Engineering control of eukaryotic translation with application to the malaria parasite *Plasmodium falciparum*

by

Stephen J. Goldfless

M.S. Biochemistry (2005)
B.S. Biochemistry and Biology (2005)
Brandeis University, Waltham, MA

Submitted to the Department of Biological Engineering in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Biological Engineering
at the
Massachusetts Institute of Technology

February 2014

Signature redacted

Author .....
Stephen J. Goldfless
Department of Biological Engineering, MIT

Signature redacted

Certified by .....
Jacquin C. Niles
Associate Professor, Department of Biological Engineering, MIT
Thesis Advisor

Signature redacted

Approved by .....
Forest M. White
Associate Professor, Department of Biological Engineering, MIT
Co-Chair, Course XX Graduate Program Committee

This doctoral thesis has been examined by a committee of the Department of Biological Engineering as follows:

Signature redacted

Certified by .....

John M. Essigmann
Professor, Departments of Biological Engineering and Chemistry, MIT
Thesis Committee Chair

Signature redacted

Certified by .....

Jacquin C. Niles
Associate Professor, Department of Biological Engineering, MIT
Thesis Advisor

Signature redacted

Certified by .....

Darrell J. Irvine
Associate Professor, Departments of Biological Engineering
and Materials Science and Engineering, MIT
Thesis Committee Member

Engineering control of eukaryotic translation with application to the malaria parasite *Plasmodium falciparum*

by

Stephen J. Goldfless

Submitted to the Department of Biological Engineering on September 6, 2013 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biological Engineering

ABSTRACT

Experimenter control of target gene expression is a fundamental component of molecular biology research. In many systems, tools exist that allow generalizable control of gene expression at the transcriptional or post-transcriptional level. *Plasmodium falciparum,* the protozoan parasite responsible for the majority of death and sickness due to malaria, remains challenging to manipulate in the laboratory. No robust and generalizable tool for gene expression control has been developed in the parasite. To address this need, we engineered a new system for control of protein translation in eukaryotes, and applied it to *P. falciparum.* This system is based on the ligand-regulated interaction between an RNA aptamer and the TetR repressor protein. Although such protein-RNA interactions are abundant in nature and are known to effectively mediate control of gene expression, our system is unique in its direct modulation by an exogenous chemical. By genetically encoding TetR-binding RNA aptamers in the 5′ untranslated region (5′UTR) of an mRNA, translation of a downstream coding sequence is repressed by TetR *in vivo* and induced upon adding a non-toxic tetracycline analog.

We first define the system's component molecular interactions *in vitro,* followed by optimization of the constituent parts for convenience and performance. We then further optimize the system and validate its performance in two model systems, the budding yeast *Saccharomyces cerevisiae* and cell-free rabbit reticulocyte extracts.

We show the broad utility of the system in *P. falciparum* for controlling expression of reporter and endogenous proteins trafficked to a variety of subcellular compartments. Induction and repression are rapid and homogeneous across the cell population. Placing a drug resistance determinant under inducible control, we are able to modulate *P. falciparum* drug sensitivity, demonstrating the usefulness of the system for controlling relevant parasite biology.

In the process of constructing and validating a novel tool for gene expression in *P. falciparum,* we built a new series of gene expression vectors for molecular biology work in the parasite. In addition to developing optimized protocols for plasmid construction, we built a standardized, sequence-defined family of plasmids for malaria research. In all, we present a generalizable, well-defined toolkit for genetic programming of *P. falciparum.*

Thesis Supervisor: Jacquin C. Niles
Title: Associate Professor of Biological Engineering

CONTENTS

ACKNOWLEDGEMENTS AND ATTRIBUTIONS

Much of the work described in this document was performed as a team or collaboration. Although it is not possible to acknowledge every individual who has played a part in the success of this work, I specifically want to acknowledge several major contributors. The original library of TetR-binding RNA aptamers was discovered by Jacquin Niles. The first characterizations of TetR-aptamer interactions were performed by Brian Belmont[1]. Chapter 2 describes work performed with Brian Belmont and several others[2]. The work described in Chapter 2 is the subject of United States and international patent applications by MIT; the co-inventors are Brian Belmont, Jacquin Niles, Michael Marletta, and myself. Chapter 3 describes work performed with Jeffrey Wagner and Suresh Ganesan[3].

Although not explicitly documented in this thesis, the unpublished contributions of several students were extremely useful in guiding and informing my work. I would particularly like to acknowledge the contributions of Alexandra Westbrook and J. Karen Wong.

Additionally, the text of Chapters 2 and 3 comprises two manuscripts submitted for publication[2,3]. These manuscripts are the product of a collaborative writing effort on the part of their respective authors.

I would like to thank the many individuals and organizations (both inside and outside MIT) that provided support in the form of reagents, equipment, and advice. Some of their key contributions are noted in the chapters that follow. I am also grateful to the funding sources that made possible my presence in the laboratory and department. In addition to various funds coordinated by MIT and the BE Department, I was the recipient of NIEHS training grant support and a fellowship from the Siebel Scholars Foundation.

CHAPTER 1. MOTIVATION

1.1. Biology is the product of dynamic gene expression

The central theory of molecular biology, commonly known as the "central dogma," states that biological information is stored in DNA, passes through an RNA intermediate, and ultimately provides the template for protein synthesis[4]. Since this theory was first articulated in the 1950s, advances in the field of molecular biology have added substantial nuance to the concept, providing evidence of the additional DNA-DNA, RNA-RNA, and RNA-DNA templating interactions anticipated in the theory's early renditions (Fig. 1.1)[5].

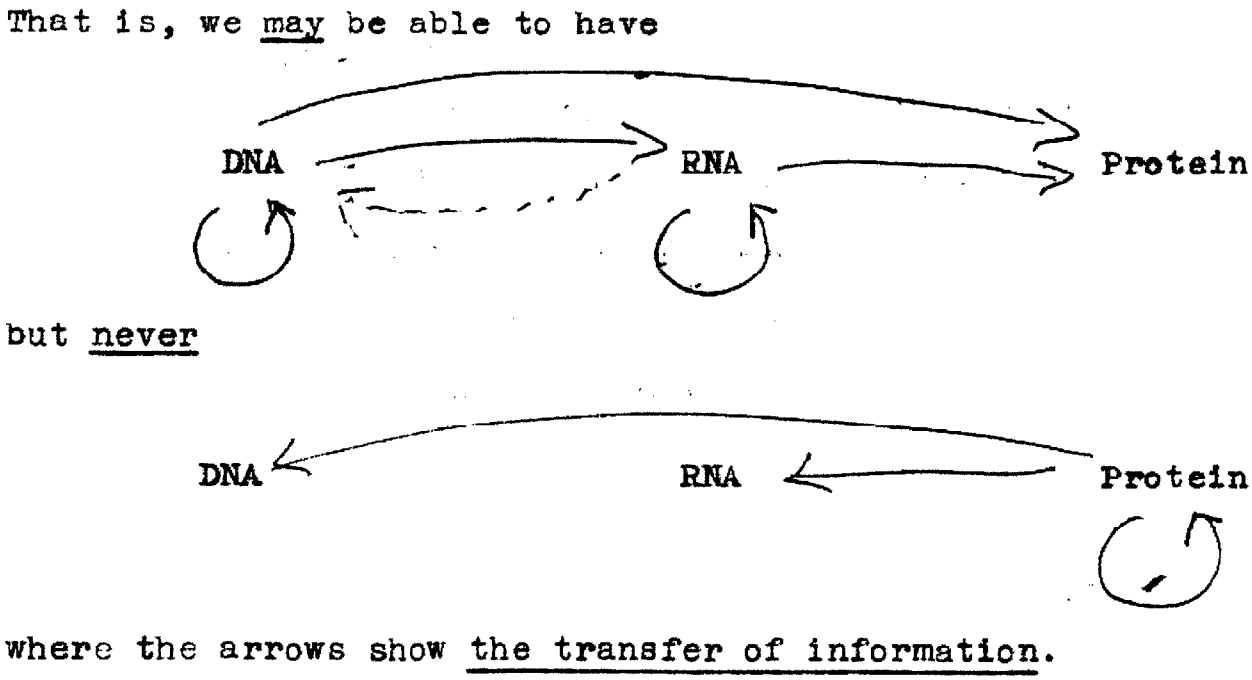

Figure 1.1. Early communication of the central theory of molecular biology. Reproduced from Crick[5].

As an organism must respond to a change in its environment with a change in its composition and behavior, dynamic control of gene expression is necessary for life. At each stage (transcription of DNA to make RNA, maintenance and metabolism of RNA, translation of RNA to make protein, and metabolism of protein, in addition to "esoteric" information transfer like RNA-templated RNA amplification[6]), there exists an opportunity for regulation of gene expression, and thus a point for control of biological behavior (Fig. 1.2). Documented examples of natural gene expression control at each of these points, and their roles in development and homeostasis, are the subject of numerous review articles[7–10].

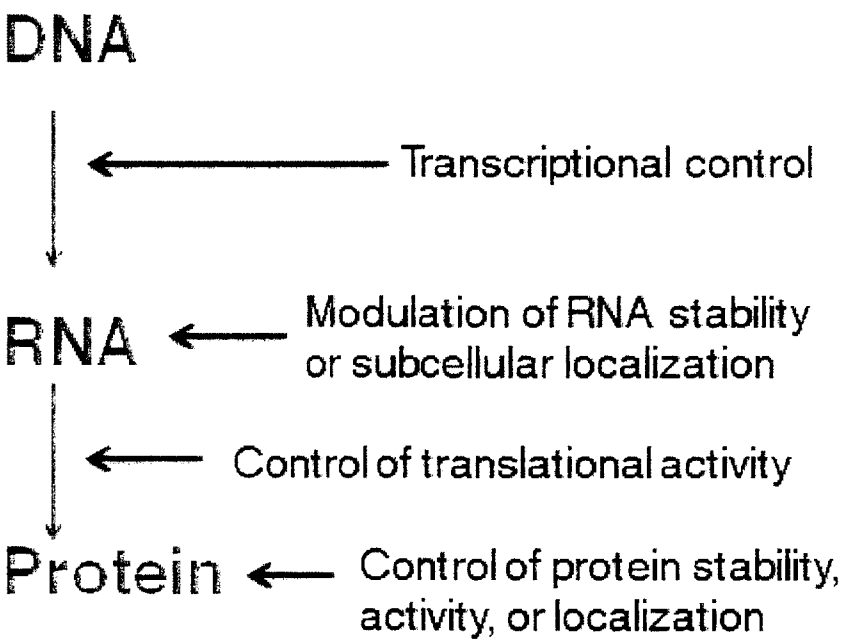

Figure 1.2. Points for control of gene expression. Red arrows indicate processes that are known to be subject to dynamic regulation.

1.2. Engineering control of target gene expression: advances in transcriptional control

Since the molecular biology revolution of the 20th century, engineered systems allowing experimenter control of target gene expression have become a mainstay of biological research. A critical early phase in this revolution was the dissection of natural molecular mechanisms that govern the bacterial transcriptional response to key metabolites such as lactose, arginine, and tryptophan. Understanding of these systems accelerated molecular biology research by providing the basic genetic parts for forward engineering of chemically-mediated control of a recombinant gene[11]. The ability to modulate the expression of a gene of interest in the context of an isogenic system allows a researcher to precisely interrogate the function of one or more genes. Of course, engineering of a native gene or transgene to respond to a particular inducer metabolite does not negate the existing natural response to the inducer. In some systems, the set of natural genes responding to the inducer is small and dispensable; the native *lac* operon of *Escherichia coli* is commonly deleted, minimizing the effect of induction with IPTG outside of the engineered target gene. Often, for the purpose of a particular experiment, it is unimportant what the pleiotropic effects of a chemical inducer might be. However, such "off-target" effects can be sizable, particularly in eukaryotic systems, which tend to possess transcriptional regulatory systems of greater complexity: the ubiquitous galactose-inducible transcriptional tool for *Saccharomyces cerevisiae* makes use of a natural transcriptional network responsible for altering the transcriptional profile of hundreds of genes in response to galactose[12]. Therefore, it is often desirable to have a tool for inducible gene expression that is practically orthogonal to the biology of the host organism. Such an optimal tool would sense the presence of an unnatural molecular signal that otherwise would pass undetected through the host organism; the tool would then effect altered gene expression via a mechanism acting specifically on the target gene(s). Motivated by this need, eukaryotic transcriptional control tools based on bacterial inducible

transcriptional repressors have been developed. In one type of approach, a number of bacterial operator DNA elements are inserted into a eukaryotic promoter, and expression of the cognate bacterial repressor protein suppresses transcription by steric hindrance of transcription factor binding[13]. This approach is limited by its requirement for a large array of repeated operator sequences, and typically low expression and/or dynamic range. A second approach, which has found widespread adoption in work with cultured mammalian cells and transgenic animals, is based on the protein fusion of a bacterial repressor to a eukaryotic transactivating or repressing domain[14–17]. These tools enable >1000-fold changes in target gene expression with virtually no effect on off-target gene expression[18]. Despite differences in the molecular mechanism of transcriptional initiation across eukaryotic model systems, the transcription-activating or transcription-repressing activities of the protein domains used in these tools are functional in a variety of contexts, making the tools useful for research in many organisms.

1.3. mRNA-related processes can be mediated by specific interactions between proteins and mRNA structural elements

After an mRNA is synthesized, it can undergo numerous molecular interactions that affect its nucleotide sequence, subcellular localization, stability, and capacity to be translated into protein[19–21] (Figure 1.3). Many of these interactions are mediated by specific nucleotide sequences within the mRNA, and many of these interactions are act not through the recognition of primary RNA sequence, but through secondary and tertiary structures formed by specific RNA sequence elements. Such structures constitute binding surfaces capable of specific, often high-affinity interactions with their targets, which are often proteins such as nucleases, molecular motors, or components of the translational machinery. Furthermore, the formation of these protein-binding structures can itself be regulated by the presence of a small molecule ligand, or an environmental stimulus such as temperature[22].

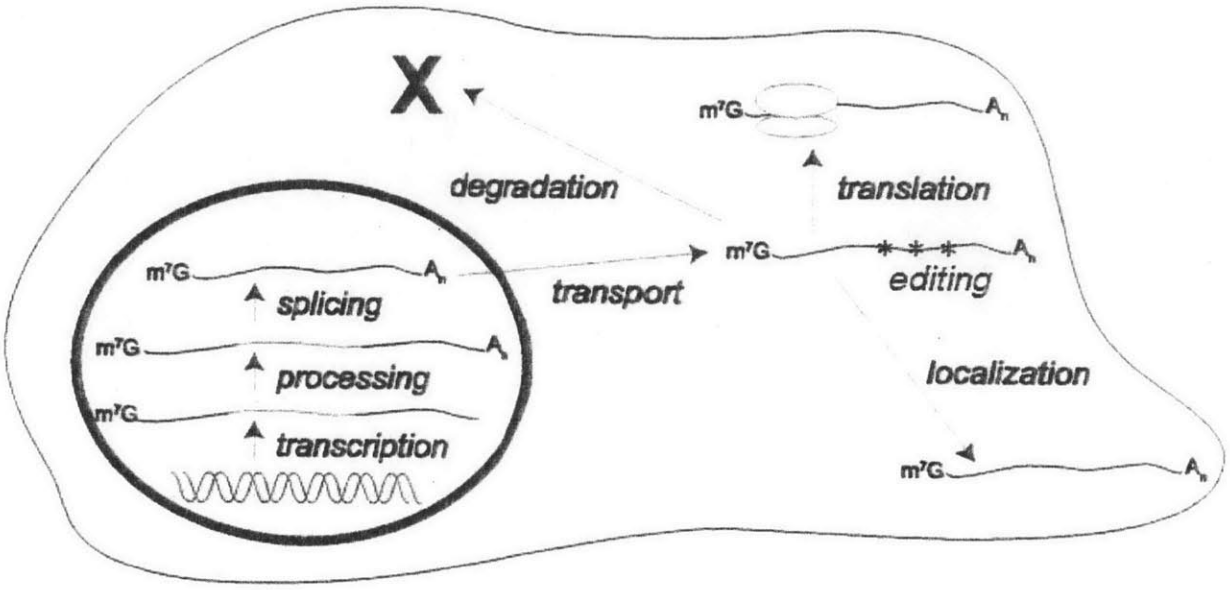

Figure 1.3. Examples of mRNA-related processes in the cell. Adapted, with permission, from Brian Belmont.

1.4. Natural post-transcriptional regulation of gene expression mediated by protein-RNA interactions in the untranslated regions of mRNA

In cellular processes as diverse as X-chromosome dosage compensation[23], embryogenesis[24], and iron homeostasis[25], regulation of gene expression is mediated by the interaction of proteins with specific, secondary-structure-forming RNA sequences in the 5′ or 3′UTR. In the well-characterized iron-responsive element/iron-responsive protein (IRE/IRP) system, mRNA coding for the iron storage protein ferritin contains a 35-nt stem-loop structure in its 5′UTR. This stem-loop structure, known as the IRE, binds the IRP with sub-nanomolar affinity under conditions of low labile iron concentration; the IRE-IRP complex represses translation severalfold by inhibiting the assembly of the 43S preribosomal complex through steric hindrance of the required interaction between eukaryotic translation initiation factor complex 4F (EIF4F) and the

40S small ribosomal subunit[25–27] (Fig. 1.4). When intracellular free iron accumulates, the IRP switches to a low-IRE-affinity state, dissociating from the ferritin mRNA and permitting synthesis of ferritin. It has been demonstrated that substituting other high-affinity RNA-protein interactions at the 5′UTR of a given transcript can yield a similar inhibition of translation[28]. Post-transcriptional regulation mediated by sequence-specific RNA-protein interactions can occur via multiple modes, sometimes mediated by a single protein: for example, while the IRP produces a steric blockage of translation initiation by binding to an IRE in the 5′UTR of ferritin mRNA, it also serves to stabilize transferrin receptor mRNA by binding to multiple IREs in the 3′UTR and blocking endonucleolytic degradation of the transcript[29]. RNA-protein interactions in the 3′UTR can also regulate translation, often through the recruitment of factors that interfere with 5′ cap/EIF4E recognition by EIF4G2. In fact, sequence-specific RNA binding proteins have been shown in many cases to serve as adaptor proteins that target RNA-localizing[30,31] or translation-enhancing or repressing[23] activities to specific mRNAs. These natural mechanisms of mRNA-specific post-transcriptional regulation can often be recapitulated in recombinant or *in vitro* systems and applied to the regulation of target transcripts.

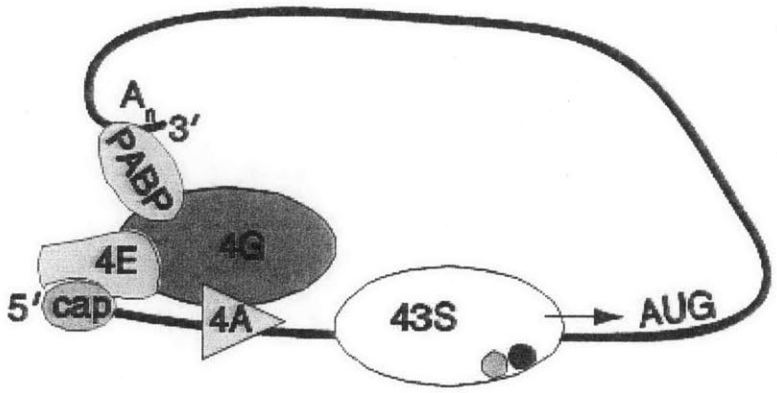

Figure 1.4. Canonical initiation of eukaryotic translation by cap-dependent ribosome assembly and scanning. Eukaryotic initiation factor 4E (EIF4E) binds the mRNA 5′^{7}Me-G_{ppp} cap and recruits the rest of the EIF4F complex (the scaffold protein EIF4G and the RNA helicase EIF4A). EIF4G recruits polyadenylate-binding protein (PABP), resulting in circularization of the mRNA. The 43S pre-initiation complex, containing the 40S small ribosomal subunit, binds EIF4G and scans the mRNA in the 5′→3′ direction until it detects the start codon AUG.

1.5. Nucleic acid aptamers are synthetic structural elements that bind a target with high affinity and specificity

In addition to occurring naturally, particularly in the genomes of bacteria, target-binding RNA elements can be discovered in the laboratory, where they are called aptamers. An aptamer is a polynucleotide (the term is infrequently used to refer to a polypeptide) that has been selected from a large, random library for its ability to bind a desired target ligand[32]. In the case of RNA aptamers, an *in vitro* process known as systematic evolution of ligands by exponential enrichment (SELEX) is often used to specifically expand and purify tightly-binding RNA molecules from a large library (>10^{12} sequences) through successive cycles of transcription, positive or negative selection for binding, elution, reverse transcription and amplification by

PCR[33,34]. Aptamers can then be further optimized for affinity and specificity by directed evolution or rational design. Aptamers have been developed that are capable of binding small molecules such as tetracycline (Tc)[35] or theophylline[36], proteins such as bacteriophage T4 DNA polymerase[33], or complex targets like the outer surface of pathogenic organisms[37]. Once an RNA aptamer has been developed, it can be genetically inserted into the context of a larger RNA (such as into the UTR of an mRNA), where it may retain its ligand-binding activity.

1.6. Placement of aptamer sequences in the UTR enables engineered gene expression control

Aptamer-bearing mRNAs have been used to effect several types of post-transcriptional gene regulation in response to ligand binding[1,35,36,38,39]. One mechanism for such regulation involves the inhibition of translation by an aptamer in the 5′UTR. In a one study, inducible post-transcriptional regulation in *S. cerevisiae* was achieved by placing one or more Tc-binding aptamers in the 5′UTR of target transcripts[39]. It was suggested that aptamer-ligand binding in this system inhibited translation by stabilizing the stem-loop structure of the aptamer, thereby blocking procession of the 43S scanning complex. The authors reported that they were unable to use the system to achieve gene regulation in mammalian cells, a biological environment known to possess a stronger 43S scanning activity capable of efficiently translating mRNAs containing 5′UTR structures that block translation in *S. cerevisiae*[40].

