutant versions, colonies were picked and grown overnight in Luria Broth with spectinomycin (50 mg/mL). Cell cultures were then diluted 1:50 into 200 mL of Luria Broth with spectinomycin (50 mg/mL) and grown in a shaking incubator at 180 RPM and 37 °C until reaching an OD_{600} of 0.6, at which point they were induced with 4mM arabinose and left to incubate an additional 6 hours. After incubation, cultures were split into centrifuge vessels and placed on ice for 10 minutes before being pelleted via centrifuge at 4200 RPM for ten minutes. The cell pellets were resuspended in Tris-HCl Buffer (50 mM Tris-HCl, 5 mM EDTA, pH 8) and lysozyme was added to 1 mg/mL, then left to incubate on ice for one hour. Cells were then disrupted via sonication in ice bath using a Bronson 450W homogenizer with a 6mm probe tip. Sonication was conducted using 1 second pulses and 2 second rests at 10% power for 20 minutes of run time. Cultures were then aliquoted into 1.5 ml Eppendorf tubes and centrifuged at 16000 G for 20 minutes to separate cell debris. Supernatant was then combined with Equilibration Buffer (20mM sodium phosphate, 300mM sodium chloride (PBS) with 10mM imidazole; pH 7.4) and purified using HisPur™ Ni-NTA Spin Columns, 3 mL (ThermoFisher Scientific, 88226).

3.6 References

1.   Wrist, A., Sun, W. & Summers, R. M. The Theophylline Aptamer: 25 Years as an Important Tool in Cellular Engineering Research. *ACS Synthetic Biology* vol. 9 682–697 (2020).

2.   Sharma, V., Nomura, Y. & Yokobayashi, Y. Engineering complex riboswitch regulation by dual genetic selection. *J Am Chem Soc* **130**, 16310–16315 (2008).

3.   Climent-Catala, A., Ouldridge, T. E., Stan, G. B. v. & Bae, W. Building an RNA-Based Toggle Switch Using Inhibitory RNA Aptamers. *ACS Synthetic Biology* **11**, 562–569 (2022).

4.   Tinafar, A. *et al.* Cell-Free Biosensors: Synthetic Biology Without Borders 12. (2022) doi:10.1007/978-3-030-23217-7_130.

5.   Pardee, K. *et al.* Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell* **165**, 1255–1266 (2016).

6.   Ma, D., Shen, L., Wu, K., Diehnelt, C. W. & Green, A. A. Low-cost detection of norovirus using paper-based cell-free systems and synbody-based viral enrichment. *Synthetic Biology* **3**, (2018).

7.   Hong, F. *et al.* Precise and Programmable Detection of Mutations Using Ultraspecific Riboregulators. *Cell* **180**, 1018-1032.e16 (2020).

8.   Weinberg, B. H. *et al.* Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. *Nature Biotechnology* **35**, 453–462 (2017).

9.   Grieger, J. C. & Samulski, R. J. Packaging Capacity of Adeno-Associated Virus Serotypes: Impact of Larger Genomes on Infectivity and Postentry Steps. *Journal of Virology* **79**, 9933–9944 (2005).

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

11.   Green, A. A. *et al.* Complex cellular logic computation using ribocomputing devices. *Nature* **548**, 117–121 (2017).

12.   Kim, J. *et al.* De novo-designed translation-repressing riboregulators for multi-input cellular logic. *Nature Chemical Biology* **15**, 1173–1182 (2019).

13.   Castro, N. J., O'Brien, J. & Zhang, L. G. Integrating biologically inspired nanomaterials and table-top stereolithography for 3D printed biomimetic osteochondral scaffolds. *Nanoscale* **7**, 14010–14022 (2015).

14.   Weber, A. M. *et al.* A blue light receptor that mediates RNA binding and translational regulation. *Nature Chemical Biology* **15**, 1085–1092 (2019).

15.   Li, X. *et al.* Imaging Intracellular S-Adenosyl Methionine Dynamics in Live Mammalian Cells with a Genetically Encoded Red Fluorescent RNA-Based Sensor. *J Am Chem Soc* **142**, 14117–14124 (2020).

16.   Litke, J. L., You, M. & Jaffrey, S. R. Developing Fluorogenic Riboswitches for Imaging Metabolite Concentration Dynamics in Bacterial Cells. in *Methods in Enzymology* vol. 572 315–333 (Academic Press Inc., 2016).

17.   Filonov, G. S., Moon, J. D., Svensen, N. & Jaffrey, S. R. Broccoli: Rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J Am Chem Soc* **136**, 16299–16308 (2014).

18.   Wolfe, B. R., Porubsky, N. J., Zadeh, J. N., Dirks, R. M. & Pierce, N. A. Constrained Multistate Sequence Design for Nucleic Acid Reaction Pathway Engineering. *J Am Chem Soc* **139**, 3134–3144 (2017).

19.   Zadeh, J. N., Wolfe, B. R. & Pierce, N. A. Nucleic acid sequence design via efficient ensemble defect optimization. *Journal of Computational Chemistry* **32**, 439–452 (2011).

20.   Wolfe, B. R. & Pierce, N. A. Sequence Design for a Test Tube of Interacting Nucleic Acid Strands. *ACS Synthetic Biology* **4**, 1086–1100 (2015).

21.   Dirks, R. M., Lin, M., Winfree, E. & Pierce, N. A. Paradigms for computational nucleic acid design. *Nucleic Acids Research* **32**, 1392–1403 (2004).

22.   Andersen, J. B. *et al.* New Unstable Variants of Green Fluorescent Protein for Studies of Transient Gene Expression in Bacteria. *Applied and Environmental Microbiology* **64**, 2240 (1998).

23.   You, M., Litke, J. L. & Jaffrey, S. R. Imaging metabolite dynamics in living cells using a Spinach-based riboswitch. *Proceedings of the National Academy of Sciences of the United States of America* vol. 112 E2756–E2765 (2015).

24.   Paige, J. S., Nguyen-Duc, T., Song, W. & Jaffrey, S. R. Fluorescence imaging of cellular metabolites with RNA. *Science* **335,** 1194 (2012).

25.   New England Biolabs. PURExpress FAQs. https://www.neb.com/-/media/nebus/files/pdf-faq/purexpress-faqs.pdf.

26.   Huang, H. *et al.* A G-Quadruplex-Containing RNA Activates Fluorescence in a GFP-Like Fluorophore. *Nat Chem Biol* **10**, 686 (2014).

27.   Oliveira, R. *et al.* Modelling aptamers with nucleic acid mimics (NAM): From sequence to three-dimensional docking. *PLOS ONE* **17**, e0264701 (2022).

28.   Buglak, A. A., Samokhvalov, A. v., Zherdev, A. v. & Dzantiev, B. B. Methods and applications of in silico aptamer design and modeling. *International Journal of Molecular Sciences* vol. 21 1–25 (2020).

29.   Li, P. T. X., Vieregg, J. & Tinoco, I. How RNA unfolds and refolds. *Annual Review of Biochemistry* vol. 77 77–100 (2008).

30.   Liu, R. *et al.* Optogenetic control of RNA function and metabolism using engineered light-switchable RNA-binding proteins. *Nature Biotechnology* **40**, 779-786 (2022).

REFERENCES

Chapter 1 References

1.   Casjens, S. R. & Hendrix, R. W. Bacteriophage lambda: Early pioneer and still relevant. *Virology* vols. 479–480 310–330 (2015).

2.   Keen, E. C. A century of phage research: Bacteriophages and the shaping of modern biology. *BioEssays* **37**, 6–9 (2015).

3.   Smith, G. P. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science (1979)* **228**, 1315–1317 (1985).

4.   Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* vol. 482 331–338 (2012).

5.   Atanasova, P. *et al.* Virus-templated synthesis of ZnO nanostructures and formation of field-effect transistors. *Advanced Materials* **23**, 4918–4922 (2011).

6.   Gerasopoulos, K., McCarthy, M., Royston, E., Culver, J. N. & Ghodssi, R. Microbatteries With Tobaco Mosaic Virus Templated Electrodes. in *2008 IEEE 21st International Conference on Micro Electro Mechanical Systems*.

7.   Nam, K. T. *et al.* Virus-Enabled Synthesis and Assembly of Nanowires for Lithium Ion Battery Electrodes. *Science* **312**, 885-888 (2008).

8.   Nam, K. T., Reelle, B. R., Lee, S. W. & Belcher, A. M. Genetically Driven Assembly of Nanorings Based on the M13 Virus. *Nano Letters* **4**, 23–27 (2004).

9.   Luo, Q., Hou, C., Bai, Y., Wang, R. & Liu, J. Protein Assembly: Versatile Approaches to Construct Highly Ordered Nanostructures. *Chemical Reviews* vol. 116 13571–13632 (2016).