Another mode of synthetic gene expression control involves the aptamer-mediated reconstitution of protein-RNA interactions that modulate translation (Section 1.4). These interactions are generally constitutive in nature, and require a separate mode of expression control (e.g. transcriptional regulation of protein production) if one wishes to modulate their effect on translation. In a notable advance, our laboratory was the first to construct a synthetic protein-

RNA interaction system that controls gene expression in response to an exogenous small molecule (tetracycline and its analogs)[1].

1.7. *Plasmodium falciparum* is an important human pathogen with an incomplete molecular biology toolset

Malaria kills 500,000–1,000,000 people each year, and the vast majority of these deaths are attributable to *P. falciparum*, one of the five species known to infect humans[41]. *P. falciparum* is a single-celled eukaryote parasite belonging to the phylum apicomplexa, a group of obligate parasites. The apicomplexa are much less recently diverged from most model organisms (such as animals, fungi and ameboazoa) than those model organisms are from each other[42]. The *P. falciparum* life cycle consists of numerous differentiation events triggered by the diverse biological environments which it must inhabit. Upon injection into a human host's bloodstream by a bite from the *Anopheles* mosquito, the asexual haploid sporozoite form of the parasite must traffic to the parenchyma of the liver, where massive parasite expansion occurs and merozoites are released into the bloodstream[43]. At this point the parasite population undergoes the asexual intraerythrocytic development cycle (IDC), a 48-hr development cycle consisting of gross morphological and gene expression changes, endoreduplication of the haploid genome, and the release of up to dozens of daughter merozoites competent for erythrocyte infection[44]. At a low frequency, parasites in the IDC differentiate into male or female gametocytes, which if taken up in a bloodmeal, can be stimulated by environmental factors in the mosquito midgut to develop into gametes competent for the formation of a diploid zygote[45]. Several developmental transitions occur within the digestive tract of the mosquito, ultimately resulting in sporozoite production and migration to the salivary gland[46]. Five distinct DNA synthesis events must occur to complete the parasite life cycle, although the molecular mechanisms of cell cycle regulation and differentiation are either poorly understood (in the case of the IDC) or nearly

uncharacterized[46–48]. Understanding these processes in greater detail will be essential to the development of antimalarial therapies that address both infection and transmission.

Despite its clinical relevance, studying the organism remains highly technically challenging due to a number of factors. Although asexual blood-stage parasites can be indefinitely passaged in culture, the parasite must be cultivated in human erythrocytes under microaerobic conditions[49]. The parasite's genome is 80.6% (A+T), with many noncoding regions exceeding 90% (A+T)[50], a property that confounds many sequence-based approaches to decoding biological function in the organism. Another challenge to sequence-based bioinformatics studies is the propensity of the parasite for "domain shuffling" whereby conserved protein domains within the parasite are often located in contexts not observed in model eukaryotes[51]. Some components of parasite transcription have been identified[52,53], but relatively little progress has been made in deciphering the parasite's promoter structure or the molecular factors required for transcription initiation or elongation. It has even been suggested that *P. falciparum* may regulate protein expression primarily through post-transcriptional means[54]. As a result, there exists no robust or generally useful inducible gene expression system for *P. falciparum*, making investigation of its biology challenging.

Available methods for genetic manipulation of *P. falciparum* include engineering of chromosomal DNA by homologous and site-specific recombination[55,56], as well as stable and transient transfection of episomal DNA[57,58]. Tools for conditional gene expression comprise a ligand-regulated protein degradation domain[59], regulation of gene expression by drug selection of increased episomal copy number[60], and a tetracycline-inducible transcriptional transactivator

system[61]. These tools for conditional expression are either unreliable or not generalizable (Section 4.3).

At a practical level, genetic manipulation of *P. falciparum* is often challenging and time-consuming. *In vitro* manipulation of parasite-derived DNA sequences (particularly at the stage of molecular cloning and propagation in *E. coli*) is hampered by the extent of low-complexity, high-(A+T) DNA, which is often unstable during the cloning process (Section 3.3). In addition, regulatory elements (e.g. UTRs and elements mediating DNA segregation and maintenance) that can be used in recombinant DNA vectors, in addition to cloning vectors themselves, are scarce and poorly defined, both in function and sequence. A well-defined, standardized toolkit for construction of synthetic genetic programs in the parasite, as is available for many model organisms, should significantly accelerate the pace of research in the field, particularly by enabling new research groups to begin molecular biology investigations with minimal time spent on building proprietary technology.

1.7. Summary of rationale and work presented

Our research group has been broadly interested in engineering approaches to challenges in malaria research. As there was a clear need for a generalizable, robust tool for conditional gene expression in *P. falciparum,* we sought to develop such a tool. Chapter 1 details the development of a synthetic, naturally-inspired tool for experimenter control of target gene translation in eukaryotes, with optimization in model systems. Chapter 2 details the construction of a molecular genetic toolkit for *P. falciparum* research, which is to be openly and freely shared with the research community. Chapter 3 details the application of our new tool, developed in Chapter 1, to gene expression control in *P. falciparum.* The work in Chapter 3 makes heavy use of molecular tools and techniques that we developed in the course of the last 6 years. In all, this

thesis documents the progression, from early concepts to validated resources, of key technologies developed in our laboratory.

CHAPTER 2. DIRECT AND SPECIFIC CHEMICAL CONTROL OF EUKARYOTIC TRANSLATION WITH A SYNTHETIC PROTEIN-RNA INTERACTION

2.1. Note

This chapter comprises the text of a manuscript now published in *Nucleic Acids Research*[2].

2.2. Abstract

Sequence-specific RNA-protein interactions, while commonly used in biological systems to regulate translation, are challenging to selectively modulate. Here, we demonstrate the use of a chemically-inducible RNA-protein interaction to regulate eukaryotic translation. By genetically encoding Tet Repressor protein (TetR)-binding RNA elements into the 5′ untranslated region (5′UTR) of an mRNA, translation of a downstream coding sequence is directly controlled by TetR and tetracycline analogs. In endogenous and synthetic 5′UTR contexts, this system efficiently regulates the expression of multiple target genes, and is sufficiently stringent to distinguish functional from non-functional RNA-TetR interactions. Using a reverse TetR variant, we illustrate the potential for expanding the regulatory properties of the system through protein engineering strategies.

2.3. Introduction

Translational regulation mediated by specific RNA-protein interactions is pervasive in shaping diverse biological processes including metabolism[62], early development[62,63], neuronal plasticity[63] and immunity[64]. However, it remains challenging to recapitulate this mode of regulation in a way that is transcript-specific and readily modulated experimentally. Such selective control represents a valuable tool for studying and designing translational regulation mechanisms based on RNA-protein interactions. Using naturally occurring RNA-protein interactions for this purpose is problematic because many are directly toggled by protein modifications such as

phosphorylation[63-65], which are in turn shaped by complex cellular signaling events that cannot be precisely manipulated.

Several systems for regulating eukaryotic gene expression post-transcriptionally have been previously described, and include RNAi[66], small molecule-regulated cis-acting ribozymes and riboswitches[67]. While RNAi has proven a highly useful tool in many contexts, its use is limited to organisms expressing the necessary machinery. Even in contexts where RNAi is well-established as a routine tool, the introduction of interfering RNAs can produce unintended off-target effects[68]. Furthermore, gene expression knockdown with RNAi cannot be rapidly switched off, and the targeted protein may not return to its initial level for days as the catalytic siRNA is slowly depleted[69]. Complementary to RNAi, small molecule-regulated cis-acting ribozymes and riboswitches can act in the absence of organism-specific RNA degradation pathways, allowing their application in a broader variety of contexts. Riboswitches, in particular, circumvent target transcript degradation and are likely to provide highly dynamic regulation. However, these ligand-regulated systems do not immediately support a direct path toward integration with endogenous translation regulation mechanisms. Such natural systems frequently couple the use of specific primary RNA-protein interactions with secondary protein-protein interactions to modulate translation in different contexts. For example, the cellular programs controlling inflammation and its resolution[64], as well as the establishment of asymmetric protein distributions critical for neuronal development[65] and cell migration[70], make use of this paradigm to achieve translational control over effector genes. For this reason, it is highly desirable to construct translational regulation schemes that can interface directly with biologically relevant, protein-based signaling networks. Therefore, we sought to construct a system for eukaryotic

translational regulation using well-defined, ligand-regulated and specific RNA-protein interactions that can be directly controlled in a straightforward and cell type-independent way.

Previous attempts at achieving such a ligand-regulated RNA-protein interaction have been described[71–76]. However, in these examples, the interaction either cannot be modulated in vivo[74], requires a ubiquitous and stringently regulated metabolite ligand such as iron[73], or relies on inducible transcription to control the RNA-binding protein[71,76]. Thus, despite the widespread nature of RNA-protein interactions in regulating eukaryotic translation, no methods exist for reversibly and precisely reproducing such regulation in vivo. We expect such methods to be broadly applicable and invaluable for reconstructing and interrogating native translation control mechanisms, as well as for designing novel cellular functions.

2.4. Introduction of TetR aptamers into a eukaryotic 5′UTR context

Previously, we described a series of RNA aptamers that bind TetR tightly, but only in the absence of tetracycline[1]. Here, we report using TetR aptamers to regulate eukaryotic translation in a mammalian cell-free extract (rabbit reticulocyte lysate, RRL) and *Saccharomyces cerevisiae* (yeast). Translation repression based on the TetR-aptamer system is directly relieved by aTc or doxycycline (Dox), which disrupts the TetR-aptamer interaction (Fig. 2.1A), and no manipulation at the transcriptional level is required. To determine the regulatory effect of TetR-aptamer interaction in the 5′UTR, we chose five RNA aptamers derived from a library selected for tight, Tc-regulated TetR binding[1]. We eliminated an AUG codon present in a dispensable region of these aptamers to obtain **5-1.13, 5-11.13, 5-14.13, 5-18.13** and **5-29.13** (Table 2.S1). We were unable to modify the AUG located within the conserved Motif 1, known to be crucial for TetR binding[1], without sacrificing high binding affinity (Fig. 2.1B). The modified aptamers bound TetR with low nanomolar affinity, similarly to their parents (Fig. 2.2A).

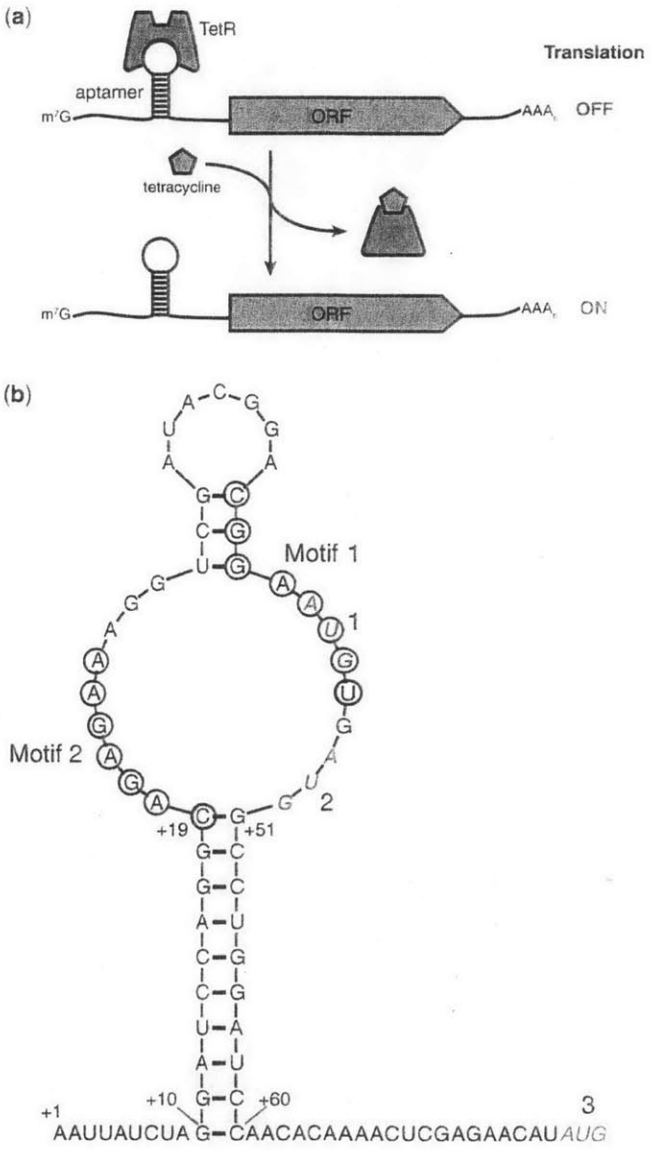

Figure 2.1. Placement of TetR-binding RNA aptamers within the 5′-UTR. (**A**) Illustration of the regulation scheme used to demonstrate functionality of the TetR-aptamer module. TetR binding to the aptamer element within the 5′UTR of the regulated ORF inhibits its translation. Tetracycline analogs (e.g. aTc and Dox) induce translation by disrupting the TetR-aptamer interaction. (**B**) The primary and predicted secondary structures of aptamer **5-1.2** (bases 10–60) within the 5′UTR context used in this study are shown. Residues comprising two conserved regions (Motifs 1 and 2) indispensable to TetR binding are circled. Three potential start codons, 1–3, are shown in green bold

italics. Start codon 1 (within Motif 1) is out of frame, whereas 2 is in frame with 3, the downstream ORF's native start codon.

In order to test translation repression with the system, we inserted TetR aptamers into a short 5′UTR upstream of the firefly luciferase gene (FLuc), a context used previously to study protein-UTR interactions[77]. We co-transformed yeast with a plasmid encoding TetR under the control of a galactose-inducible promoter, and a plasmid containing FLuc regulated by either a TetR aptamer or the iron-responsive element (IRE) as a structured, non-TetR-binding control. We then measured FLuc activity after growth in the presence or absence of aTc. We also prepared each mRNA construct by in vitro transcription and tested for translational regulation in RRL supplemented with recombinant TetR in the presence and absence of aTc (Fig. 2.2B). Although all five aptamers bound TetR with similarly high affinity (Fig. 2.2A), they did not similarly regulate translation. In the presence of TetR, only **5-1.13** inhibited FLuc synthesis in both RRL (~50% repression) and yeast (~80%), while **5-11.13** inhibited FLuc synthesis only in yeast (~67%). Regulation, when observed, was always fully aTc-modulated as expected based on Fig. 2.1A.

(a)

Aptamer	K_d ± S.D. (nM)
5-1.13	0.37 ± 0.04
5-11.13	1.5 ± 0.51
5-14.13	1.4 ± 0.52
5-18.13	0.35 ± 0.28
5-29.13	2.9 ± 1.5
5-1.2	1.5 ± 0.85
5-1.2m2	No binding detected

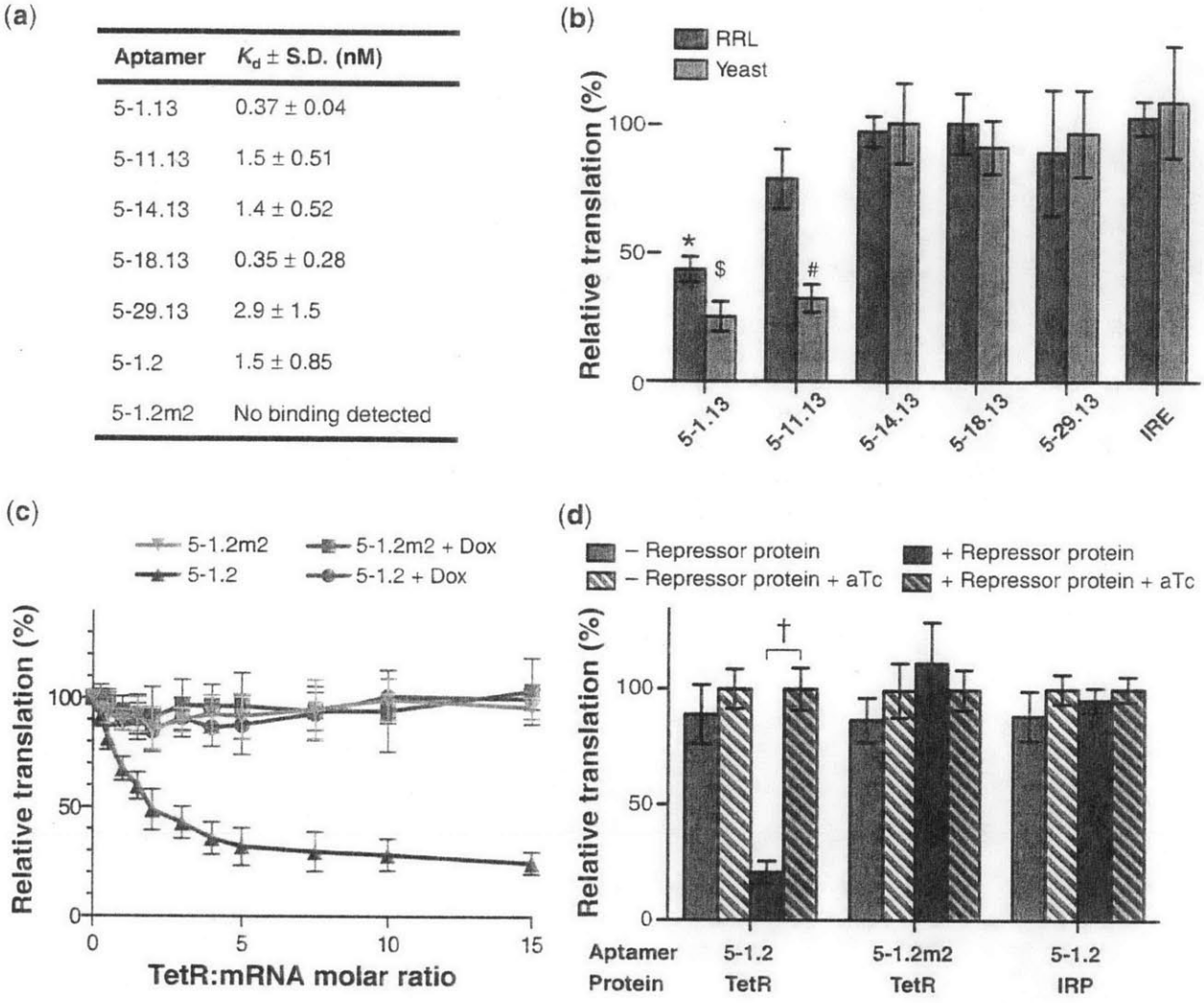

Figure 2.2. Genetically encoded TetR aptamers enable translational regulation in RRL and in yeast. (**A**) Aptamers bind TetR with high affinity. The aptamer **5-1.2m2** contains two point mutations that eliminate specific binding to TetR. (**B**) Only aptamers **5-1.13** and**5-11.13** function to repress translation in RRL *in vitro* and in yeast *in vivo*. Relative translation was calculated by dividing FLuc signal in the absence of aTc by FLuc signal in the presence of aTc. (**C**) TetR dose-dependently represses RRL translation of synthetic FLuc mRNA containing the minimized aptamer **5-1.2**, but not the mutant **5-1.2m2**. Addition of 1μM Dox completely relieves repression. (**D**) Inducible expression in yeast requires a functional TetR-binding aptamer and TetR expression. In all figures, data represent the mean ± s.d. of at least four experiments. A two-tailed, unpaired t-test was used to calculate the significance ($\alpha = 0.005$) of the difference between induced and uninduced conditions. $^{*}P = 5.1 \times 10^{-7}$; $^{\$}P = 2.0 \times 10^{-7}$; $^{\#}P = 4.9 \times 10^{-12}$ $^{\dagger}P = 2.2 \times 10^{-11}$.

2.5. Aptamer minimization and functional validation

Since **5-1.13** regulated translation most effectively in both test systems, we selected it for further optimization. We constructed the more minimal aptamer **5-1.2** by retaining the conserved loop region and flanking it with a predicted RNA stem that lowered the folding stability relative to the starting aptamer (ΔG_{fold} = -22.6 versus -25.2 kcal/mol). This potentially reduces the negative impact that structured RNA within the 5'UTR can have on translational efficiency[78]. We also introduced two point mutations in the conserved Motif 2 to produce **5-1.2m2**, which has an identical predicted secondary structure, but no longer binds TetR.

We performed RRL translation regulation experiments using fixed mRNA and titrated TetR concentrations. Translation of mRNA containing aptamer **5-1.2**, but not **5-1.2m2**, was dose-dependently repressed by TetR, reaching ~90% repression at a 5:1 TetR:mRNA ratio (Fig. 2.2C). This indicates that TetR-dependent translational repression is specific to mRNA containing an aptamer competent for binding TetR and definitively occurs post-transcriptionally. Repression was fully relieved by aTc (1 μM), consistent with inducibility of the TetR-aptamer interaction. Next, we tested the specificity of inducible translation in yeast using the experimental design described earlier, but including either a non-functional RNA (**5-1.2m2**) or an unrelated protein (the iron-responsive element binding protein, IRP) as substitutes for **5-1.2** and TetR, respectively (Fig. 2.2D). We observed aTc-inducible regulation of FLuc synthesis (~80%) only in strains simultaneously carrying **5-1.2** and expressing TetR. Importantly, TetR was only expressed in galactose-containing media and its abundance was not decreased by aTc, as confirmed by western blot (Fig. 2.S1). For the above experiments, we used galactose-inducible transcription to control TetR expression, allowing us to test isogenic strains. However, even when TetR was expressed constitutively from the TDH3 promoter, we found identical aTc-dependent

translational regulation (Fig. 2.S2). These data underscore the potential for using this system in biological contexts where transcriptional control is not accessible. Quantitative PCR analysis showed that increased FLuc reporter expression in the presence of aTc was not accompanied by an increase in FLuc mRNA, as would be expected if a transcriptional response or a significant change in mRNA stability were responsible for inducible expression (Fig. 2.S3). Altogether, these data demonstrate that the observed *in vivo* regulation occurs at the translational level and is due to a specific interaction between TetR and an RNA aptamer competent for binding TetR. Lastly, we replaced FLuc with the Venus yellow fluorescent protein (vYFP), and this reporter was either expressed episomally or integrated at the *TRP1* locus. Flow cytometry showed quantitatively that inducible expression is homogeneous across a yeast cell population, and similar in dynamic range irrespective of the gene being regulated (Fig. 2.3).

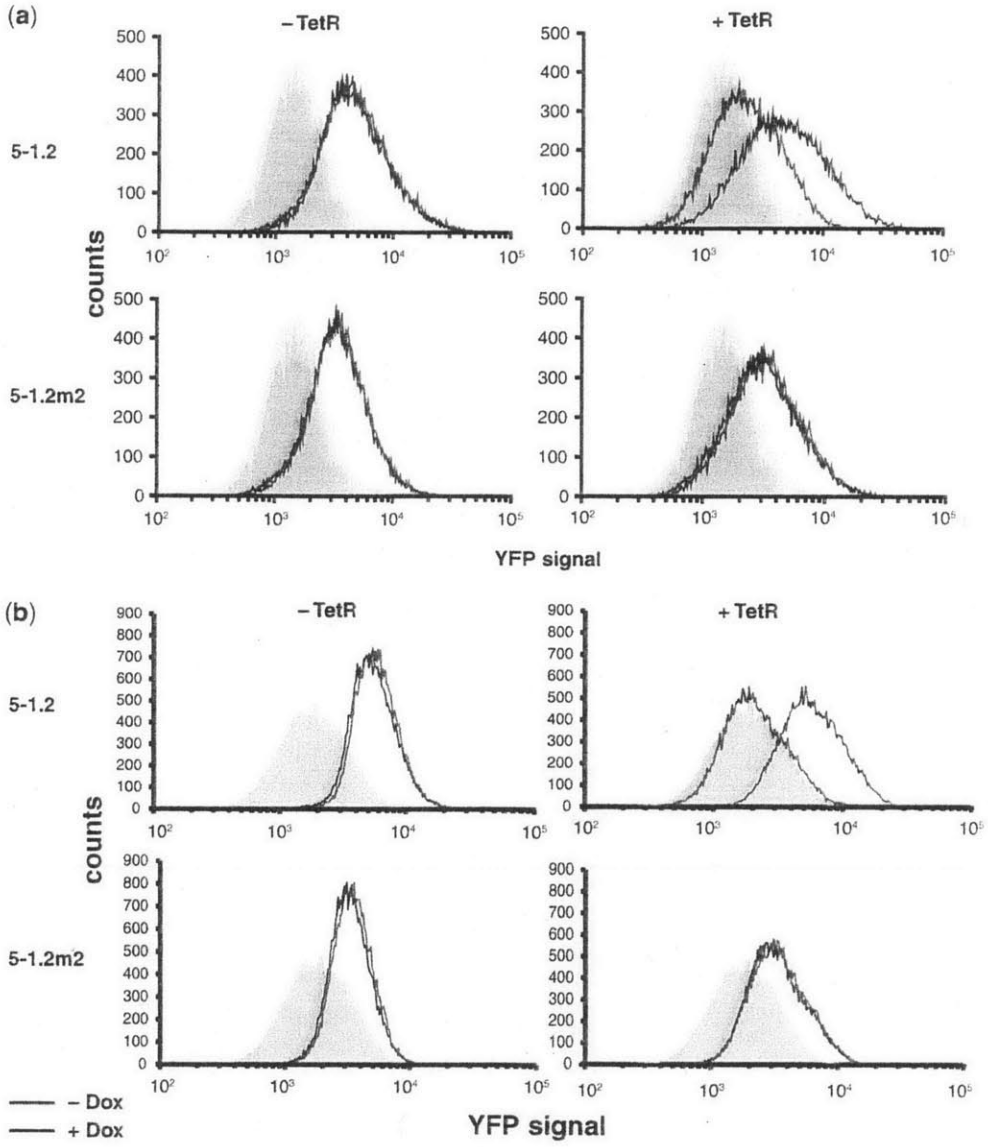

Figure 2.3. Population measurements of yeast gene expression regulated by TetR aptamers. Flow cytometry histograms show population-wide expression levels of aptamer-regulated vYFP. The aptamer located within the 5′UTR of vYFP and the expression status of TetR are indicated. In (**B**), the vYFP reporter is episomal, and in (**B**) integrated at the *TRP1* locus. Shaded gray histograms represent the auto-fluorescence of the background yeast strain. Measurements for yeast grown in the absence (red) and presence (blue) of Dox are indicated. Data for each histogram are representative of four independent experiments.

2.6. *In vivo* selection of functional aptamer-protein interactions

Different implementations of the translation regulatory system may necessitate discovering TetR aptamers with unique binding characteristics. Therefore, it is important to define a strategy for rapidly identifying new functional aptamer variants. To support this effort, we devised a positive/negative selection scheme capable of distinguishing between TetR-binding and non-binding RNA sequences (Fig. 2.4A). To establish proof-of-concept, we constructed plasmids containing the yeast URA3 gene (encoding orotidine-5′-phosphate decarboxylase, Ura3p) with either **5-1.2** or **5-1.2m2** within the 5′UTR context used earlier. We co-transformed each separately with a TetR-encoding plasmid into a yeast *ura3* mutant auxotrophic for uracil. These strains exhibited similar growth on nutrient-rich YPD media (Fig. 2.4B). For the negative selection, we plated cells on media containing 5-fluoroorotic acid (5-FOA), which is converted to the cytotoxic 5-fluorouracil by Ura3p. These conditions only permit growth of cells containing a functional TetR/**5-1.2** interaction which can repress Ura3p synthesis (Fig. 2.4C). For the positive selection, we plated cells in the presence of aTc on uracil-deficient media. Cells surviving this selection step have been induced to synthesize Ura3p to complement the uracil auxotrophy. This ensures that interactions identified by negative selection are aTc inducible (Fig. 2.4D). As expected, the cells containing a functional TetR-aptamer system exhibit a significant growth defect when plated in the absence of aTc on uracil-deficient media (Fig. 2.4E). To further establish the utility of this selection scheme, we mixed TetR-expressing cells containing *URA3* controlled by either **5-1.2** or **5-1.2m2** in a ratio of $1{:}10^4$, respectively. From this mixture, we plated $\sim 1.5\times10^5$ cells on media containing 5-FOA, and seventeen large colonies grew. Sequencing the 5′UTR of plasmid DNA isolated from ten of these colonies revealed that all carried the **5-1.2** sequence. These data demonstrate that this selection strategy can specifically

recover functional TetR aptamers from a large non-functional background, which should prove useful for identifying aptamer or TetR variants with novel regulatory characteristics.

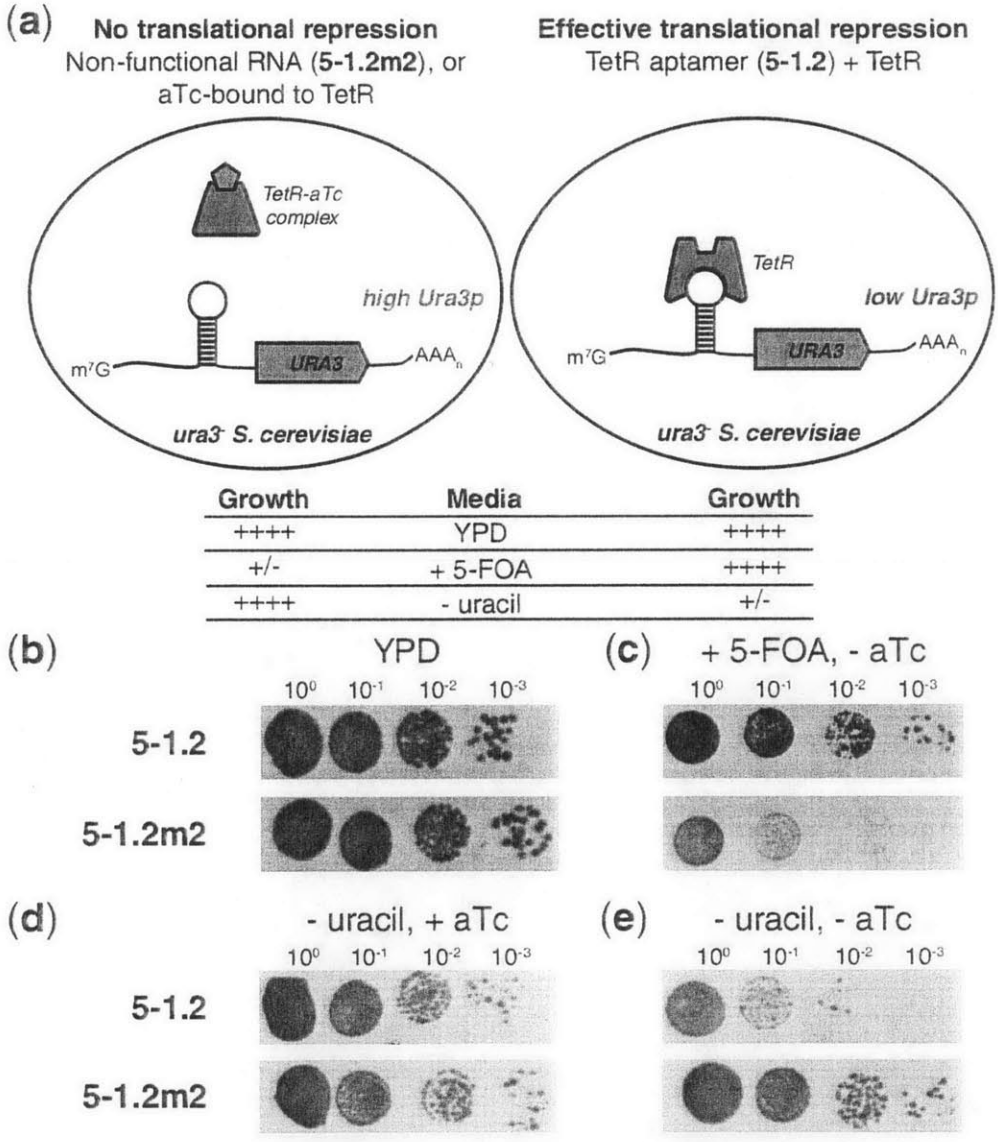

Figure 2.4. Positive/negative selection of yeast mediated by TetR aptamers. (**A**) Illustration of the *URA3* positive/negative selection scheme showing the predicted growth phenotypes of the strains under indicated conditions. (**B**) TetR-expressing strains in which Ura3p synthesis is controlled by either **5-1.2**or **5-1.2m2** grow similarly on YPD.

(**C**) Negative selection on 5-FOA only permits growth of strains capable of repressing Ura3p translation. (**D, E**) When **5–1.2** controls Ura3p translation, aTc is required for growth in the absence of uracil. In each condition, 10-fold dilutions are shown from left to right. Plates were imaged after growth at 30°C for 2 days (YPD) or 3 days (synthetic defined media).

2.7. Inversion of Tc dependence with a TetR variant

Several engineered RNA regulatory schemes are based on direct interactions between small molecule ligands and RNA[36,39,79]. However, a unique and compelling basis for using protein-RNA interactions is the potential to take advantage of protein engineering strategies to expand the scope of regulatory behavior achievable while maintaining the aptamer as a validated and defined component. To illustrate this principle, we chose revTetR-S2, a TetR variant previously derived through protein mutagenesis[80] that binds the cognate tetO DNA operator, but only in the presence of Dox. Our earlier work established that TetR aptamers compete with tetO for binding to TetR, indicating that both likely interact with the TetR nucleic acid binding domain[1]. Therefore, provided that the amino acid residues involved in the interaction between TetR and its aptamers are retained in revTetR-S2, the latter is reasonably expected to bind these aptamers, but with an inverse dependence on Dox. We first determined that purified revTetR-S2 bound **5-1.2** tightly in vitro (K_d = 3.1 ± 1.0 nM). Similarly to TetR, revTetR-S2 repressed translation of aptamer-containing mRNA in RRL, but as expected, only in the presence of Dox (Fig. 2.5A). Upon expressing revTetR-S2 in yeast with **5-1.2**-regulated vYFP, Dox enabled 50% repression of vYFP (Fig. 2.5B). Expectedly, **5-1.2m2** did not regulate vYFP expression, emphasizing the retained requirement for a specific revTetR-S2-aptamer interaction. Interestingly, another previously reported revTetR based on the TetR(BD) hybrid repressor, revTetR r1.7[81], did not demonstrate aptamer- and aTc-dependent translation repression activity (data not shown). The mutations present in the nucleic acid binding domains of revTetR-S2 and revTetR r1.7 are

distinct. While both proteins support tetO binding, our data suggest that aptamer **5-1.2** discriminates between these two TetR variants. Thus, in cases where modifying TetR can potentially disrupt the binding interface with the aptamer, it is important to confirm that the protein-aptamer interaction is preserved.

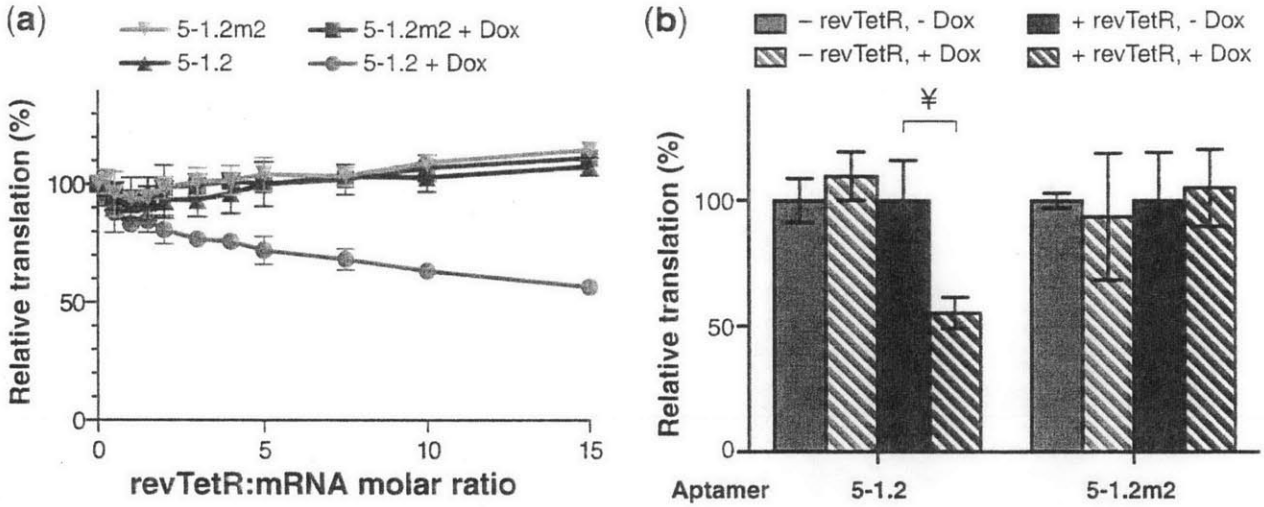

Figure 2.5. Inversion of Dox-inducible phenotype with a mutant TetR. (**A**) A revTetR variant dose-dependently represses RRL translation of synthetic FLuc mRNA as in Fig. 2.2C, but only in the presence of Dox. (**B**) A revTetR variant specifically represses translation as in Fig. 2.2D, but only in the presence of Dox. In all figures, data represent the mean ± SD of at least four experiments. A two-tailed, unpaired t-test was used to calculate the significance ($\alpha = 0.005$) of the difference between induced and uninduced conditions. $^{¥}P = 8.1 \times 10^{-5}$.

2.8. Aptamer engineering to reduce inhibition of translation in the induced state

A noteworthy challenge associated with placing an aptamer in the 5′UTR is that this may decrease the maximal protein expression levels of the regulated ORF. Indeed, we observe that

aptamer **5-1.2** causes a significant decrease in maximal reporter gene expression levels compared to when no aptamer is present. Therefore, in optimizing our system to ensure its broadest utility, it is desirable to define functional aptamer elements that minimally impact the maximum protein output attainable for any given 5′UTR while preserving regulatory dynamic range. In addressing this need, we surmised that the observed translation inhibition caused by aptamer **5-1.2** could be due to: (i) the stability of its stem structure; and/or (ii) the putative start codon and its sequence context within the aptamer[40]. To test the effect of secondary structure stability on translation levels, we scrambled the first 25 bases of **5-1.2** to obtain **5-1.2half** (Table 2.S4). This removed substantial predicted secondary structure within the 5′UTR, but left the start codon sequence context intact. Placing **5-1.2half** upstream of vYFP minimally increased basal vYFP expression compared to **5-1.2**, suggesting that the stability of the aptamer stem was not the major determinant of reduced expression of the downstream reporter (Fig. 2.S4).

The **5-1.2** aptamer contains an AUG start codon followed immediately by a stop codon and a second start codon that is out of frame with the first, but in frame with the downstream reporter gene. Because previous studies have demonstrated that short upstream ORFs can inhibit translation of a downstream ORF[40], we investigated whether the sequence ^{43}AUGUGAUG^{50} within a predicted loop region of **5-1.2** (Fig. 2.1A) could be primarily responsible for the translation efficiency of a downstream ORF. While keeping the rest of aptamer **5-1.2** constant, we introduced an A→G mutation at position 48 in the loop region above to generate **aptamer 5-1.4d** containing the sequence ^{43}AUGUGGUG^{50}. This change simultaneously eliminated the stop codon and the second start codon within the aptamer loop. Replacing **5-1.2** with **5-1.4d** (and placing the regulated ORF in frame with the single start codon in **5-1.4d**) resulted in modestly higher expression levels, but maintained TetR-dependent regulation (Fig. 2.6A). To further

increase expression of the aptamer-regulated ORF, we systematically reduced aptamer stem strength by successively eliminating base-pair interactions at the stem base while retaining sequence downstream of the initiator AUG within the aptamer (Table 2.S4). When we used these aptamers (**5-1.30, 5-1.31, 5-1.32, 5-1.33**) to control translation as described previously, we measured a large increase in maximal expression level. Furthermore, TetR-dependent regulation was preserved at the ≥ 80% repression level previously observed (Fig. 2.6A). To determine the expected upper limit of expression when using 5-1.4d, we scrambled the first 25 bases of the aptamer to remove substantial predicted secondary structure (**5-1.4dhalf**). When used to control gene expression, this modification produced a maximal expression level comparable to that of **5-1.31**, but with no TetR-dependent regulation (Fig. 2.6A), indicating that further destabilization of the aptamer was unlikely to yield further increases in maximal expression levels. Overall, replacing **5-1.2** with **5-1.31** increases maximal expression by ~25-fold, and with no adverse impact on the magnitude of Dox inducible, TetR-dependent regulation.

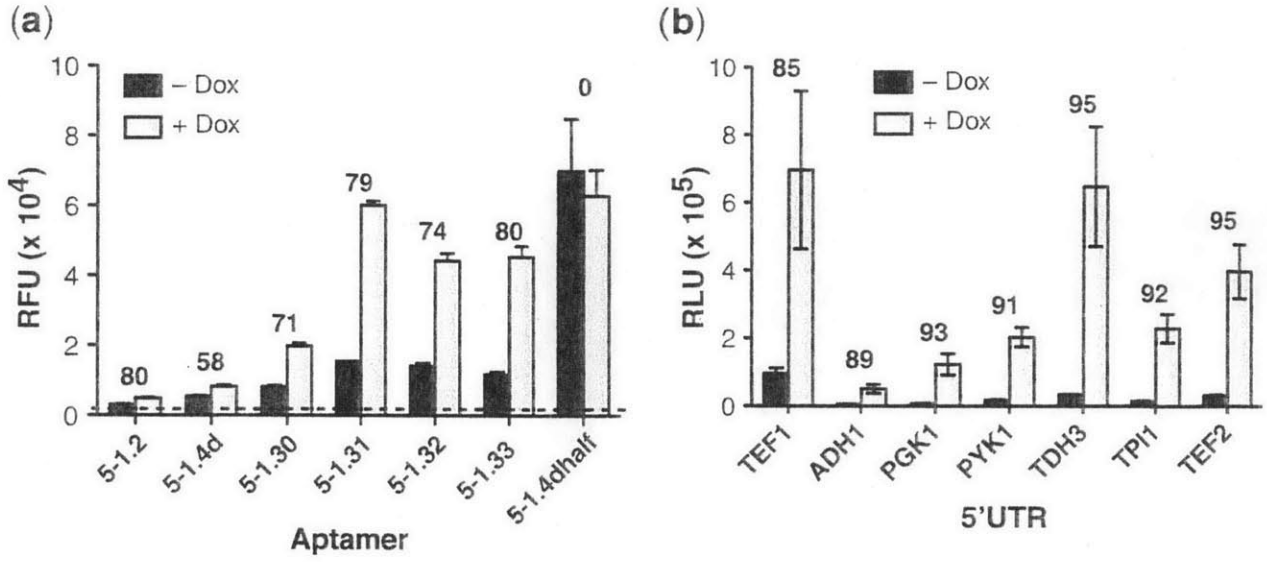

Figure 2.6. Systematic optimization of an aptamer variant that improves maximal expression levels and demonstration of robust regulation when an aptamer is placed in different 5′UTR contexts. (**A**) Several aptamer variants generated by introducing a point mutation within Motif 1 and reducing the stem length of **5-1.2** were tested for their ability to regulate vYFP. These modifications allowed for expression level maximization while preserving regulatory range. (**B**) Aptamer **5-1.31** was placed within the context of several endogenous 5′-UTRs and used to control FLuc expression. In (**A**) and (**B**), numbers above the bars indicate the percent repression observed in the - Dox condition relative to the + Dox condition. In all cases, cells expressed TetR. These data show that regulation is achieved in all contexts tested.

2.9. Assessing the modularity of the system

Lastly, to validate the utility of the TetR-aptamer system in different native 5′UTR contexts, we transcriptionally fused **5-1.31** with several yeast promoters of varied strength. We retained the sequence downstream of the aptamer used in previous experiments. Despite the expected variation in the maximal expression levels due to differences in promoter strength, luciferase expression was regulated with over 85% repression in all contexts (Fig. 2.6B and Table 2.S5).