10.   Sugimoto, K., Kanamaru, S., Iwasaki, K., Arisaka, F. & Yamashita, I. Construction of a Ball-and-Spike Protein Supramolecule. *Angewandte Chemie International Edition* **45**, 2725–2728 (2006).

11.   Huang, Y. *et al.* Programmable assembly of nanoarchitectures using genetically engineered viruses. *Nano Letters* **5**, 1429–1434 (2005).

12.   Soto, C. M. & Ratna, B. R. Virus hybrids as nanomaterials for biotechnology. *Current Opinion in Biotechnology* vol. 21 426–438 (2010).

13.   Holder, P. G. *et al.* Dramatic thermal stability of virus-polymer conjugates in hydrophobic solvents. *Langmuir* **26**, 17383–17388 (2010).

14.   Jackson, K., Peivandi, A., Fogal, M., Tian, L. & Hosseinidoust, Z. Filamentous Phages as Building Blocks for Bioactive Hydrogels. *ACS Applied Bio Materials* **4**, 2262–2273 (2021).

15.   Cao, B. *et al.* Bacteriophage-based biomaterials for tissue regeneration. *Adv Drug Deliv Rev* **145**, 73 (2019).

16.   Hoffmann, S. *et al.* Competitively selected protein ligands pay their increase in specificity by a decrease in affinity. *Molecular BioSystems* **6**, 126–133 (2009).

17.   Zhang, J. *et al.* Phage-Derived Fully Human Antibody scFv Fragment Directed Against Human Vascular Endothelial Growth Factor Receptor 2 Blocked Its Interaction with VEGF. (2012).

18.   Wiebe, J., Kramer, K. & Skerra, A. A System for Repertoire Cloning and Phage Display of Murine and Leporid Antibody Fragments Drug Development of PASylated Interleukin-1 Receptor Antagonist View project Enzyme engineering View project. *Article in Journal of AOAC International* (2010.

19.   Pires, D. P., Cleto, S., Sillankorva, S., Azeredo, J. & Lu, T. K. Genetically Engineered Phages: a Review of Advances over the Last Decade. *Microbiology and Molecular Biology Reviews* **80**, 523–543 (2016).

20.   Briggs, B. D. & Knecht, M. R. Nanotechnology meets biology: Peptide-based methods for the fabrication of functional materials. *Journal of Physical Chemistry Letters* vol. 3 405–418 (2012).

21.   Love, A. J. *et al.* A genetically modified tobacco mosaic virus that can produce gold nanoparticles from a metal salt precursor. *Frontiers in Plant Science* **6**, (2015).

22.   Rajagopala, S. v., Casjens, S. & Uetz, P. The protein interaction map of bacteriophage lambda. *BMC Microbiology* **11**, (2011).

23.   Katsura, I. & Hendrixt, R. W. Length Determination in Bacteriophage Lambda Tails. *Cell* vol. 39 (1984).

24.   Katsura, I. Mechanism of Length Determination In Bacteriophage Lambda Tails. *Adv. Biophys* vol. 26 (1990).

25.   Xu, J., Hendrix, R. W. & Duda, R. L. Chaperone-protein interactions that mediate assembly of the bacteriophage lambda tail to the correct length. *Journal of Molecular Biology* **426**, 1004–1018 (2014).

26.   Xu, J., Hendrix, R. W. & Duda, R. L. A balanced ratio of proteins from gene gand frameshift-extended gene GT is required for phage lambda tail assembly. *Journal of Molecular Biology* **425**, 3476–3487 (2013).

27.   Pell, L. G. *et al.* The X-Ray Crystal Structure of the Phage λ Tail Terminator Protein Reveals the Biologically Relevant Hexameric Ring Structure and Demonstrates a Conserved Mechanism of Tail Termination among Diverse Long-Tailed Phages. *Journal of Molecular Biology* **389**, 938–951 (2009).

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

29.   Green, A. A. *et al.* Complex cellular logic computation using ribocomputing devices. *Nature* **548**, 117–121 (2017).

30.   Kim, J. *et al.* De novo-designed translation-repressing riboregulators for multi-input cellular logic. *Nature Chemical Biology* **15**, 1173–1182 (2019).

31.   Hong, F. *et al.* Precise and Programmable Detection of Mutations Using Ultraspecific Riboregulators. *Cell* **180**, 1018-1032.e16 (2020).

32.   Minuesa, G., Alsina, C., Garcia-Martin, J. A., Oliveros, J. C. & Dotu, I. MoiRNAiFold: A novel tool for complex in silico RNA design. *Nucleic Acids Research* **49**, 4934–4943 (2021).

33.   Wolfe, B. R. & Pierce, N. A. Sequence Design for a Test Tube of Interacting Nucleic Acid Strands. *ACS Synthetic Biology* **4**, 1086–1100 (2015).

34.   Wolfe, B. R., Porubsky, N. J., Zadeh, J. N., Dirks, R. M. & Pierce, N. A. Constrained Multistate Sequence Design for Nucleic Acid Reaction Pathway Engineering. *J Am Chem Soc* **139**, 3134–3144 (2017).

35.   Zadeh, J. N., Wolfe, B. R. & Pierce, N. A. Nucleic acid sequence design via efficient ensemble defect optimization. *Journal of Computational Chemistry* **32**, 439–452 (2011).

36.   Dirks, R. M., Lin, M., Winfree, E. & Pierce, N. A. Paradigms for computational nucleic acid design. *Nucleic Acids Research* **32**, 1392–1403 (2004).

37.   Dotu, I. *et al.* Complete RNA inverse folding: computational design of functional hammerhead ribozymes. *Nucleic Acids Research* **42**, 11752–11762 (2014).

38.   Green, A. A. *et al.* Complex cellular logic computation using ribocomputing devices. *Nature* **548**, 117–121 (2017).

39.   Ma, D., Shen, L., Wu, K., Diehnelt, C. W. & Green, A. A. Low-cost detection of norovirus using paper-based cell-free systems and synbody-based viral enrichment. *Synthetic Biology* **3**, (2018).

40.   Pardee, K. *et al.* Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell* **165**, 1255–1266 (2016).

41.   Sadat Mousavi, P. *et al.* A multiplexed, electrochemical interface for gene-circuit-based sensors. *Nature Chemistry* **12**, 48–55 (2020).

42.   Hong, F. *et al.* Precise and Programmable Detection of Mutations Using Ultraspecific Riboregulators. *Cell* **180**, 1018-1032.e16 (2020).

43.   Ellington, A. D. & Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature 1990 346:6287* **346**, 818–822 (1990).

44.   Paige, J. S., Nguyen-Duc, T., Song, W. & Jaffrey, S. R. Fluorescence imaging of cellular metabolites with RNA. *Science* **335**, 1194 (2012).

45.   Wu, J. *et al.* Live imaging of mRNA using RNA-stabilized fluorogenic proteins. *Nat Methods* **16**, 862–865 (2019).

46.   Dolgosheina, E. v. *et al.* RNA Mango aptamer-fluorophore: A bright, high-affinity complex for RNA labeling and tracking. *ACS Chemical Biology* **9**, 2412–2420 (2014).

47.   Nakatsuka, N. *et al.* Aptamer–field-effect transistors overcome Debye length limitations for small-molecule sensing. *Science (1979)* **362**, 319–324 (2018).

48.   Baker, B. R. *et al.* An electronic, aptamer-based small-molecule sensor for the rapid, label-free detection of cocaine in adulterated samples and biological fluids. *J Am Chem Soc* **128**, 3138–3139 (2006).

49.   Wrist, A., Sun, W. & Summers, R. M. The Theophylline Aptamer: 25 Years as an Important Tool in Cellular Engineering Research. *ACS Synthetic Biology* vol. 9 682–697 (2020).

50.   Liu, S., Xu, Y., Jiang, X., Tan, H. & Ying, B. Translation of aptamers toward clinical diagnosis and commercialization. *Biosensors and Bioelectronics* **208**, 114168 (2022).

51.   Osborne, S. E., Matsumura, I. & Ellington, A. D. Aptamers as therapeutic and diagnostic reagents: problems and prospects. *Current Opinion in Chemical Biology* **1**, 5–9 (1997).

52.   Marnissi, B., Kamali-Moghaddam, M., Ghram, A. & Hmila, I. Generation of ssDNA aptamers as diagnostic tool for Newcastle avian virus. *PLOS ONE* **15**, e0237253 (2020).

53.   Weber, A. M. *et al.* A blue light receptor that mediates RNA binding and translational regulation. *Nature Chemical Biology* **15**, 1085–1092 (2019).

Chapter 2 References

1.   Iannolo, G., Minenkova, O., Petruzzelli, R. & Cesareni, G. *Modifying Filamentous Phage Capsid: Limits in the Size of the Major Capsid Protein. J. Mol. Biol* vol. 248 (1995).