We found similar, albeit slightly reduced, regulation when using vYFP as the reporter. Similarly, **5-1.2** in the same 5′UTR contexts yielded strong, inducible repression. In these experiments, the TDH3 5′UTR context produced 97% repression when paired with the **5-1.2** aptamer. This led us to hypothesize that the TDH3 5′UTR/**5-1.2** combination might be a superior sequence context module that increases the regulatory efficiency of our TetR-aptamer system. To test this, we fused the combined TDH3 5′UTR–**5-1.2** sequence downstream of the other 5′UTR contexts tested above. While we did not achieve equally high repression levels as in the original TDH3 5′UTR context, robust regulation was nonetheless seen in every case (Table 2.S5). Altogether, these data strongly support the feasibility of using our optimized TetR-aptamer regulatory system in endogenous 5′UTR contexts of diverse RNA sequence, even when little information about the targeted 5′UTR's size, sequence and structural characteristics is available.

2.10. Polysome analysis

Our data from cell-free translation and qPCR experiments firmly support a post-transcriptional regulatory mechanism that does not act via a decrease in mRNA levels. Therefore, using polysome analysis, we sought to define whether the aptamer-TetR interaction modulates initiation or some downstream step in the translation process. For these experiments, vYFP regulated by **5-1.2** was used as a representative target transcript. If TetR interaction with **5-1.2**-vYFP mRNA predominantly inhibits translation initiation, in the absence of Dox this should reduce **5-1.2**-vYFP mRNA ribosome occupancy and lead to the transcript's accumulation in non-polysomal fractions. Conversely, disrupting the TetR-**5-1.2** interaction by adding Dox would result in more efficient translation initiation and increased accumulation of **5-1.2**-vYFP mRNA in the polysomal fractions. However, we consistently found no significant difference in **5-1.2**-vYFP mRNA ribosome occupancy between the condition where the TetR-**5-1.2** interaction is

intact (- Dox) or disrupted (+ Dox) (Figs. 2.7 and 2.S5). This suggests that either: (i) standard polysome profiling is insufficiently sensitive to detect a small but functionally important shift in ribosome occupancy that may be occurring; and/or (ii) the aptamer-TetR interaction inhibits translation mainly downstream of initiation. Our polysome profiling results indicate that both the translationally repressed and actively translated **5-1.2**-vYFP mRNA are similarly associated with the polysomal fractions. While polysome-associated mRNAs are generally considered to be actively translated, some of these mRNAs are known to be translationally repressed[82–86]. The specific molecular details underlying repression of polysome-associated mRNA are still generally unclear, but they could involve mRNA decapping, mRNA deadenylation, altered elongation kinetics, nascent polypeptide degradation and impaired ribosome release[87]. Understanding exactly how the aptamer-TetR interaction and its disruption by Dox facilitate differential partitioning of aptamer-containing target transcripts between translationally repressed and actively translated pools is an intriguing problem that will require more detailed study beyond the scope of the present work. However, such efforts could provide additional insight into regulation mechanisms downstream of translation initiation, which are increasingly being recognized to be of broad biological importance[83,88–90]. Furthermore, this knowledge can further enable engineering of improved versions of our presently described system for inducibly regulating protein expression.

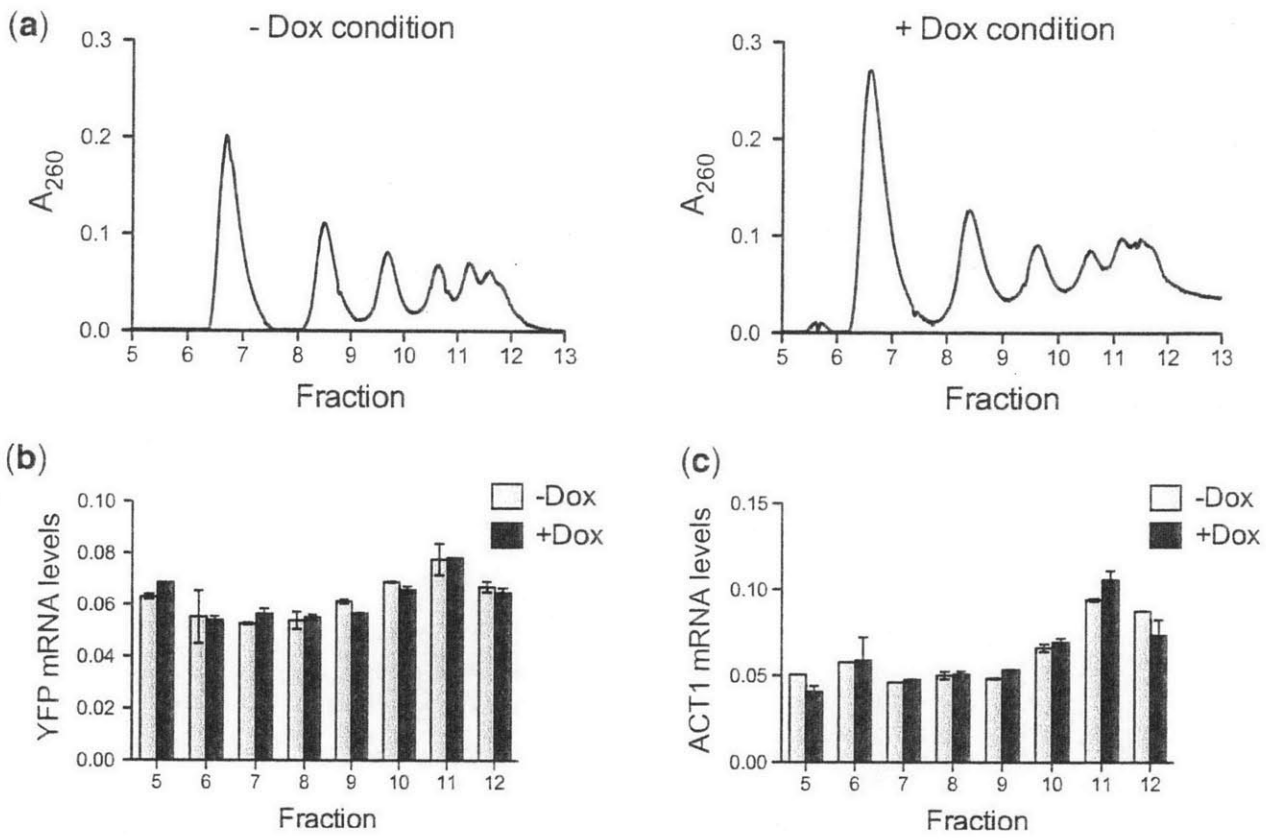

Figure 2.7. Polysome profiles of aptamer-containing mRNA indicate regulation is independent of translation initiation. (**A**) Polysomes were fractionated from yeast expressing both TetR and a **5-1.2**-containing vYFP reporter mRNA, which were grown in the absence or presence of Dox. Polysome profiles for both the - Dox and + Dox growth conditions are shown. (**B**) qPCR measurements of the relative amounts of reporter mRNA within each polysome fraction, both for the - Dox and + Dox conditions. (**C**) qPCR measurements of relative amount of ACT1 mRNA in each polysome fraction, under - Dox and + Dox conditions. For both (**B**) and (**C**) error bars indicate the range of values for technical replicates. The data are representative of two independent, biological replicates.

2.11. Summary

We have developed a new system for directly and transcript-specifically controlling protein expression that requires a small number of defined, genetically-encoded components and no

knowledge of transcriptional regulation. Induction is achieved using inexpensive, cell permeable, and well-tolerated tetracycline analogs. Both the magnitude and stringency of the regulation attainable permit modulation of a survival phenotype in yeast under highly selective growth conditions. indicating that this approach is sufficiently robust so as to be biologically useful. Furthermore, this system functions equally well within the context of several natural 5′UTRs, and we have optimized it such that translational repression due to the presence of structured RNA within the 5′UTR is minimal. As key aspects of translation are well conserved among eukaryotes, we envision this system being broadly applicable in controlling gene expression. This is evidenced by our demonstration of translation control in both yeast and mammalian contexts. We anticipate this system will be especially useful in organisms with poorly understood transcriptional regulatory mechanisms and few inducible gene expression options. Overall, we have described a modular, inducible framework for biological control that provides a direct interface with protein synthesis. We foresee that the minimal nature of this system will enable investigation of previously intractable problems in cell biology, such as the role of gene-specific translational regulation in early development and neurobiology. Additionally, the modularity and host cell independence of the system make it particularly suited for use in the construction of synthetic biological circuits that operate independently of transcription.

2.12. Experimental procedures

RNA aptamers. Secondary structure predictions, such as the one shown in Fig. 2.1B, were made using mfold[91]. Equilibrium binding constants were determined using a cytometric bead binding assay as described[1].

Plasmid construction. Unless otherwise noted, cloning procedures were performed by standard techniques. Reporter plasmids were based on YCp22FL1, which contains an *S. cerevisiae*

TEF1 promoter driving firefly luciferase (FLuc)[92]. Alternative reporter plasmids containing Venus yellow fluorescent protein (vYFP)[93] were created by replacing the XhoI/XbaI FLuc fragment of YCp22FL1 with vYFP PCR amplified using primers BJBOL268/269 (Table 2.S2). PCR (25 cycles) was performed with Phusion DNA polymerase (New England Biolabs) according to the manufacturer's instructions, annealing at 55 °C for 30 seconds. Aptamer-encoding sequences were inserted by annealing oligonucleotide pairs (Table 2.S1), extending with Klenow fragment (3′→5′ exo$^-$, New England Biolabs), digesting with BglII/XhoI and ligating to BamHI/XhoI-digested plasmid.

The *URA3*-marked TetR(B) plasmid, YCpSUP-TetR, was created by PCR amplifying *tetR(B)* fused to sequences encoding 5′T7 and 3′His_6 tags from pET24-TetR[1], using primers BJBOL111/112, digesting with SalI/AvrII and ligating to XhoI/XbaI-digested YCpSUP-IRF1[92]. The *HIS3*-marked TetR plasmid pRS413-TetR was created by subcloning the ClaI fragment from YCpSUP-TetR containing the *PGK1/GAL* promoter through the *PGK1* terminator into ClaI-digested pRS413 (American Type Culture Collection, ATCC). The *URA3*-marked revTetR plasmid pSG116 was created by PCR amplifying *revTetR-S2*[80] (without affinity tags or transcriptional activator domains) using primers BJBOL295/296 and cloning as described for YCpSUP-TetR. The *LEU2*-marked plasmid pSG95, encoding *tetR(B)* under the control of the *TDH3* (GPD) promoter, was created by inserting *tetR(B)* into pRS415-GPD (ATCC) by yeast gap repair cloning[94]. Yeast strains expressing chromosomally-integrated reporters were constructed by cloning aptamer-regulated reporter constructs into pRS404 (ATCC), linearizing by digestion with EcoRV, transforming into W303-1B yeast and selecting on media lacking tryptophan. Reporter integration at the *TRP1* locus was confirmed by PCR with primers BJBOL411/412 and BJBOL410/413, which flank the 5′ and 3′ integration sites, respectively.

***In vitro* translation experiments.** Templates for *in vitro* transcription containing **5-1.2** and **5-1.2m2** were amplified from yeast reporter plasmids using SG151/SG137 and SG291/SG137, respectively (Table 2.S3). All other DNA templates were constructed by PCR amplifying the FLuc gene from YCp22FL1 with the reverse primer SG137 (encoding an A_{60} tail) and a combination of overlapping forward primers (Table 2.S3) to generate the desired 5'UTR with an upstream T7 promoter. PCR mixtures contained the outermost primers at 0.5 μM and all additional overlapping primers at 0.05 μM. Assembly PCR (30 cycles) was performed with Phusion DNA polymerase (New England Biolabs) according to the manufacturer's instructions, annealing at 65 °C for 20 seconds and extending at 72 °C for 25 seconds. Unpurified PCR products were transcribed with the MEGAScript T7 kit (Ambion), precipitated with an equal volume of 7.5 M LiCl/50 mM EDTA, washed with 70% ethanol and redissolved in water. mRNA products were capped with the ScriptCap m^7G capping system (Epicentre Biotechnologies) or the Vaccinia Capping System (New England Biolabs), precipitated with an equal volume of 7.5 M LiCl/50 mM EDTA, washed with 70% ethanol and redissolved in RBB (50 mM Tris-HCl, pH 8.0, 50 mM KCl, 5 mM $MgCl_2$, 5% (v/v) glycerol, 0.05% (v/v) Tween-20). mRNA concentrations were determined by measuring absorbance at 260 nm on a Nanodrop ND-1000 spectrophotometer (ThermoFisher Scientific).

For cell-free translation with rabbit reticulocyte lysate (RRL), the mRNA concentration was adjusted to 133 nM in RBBD (RBB plus 1 mM dithiothreitol and 10 μg/mL bovine serum albumin, New England Biolabs) and refolded by heating at 65°C for 2 minutes and incubating at room temperature for 10 minutes. TetR(B) with N-terminal T7 and C-terminal His_6 tags, and revTetR-S2 with a C-terminal His_6 tag were purified from *E. coli* as previously described[1]. TetR and revTetR-S2 dilutions were prepared in RBBD with or without 22 μM Dox. For each

translation reaction, 1.5 μL mRNA was mixed with 2 μL TetR or revTetR-S2, allowed to equilibrate for 30 minutes at room temperature, and then mixed with 0.5 μL of a complete amino acid solution (0.5 mM each amino acid, Promega) and 6 μL of nuclease-treated RRL (Promega). The final reactions contained 20 nM mRNA, 0–300 nM repressor protein and 25 μM amino acids in 0.6X RRL. Reactions were incubated at 30 °C for 20 minutes. For the screening experiment in Fig. 2.2B, reactions were stopped by adding 200 μL Stop Buffer (20 mM GlyGly-NaOH, pH 7.8, 8 mM $Mg(OAc)_2$, 0.13 mM EDTA, 500 μM cycloheximide). FLuc activity was determined by mixing 90 μL of the reaction mixture with 40 μL FLuc Assay Buffer (20 mM GlyGly-NaOH, pH 7.8, 90 mM DTT, 8 mM $Mg(OAc)_2$, 0.13 mM EDTA, 1.5 mM ATP, 0.8 mM coenzyme A, 1.4 mM D-luciferin) in a 96-well microplate and measuring luminescence with a Spectramax L plate reader (Molecular Devices). Translation activity was calculated as the intensity of FLuc signal. For the TetR and RevTetR(S2) titration experiments in Figs. 2.2C and 2.5A, a *Renilla* luciferase mRNA lacking an aptamer was included in each translation reaction as a reference. Reactions were performed as described above, and stopped with the addition of 200 μL of 500 μM cycloheximide in water. 4 μL of the reaction mixture was mixed with 20 μL of passive lysis buffer (Promega) and dual luciferase activity was measured by the sequential addition of 100 μL DLB1 (75 mM HEPES-K, 20 mM DTT, 4 mM $MgSO_4$, 0.1 mM EDTA, 0.53 mM ATP, 0.27 mM coenzyme A, 0.47 mM D-luciferin, pH 8.0) and 100 μL DLB2 (15 mM $Na_4P_2O_7$, 7.5 mM sodium acetate, 10 mM EDTA, 400 mM Na_2SO_4, 1% (v/v) methanol, 10 μM 2-(4′-(dimethylamino)phenyl)-6-methyl-benzothiazole, 5 mM KI, 12 μM benzyl coelenterazine, pH 5.0). Translation was calculated as the ratio of FLuc (aptamer-regulated) to RLuc (aptamer-independent) signal.

Yeast inducible expression assays. *S. cerevisiae* W303-1B cells harboring both a repressor (YCpSUP-TetR or pSG116) and reporter plasmid were grown to saturation at 30 °C in Synthetic Defined Media #1 (SD1) (6.7 g/L YNB, 20 mg/L adenine, 30 mg/L lysine, 100 mg/L leucine, 20 mg/L histidine) + 20 g/L glucose. Cells were diluted 1:80 into SD1 + 20 g/L raffinose and grown for four hours. Glucose (to repress TetR or revTetR-S2 expression) or galactose (to induce TetR or revTetR-S2 expression) was added to 20 g/L, and cells were grown 16 hours at 30 °C with shaking before measurement. For FLuc activity measurements, luminescence values were normalized to the OD_{600} of each culture as determined using a Spectramax M2 plate reader (Molecular Devices). Eighty microliters of yeast culture was added to 20 μL of 5X Passive Lysis Buffer (Promega) and incubated for 10 seconds. Ten microliters of this suspension was added to 100 μL of FLuc Assay Buffer in a 96-well microplate and measured as described for cell-free translation experiments. For vYFP measurement by flow cytometry, cells were grown as above and analyzed on a C6 Flow Cytometer (Accuri). For each sample, 5×10^4 events were captured and vYFP fluorescence was measured in the FL1 channel.

Plate-based yeast selection to identify functional aptamers. *S. cerevisiae* W303-1B cells harboring both pRS413-TetR and a *URA3* plasmid were grown to saturation at 30 °C in Synthetic Defined Media #2 (SD2) (6.7 g/L YNB, 20 mg/L adenine, 30 mg/L leucine, 20 mg/L lysine, 50 mg/L uracil) + 20 g/L glucose. Cells were diluted 1:40 into SD2 + 20 g/L raffinose and grown for 4 hours. Galactose (to induce TetR expression) was added to 20 g/L and cells were grown for 4 hours. Cultures were washed once in SD2 without uracil and serially diluted 10-fold into SD2 without uracil. Cell dilutions were spotted onto agar plates as indicated and grown for three days before visualization. Uracil dropout plates for positive selection contained

SD2 without uracil, 20 g/L agar and 20 g/L galactose, with or without 1 μM aTc. Negative selection plates contained SD2, 20 g/L agar, 20 g/L galactose and 0.25 g/L 5-FOA.

cDNA preparation and quantitative PCR (qPCR). Cells were grown to mid-exponential phase as described above. Total RNA was extracted using an RNeasy Mini Kit (Qiagen). Contaminating genomic and plasmid DNA were removed by treating with 10 U of TURBO DNase (Ambion) for 2 hours at 37 °C, followed by phenol/chloroform extraction and LiCl precipitation. Two micrograms of RNA was reverse transcribed to cDNA using random hexamer primers (Fermentas) and RevertAid M-MuLV reverse transcriptase (Fermentas). The reaction mixture was incubated at 25 °C for 10 min, 42 °C for 60 min, then 70 °C for 10 min. cDNA was diluted 1:100 in water to produce a template solution. All qPCR reactions were performed in a PTC-200 Peltier Thermal Cycler (MJ Research) equipped with a Chromo4 Detector (Bio-Rad). FLuc cDNA levels were quantitated and compared to levels of the endogenous *ACT1* gene. The primer pairs used for FLuc and *ACT1* amplification were BJBOL233/234 and ACT1F/ACT1R, respectively (Table 2.S2). The qPCR reactions were performed in 20 μL and contained: 1x Standard Taq Buffer (New England Biolabs), 2.5 mM $MgCl_2$, 200 μM each dNTP, 100 nM each primer, 0.4x SYBR Green I (Invitrogen), 0.4 U Taq DNA polymerase (New England Biolabs) and 10 μL template solution. Reactions were performed by denaturing at 95 °C for 2 minutes, followed by 40 cycles of: 95 °C for 30 s, 60 °C for 30 s and 72 °C for 30 s. SYBR Green I fluorescence was measured at the end of each cycle, and relative quantitation was performed as described[95].

Preparation of polysome fractions. *S. cerevisiae* W303-1B cells harboring both a repressor (YCpSUP-TetR) and an integrated **5-1.2**-vYFP reporter were grown to saturation at 30 °C in

SD1 + 20 g/L glucose. Cells were diluted 1:300 into 250 mL SD1 + 20 g/L galactose, in the presence or absence of 22 μM Dox, and grown until the cultures reached OD_{600} = 0.6. Prior to cell lysis, flow cytometry analysis was performed to confirm that vYFP expression under induced and uninduced conditions was identical to previous experiments. Once the proper cell density was reached, cyclohexamide was added to 0.1 mg/mL and incubated at 30 °C with shaking for two minutes. Cells were pelleted by centrifugation (12,000*g*, 4 °C, 5 minutes) and resuspended in 40 mL of cold Polysome Lysis Buffer (PLB, 20 mM HEPES-K, pH 7.4, 2 mM magnesium acetate, 100 mM potassium acetate, 0.1 mg/ml cycloheximide, 3 mM DTT, 10 ml/L Triton X-100). Cells were pelleted again (2,200*g*, 4 °C, 5 minutes), resuspended in 30 mL PLB, and pelleted again. The supernatant was removed, and cells were weighed. For each gram of cell mass, 1.5 mL of PLB and 5 g of 0.5-mm glass beads were added. Cells were lysed by vortexing for two minutes. The crude lysate was centrifuged (2,200*g*, 4 °C, 5 minutes) and the supernatant subsequently centrifuged again (15,800*g*, 4 °C, 20 minutes). The final supernatant was frozen in liquid nitrogen, and stored at -80 °C until fractionation.