2.   Katsura, I. & Hendrixt, R. W. Length Determination in Bacteriophage Lambda Tails. *Cell* vol. 39 (1984).

3.   Hoffmann, S. *et al.* Competitively selected protein ligands pay their increase in specificity by a decrease in affinity. *Molecular BioSystems* **6**, 126–133 (2009).

4.   Zhang, J. *et al.* Phage-Derived Fully Human Antibody scFv Fragment Directed Against Human Vascular Endothelial Growth Factor Receptor 2 Blocked Its Interaction with VEGF. (2012) doi:10.1002/btpr.1559.

5.   Nam, K. T. *et al.* Virus-Enabled Synthesis and Assembly of Nanowires for Lithium Ion Battery Electrodes. *Science* **312**, 885-888 (2008).

6.   Atanasova, P. *et al.* Virus-templated synthesis of ZnO nanostructures and formation of field-effect transistors. *Advanced Materials* **23**, 4918–4922 (2011).

7.   Gerasopoulos, K., McCarthy, M., Royston, E., Culver, J. N. & Ghodssi, R. Microbatteries With Tobaco Mosaic Virus Templated Electrodes. in *2008 IEEE 21st International Conference on Micro Electro Mechanical Systems*.

8.   Nam, K. T., Reelle, B. R., Lee, S. W. & Belcher, A. M. Genetically Driven Assembly of Nanorings Based on the M13 Virus. *Nano Letters* **4**, 23–27 (2004).

9.   Jackson, K., Peivandi, A., Fogal, M., Tian, L. & Hosseinidoust, Z. Filamentous Phages as Building Blocks for Bioactive Hydrogels. *ACS Applied Bio Materials* **4**, 2262–2273 (2021).

10.   Cao, B. *et al.* Bacteriophage-based biomaterials for tissue regeneration. *Adv Drug Deliv Rev* **145**, 73 (2019).

11.   Luo, Q., Hou, C., Bai, Y., Wang, R. & Liu, J. Protein Assembly: Versatile Approaches to Construct Highly Ordered Nanostructures. *Chemical Reviews* vol. 116 13571–13632 (2016).

12.   Sugimoto, K., Kanamaru, S., Iwasaki, K., Arisaka, F. & Yamashita, I. Construction of a Ball-and-Spike Protein Supramolecule. *Angewandte Chemie International Edition* **45**, 2725–2728 (2006).

13.   Huang, Y. *et al.* Programmable assembly of nanoarchitectures using genetically engineered viruses. *Nano Letters* **5**, 1429–1434 (2005).

14.   Soto, C. M. & Ratna, B. R. Virus hybrids as nanomaterials for biotechnology. *Current Opinion in Biotechnology* vol. 21 426–438 (2010).

15.   Holder, P. G. *et al.* Dramatic thermal stability of virus-polymer conjugates in hydrophobic solvents. *Langmuir* **26**, 17383–17388 (2010).

16.   Xu, J., Hendrix, R. W. & Duda, R. L. Chaperone-protein interactions that mediate assembly of the bacteriophage lambda tail to the correct length. *Journal of Molecular Biology* **426**, 1004–1018 (2014).

17.   Smith, G. P. Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science (1979)* **228**, 1315–1317 (1985).

18.   Keen, E. C. A century of phage research: Bacteriophages and the shaping of modern biology. *BioEssays* **37**, 6–9 (2015).

19.   Kamiyama, D. *et al.* Versatile protein tagging in cells with split fluorescent protein. *Nature Communications* **7**, (2016).

20.   Kakimoto, Y. *et al.* Visualizing multiple inter-organelle contact sites using the organelle-Targeted split-GFP system. *Scientific Reports* **8**, (2018).

21. Feng, S. *et al.* Improved split fluorescent proteins for endogenous protein labeling. *Nature Communications* **8**, (2017).

22. Katsura, I. Mechanism of Length Determination In Bacteriophage Lambda Tails. *Adv. Biophys* vol. 26 (1990).

23. Dickerson, M. B., Sandhage, K. H. & Naik, R. R. Protein- and peptide-directed syntheses of inorganic materials. *Chemical Reviews* **108**, 4935–4978 (2008).

24. Briggs, B. D. & Knecht, M. R. Nanotechnology meets biology: Peptide-based methods for the fabrication of functional materials. *Journal of Physical Chemistry Letters* vol. 3 405–418 (2012).

25. Kožíšek, M. *et al.* Molecular design of specific metal-binding peptide sequences from protein fragments: Theory and experiment. *Chemistry - A European Journal* **14**, 7836–7846 (2008).

26. Satofuka, H., Fukui, T., Takagi, M., Atomi, H. & Imanaka, T. *Metal-binding properties of phytochelatin-related peptides*. *Journal of Inorganic Biochemistry* vol. 86 www.elsevier.com/locate/jinorgbio (2001).

27. Park, T. J., Lee, S. Y., Heo, N. S. & Seo, T. S. In vivo synthesis of diverse metal nanoparticles by recombinant escherichia coli. *Angewandte Chemie - International Edition* **49**, 7019–7024 (2010).

28. Huang, J. *et al.* Bio-inspired synthesis of metal nanomaterials and applications. *Chemical Society Reviews* **44**, 6330–6374 (2015).

29. Lee, J. Y. *et al.* Assembly of collagen-binding peptide with collagen as a bioactive scaffold for osteogenesis in vitro and in vivo. *Biomaterials* **28**, 4257–4267 (2007).

30. Love, A. J. *et al.* A genetically modified tobacco mosaic virus that can produce gold nanoparticles from a metal salt precursor. *Frontiers in Plant Science* **6**, (2015).

31. Xu, J., Hendrix, R. W. & Duda, R. L. A balanced ratio of proteins from gene gand frameshift-extended gene GT is required for phage lambda tail assembly. *Journal of Molecular Biology* **425**, 3476–3487 (2013).

Chapter 3 References

1.   Wrist, A., Sun, W. & Summers, R. M. The Theophylline Aptamer: 25 Years as an Important Tool in Cellular Engineering Research. *ACS Synthetic Biology* vol. 9 682–697 (2020).

2.   Sharma, V., Nomura, Y. & Yokobayashi, Y. Engineering complex riboswitch regulation by dual genetic selection. *J Am Chem Soc* **130**, 16310–16315 (2008).

3.   Climent-Catala, A., Ouldridge, T. E., Stan, G. B. v. & Bae, W. Building an RNA-Based Toggle Switch Using Inhibitory RNA Aptamers. *ACS Synthetic Biology* **11**, 562–569 (2022).

4.   Tinafar, A. *et al.* Cell-Free Biosensors: Synthetic Biology Without Borders 12. (2022) doi:10.1007/978-3-030-23217-7_130.

5.   Pardee, K. *et al.* Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell* **165**, 1255–1266 (2016).

6.   Ma, D., Shen, L., Wu, K., Diehnelt, C. W. & Green, A. A. Low-cost detection of norovirus using paper-based cell-free systems and synbody-based viral enrichment. *Synthetic Biology* **3**, (2018).

7.   Hong, F. *et al.* Precise and Programmable Detection of Mutations Using Ultraspecific Riboregulators. *Cell* **180**, 1018-1032.e16 (2020).

8.   Weinberg, B. H. *et al.* Large-scale design of robust genetic circuits with multiple inputs and outputs for mammalian cells. *Nature Biotechnology* **35**, 453–462 (2017).

9.   Grieger, J. C. & Samulski, R. J. Packaging Capacity of Adeno-Associated Virus Serotypes: Impact of Larger Genomes on Infectivity and Postentry Steps. *Journal of Virology* **79**, 9933–9944 (2005).

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

11.   Green, A. A. *et al.* Complex cellular logic computation using ribocomputing devices. *Nature* **548**, 117–121 (2017).

12.   Kim, J. *et al.* De novo-designed translation-repressing riboregulators for multi-input cellular logic. *Nature Chemical Biology* **15**, 1173–1182 (2019).

13.   Castro, N. J., O'Brien, J. & Zhang, L. G. Integrating biologically inspired nanomaterials and table-top stereolithography for 3D printed biomimetic osteochondral scaffolds. *Nanoscale* **7**, 14010–14022 (2015).

14.   Weber, A. M. *et al.* A blue light receptor that mediates RNA binding and translational regulation. *Nature Chemical Biology* **15**, 1085–1092 (2019).

15. Li, X. *et al.* Imaging Intracellular S-Adenosyl Methionine Dynamics in Live Mammalian Cells with a Genetically Encoded Red Fluorescent RNA-Based Sensor. *J Am Chem Soc* **142**, 14117–14124 (2020).