For polysome fractionation, 15 A_{260} units of yeast lysate were loaded onto a 10–50% sucrose gradient in Polysome Gradient Buffer (20 mM HEPES-K, pH 7.4, 2 mM magnesium acetate, 100 mM potassium acetate, 0.1 mg/ml cycloheximide, 3 mM DTT). Samples were centrifuged in a Beckman SW-41 rotor for three hours at 35,000 rpm at 4 °C. Individual fractions were collected on a Gradient Station (BioComp Instruments) and frozen at -20 °C until further analysis.

cDNA preparation and qPCR analysis of polysome fractions. RNA from each polysome fraction was isolated with an RNeasy kit (QIAGEN) according to the manufacturer's protocol. Prior to RNeasy purification, 1 fmol of *in vitro* transcribed FLuc RNA was added to each sample

as an internal control. RNA was DNase treated (TURBO DNA-free kit, Ambion) and reverse transcribed as described above for other qPCR experiments. Quantitative PCR measurements were performed with the PrimeTime primer and 5′ hydrolysis probe sets listed in Table 2.S6 (Integrated DNA Technologies). qPCR measurement of FLuc was duplexed with measurement of either vYFP or ACT1. Each reaction included 1x Thermopol Buffer (New England Biolabs), 0.2 mM dNTPs, 500 nM each primer, 250 nM each probe, 1 μL cDNA, and 0.1 μL Taq DNA polymerase (New England Biolabs). Thermocycling was performed on a Roche LightCycler 480 II for 40 cycles according to the following protocol: initial denaturation: 95 °C for 20 seconds; denature: 95 °C for 3 seconds; anneal/extend: 60 °C for 30 seconds; fluorescence measurement. Raw data was color compensated, and threshold cycle values were determined. Standard curves were generated from purified vYFP, ACT1, or FLuc DNA to calculate amplification efficiencies and to ensure that FLuc internal standard and cDNA template concentrations were within the linear detection range. The vYFP or ACT1 qPCR measurement from each sample was then normalized to the duplexed FLuc qPCR measurement from the same sample.

Western blot analysis. Yeast lysates were prepared by heating cell pellets in 2x Laemmli sample buffer for 10 minutes at 95 °C. Proteins were then separated by 12% SDS-PAGE and transferred to a PVDF membrane. TetR was detected using an anti-His_6 tag antibody (Abcam #ab18184). An anti-GAPDH antibody (Genscript #A00191) was used to quantify glyceraldehyde 3-phosphate dehydrogenase levels as a loading control.

2.13. Tables

Table 2.S1. Aptamer sequences.

Aptamer	**Parent aptamer**	**RNA sequence**
5-1.13	5-1[a]	5′GGGAGCUCAGAAUAAACGCUCAACUCCUGUAGUGAAGGCAGAGA AAGGUCGAUACGGACGGAAUGUGAUGGCCUUCGACAUCAGGCCCGG AUCCGGC
5-11.13	5-11[a]	5′GGGAGCUGAGAAUAAACGCUCAAAACUAGCAGGCAGAGAAGAGU GGGUGCGACCACAGGAUGUUAUGGCCUGUUCGACAUCAGGCCCGGA UCCGGC
5-14.13	5-14[a]	5′GGGAGCUGAGAAUAAACGCUCAACAGGAAACAGCAAGACAAACG AUGGGGAGCGUAAGACUGCGAGUGUCGGAUUCGACAUCAGGCCCG GAUCCGGC
5-18.13	5-18[a]	5′GGGAGCUGAGAAUAAACGCUCAAUAGGGAGAGAACUGUGUCAGA AUGUAGUGAACCAGACACGGAGUGGAGUAUUCGACAUCAGGCCCG GAUCCGGC
5-29.13	5-29[a]	5′GGGAGCUCAGAAUAAACGCUCAACUUGCUGCAGAGGGUCGAGAA UAUGUGUGACACUGCGUCGACGGGUUAAGUUCGACAUCAGGCCCG GAUCCGGC
5-1.2	5-1.13	5′GGAUCCAGGCAGAGAAAGGUCGAUACGGACGGAAUGUGAUGGCC UGGAUCCAAA
5-1.2m2	5-1.2	5′GGAUCCAGGCAGUGUAAGGUCGAUACGGACGGAAUGUGAUGGCC UGGAUCCAAA

[a] From Belmont and Niles[1].

Table 2.S2. Oligonucleotides used for cloning and qPCR.

Oligonucleotide	**DNA sequence**
BJBOL111	5′GCGGTCGACTACAACATGGCTAGCATGACTGGTGGAC
BJBOL112	5′GCGCCTAGGTCAGTGGTGGTGGTGGTGGTGC
BJBOL233	5′ATTTATCGGAGTTGCAGTTGCGCC
BJBOL234	5′AACAAACACTACGGTAGGCTGCGA
BJBOL268	5′GCGCTCGAGAACATATGTCTAAAGGTGAAGAATTAT
BJBOL269	5′GCGTCTAGATTATTTGTACAATTCATCCATA
BJBOL295	5′GCGGTCGACTACAACATGTCTAGATTAGATAAAAGTA
BJBOL296	5′GCGCCTAGGTCAAGACCCACTTTCACATTTA
BJBOL410	5′GAGCAGATTGTACTGAGAGTGCACC
BJBOL411	5′TACGCATCTGTGCGGTATTTCACAC
BJBOL412	5′CAGAGACCAATCAGTAAAAATCAACGG
BJBOL413	5′GATCTTTTATGCTTGCTTTTCAAAAGGCC
ACT1F	5′GCCTTGGACTTCGAACAAGA
ACT1R	5′CCAAACCCAAAACAGAAGGA

Table 2.S3. Overlapping oligonucleotides used for PCR assembly of DNA templates for *in vitro* transcription.

Product	**Oligonucleotide**	**DNA sequence**
5-1.13-**FLuc-A60**	SG178	5′TGGCCTTCGACATCAGGCCCGGATCCGGCAACACAAAACTCGAGAACATATGGAAGACGC
	SG179	5′GTAGTGAAGGCAGAGAAAGGTCGATACGGACGGAATGTGATGGCCTTCGACATCAGGCC
	SG180	5′ATCTAGGGAGCTCAGAATAAACGCTCAACTCCTGTAGTGAAGGCAGAGAAAGGTCGATAC
	SG181 [a]	5′*CTAATACGACTCACTATAGGG*AATTATCTAGGGAGCTCAGAATAAACGCTCAA
5-11.13-**FLuc-A60**	SG184	5′GGCCTGTTCGACATCAGGCCCGGATCCGGCAACACAAAACTCGAGAACATATGGAAGACG
	SG185	5′AGCAGGCAGAGAAGAGTGGGTGCGACCACAGGATGTTATGGCCTGTTCGACATCAGGCCC
	SG186	5′GGGAATTATCTAGGGAGCTGAGAATAAACGCTCAAAACTAGCAGGCAGAGAAGAGTGGG
	SG187 [a]	5′*CTAATACGACTCACTATAGGG*AATTATCTAGGGAGCTGAGAATAAACGCT
5-14.13-**FLuc-A60**	SG191	5′GGATTCGACATCAGGCCCGGATCCGGCAAAACACAAAACTCGAGAACATATGGAAGACGC
	SG192	5′AACAGCAAGACAAACGATGGGGAGCGTAAGACTGCGAGTGTCGGATTCGACATCAGGCCC
	SG193	5′AATTATCTAGGGAGCTGAGAATAAACGCTCAACAGGAAACAGCAAGACAAACGATGGGGA
	SG194 [a]	5′*CTAATACGACTCACTATAGGG*AATTATCTAGGGAGCTGAGAATAAACGCTCAA
5-18.13-**FLuc-A60**	SG251	5′CGGCAAAACACAAAACTCGAGAACATATGGAAGACGCCAAAAACATAAAGAAAGGCCCGG
	SG253	5′AACTGTGTCAGAATGTAGTGAACCAGACACGGAGTGGAGTATTCGACATCAGGCCCGGAT
	SG254	5′GGGAGCTGAGAATAAACGCTCAATAGGGAGAGAACTGTGTCAGAATGTAGTGAACCAGAC
	SG255 [a]	5′*CTAATACGACTCACTATAGGG*AATTATCTAGGGAGCTGAGAATAAACGCTCAATAGG
5-29.13-**FLuc-A60**	SG256	5′AGTTCGACATCAGGCCCGGATCCGGCAACACAAAACTCGAGAACATATGGAAGACGCCAA
	SG257	5′GAGGGTCGAGAATATGTGTGACACTGCGTCGACGGGTTAAGTTCGACATCAGGCCCGGAT
	SG259	5′AGGGAGCTCAGAATAAACGCTCAACTTGCTGCAGAGGGTCGAGAATATGTGTGACACTGC
	SG261 [a]	5′*CTAATACGACTCACTATAGGG*AATTATCTAGGGAGCTCAGAATAAACGCTCAACT
5-1.2-**FLuc-A60**	SG151 [a]	5′*CTAATACGACTCACTATAGGG*AATTATCTAGGATCCAGGCAGAGAAAGG

5-1.2m2-**FLuc-A60**	SG291 [a]	5′*CTAATACGACTCACTATAGGG*AATTATCTAGGATCCAGGCAGTGTAAGGTCGA
(Universal reverse primer)	SG137	5′TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTACAATTTGGACTTTCCGCCCTTCTTGG

[a] Indicates outermost forward primer containing T7 promoter sequence (in italics).

Table 2.S4. 5′UTR sequences resulting when 5-1.2 **and its variants are transcriptionally fused to the TEF1 promoter.**

Aptamer	**RNA sequence**
5-1.2	AAUUAUCUAGGAUCCAGGCAGAGAAAGGUCGAUACGGACGGAAUGUGAUGGCCUGGAUCCAACACAAAACUCGAGAACAU*AUG*
5-1.2half	AAUUAUCUAAGAGAGAGACACCAAAGCGUGUGUGCGGACGGAAUGUGAUGGCCUGGAUCCAACACAAAACUCGAGAACAU*AUG*
5-1.4d	AAUUAUCUAGGAUCCAGGCAGAGAAAGGUCGAUACGGACGGAAUGUGGUGGCCUGGAUCCAAACACAAAACUCGAGAACAU*AUG*
5-1.4dhalf	AAUUAUCUAAGAGAGAGACACCAAAGCGUGUGUGCGGACGGAAUGUGGUGGCCUGGAUCCAAACACAAAACUCGAGAACAU*AUG*
5-1.30	AAUUAUCUAAGAUCCAGGCAGAGAAAGGUCGAUACGGACGGAAUGUGGUGGCCUGGAUCCAAACACAAAACUCGAGAACAU*AUG*
5-1.31	AAUUAUCUAAAAUCCAGGCAGAGAAAGGUCGAUACGGACGGAAUGUGGUGGCCUGGAUCCAAACACAAAACUCGAGAACAU*AUG*
5-1.32	AAUUAUCUAAAUUCCAGGCAGAGAAAGGUCGAUACGGACGGAAUGUGGUGGCCUGGAUCCAAACACAAAACUCGAGAACAU*AUG*
5-1.33	AAUUAUCUAAAUACCAGGCAGAGAAAGGUCGAUACGGACGGAAUGUGGUGGCCUGGAUCCAAACACAAAACUCGAGAACAU*AUG*

Table 2.S5. Summary of the regulatory behavior observed when using 5-1.2 **and** 5-1.31 **within the endogenous yeast 5′UTRs indicated. FLuc and vYFP were used as reporters.**

	%Repression					
Promoter/ 5′UTR	**5-1.2 FLuc**	**5-1.31 vYFP**	**5-1.31 FLuc**	**+TDH3 5′UTR 5-1.2 FLuc**	**Length before aptamer (nt)**	**%(G+C)**
TEF1	87	85	85	95	9	11
ADH1	77	77	89	84	44	34
PGK1	81	84	93	85	46	24
PYK1	94	84	91	80	33	27
TDH3	97	84	95	N.D.	45	29
TEF2	96	88	92	89	28	29
TPI1	61	71	95	94	36	25

Table 2.S6. Oligonucleotide primers and probes used in qPCR of polysome fractions.

	Sequence
Firefly luciferase	
Primer #1	5′TCCTCTGACACATAATTCGCC
Primer #2	5′GCTATTCTGATTACACCCGAGG
Probe	5′HEX/TCCAGATCC/ZEN/ACAACCTTCGCTTCAAA/IABkFQ
vYFP	
Primer #1	5′CACCTTCAAACTTGACTTCAGC
Primer #2	5′TGTGTTTTGCTAGATACCCAGATC
Probe	5′6-FAM/TTTCTTGAA/ZEN/CATAACCTTCTGGCATGGC/IABkFQ
ACT1	
Primer #1	5′GGCAGATTCCAAACCCAAAAC
Primer #2	5′TCGAACAAGAAATGCAAACCG
Probe	5′Cy5/ACGAAAGATTCAGAGCCCCAGAAGC/IAbRQSp

2.14. Supplementary figures

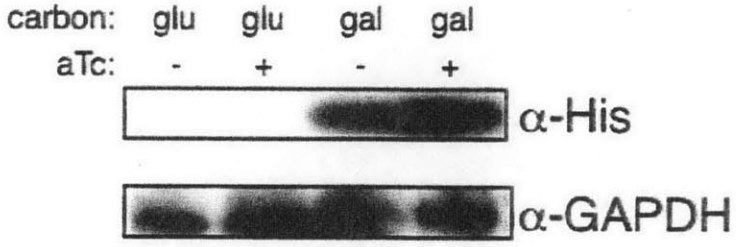

Figure 2.S1. Western blot measurement of TetR protein. The addition of aTc does not decrease TetR expression. Yeast cells were grown in the presence or absence of aTc plus either glucose or galactose (to repress or induce TetR expression, respectively). Anti-His_6 tag antibody was used to detect His_6-tagged TetR. Sample loading was verified by GAPDH detection with an anti-GAPDH antibody.

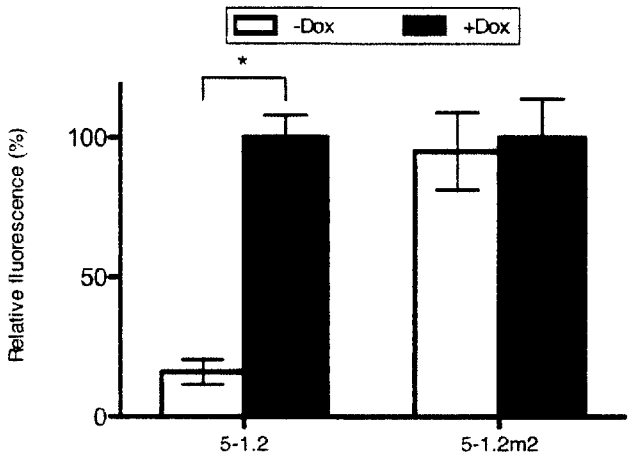

Figure 2.S2: Constitutively expressed TetR represses translation. TetR expressed constitutively in yeast from the *TDH3* promoter is able to repress translation of vYFP controlled by **5-1.2**, but not **5-1.2m2**. The addition of Dox relieves vYFP repression. Data represent the mean ± s.d. of six experiments. A two-tailed, unpaired t-test was used to calculate the significance ($\alpha = 0.005$) of the difference between induced and uninduced conditions. *, $P = 7.1 \times 10^{-10}$.

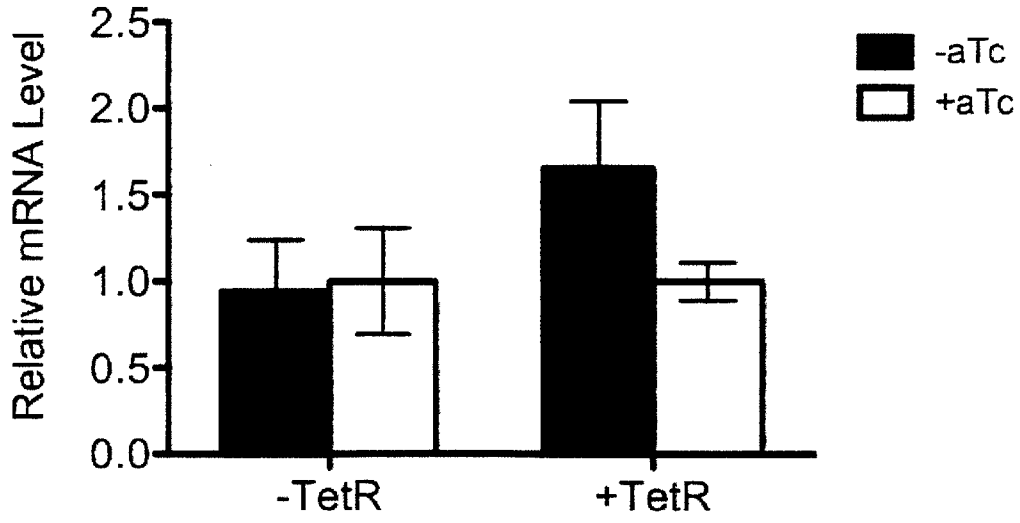

Figure 2.S3: Quantitative PCR measurement of FLuc mRNA. The addition of aTc does not increase the steady-state level of FLuc mRNA, indicating that regulation by

TetR/aTc occurs at the level of translation. For each condition (with or without TetR), values indicate the amount of FLuc mRNA in the presence of 1 μM aTc relative to the amount of mRNA in the absence of aTc. The data represent the mean ± s.d. of six experiments.

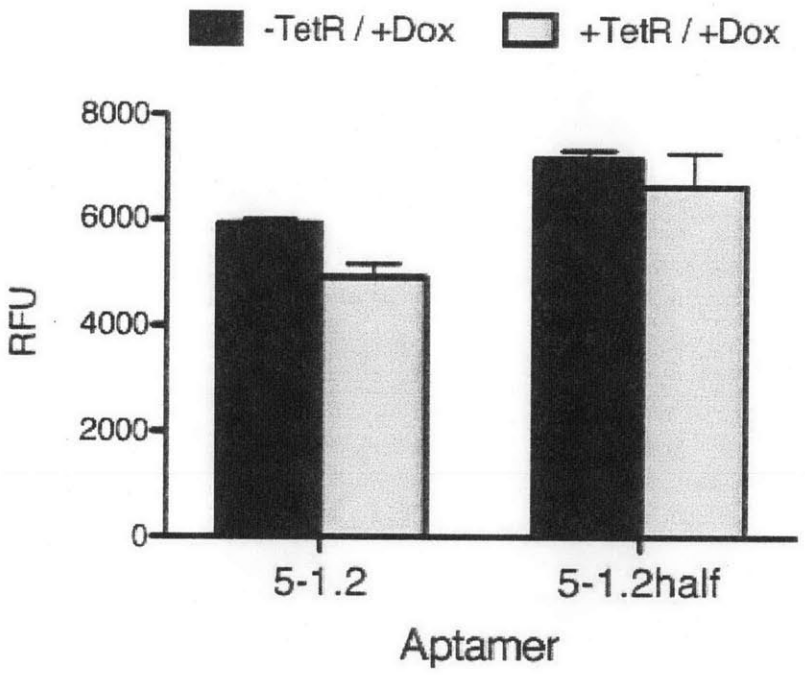

Figure 2.S4 Removal of 5-1.2 aptamer structure has little effect on basal expression. Flow cytometry measurements show expression levels of aptamer-regulated vYFP. The aptamer located within the 5′UTR of vYFP and the expression status of TetR are indicated. In all cases, cells were grown in the presence of Dox.

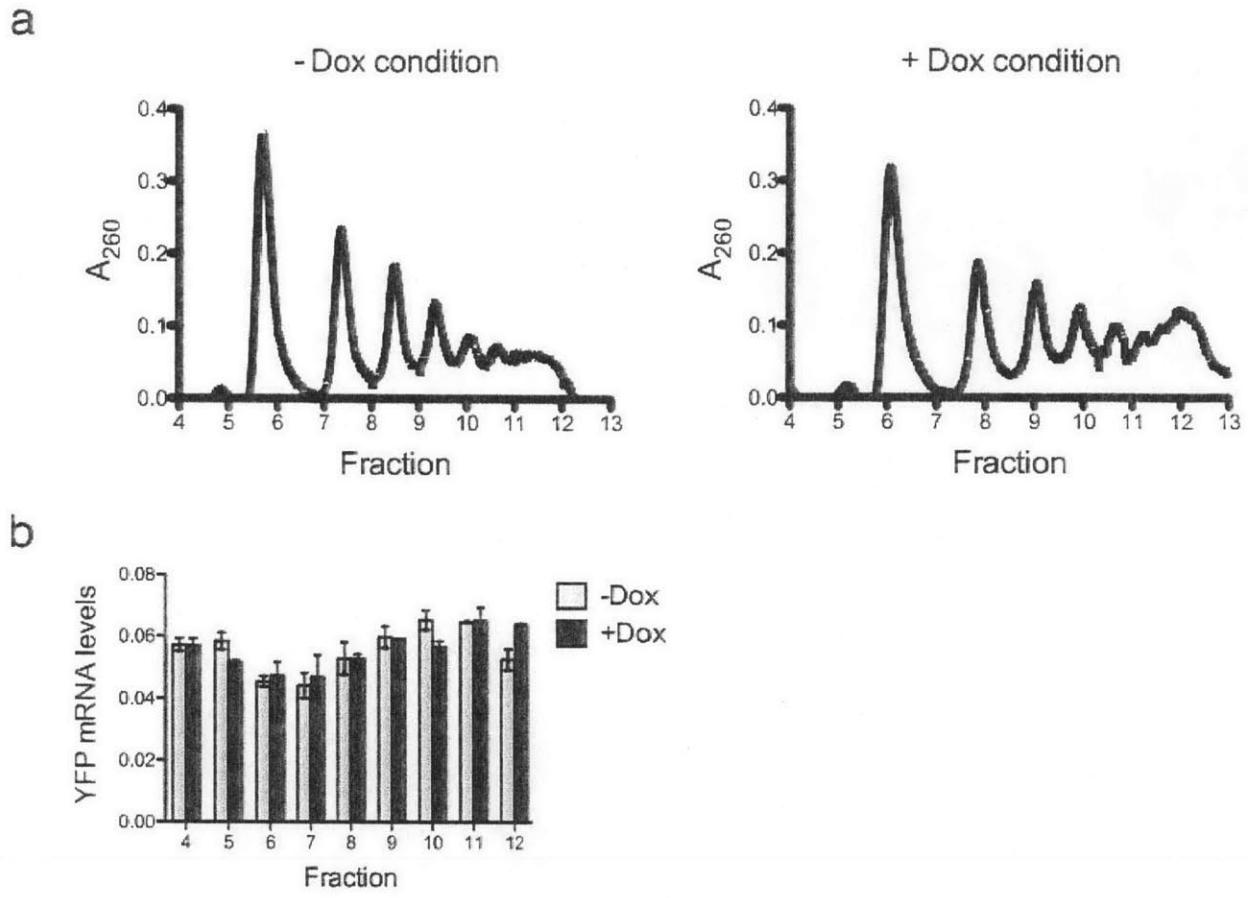

Figure 2.S5. Polysome profiles of aptamer-containing mRNA are consistent across separate experiments. Growth of yeast and subsequent polysome separation and analysis were performed independently of the experiment shown in Figure 7. **(A)** Polysomes were fractionated from yeast expressing both TetR and a **5-1.2**-containing vYFP reporter mRNA, which were grown in the absence and presence of Dox. Polysome profiles for both the - Dox and + Dox growth conditions are shown. **(B)** Quantitative PCR measurements of the relative amount of reporter mRNA within each polysome fraction under - Dox and + Dox conditions. Error bars indicate the range of values for technical duplicates.

CHAPTER 3. AN INTEGRATED STRATEGY FOR EFFICIENT VECTOR CONSTRUCTION AND MULTI-GENE EXPRESSION IN *PLASMODIUM FALCIPARUM*

3.1. Note

This chapter comprises the text of a manuscript submitted for publication and currently in review[3].

3.2. Abstract

Construction of plasmid vectors for transgene expression in the malaria parasite, *Plasmodium falciparum*, presents major technical hurdles. Traditional molecular cloning by restriction and ligation often yields deletions and rearrangements when assembling low-complexity, (A+T)-rich parasite DNA. Furthermore, the use of large 5′- and 3′ untranslated regions of DNA sequence (UTRs) to drive transgene transcription limits the number of expression cassettes that can be incorporated into plasmid vectors.

To address these challenges, we evaluated two cloning methods–yeast homologous recombination and the Gibson assembly method–as high fidelity strategies for reliably constructing *P. falciparum* vectors. Additionally, we assessed some general rules for reliably using the viral 2A-like peptide to express multiple proteins from a single expression cassette while preserving their proper trafficking to various subcellular compartments.

We demonstrate that both yeast homologous recombination and Gibson assembly are highly effective strategies for successfully constructing *P. falciparum* plasmid vectors. Using these cloning methods, we have created and validated a family of expression vectors that provide a flexible starting point for user-specific applications. These are also compatible with traditional cloning by restriction and ligation, and contain useful combinations of all the commonly used

utility features for enhancing plasmid segregation and site-specific integration in *P. falciparum.* We also illustrate application of a 2A-like peptide for synthesis of multiple proteins from a single expression cassette, and detail its utility in combinatorially directing proteins to discrete subcellular compartments.

3.3. Background and introduction

Malaria continues to be a leading cause of morbidity and mortality worldwide. Nearly 50% of the global population is at risk, and in 2010 there were an estimated 219 million cases and 660,000 deaths[96]. Plasmodium falciparum is the parasite pathogen responsible for the most virulent disease. No vaccine is clinically approved to prevent malaria. Treatment relies heavily on the use of a limited number of antimalarial drugs to which resistance is increasingly widespread[97], which makes it critical to identify new and effective drugs. Using genetic approaches to validate potential drug targets in P. falciparum is pivotal to this effort. However, the process of constructing the plasmid vectors needed for these studies is time-consuming and inefficient, and imposes a significant barrier to genetically manipulating the parasite.

Several aspects of parasite biology interact to create this challenge. First, the parasite's genome is extremely (A+T)-rich (80-90%)[50], and extended regions of low complexity sequence are common[98,99]. Second, regulatory regions upstream (5′UTR) and downstream (3′UTR) of coding sequences are poorly defined in *P. falciparum,* and large regions of putative regulatory DNA are needed to facilitate robust transgene expression[57]. Very few 5′- and 3′UTRs have been precisely mapped. As a result, 1-2 kb 5′- and 3′UTRs are frequently selected on the assumption that these comprise the information necessary to support efficient transcription[57,100–102]. These long UTRs are close to 90% in (A+T) composition. Third, the mean coding sequence (CDS) length in *P. falciparum* (excluding introns) is 2.3 kb, nearly twice that of many model organisms[50].

Gene complementation is a powerful strategy used extensively in forward genetics studies in other organisms, but this approach is under-utilized in *P. falciparum* due in large part to the challenges associated with efficiently assembling the necessary complementing constructs[103]. The ability to more routinely construct expression vectors for complementation studies is highly synergistic with the increasing rate at which genome-wide insertional mutagenesis studies are identifying candidate genes associated with growth, cell cycle and other phenotypic defects in *P. falciparum*[104,105]. In constructing overexpression, complementation and gene targeting vectors in *P. falciparum*, long (A+T)-rich regions must be cloned into final plasmids that can exceed 10 kb. It is the experience of our group and others that the traditional and commonly used restriction/ligation-based cloning method is inefficient for assembling these vectors, and often yields plasmids with regions that are deleted and/or rearranged[106]. Consequently, time-consuming screening of large numbers of bacterial clones is needed to increase the probability of recovering the intact target vector, if it is at all present.

In addition to the vector assembly challenges, typical overexpression vectors are limited in the number of transgenes that can be simultaneously expressed. In the most common format, two expression cassettes are available, and one of these is dedicated to expressing a selectable marker[102]. Increasing the expression capacity of a single plasmid can be accomplished by introducing additional 5'UTR-CDS-3'UTR cassettes, but this further complicates vector construction for reasons described above. This problem has been circumvented in several eukaryotes through the use of a viral 2A-like peptide that prevents peptide bond formation between two specific and adjacent amino acids during translation and results in the production of two separate proteins from a single expression cassette[107]. Recently, the 2A signal has been

shown to be functional in *P. falciparum*[108], but its broader utility with respect to proteins that are trafficked to different subcellular parasite compartments has not been examined.

Here, we introduce an inexpensive and straightforward strategy for more robustly and flexibly assembling *P. falciparum* vectors, while simultaneously maximizing the amount of transgenic information expressible from a single plasmid without using additional 5′UTR-CDS-3′UTR expression cassettes. We have achieved this by developing a family of vectors that integrate use of high fidelity and robust DNA assembly by yeast homologous recombination[109] and *in vitro* assembly by the isothermal chew-back-anneal "Gibson" method[110] with traditional restriction/ligation-based cloning. Additionally, we consolidate several desirable utility features in this vector family, such as: site-specific integration mediated by the Bxb1 integrase[55]; improved plasmid segregation mediated by either Rep20 elements[111] or a *P. falciparum* minicentromere (pfcen5-1.5)[58]; and all of the currently used *P. falciparum* selection markers. Lastly, we demonstrate the broader utility of using a viral 2A-like peptide to achieve expression from a single cassette of multiple genes targeted to distinct parasite subcellular compartments. This resource is freely available through the Malaria Research and Reference Reagent Resource Center (MR4; www.mr4.org).

3.4. Vector family design and features

In creating this plasmid vector resource we have incorporated several useful design criteria, namely: (1) access to multiple, orthogonal and high-fidelity strategies for cloning a target fragment into the identical context; (2) pre-installed utility features including access to all commonly used *P. falciparum* selection markers (*bsd, hDHFR, yDHODH* and *nptII*)[102,112,113], plasmid integration sequences (attP sites)[55], and plasmid segregation/maintenance features such as Rep20[111] and the mini-centromere, pfcen5-1.5[58]; (3) sufficient modularity to permit

straightforward tailoring for user-specific needs; and (4) ease of manipulation using reagents that are readily prepared in-house or commercially available at low cost. Our vector family framework includes access to yeast homologous recombination (HR)[109], Gibson assembly[110] and restriction/ligation as central cloning strategies (Fig. 3.1). The challenges associated with the traditionally used restriction/ligation method when cloning *P. falciparum* sequences have been described by others[106] and observed during plasmid manipulations by our group. It is thought that the observed genomic deletions and rearrangements are related to the long (A+T)-rich regions in combination with the restriction and ligation process and the instability of these constructs in *Escherichia coli*. Though inefficient overall, this strategy is used successfully, and so we wished to preserve it as an option that interfaced directly with the more efficient yeast HR and Gibson strategies.

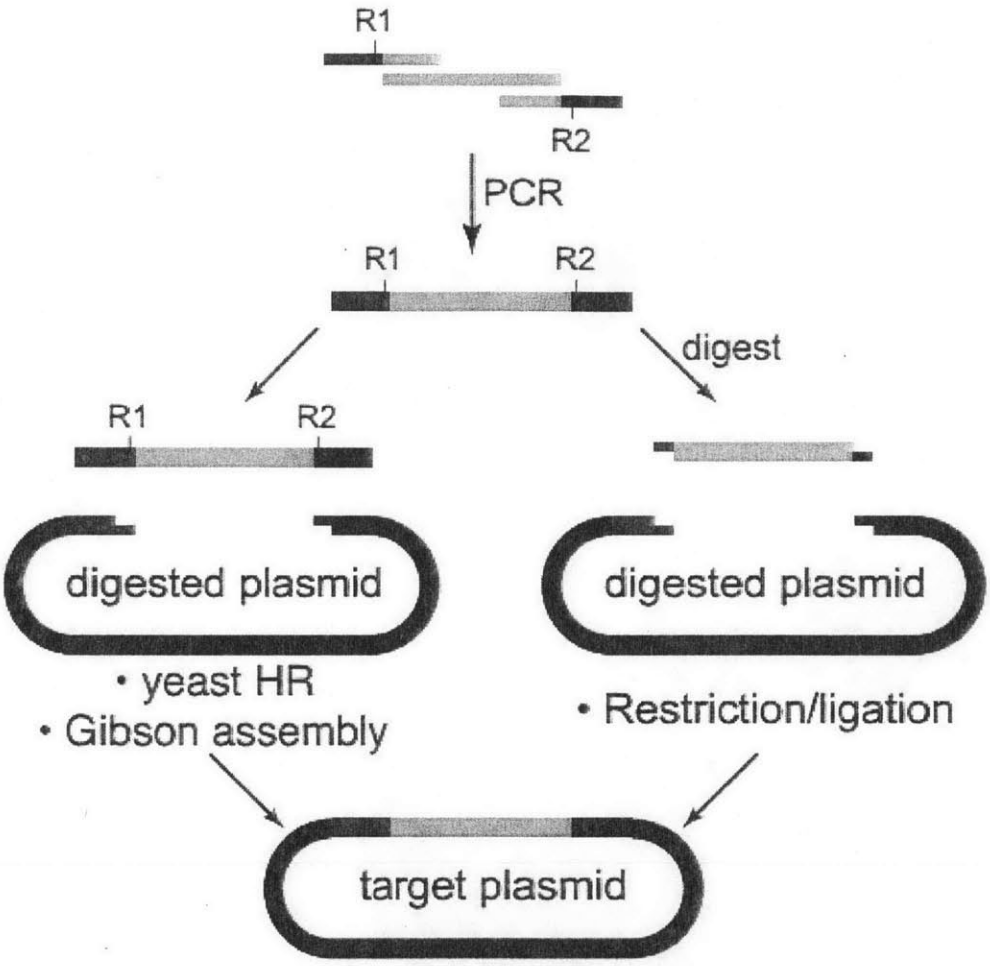

Figure 3.1. Schematic of the homology-based (yeast HR and Gibson assembly) and traditional restriction/ligation cloning strategies selected as part of an integrated framework for the orthogonal assembly of P. falciparum constructs. Beginning with a common primer set, PCR products and the desired vector backbone (Fig. 3.2), the identical target plasmid can be assembled using any of these approaches individually or in parallel.

A major advantage of both Gibson and yeast HR strategies over traditional restriction/ligation based cloning is that they do not require enzymatic digestion of the inserted fragment, which can impose constraints on cloning target DNA that contains these restriction sites internally. Rather,

as they depend on homologous ends overlapping with a digested vector, the insert does not need to be digested. This allows greater flexibility by permitting a larger set of restriction sites on the vector to be used. Yeast HR requires more overall time compared to Gibson and restriction/ligation cloning, as *S. cerevisiae* grows more slowly than *E. coli*. However, this strategy efficiently yields target constructs and all the key components can be inexpensively generated in-house[114]. The Gateway® strategy (Life Technologies) has also been used to construct *P. falciparum* vectors[115]. We have not included this approach in the current study, as it is significantly more expensive than the methods described here. However, when needed, the features required for enabling Gateway® cloning should be straightforward to introduce into the framework described below.

The overall architecture of our vector family and the built-in utility features are summarized in Fig. 3.2. We began with the pfGNr plasmid, previously deposited as MRA-462 in MR4. This plasmid contains bacterial (pMB1) and yeast (*CEN6/ARS4*) origins of replication, and the *kanMX4* gene under the control of a hybrid bacterial/yeast promoter to facilitate selection of bacterial or yeast colonies on kanamycin or G-418, respectively. This plasmid contains *two P. falciparum* gene expression cassettes, consisting of the commonly used 5′/3′UTR pairs PfCaM/PfHsp86 and PcDT/PfHRPII, arranged head-to-head to improve transcriptional efficiency[116]. In *P. falciparum*, plasmid selection using G-418 is enabled by a gfp-nptII gene fusion expressed from the PcDT/PfHRPII cassette, and a 2×Rep20 element to enhance plasmid segregation during replication is also present.

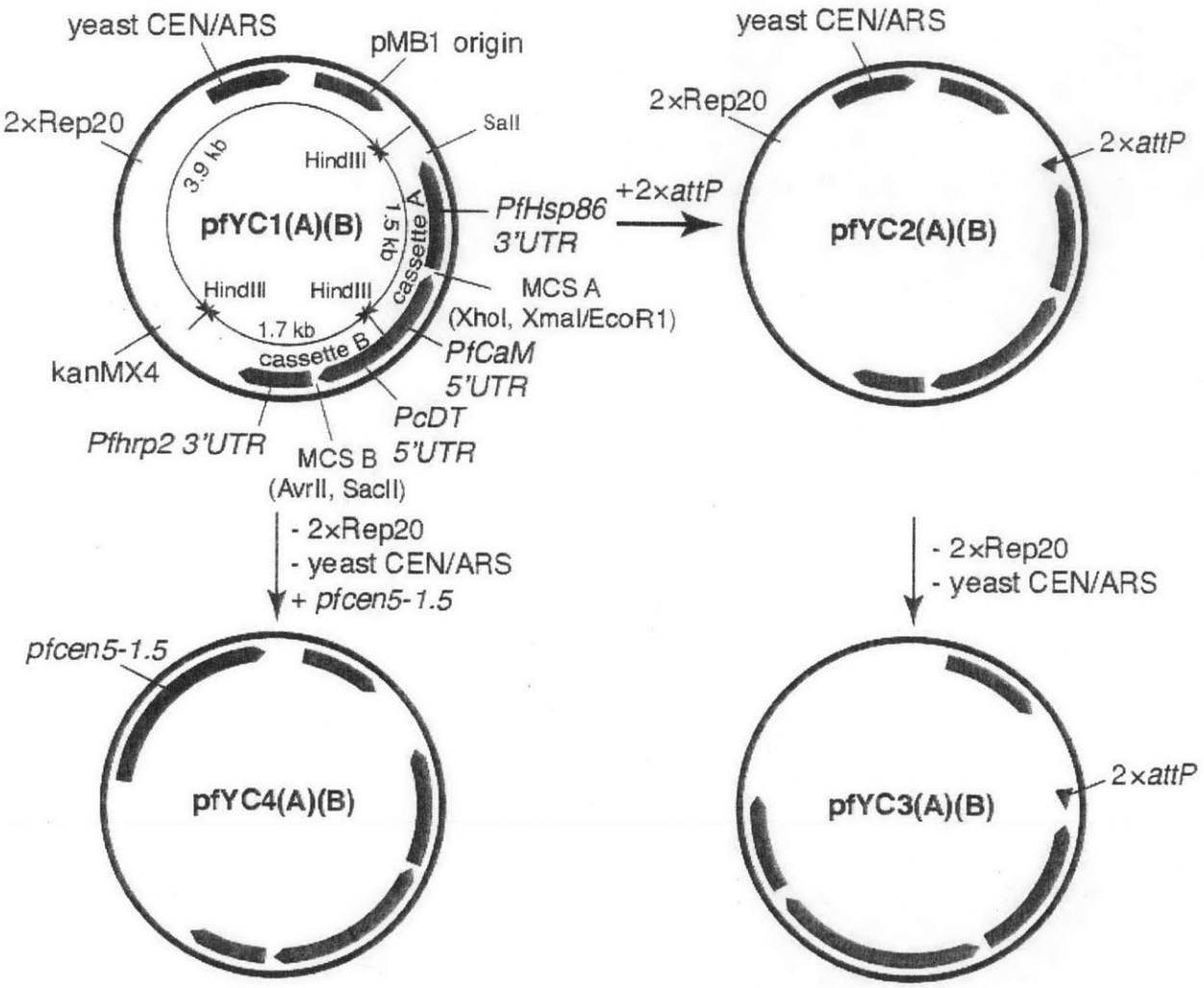

Figure 3.2. Schematic summary of the new family of plasmid vectors. Convenient restriction sites are indicated.

From this vector, we first created a library of eight base plasmids in which each of the four frequently used *P. falciparum* selection markers was cloned into one of the two *P. falciparum* expression cassettes. For ease of reference, we devised a nomenclature to describe the various vector family members. Plasmids are designated as pfYC*x*AB, where *x* is a series number indicating the presence of a specific set of utility features (1 = Rep20/yCEN, 2 =

Rep20/yCEN/2×attP, 3 = 2×attP and 4 = pfcen5-1.5) and *A* and *B* denote the resistance marker expressed from the PfCaM/PfHsp86 (cassette A) and PcDT/PfHRPII (cassette B) UTR pairs, respectively (0 = no marker; 1= *nptII*; 2 = *bsd*; 3 = *hDHFR*; and 4 = *yDHODH)*. Introducing the 2×attP site, which facilitates site-specific integration mediated by the Bxb1 integrase into compatible attB strains[55], at the SalI site yields the pfYC2 plasmid series. We have generated two representative members, namely pfYC220:FL and pfYC240:FL (Table 3.S1), in this study and provide a standardized approach for easily generating the entire set. Both the pfYC1 and pfYC2 plasmid series facilitate manipulation through yeast homologous recombination, Gibson assembly and traditional restriction/ligation cloning to provide the greatest flexibility in assembling a specific construct.

We have also generated a limited set of pfYC3 plasmids (pfYC320:FL and pfYC340:FL) that retain the attP sites but not the Rep20 or yeast centromere elements from the pfYC2 plasmid series. Elimination of the Rep20 and *CEN6/ARS4* elements from plasmids intended for integration into the *P. falciparum* genome may be desirable, as the Rep20 element has the potential to induce transcriptional silencing in a subtelomeric chromosomal context[117]. Likewise, the *S. cerevisiae*-derived *CEN6/ARS4* element could possibly behave aberrantly when integrated into a *P. falciparum* chromosome. We generated a limited set of pfYC4 plasmids (pfYC402:FL and pfYC404:FL) in which the Rep20 and *CEN6/ARS4* elements in the pfYC1 series have been replaced by the mini-centromere pfcen5-1.5. The option to use yeast homologous recombination in the pfYC3 and pfYC4 series is eliminated. However, Gibson assembly and/or traditional restriction/ligation can be used to generate final constructs that are immediately ready for integration. We also provide validated procedures for generating the complete set as dictated by user needs.

3.5. Vector construction using various cloning methods

We constructed several vectors to illustrate our ability to successfully clone firefly and *Renilla* luciferase reporter genes, and two native *P. falciparum* genes (*ama1* and *trxR*, both ~1.85 kb and ~70% in (A+T) content) into this vector family using all three cloning strategies. Using yeast HR or Gibson assembly, we cloned firefly or *Renilla* luciferase into the available expression cassette of the entire pfYC1 series (Table 3.S1). All vectors were sequenced and topologically mapped by HindIII restriction digestion. As shown in Fig. 3.3A, final plasmids with the expected topology can be assembled using these methods. Similarly, we inserted the candidate *P. falciparum* genes *ama1* and *trxR* into pfYC120 using the three vector assembly methods in parallel. Cloning reactions were carried out using the same insert and vector preparations to minimize differences between the materials used in each reaction. Five colonies derived from each cloning method were screened for each gene target and mapped by HindIII digestion to establish proper assembly of the target vector (Fig. 3.3B). Gibson assembly yielded topologically correct plasmids for both gene targets. However, under the conditions tested, yeast HR and restriction/ligation yielded the expected plasmid for trxR only. Overall, these data show that all three methods can be used to successfully clone native *P. falciparum* genes into this new vector family. Importantly, these independent cloning strategies allow use of the same plasmid backbone and insert combinations to assemble the identical final construct, thus improving the flexibility and overall ease with which *P. falciparum* vectors are made. Plasmids in this vector family can both be maintained as stable episomes and chromosomally integrated in *P. falciparum*. Toward establishing our vector family as a verified resource and a framework for routine use in *P. falciparum* transgenic experiments, we validated using it to select stable episomal and integrated *P. falciparum* lines.

We transfected the entire pfYC1AB:FL vector set either singly or in a single paired combination (pfYC110/pfYC120) into *P. falciparum* strain 3D7. Transfected parasites were selected using the appropriate drug(s), and growth was monitored by following luciferase activity. As shown in Fig. 3.4A, parasites transfected with these plasmids were successfully selected with typical kinetics[118,119]. Interestingly, under the conditions tested, the pfYC104:FL-and pfYC140:FL-transfected parasites selected with DSM-1 emerged more rapidly than parasites selected with blasticidin, WR99210 or G-418. Dual plasmid transfected parasites emerged at rates similar to those observed in single plasmid transfections (Fig. 3.4B). We also determined copy numbers for the various pfYC1 plasmids by quantitative PCR using the single-copy chromosomal β-actin gene as a reference. These data indicate that plasmids selected with BSD, DSM-1 and G-418 are maintained at an average ~ 5 copies, and for WR99210 at ~10 copies per parasite genome (Fig. 3.4C). This is consistent with results using other *P. falciparum* vectors[58,102], indicating that the pfYC vector family behaves similar to currently used plasmids and is suitable for use in transgenic experiments.