16. Litke, J. L., You, M. & Jaffrey, S. R. Developing Fluorogenic Riboswitches for Imaging Metabolite Concentration Dynamics in Bacterial Cells. in *Methods in Enzymology* vol. 572 315–333 (Academic Press Inc., 2016).

17. Filonov, G. S., Moon, J. D., Svensen, N. & Jaffrey, S. R. Broccoli: Rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J Am Chem Soc* **136**, 16299–16308 (2014).

18. Wolfe, B. R., Porubsky, N. J., Zadeh, J. N., Dirks, R. M. & Pierce, N. A. Constrained Multistate Sequence Design for Nucleic Acid Reaction Pathway Engineering. *J Am Chem Soc* **139**, 3134–3144 (2017).

19. Zadeh, J. N., Wolfe, B. R. & Pierce, N. A. Nucleic acid sequence design via efficient ensemble defect optimization. *Journal of Computational Chemistry* **32**, 439–452 (2011).

20. Wolfe, B. R. & Pierce, N. A. Sequence Design for a Test Tube of Interacting Nucleic Acid Strands. *ACS Synthetic Biology* **4**, 1086–1100 (2015).

21. Dirks, R. M., Lin, M., Winfree, E. & Pierce, N. A. Paradigms for computational nucleic acid design. *Nucleic Acids Research* **32**, 1392–1403 (2004).

22. Andersen, J. B. *et al.* New Unstable Variants of Green Fluorescent Protein for Studies of Transient Gene Expression in Bacteria. *Applied and Environmental Microbiology* **64**, 2240 (1998).

23. You, M., Litke, J. L. & Jaffrey, S. R. Imaging metabolite dynamics in living cells using a Spinach-based riboswitch. *Proceedings of the National Academy of Sciences of the United States of America* vol. 112 E2756–E2765 (2015).

24. Paige, J. S., Nguyen-Duc, T., Song, W. & Jaffrey, S. R. Fluorescence imaging of cellular metabolites with RNA. *Science* **335,** 1194 (2012).

25. New England Biolabs. PURExpress FAQs. https://www.neb.com/-/media/nebus/files/pdf-faq/purexpress-faqs.pdf.

26. Huang, H. *et al.* A G-Quadruplex-Containing RNA Activates Fluorescence in a GFP-Like Fluorophore. *Nat Chem Biol* **10**, 686 (2014).

27. Oliveira, R. *et al.* Modelling aptamers with nucleic acid mimics (NAM): From sequence to three-dimensional docking. *PLOS ONE* **17**, e0264701 (2022).

28. Buglak, A. A., Samokhvalov, A. v., Zherdev, A. v. & Dzantiev, B. B. Methods and applications of in silico aptamer design and modeling. *International Journal of Molecular Sciences* vol. 21 1–25 (2020).

29.   Li, P. T. X., Vieregg, J. & Tinoco, I. How RNA unfolds and refolds. *Annual Review of Biochemistry* vol. 77 77–100 (2008).

30.   Liu, R. *et al.* Optogenetic control of RNA function and metabolism using engineered light-switchable RNA-binding proteins. *Nature Biotechnology* **40**, 779-786 (2022).

APPENDIX A

SUPPLEMENTAL INFORMATION TO CHAPTER 2: RE-IMAGINING LEGACY
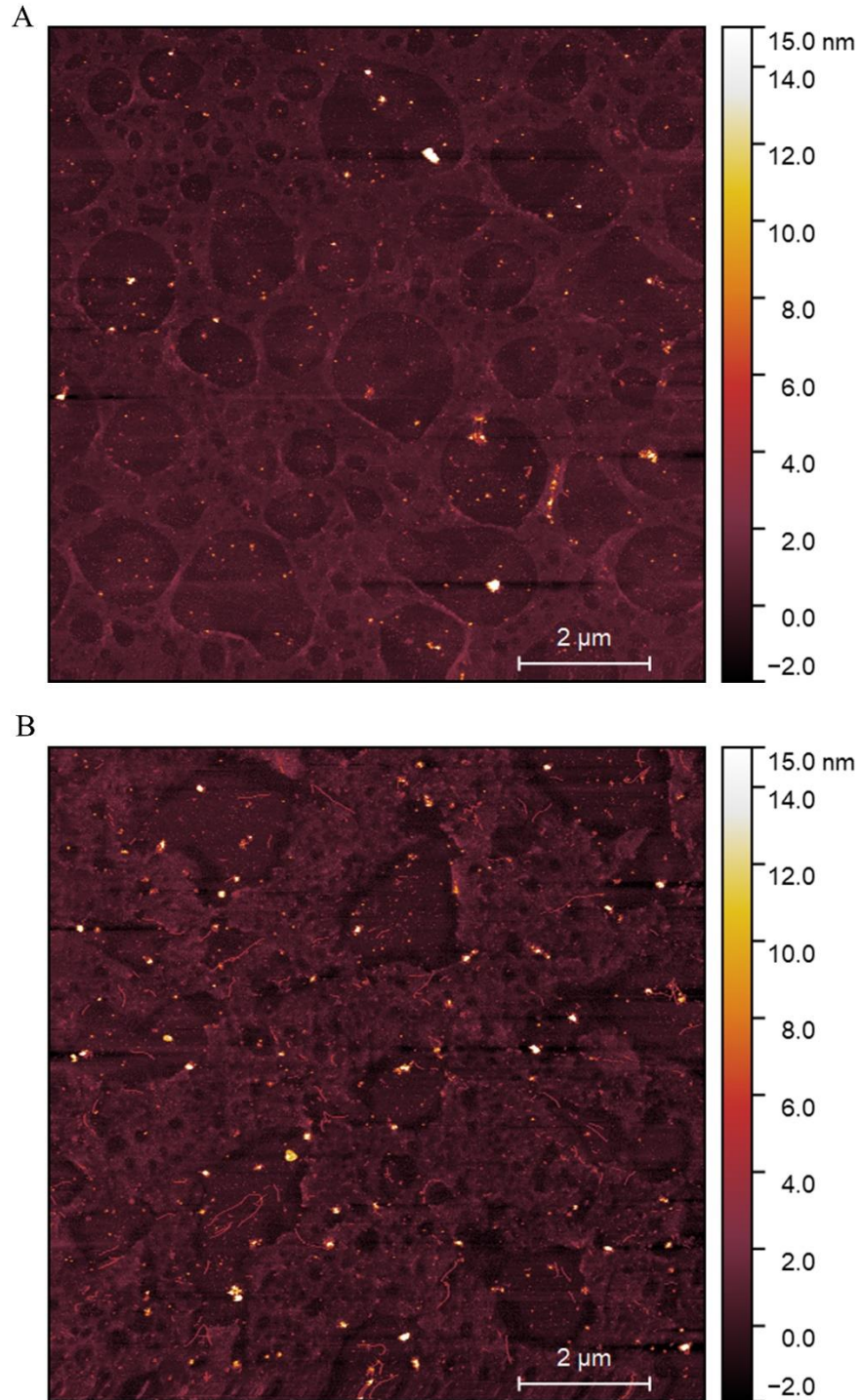
PHAGES FOR VIRAL SELF ASSEMBLY

Figure S2.1 Validation of gpV Knockout. A) AFM imaging of purified product from expression of only pλΔgpV plasmid. B) AFM imaging of tail expression with pλΔgpV and the pλ-gpVonly rescue plasmid. The successful tails seen in (B) exhibit a distribution of length into the µm range, whereas a standard tail is 150 nm on average.

Table S2.1 Amino Acid Sequences of Tested Peptides

| Peptide ID	AA sequence
Ge34	TGHQSPGAYAAH
Tetraglutamate	EEEE
dTiRRK	RKKRKKRKKRKKGGGW
Ge8	SLKMPHWPHLLP
KDTK	KTEYVDERSKSLTVDLTK
EC7G	ECECECECECECECG
MT_AGABI	GDCGCSGASSCTCASGQCTCSGCGK
Ferritin	MLSERMLKALNDQLNRELYSAYLYFAMAAYFEDLGLE GFANWMKAQAEEEIGHALRFYNYIYDRNGRVELDEIPK PPKEWESPLKAFEAAYEHEKFISKSIYELAALAEEEKDY STRAFLEWFINEQVEEEASVKKILDKLKFAKDSPQILFM LDKELSARAPKLPGLLMQGGE
OPN_S1	GLRSKSKKFRRPDIQYPDATDEDITSHM
MT3_Mouse	MDPETCPCPT GGSCTCSDKC KCKGCKCTNC KKSCCSCCPA GCEKCAKDCV CKGEEGAKAE AEKCSCCQ
MT4B_ARATH	MADTGKGSASASCNDRCGCPSPCPGGESCRCKMMSEA SGGDQEHNTCPCGEHCGCNPCNCPKTQTQTSAKGCTCG EGCTCATCAA
TBP1_mini	RKLPDA
HHTC	HNLGMNHDLQGERPYVTEGC
DC7G	DCDCDCDCDCDCDCG
PCS1_ARATH	MAMASLYRRSLPSPPAIDFSSAEGKLIFNEALQKGTMEG FFRLISYFQTQSEPAYCGLASLSVVLNALSIDPGRKWKG PWRWFDESMLDCCEPLEVVKEKGISFGKVVCLAHCSG AKVEAFRTSQSTIDDFRKFVVKCTSSENCHMISTYHRGV FKQTGTGHFSPIGGYNAERDMALILDVARFKYPPHWVP LKLLWEAMDSIDQSTGKRRGFMLISRPHREPGLLYTLSC KDESWIEIAKYLKEDVPRLVSSQHVDSVEKIISVVFKSLP SNFNQFIRWVAEIRITEDSNQNLSAEEKSRLKLKQLVLK EVHETELFKHINKFLSTVGYEDSLTYAAAKACCQGAEIL SGSPSKEFCCRETCVKCIKGPDDSEGTVVTGVVVRDGN EQKVDLLVPSTQTECECGPEATYPAGNDVFTALLLALPP QTWSGIKDQALMHEMKQLISMASLPTLLQEEVLHLRRQ LQLLKRCQENKEEDDLAAPAY
Gold Binding Motif	LKAHLPPSRLPS
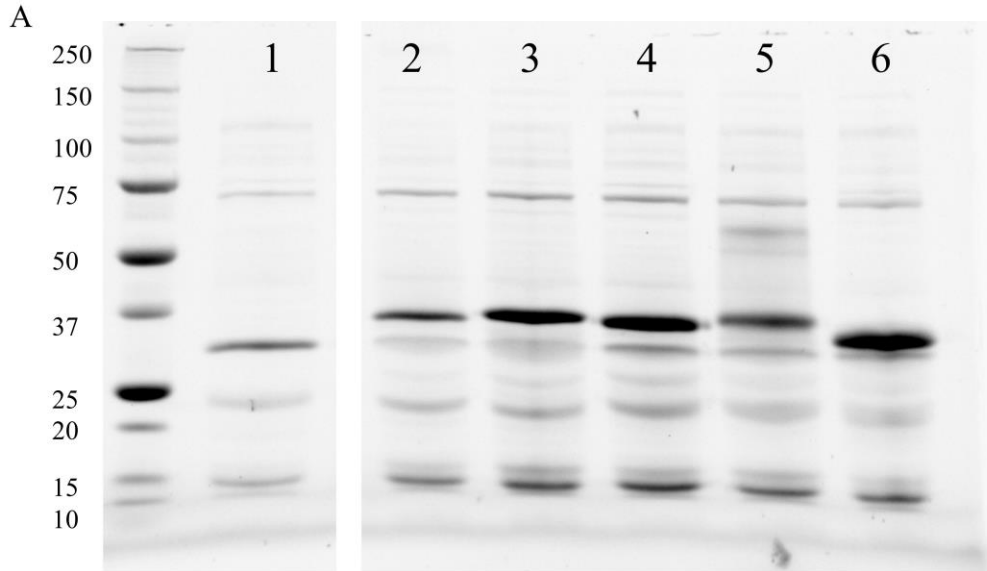
Figure S2.2 Gel Validation of λ-gpVonly Incorporation. A) SDS-PAGE Gel of Lambda Tails. To verify the presence of gpV containing tail fusions, successfully formed tails were also run on TGX Stain-Free™ FastCast™ Acrylamide 10% gels (BioRad, 1610183) in Mini-PROTEAN gel electrophoresis chamber (BioRad, 1658001FC). The leftmost lane contains a Precision Plus Protein Dual Color Standards protein ladder (BioRad, 1610374). Lanes one through six contain PEG-precipitated tails created by expressing the pλ plasmid alongside modified or unmodified pλ-gpVonly plasmid containing larger protein fusions to the gpV protein. The expected mass of unmodified gpV proteins is 25.81 kilodaltons. Samples are as follows: 1) Unmodified control tails, 2) gpV-KDTK peptide fusion, 3) gpV-HHTC peptide fusion, 4) gpV-EAK peptide fusion, 5) gpV-MT_AGABI metallothionine protein fusion, and 5) gpV-tetraglutamate peptide fusion. Standard molecular weight tails can be seen in addition to the modified tails on the gel. This suggests that in cases where major coat protein modification disrupts tail formation, sometimes tails can still be formed when supplementary unmodified major coat protein is present.
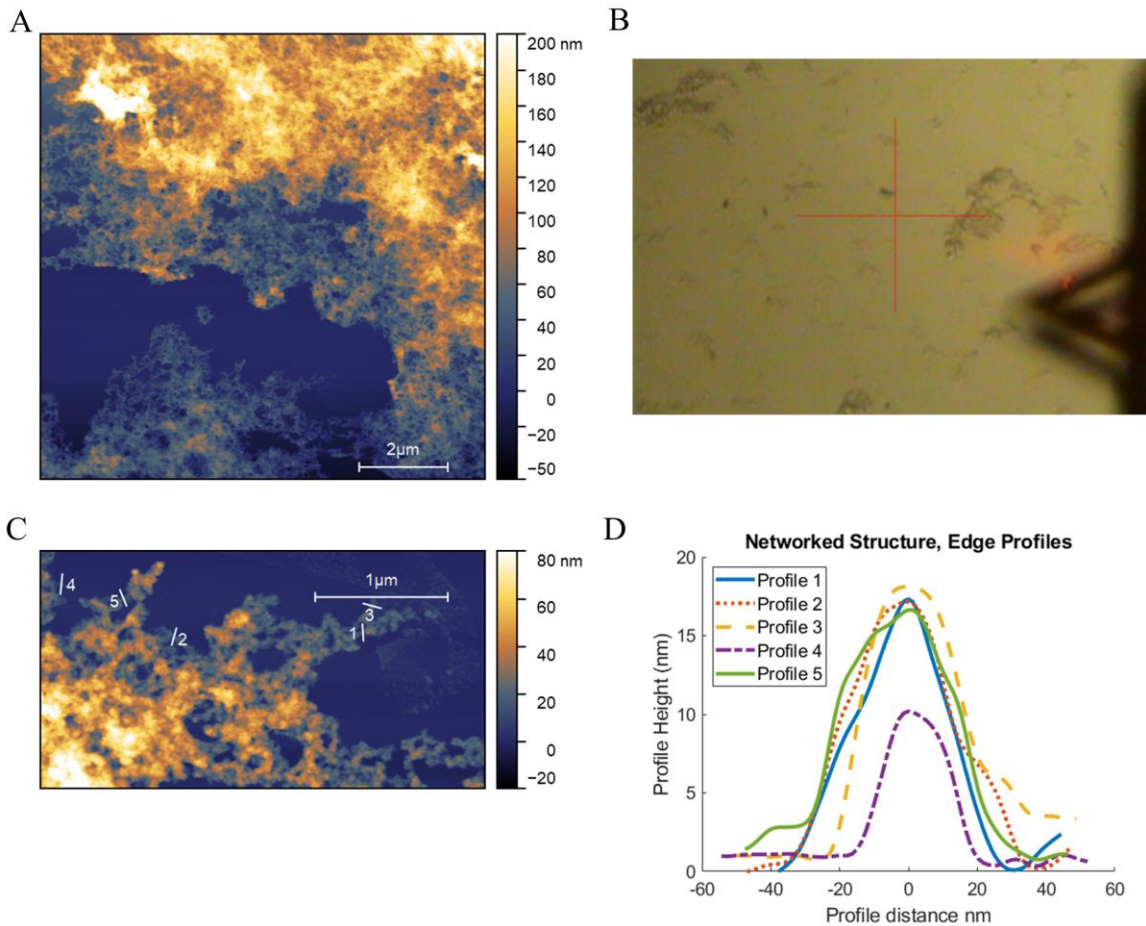
Figure S2.3 Networked Crystalline Tails. Incubation of lambda tails displaying the tetraglutamate peptide with 20 mM CoCl2 and reduction on ice with NaBH4 led to the formation of networked precipitates as seen under AFM in (A) and (C). (B) Picture of mica surface from AFM guide camera. Visual inspection of the sample as deposited on mica slides for AFM shows clearly visible precipitate. The scans taken are from areas adjacent to visible precipitate. C) Enhanced view of (A). A closer look at the edge of these networked precipitates showed more loosely connected, potentially porous structures. D) Profilometry of the regions indicated in (C). As the standard tail diameter registers as 6 to 10 nm, and disperse tails formations show an average thickness of 8-14 nm, the 10-17 nm range shown at the edge of these precipitates suggests that they are a densely interconnected mesh of tails that were linked during the reduction of the metal ions.
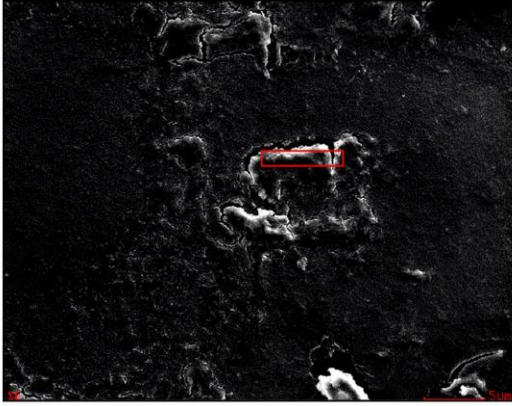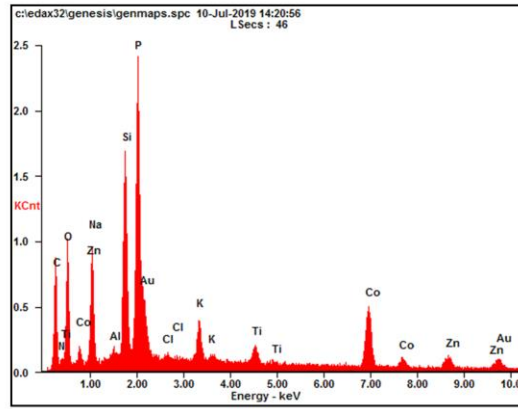
Figure S2.4 SEM-EDAX Analysis of dTIRRK. A) SEM visual display. Elemental analysis was conducted of lambda phage tails expressing the dTiRRK peptide that were reacted with CoCl2. Comparing the tails to unreacted and standard tails showed shifts in the cobalt peak, indicating bound cobalt with the tails. The red highlighting box indicates the scan position, which is on top of an aggregate of tail structures. B) EDAX spectral readout of dTiRRK-modified tails reduced with $CoCl_2$.