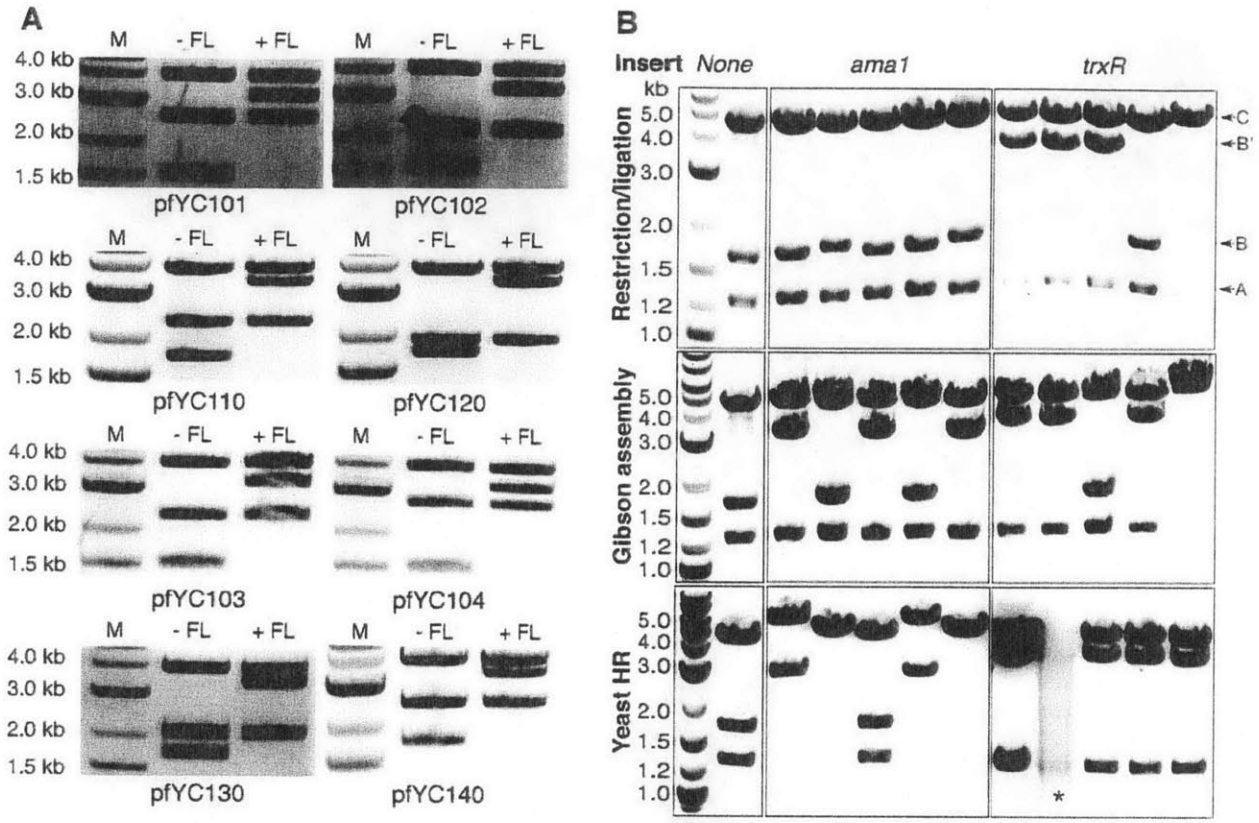

Figure 3.3. Heterologous and native P. falciparum genes can be successfully assembled into pfYC vectors using all three cloning strategies. (**A**) Firefly lucferase (FL, 1.65 kb) was cloned into the pfYC1 and pfYC3 series (Table 3.S1) using either yeast HR or Gibson assembly. Topological mapping by HindIII digestion yields three fragments, as FL and the selection markers do not contain HindIII sites. A 3.9 kb fragment is released from the pfYC1 series whether FL is present or not (Fig. 3.2). The fragments containing cassettes A and B from pfYC10x:FL plasmids are (1.5 kb + FL) = 3.2 kb and (1.7 + selection marker size) kb, respectively. Similarly, the fragments containing cassettes A and B from pfYC1x0:FL plasmids are (1.5 + selection marker size) kb and (1.7 + FL size) = 3.4 kb, respectively. The sizes of the different selection markers are: *nptII* (0.8 kb); *hDHFR* (0.6 kb); *bsd* (0.4 kb) and *yDHODH* (0.95 kb). This analysis confirms correct insertion of FL without gross plasmid rearrangements or insertions/deletions. (**B**) Two native *P. falciparum* genes, *ama1* (apical membrane antigen 1; PF3D7_1133400; 1.87 kb) and *trxR* (thioredoxin reductase; PF3D7_0923800.1; 1.85 kb) were cloned in parallel using restriction/ligation, Gibson assembly and yeast HR, and the same PCR products and digested pfYC120 vector. Successful gene insertion is expected to yield

three HindIII digestion products comprising: a backbone fragment (denoted as C); cassette B with the ama1 or trxR gene inserted (denoted as B′); and cassette A containing the bsd gene (denoted as A′). As a reference, the parent pfYC120 plasmid yields products denoted as A, C and B upon HindIII digestion. Altogether, the three strategies yielded the desired final plasmid, with Gibson assembly successfully yielding both *ama1* and *trxR* constructs. The asterisk in the yeast HR trxR panel denotes sample degradation that occurred during storage prior to analysis by gel electrophoresis.

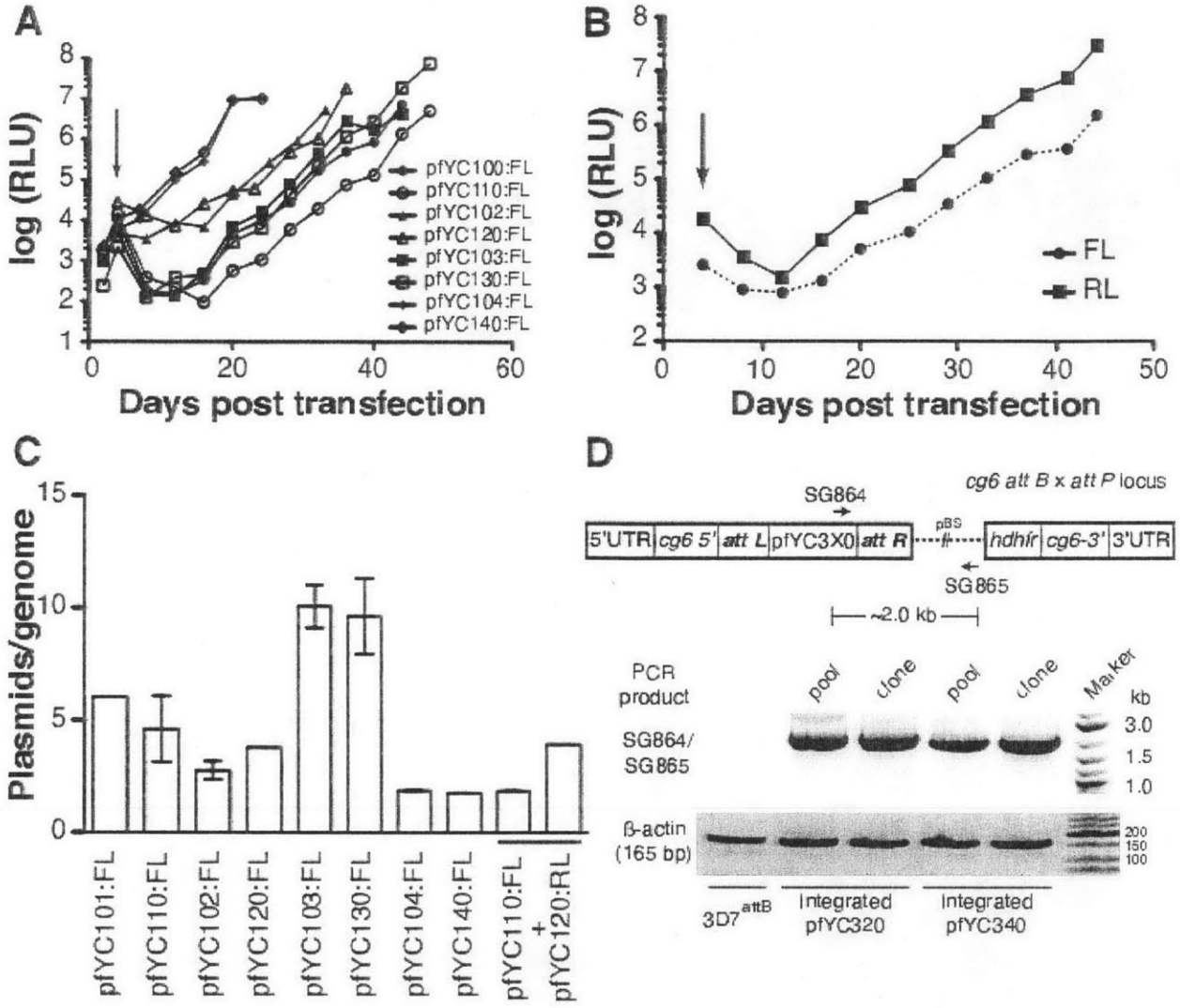

Figure 3.4. The pfYC plasmid family exhibits typical behavior during *P. falciparum* transfection, and can be maintained episomally and chromosomally integrated. The entire pfYC1xx:FL plasmid series was either transfected individually (**A**) or as a single pair (pfYC110:FL + pfYC120:RL) (**B**) under the appropriate drug selection initiated on day 4 post-transfection (arrow). Firefly and *Renilla* luciferase levels were monitored to assess parasite population growth kinetics until 1% parasitemia was reached. (**C**) The copy number of each plasmid per parasite genome was determined for both the single and double transfections. (**D**) PCR confirmation of chromosomal integration of pfYC320 and pfYC340 at the *cg6* locus in *P. falciparum* 3D7-attB. The β-actin gene was used as a positive control.

Frequently, the ability to site-specifically integrate constructs is preferred to ensure stable, homogeneous transgene expression at single copy. The pfYC3 plasmid series is designed to accomplish this by combining cloning strategy flexibility and a site-specific integration attP utility feature, while eliminating plasmid elements that are potentially deleterious when chromosomally integrated (Rep20 and *CEN6/ARS4*). As validation of this desired behavior, we transfected 3D7-attB parasites with pfYC320:FL and pfYC340:FL . Stable parasite lines expressing firefly luciferase were selected under blasticidin or DSM-1 pressure, respectively, and site-specific integration at the *cg6* locus was detected by PCR both at the population level and in isolated clones (Fig. 3.4D). Overall, these data collectively show that the pfYC vector family provides a robust and complementary set of high efficiency and timesaving cloning strategies for enabling routine assembly of DNA constructs that can be successfully used in *P. falciparum* transgenic experiments.

3.6. Expanded transgene expression from a single plasmid that is compatible with proper subcellular trafficking

While improving the ease of constructing new expression vectors, we aimed to simultaneously maximize the amount of expressible information that can be encoded on a single plasmid using a minimal set of 5′/3′UTR cassettes. From a technical standpoint, this is especially useful, as it simplifies the vector construction process by reducing overall plasmid size and instability during propagation in *E. coli*. Practically, this provides more efficient avenues for addressing questions in parasite biology requiring the co-expression of multiple genes. For example, several antigenically variant, multi-gene families, such as PfEMP1, STEVORs and RIFINs, are combinatorially expressed by the parasite to modulate host-parasite interactions, such as immune

recognition and evasion[120,121]. Therefore, the ability to achieve predetermined expression of specific combinations of these proteins could prove useful in understanding their combined contributions to these outcomes. Furthermore, subsets of proteins involved in multi-gene pathways must often be trafficked to distinct subcellular compartments. The tricarboxylic acid cycle[122,123], lipid and isoprenoid biosynthesis[124] and heme biosynthesis[125] involve multiple proteins distributed between the cytosol, mitochondrion and apicoplast, or exclusively targeted to one of these organelles. Also, a substantial fraction of the parasite-encoded proteome is trafficked to the host red blood cell (RBC), including the antigenically variant STEVOR and RIFIN families, and a significant number of these trafficked proteins play essential but poorly understood roles[126]. Therefore, it is important to establish that multi-cistronic expression and proper protein trafficking can be simultaneously achieved for this approach to be most broadly useful.

Towards this goal, we used virus-derived 2A-like peptide sequences (2A tags), which have successfully been used in mammalian, yeast, plant and protozoan contexts to enable polycistronic expression from a single eukaryotic mRNA[107,108]. 2A tags mediate peptide bond "skipping" between conserved glycine and proline residues, yielding one protein with a short C-terminal extension encoded by the tag, and the other with an N-terminal proline. The small size (8 conserved amino acid positions) and broad cross-species functionality of the 2A tag makes it an attractive candidate for application to *P. falciparum*, an organism in which this technology has not been extensively explored. As an entire expression cassette is usually committed exclusively to expressing a selection marker, we designed our initial experiment to address whether we could use the *Thosea asigna* virus 2A-like sequence (T2A) to expand the number of genes expressed from this cassette without compromising our ability to select transfected parasites. We inserted

T2A with a short, N-terminal linker region[127] between the firefly luciferase (*FL*) and *nptII* genes in cassette A to generate pfYC101:FL-2A-nptII (pSG93). We also generated a control construct containing a non-functional tag (T2Am), in which two conserved residues are mutated to alanine[127] (pSG94, Fig. 3.5A). We transfected these plasmids into *P. falciparum* 3D7 under G-418 selection pressure and successfully obtained resistant parasites with FL activity (Fig. 3.5B), demonstrating the production of functional nptII and FL proteins in both cases. We confirmed the ability of T2A to produce distinct FL and nptII proteins from a single mRNA by western and northern blot (Fig. 3.5C and 3.5D, respectively). As expected, mutating T2A to T2Am eliminates the formation of discrete proteins, but does not alter the size of the *FL-nptII* mRNA. This initial characterization, in addition to demonstrating T2A functionality in *P. falciparum*, highlights the potential for using T2A to recover valuable expression capacity by encoding additional information into existing selection marker cassettes while eliminating the unpredictability of how a protein fusion will function. Next, we examined the flexibility with which T2A can be used to produce dicistronic messages encoding proteins destined for distinct subcellular compartments within the parasite and its RBC host. We engineered several dicistronic constructs encoding an N-terminal Venus yellow fluorescent protein (vYFP) and a C-terminal tdTomato protein (tdTom) separated by T2A in the pfYC120 vector. We used previously validated apicoplast, mitochondrial and RBC export targeting sequences derived, respectively, from: acyl carrier protein (PF13_0208500; aa 1-60 = ATS)[128], Hsp60 (PF13_1015600; aa 1-68 = MTS)[129], and knob-associated histidine-rich protein (PF13_0202000; aa 1-69 = PEX)[130]. We created seven contexts in which a different protein targeting signal (or none at all) was placed immediately upstream of vYFP and/or tdTom as follows: (a) vYFP-2A-tdTom; (b) vYFP-2A-ATS-tdTom; (c) vYFP-2A-MTS-tdTom; (d) vYFP-2A-PEX-tdTom; (e) MTS-vYFP-2A-MTS-tdTom; (f) PEX-

vYFP-2A-PEX-tdTom and (g) ATS-vYFP-2A-tdTom.We evaluated vYFP and tdTom trafficking by fluorescence microscopy, and distinguished production of vYFP versus a possible fusion to tdTom by Western blot (Fig. 3.6). Overall, when no targeting sequence is upstream of vYFP, the downstream tdTom is faithfully trafficked to the subcellular compartment expected based on the associated targeting sequence. Similarly, when vYFP and tdTom are associated with the same targeting sequence (parasite cytosol, mitochondrion and RBC cytosol tested), both are trafficked as separate proteins to the same subcellular compartment. For the ATS-vYFP-2A-tdTom construct, vYFP is trafficked to the apicoplast as expected. Interestingly, a substantial fraction of the tdTom is mislocalized to the apicoplast with some signal distributed in the parasite's cytoplasm. By western blot, vYFP is detected as both the isolated protein and the tdTom fusion (~ 100 kDa). Presumably, the fusion product accounts for the majority of the mis-localized tdTom, while the cytosolic fraction arises due to the expected T2A behavior. These data suggest that "ribosome skipping" might be less efficient and/or the downstream protein is more often misdirected when the upstream protein is apicoplast-targeted, at least within the context tested by our constructs. This outcome is reminiscent of "slipstreaming" observed when using T2A for multi-cistronic expression of secreted proteins in mammalian cells, though this phenomenon is thought to be primarily influenced by the C-terminal portion of the upstream protein, which does not vary across our constructs[131]. However, our observation that the vYFP-2A-ATS-tdTom construct exhibits the expected subcellular targeting patterns, together with the other combinations in which we observed proper subcellular trafficking, defines a set of rules for using T2A to successfully achieve multi-cistronic protein expression with proper subcellular targeting.

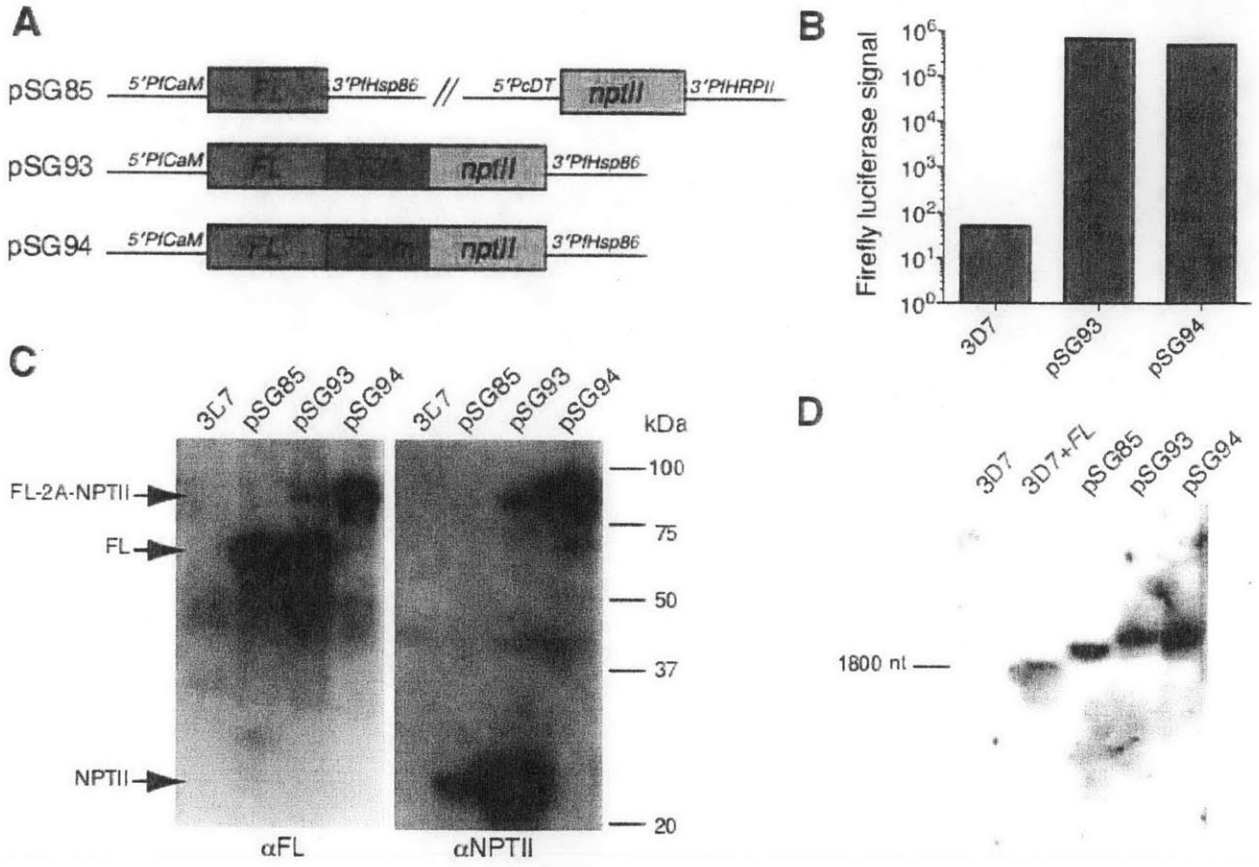

Figure 3.5. The *Thosea asigna* virus 2A-like peptide (T2A) enables expression of two functional proteins in *P. falciparum* from a single expression cassette, with the expected subcellular targeting in many cases. (**A**) Schematic of FL-nptII and control constructs. (**B**) Both T2A-and T2Am-containing constructs produce active FL. (**C**) Western blot detection of FL-and nptII-containing proteins. (**D**) Northern blot analysis of FL-containing transcripts in transfected parasites. **3D7 + *FL*** indicates the inclusion of a synthetic *FL* mRNA produced by *in vitro* transcription.