APPENDIX B

SUPPLEMENTAL INFORMATION TO CHAPTER 3: INTEGRATION OF

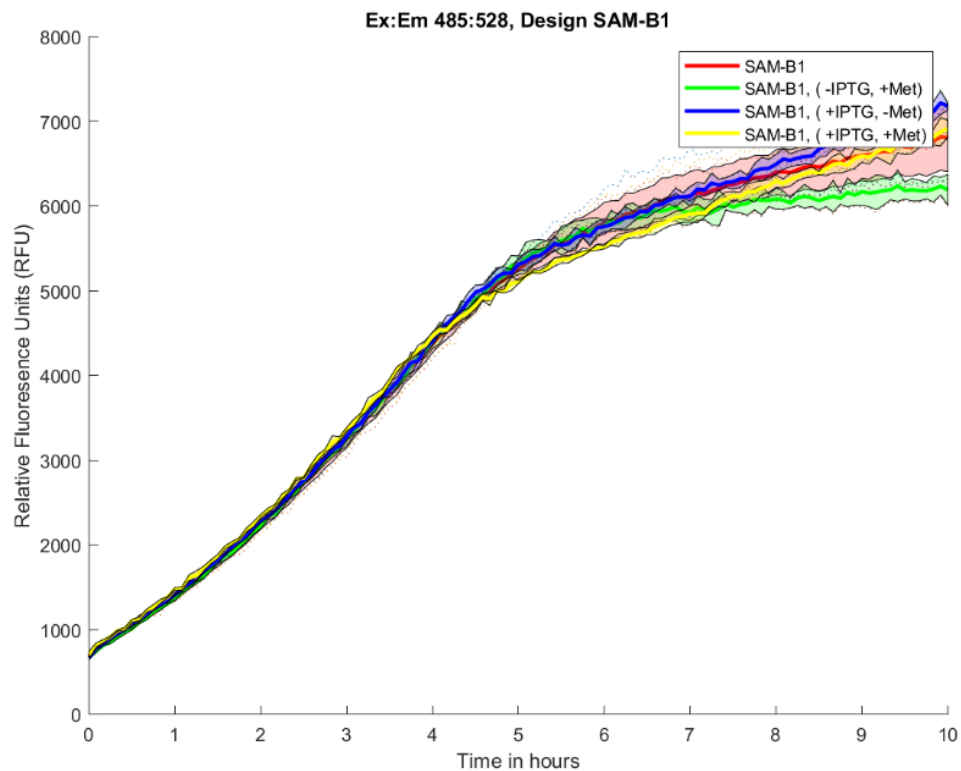CHEMICAL AND OPTICAL SENSING INTO TOEHOLD SWITCH-LIKE

RIBODEVICES

Figure S3.1 In Vivo Testing of SAM SAS Devices. Sample data of SAM-SAS performance *in vivo*. The design displayed displayed reactivity to SAM concentrations in cell free conditions, but *in vivo* conditions showed no difference in expression with or without IPTG induction or addition of excess methionine to induce greater SAM levels.
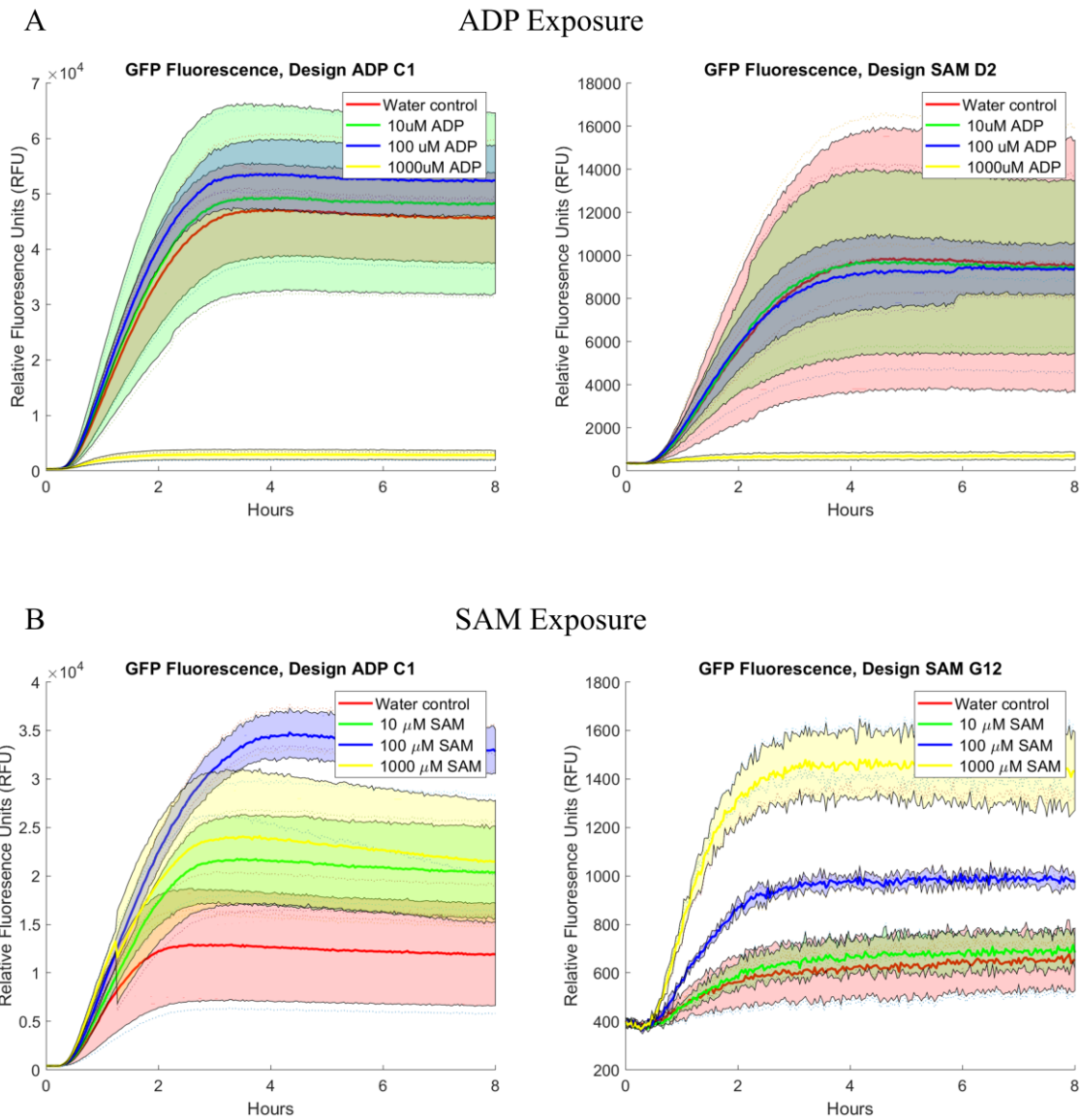
Figure S3.2 ADP Inhibitory Effect at High Concentrations. The impact of ADP and SAM metabolites on each other's SAS devices was tested. As (A) demonstrates, high concentrations of ADP can cause an inhibitory effect on switches designed for SAM targets. SAM can have a disruptive effect on some ADP aptamers, as seen in (B), but its effect is not a universally depressive one, as the positive activation of SAM-SAS G12 shows.
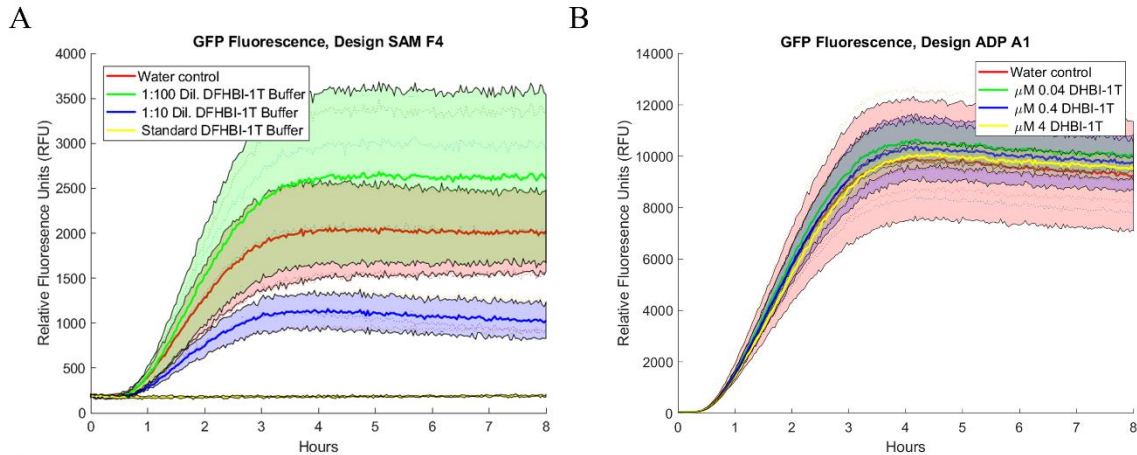
Figure S3.3 Impact of DFHBI-1T Buffer on Plasmid Expression. A) The SAM-SAS
F4 device was shown to be unresponsive to SAM in prior testing and was used as a
control for testing the impact of DFHBI-1T buffer (40 mM HEPES, 100mM NaCl, 1mM
MgCl$_2$) on the PureExpress system. The standard concentration of the DFHBI-1T buffer
cause over 100-fold inhibition of GFP expression in the PURExpress system. Based on
these results, a 1:10 dilution of the buffer (4 mM HEPES, 10 mM NaCl, 100 µM MgCl$_2$)
was used for later DFHBI-1T to provide the magnesium ions for Broccoli fluorescence
without unduly disrupting the PURExpress components. B) Demonstration of SAS-ADP
A1 remaining unaffected by increasing DFHBI-1T concentrations when buffer
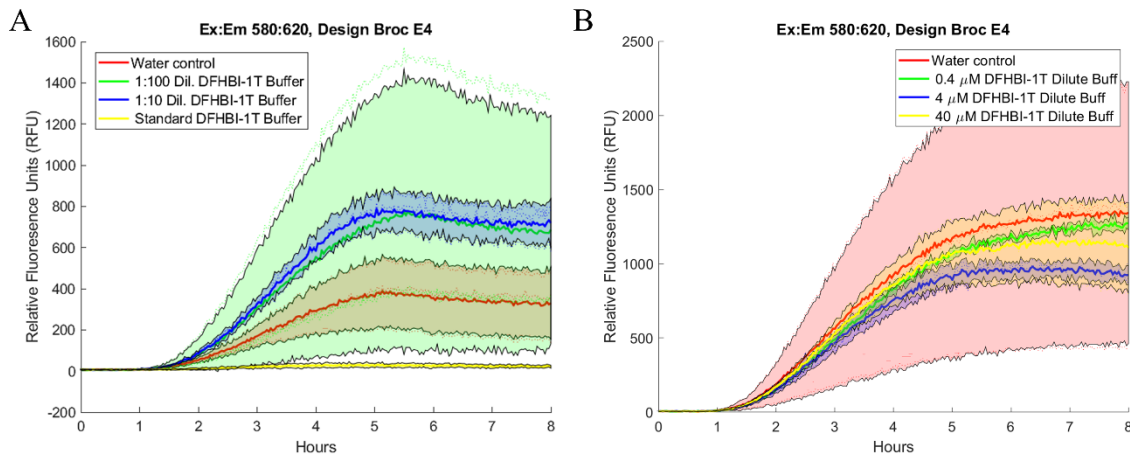concentration used at 1:10 of standard.

Figure S3.4 State Switching Caused by DFHBI-1T Buffer. A) Initial testing conducted
with non-constant DFHBI-1T concentration. During this testing, the DFHBI-1T stock
was diluted with water rather than additional buffer. The shift in buffer concentration,
along with the DFHBI-1T concentration shift, enacted activation of the switch and
increased fluorescence by 2-fold at 1:10 dilution, though the signal was fully repressed at
the standard concentration. B) Retesting with constant Buffer conditions at 1:10 dilution
demonstrated that the binding of DFHBI-1T and the broccoli aptamer is not sufficient
alone to enact switch activation.

76

A

Broccoli MFE at 37 °C
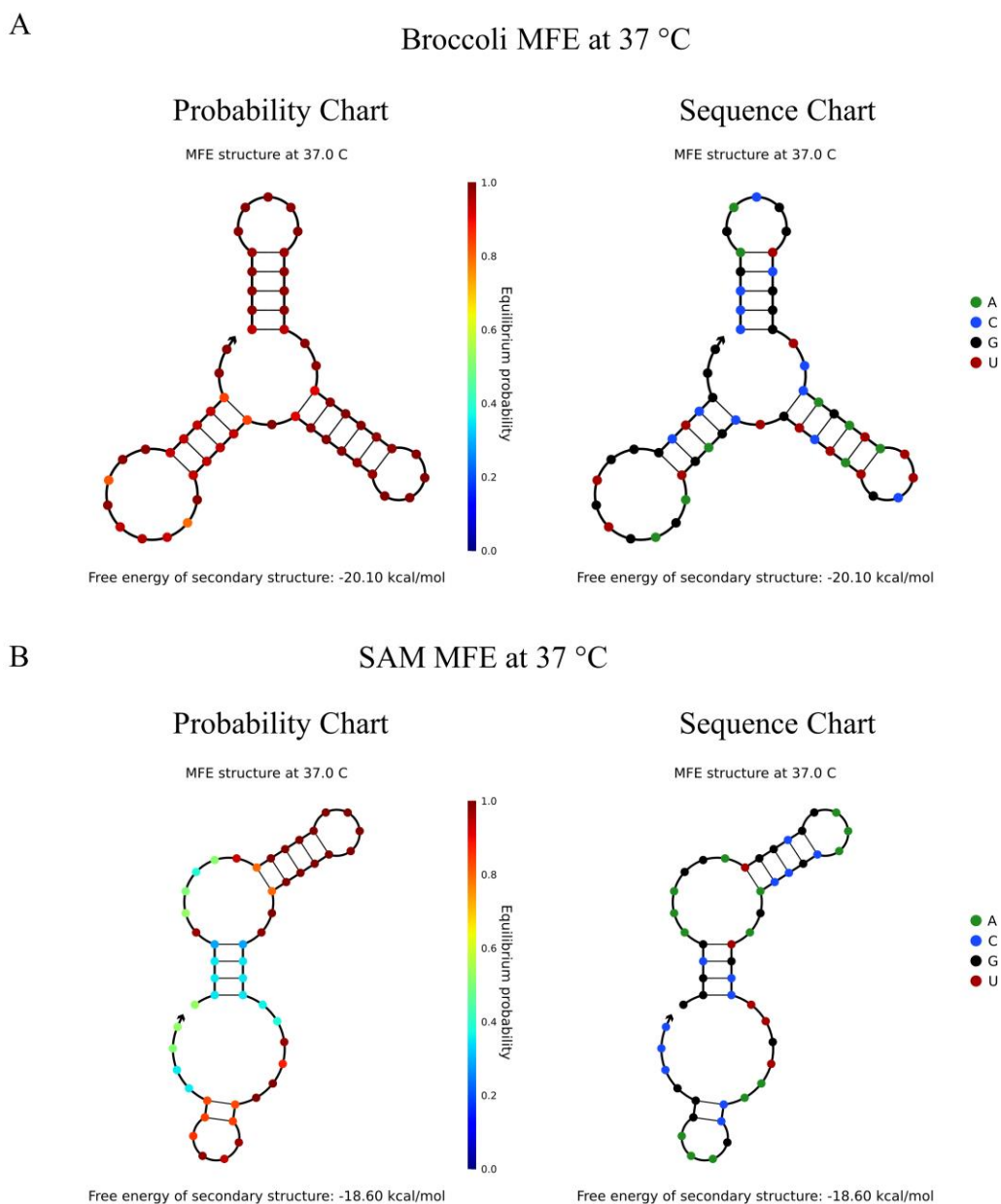
Probability Chart

Sequence Chart

Figure S3.5 Aptamer Site Minimum Free Energy Predictions. (A) Minimum Free energy of Broccoli RNA aptamer at 37°C according to NUPACK calculations. Pair probability and sequence identity color coding displayed. (B) Minimum Free energy of S-adenosyl-methionine RNA aptamer at 37°C according to NUPACK calculations. Pair probability and sequence identity color coding displayed.
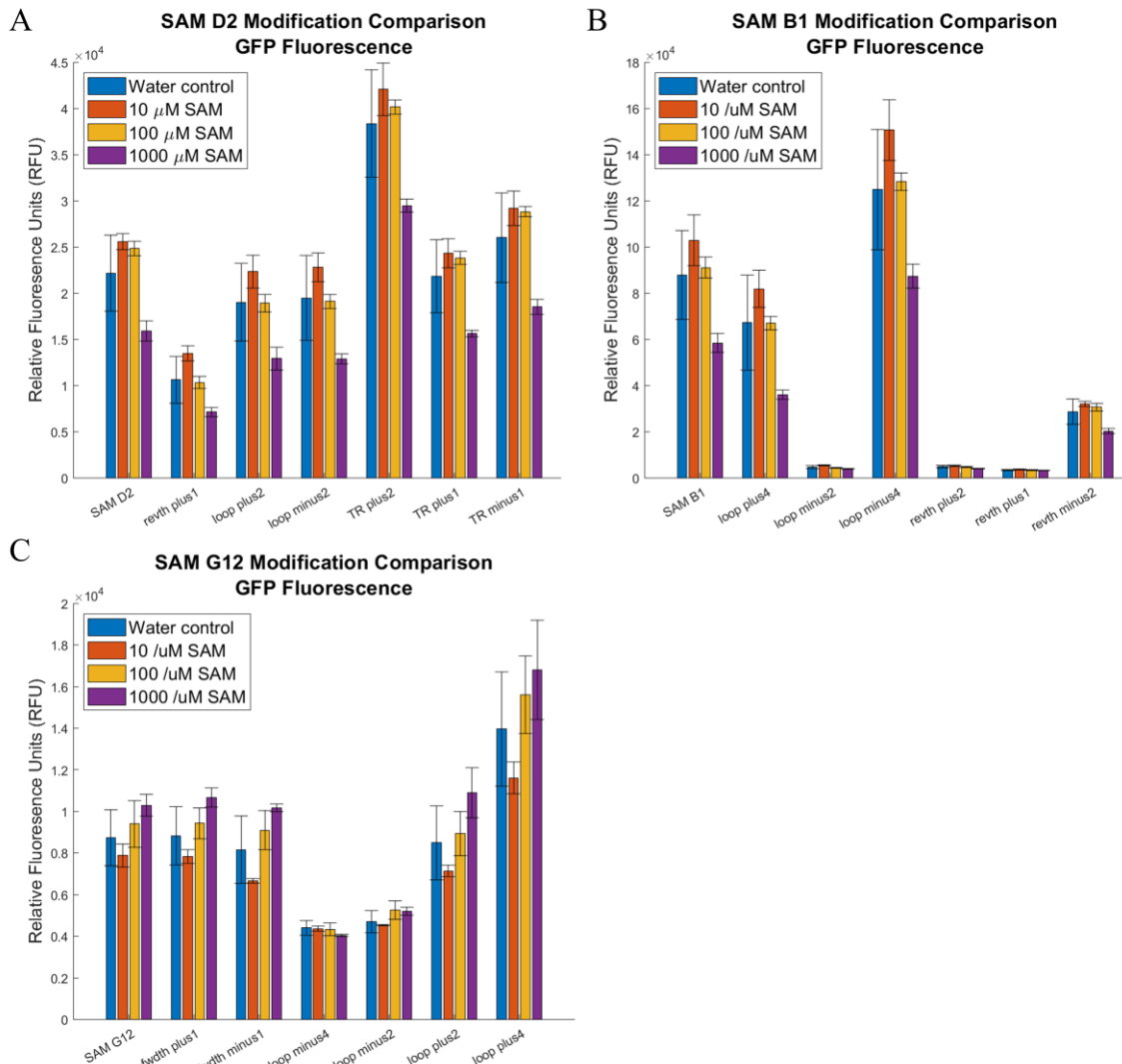
Figure S3.6 Piecewise Modification Results of SAM B1, SAM D2, and SAM G12.
Sample results displaying the shifts in ON:OFF ratio resulting from some of the structural
parameter modifications. A) Modification results for parameter modifications of SAM
D2 SAS switch. Notable modification effects the increase in fluorescence with 2 basepair
extension of transducer region (TR plus2), and increased sensitivity to the aptamer target
at 100 µM when increasing the reverse toehold region by one basepair (revth plus1) as
well as when increasing decreasing the aptamer site loop length by two nucleotides (loop
plus2, loop minus2). However, none of the modifications resulted in significant
improvements in the dynamic range. B) Modification results for parameter modifications
of SAM D2 SAS switch. Modifications to the aptamer loop site had unpredictable results,
with an increase of four nucleotides (loop plus4) improving the dynamic range by 29%,
while shortening the loop alternately broke the switch activity completely (loop minus2)
or boosted the fluorescent output without changing the dynamic range (loop plus4).
Unlike modifications to SAM D2, all changes in the reverse toehold region severely
decreased switch dynamic range and sensitivity. C) Modification results for parameter
modifications of SAM G12 SAS switch. Shortening the forward toehold region by one

basepair (fwdth minus1) yielded a 17% improvement in dynamic range. Aptamer loop site modifications continued to yield noticeable, though unpredictable shifts in behavior, with loop subtraction inhibiting switch function, and loop extension showing mild improvements (16%) to the dynamic range and sensitivity.

Table S3.1 Beta Coefficients of Multivariate Linear Model – SAM SAS Switches.
Table lists the calculated coefficient values of a multivariate linear model for the SAM SAS switches given their reporter fluorescence under low induction (no SAM addition) and high induction (100 µM SAM addition). Model assumes normalized values for the independent parameters. Calculated β_0 values were less than 10^{14}.

Beta-Coefficients of dependent variables	Reporter Fluorescence		
	Low Induction	High Induction	High:Low Reporter Ratio
β_0	0.000*	0.000*	0.000*
Forward Toehold	-1.244	-1.111	-0.046
Reverse Toehold	-0.214	-0.454	0.237
Transducer Length	0.415	0.514	0.033
Total Structural Defect	-0.067	-0.356	1.560
Defect of ON State	-0.011	0.208	-1.366
Defect of OFF State	1.045	1.063	-0.789
ΔG full sequence	3.394	2.751	0.375
ΔG Switch sequence	-3.302	-3.020	0.955
ΔG ON State	0.000	0.000	0.000
ΔG OFF State	0.718	1.132	-1.363
$\Delta\Delta$G ON vs OFF State	0.143	0.188	0.094
Forward Toehold GC content	0.772	0.531	0.115
Reverse Toehold GC Content	-1.047	-0.640	-0.471
Transducer Length GC Content	0.629	0.536	0.326
Migration Region GC Content	0.314	0.191	0.315

Table S3.2 P-values of Multivariate Linear Model Coefficients – SAM SAS Switches. Tables describes the calculated p-values for the beta coefficients listed in Table S3.1. Unless otherwise noted, p-values are rounded to nearest thousandth.

P-values of Coefficient Variables	Reporter Fluorescence		
	Low Induction	High Induction	High:Low Reporter Ratio
β_0	~	~	~
Forward Toehold	0.000	0.000	0.880
Reverse Toehold	0.034	0.000	0.160
Transducer Length	0.001	0.000	0.866
Total Structural Defect	0.673	0.036	0.000
Defect of ON State	0.959	0.356	0.000
Defect of OFF State	0.000	0.000	0.031
ΔG full sequence	0.000	0.000	0.670
ΔG Switch sequence	0.000	0.000	0.279
ΔG ON State	1.000	1.000	1.000
ΔG OFF State	1.000	1.000	1.000
$\Delta G \Delta G$ ON vs OFF State	1.000	1.000	1.000
Forward Toehold GC content	0.000	0.006	0.704
Reverse Toehold GC Content	0.000	0.014	0.256
Transducer Length GC Content	0.000	0.000	0.017
Migration Region GC Content	0.023	0.181	0.174

Unless otherwise noted, p-values are rounded to nearest thousandth