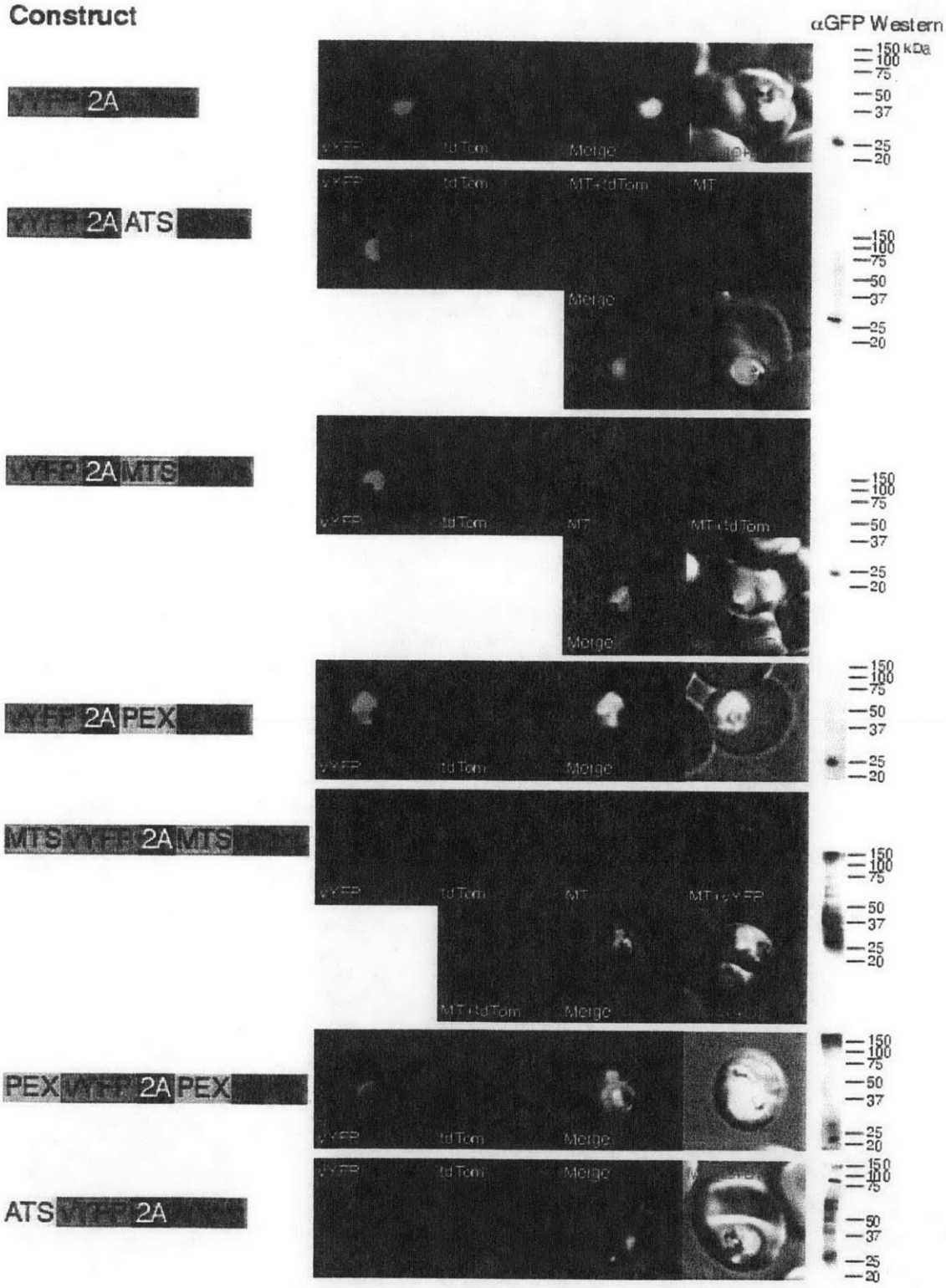

Figure 3.6. T2A can be used successfully to target proteins to distinct subcellular compartments. Various targeting sequences were N-terminally fused to an upstream

vYFP and a downstream tdTom reporter separated by T2A. The vYFP and tdTom proteins were localized using direct fluorescence microscopy imaging. Mitochondria were stained with MitoTracker (MT), and nuclei with Hoechst 33342. ATS = apicoplast targeting sequence; MTS = mitochondrial targeting sequence; PEX = PEXEL protein export element.

3.7. Conclusions

We have developed a validated set of broadly useful plasmid vectors that enable versatile assembly of *P. falciparum* constructs, a frequently time-consuming process given the (A+T)-richness and large size of final vectors. We have achieved this by integrating simultaneous access to the high efficiency and inexpensively available yeast homologous recombination, Gibson assembly and conventional restriction/ligation cloning strategies. All three strategies are technically straightforward and use the same restriction enzyme-digested vector and PCR products generated with the same primer set. In principle, all three strategies can be executed in parallel or used interchangeably without the need for new genetic reagents and can be used to successfully clone both reporter and native *P. falciparum* genes. Additionally, we have pre-installed several widely used utility features for enhancing plasmid segregation and site-specific chromosomal integration, and provided validated operations to enable user-tailored modifications to this vector family. Altogether, these openly available tools and validated methods provide a convenient route to more routinely generating *P. falciparum* vectors that enable basic biological studies.

3.8. Experimental procedures.

Molecular biology. Unless otherwise indicated, enzymes were from New England Biolabs and chemicals were from Sigma or Research Products International. High Fidelity (HF) restriction

enzymes were used when available. PCR was routinely performed with Phusion DNA polymerase in HF Buffer, or with a 15:1 (v:v) mixture of Hemo KlenTaq:Pfu Turbo (Agilent) in Hemo KlenTaq Buffer. The latter conditions permit PCR amplification directly from parasite culture samples, usually included at 5% of the total reaction volume. Plasmids were prepared for transfection with maxi columns (Epoch Life Science) or the Xtra Midi Kit (Clontech).

Vector construction. The primers used for these studies are listed in Table 3.S2.

Yeast homologous recombination (HR). Yeast HR vector construction was carried out by standard methods[109,132]. Variable amounts of vector backbone (typically 0.1–2 μg) were digested using standard methods to generate linearized vector. PCR was carried out using standard techniques to generate fragments for insertion bearing 20–40 bp homology to the desired flanking regions on the vector. Competent *S. cerevisiae* W303-1B was prepared as described (21) and frozen at -80 °C. Either unpurified or column-purified PCR product was co-transformed with either unpurified or column-purified linearized vector. A wide range of concentrations of both linearized vector and PCR product were observed to be efficacious. Transformed yeast were plated on YPD agar (10 g/L yeast extract, 20 g/L peptone, 20 g/L dextrose, 20 g/L agar) supplemented with 400 mg/L G-418 disulfate and allowed to grow for 48-72 hours at 30 °C. Typical yields were 10–100 colonies for a negative control transformation lacking PCR insert, and 50–1,000 colonies for the complete HR reaction.

Colonies were then either harvested by plate scraping or picked and grown overnight in YPD + 400 μg/mL G-418. Cells were then treated with 2 U zymolyase (Zymo Research) in 250 μL buffer ZB (10 mM sodium citrate pH 6.5, 1 M sorbitol, 25 mM EDTA and 40 mM dithiothreitol) for 1 hr at 37 °C to generate speroplasts. Yeast spheroplasts were lysed with the addition of 250

μL buffer MX2 (0.2 M NaOH and 10 g/L sodium dodecyl sulfate). Plasmid DNA was then purified either by spin column or alcohol precipitation. For spin column precipitation, the lysate was mixed with 350 μL buffer MX3 (4.2 M guanidine HCl, 0.9 M potassium acetate, pH 4.8) and applied to a mini prep column (Epoch Life Science), followed by washing with Buffer WS (80% ethanol, 10 mM Tris-Cl, pH 7.5) and elution in sterile water. For alcohol precipitation, lysate was mixed with 250 μL of 3 M potassium acetate, pH 5.5 and centrifuged at 21,000*g* for 2 min. The supernatant was mixed with an equal volume of isopropanol and centrifuged at 21,000*g* for 8 min. The pellet was washed with 70% ethanol and dissolved in sterile water.

The recovered DNA was then transformed directly into *E. coli* DH5α or EPI300 (Epicentre) prepared with a Z-Competent Transformation Kit (Zymo Research) or transformed by electroporation. Occasionally, the DNA mixture was drop dialyzed against water for 20 minutes prior to transformation to increase electroporation. After 1–2 hr recovery in LB or SOC media, cells were plated on LB agar containing 50 mg/L kanamycin and grown overnight at 37 °C. Typical yields were 5–100 colonies. Colonies were picked, grown overnight, mini-prepped, and assayed for correct vector assembly by restriction digest, diagnostic PCR and/or DNA sequencing.

Gibson assembly. Isothermal chew-back-anneal assembly, commonly known as Gibson assembly, was carried out as described[133]. Briefly, vector and PCR product were prepared in the same way as for yeast HR assembly. Fragments were combined with either a homemade or commercially available Gibson Assembly Master Mix. The homemade Master Mix was prepared by combining 699 μL water, 320 μL of 5x isothermal reaction buffer (500 mM Tris-Cl, pH 7.5, 250 mg/mL PEG-8000, 50 mM $MgCl_2$, 50 mM dithiothreitol, 1 mM each of four dNTPs, 5 mM NAD), 0.64 μL T5 Exonuclease (Epicentre, 10 U/μL), 20 μL Phusion DNA polymerase (2 U/μL)

and 160 μL Taq DNA ligase (40 U/μL). This solution was divided into 20 μL aliquots and stored at -20 °C. Generally, >100 ng of linearized vector was added to the mixture with an equal volume of PCR insert, generating a variable vector:insert ratio. The mixture was incubated at 50 °C for 1 hr and 0.5 μL was transformed into *E. coli* as described above.

Restriction/ligation cloning. Restriction/ligation cloning was carried out by standard techniques. Ligations were incubated overnight at 16°C and heat inactivated prior to transformation.

Construction of *attP*-containing (pfYC3 series) plasmids. The 2×*attP* fragment was PCR amplified from pLN-ENR-GFP (18) with primers SG702/703. pfYC120:FL and pfYC140:FL were digested with *SalI* and combined with gel-purified PCR product in a Gibson assembly reaction to obtain pfYC220:FL and pfYC240:FL, respectively. These vectors were then digested with *MluI* and *PmlI* to release the fragment containing the Rep20 and CEN/ARS elements. Approximately 100 ng of digested vector was then combined in a Gibson assembly reaction containing 50 nM each SG814 and SG815 to re-circularize the vector while adding a unique *PmeI* site between the *MluI* and *PmlI* sites. This yielded pfYC320:FL and pfYC340:FL. For integration at the *cg6* locus, these two vectors were co-transfected with pINT (18) (~50 μg each) into *P. falciparum* 3D7-*attB* (MRA-845 from MR4). Clonal populations were obtained by limiting dilution and integration verified by PCR using the SG864/865 primer pair.

Construction of the pfcen5-1.5 mini-centromere containing (pfYC4 series) plasmids. The 2×*attP* element was amplified from *P. falciparum* 3D7 genomic DNA with primers SG894/SG928. pfYC102:FL and pfYC104:FL were digested with MluI and PmlI and the gel-purified backbone lacking the 2×*Rep20* element was attached to the SG894/SG928 PCR product

by Gibson assembly to yield pfYC402:FL and pfYC404:FL, respectively. Clones were verified by restriction digest and by sequencing with SG369.

Cloning *P. falciparum ama1* and *trxR* genes into pfYC120. The *ama1* and *trxR* genes were amplified from *P. falciparum* 3D7 genomic DNA using the AMA1 pcDT F/R and Trx pcDT F/R primer pairs, respectively. Restriction/ligation, Gibson assembly and yeast HR were performed as described above.

Multi-cistronic constructs using the T2A for evaluating subcellular trafficking rules. After Western blot and microscopic imaging analysis, the identity of each strain was re-verified by PCR amplifying and sequencing a uniquely identifying fragment of the transfected construct using the SG763/764 and SG502/646 primer pairs, respectively.

Parasite culture and transfection. *P. falciparum* strain 3D7 parasites were cultured under 5% O_2 and 5% CO_2 in RPMI-1640 media supplemented with 5 g/L Albumax II (Life Technologies), 2 g/L $NaHCO_3$, 25 mM HEPES-K pH 7.4, 1 mM hypoxanthine and 50 mg/L gentamicin. Transfections used ~50 μg of each plasmid and were performed by the spontaneous DNA uptake method[134] or by direct electroporation of ring-stage cultures[135]. Transgenic parasites were selected with 2.5 mg/L Blasticidin S, 1.5 μM DSM-1, 5 nM WR9920 (Jacobus Pharmaceuticals) and/or 250 mg/L G-418 beginning 2–4 days after transfection.

Monitoring transfection progress by luciferase expression. Firefly and *Renilla* luciferase levels were measured every fourth day after transfection using the Dual-Luciferase Reporter Assay System (Promega). Samples for measurement were prepared by centrifugation of 1.25 mL of parasite culture at 2% hematocrit to generate an ~25 μL parasite-infected RBC pellet. Pellets were either used immediately or stored at -80 °C until needed for luciferase measurements.

Parasite DNA extraction and qPCR analysis. DNA was harvested from schizont stage parasites at >5% parasitemia. Infected RBCs were treated with 0.1 mg/mL saponin in PBS to release parasites, which were either immediately used or stored in liquid nitrogen for later analysis. Parasites were lysed for 1 hour at 50 °C in 200 µL of 50 mM Tris-Cl pH 8.0, 50 mM EDTA, 0.5 mg/mL SDS and 10 µL of protease solution (Qiagen). After adding RNase A (28 U; Qiagen), reactions were incubated at 37 °C for 5 minutes. After adding 20 µL of 6 M sodium perchlorate, DNA was isolated by phenol/chloroform extraction and ethanol precipitation.

qPCR reactions (20 µL each) contained 1× Thermopol Buffer , 0.2 mM dNTPs, 200 nM relevant primer pair (Table 3.S3), 0.5× SYBR Green I (Life Technologies), 0.1 µL Taq DNA polymerase and 5 µL of a DNA dilution. Thermocycling was performed on a Roche LightCycler 480 II for 40 cycles according to the following program: 95 °C for 20 s; denature: 95 °C for 3 s; anneal/extend: 60 °C for 30 s; fluorescence measurement. Vector-borne amplicons (drug resistance marker genes) and a native chromosomal locus (β-actin) were quantified by comparison with plasmid or PCR-amplified DNA standards, respectively. Primers are listed in Table 3.S3.

Western blot. Approximately 10^6 late-stage parasites were harvested by lysis of infected RBCs with 0.5 g/L saponin and then lysed by heating in urea sample buffer (40 mM dithiothreitol, 6.4 M urea, 80 mM glycylglycine, 16 g/L SDS, 40 mM Tris-Cl, pH 6.8) at 95 °C for 10 min. After separation by SDS-PAGE, proteins were transferred to a PVDF membrane and probed with an antibody against firefly luciferase (Promega G7451), neomycin phosphotransferase II (Millipore 06-747) or green fluorescent protein (Abcam ab1218). Blots were then imaged using a horseradish peroxidase-coupled secondary antibody and SuperSignal West Femto substrate (Thermo Scientific).

Northern blot. Total RNA was purified from infected RBCs with a combination of Tri Reagent RT Blood (Molecular Research Center) and an RNeasy Mini Kit (Qiagen). One mL of parasite culture at 20% hematocrit and ~10% late-stage parasitemia was frozen on liquid nitrogen and thawed with the addition of 3 mL Tri Reagent RT Blood. After phase separation with 0.2 mL BAN (Molecular Research Center), 2 mL of the upper aqueous phase was mixed with 2 mL ethanol and applied to an RNeasy Mini column for purification according to the manufacturer's instructions. Total RNA (6.5 μg) for each sample (with or without the addition of 6 pg firefly luciferase RNA generated by *in vitro* transcription) was mixed with an equal volume of denaturing sample buffer (95% formamide, 0.25 g/L SDS, 0.25 g/L bromophenol blue, 0.25 g/L xylene cyanol, 2.5 g/L ethidium bromide, 50 mM EDTA) and heated at 75 °C for 10 min before loading on a 1% TAE agarose gel. Electrophoresis was performed at 80 V for 75 min and RNA was transferred to a Nylon membrane (Pall Biodyne Plus) by downward capillary transfer in 50 mM NaOH for 90 min. After UV fixation (Stratagene Stratalinker, 1.25 mJ), the membrane was probed and imaged with the North2South Chemiluminescent Detection Kit (Thermo). The biotinylated firefly luciferase probe was prepared from a DNA template generated by PCR with primers SG311 and SG313.

Fluorescence microscopy. For live cell imaging, parasite cultures were incubated for 20 minutes with 30 nM MitoTracker Deep Red FM (Life Technologies). Infected RBCs were then washed with phosphate-buffered saline and applied to poly-L-lysine-coated glass-bottom culture dishes (MatTek, Ashland, MA). Attached cells were overlaid with RPMI media (free of phenol red and Albumax II) containing 2 μg/mL Hoechst 33342 (Sigma), and imaged immediately at room temperature using a Nikon Ti-E inverted microscope with a 100x objective and a Photometrics

CoolSNAP HQ2 CCD camera. Images were collected with the Nikon NIS Elements software and processed using ImageJ[136].

3.9. Tables

Table 3.S1. Vectors used in this study.

Vector	Selection in *P. falciparum*	Assembly method	Notes
pfYC101:FL	G-418	Yeast HR	Rep20, yeast centromere; FL and resistance marker are in cassettes A and B, respectively
pfYC110:FL	G-418	Yeast HR	Rep20, yeast centromere; Resistance marker and FL are in cassettes A and B, respectively
pfYC102:FL	Blasticidin S	Yeast HR	Rep20, yeast centromere; FL and resistance marker in cassettes A and B, respectively
pfYC120:FL	Blasticidin S	Yeast HR	Rep20, yeast centromere; Resistance marker and FL are in cassettes A and B, respectively
pfYC120:RL	Blasticidin S	Yeast HR	Rep20, yeast centromere; Resistance marker and FL are in cassettes A and B, respectively; Used in double transfection experiment (Figure 4B)
pfYC103:FL	WR99210	Yeast HR	Rep20, yeast centromere; FL and resistance marker are in cassettes A and B, respectively
pfYC130:FL	WR99210	Yeast HR	Rep20, yeast centromere; Resistance marker and FL are in cassettes A and B, respectively
pfYC104:FL	DSM-1	Gibson	Rep20, yeast centromere; FL and resistance marker are in cassettes A and B, respectively
pfYC140:FL	DSM-1	Gibson	Rep20, yeast centromere; Resistance marker and FL are in cassettes A and B, respectively
pfYC320:FL	Blasticidin S	Gibson	Contains 2xattP element; No Rep20 or yeast centromere elements; Resistance marker and FL are in cassettes A and B, respectively
pfYC340:FL	DSM-1	Gibson	Contains 2xattP element; No Rep20 or yeast centromere elements; Resistance marker and FL are in

			cassettes A and B, respectively
pfYC120:vYFP-2A-tdTom	Blasticidin S	Gibson	
pfYC120:vYFP-2A-(ATS)-tdTom	Blasticidin S	Gibson	
pfYC120:vYFP-2A-(MTS)-tdTom	Blasticidin S	Gibson	
pfYC120:vYFP-2A-(PEX)-tdTom	Blasticidin S	Gibson	
pfYC120:(MTS)-vYFP-2A-(MTS)-tdTom	Blasticidin S	Gibson	
pfYC120:(PEX)-vYFP-2A-(PEX)-tdTom	Blasticidin S	Gibson	
pfYC120:(ATS)-vYFP-2A-tdTom	Blasticidin S	Gibson	

Table 3.S2. Oligonucleotides used in vector construction.

Primer	Sequence	Notes
SG311	GAAATATATCAGACGTCTCCCCCGGGACCATG GAAGACGCCAAAAACATAAAGAAAGGCC	*FL* probe (forward)
SG313	GACCCCATTGTGAGTACATAAATATATTATATA ACTCGAGTTACAACTCGGACTTTCCGC	*FL* probe (reverse)
SG763	AGCATGTGCATGGCATCCCCTT	Amplifying uniquely identifying regions of vYFP-tdTom cassette (forward)

SG764	TGACCTCCTCGCCCTTGCTCA	Amplifying uniquely identifying regions of vYFP-tdTom cassette (reverse)
SG502	AGTAGCATCACCTTCACCTTCACC	Sequencing from vYFP in dicistronic vYFP-tdTom constructs (reverse)
SG646	CTGCCTTATCCAAAGATCCAAACG	Sequencing from vYFP in dicistronic vYFP-tdTom constructs (forward)
SG702	TAGGTGACACTATAGAATACTCAAGCTTGGCG GCCGCCCGAGCTCGAATTCCGGGTTTGT	Tandem *attP* (forward)
SG703	AGTTAATTCATCAAATAGCATGCCTGCAGGTC GACGCCAGGGTTTTCCCAGTCACGA	Tandem *attP* (reverse)
SG814	GGACATTGTTTAAACGAGCAGG ACGCGT TGAATTGTCCCCACGCCGCGCCC	Recircularization at *PmlI/MluI* sites (forward)
SG815	CCTGCTCGTTTAAACAATGTCC CACGTG ATGAAAAGGACCCAGGTGGCA	Recircularization at *PmlI/MluI* sites (reverse)
SG864	GCAGGTCGACGCCAGGGTTT	Confirming pfYC integration at *cg6-attB* (forward)
SG865	GACGCCGGGCAAGAGCAACT	Confirming pfYC integration at *cg6-attB* (reverse)
SG928	GGGCGCGGCGTGGGGACAATTCAACGCGT TAATTATTAATATATTAATTATTTAGACTTA	Amplify pfcen5-1.5 (forward)
SG894	TGCCACCTGGGTCCTTTTCATCACGTG TATGTATAATTAAATTAAATATTATAAACACAC	Amplify pfcen5-1.5 (reverse)
SG369	CCAATGCTTAATCAGTGAGGC	Sequence yCEN/or pfcen5-1.5 region
Trx pcDT F	aaatatatacacacacctaaaacttacaaagtatcctaggaaaaatgaacaat gtaattt	Amplification of TrxR to clone into pfYC120
Trx pcDT R	tttaatctattattaaataaatttaatggggtacccgcggttatccacattttccac ccc	Amplification of TrxR to clone into pfYC120
AMA1 pcDT F	aaatatatacacacacctaaaacttacaaagtatcctaggaaaaatgagaaaa ttatact	Amplification of AMA1 to clone into pfYC120

AMA1 pcDT R	tttaatctattattaaataaatttaatggggtacccgcggttaatagtatggtttttc ca	Amplification of AMA1 to clone into pfYC120

Table 3.S3. Oligonucleotides used for quantitative PCR.

Primer	**Sequence**	**Target gene**	**Source**
Neo (forward)	CATCCTGATCGACAAGACCG	*nptII*	This work
Neo (reverse)	CCTGCCGAGAAAGTATCCATC	*nptII*	This work
BSD (forward)	GCTGTCCATCACTGTCCTTC	*bsd*	This work
BSD (reverse)	TGGCAACCTGACTTGTATCG	*bsd*	This work
hDHFR (forward)	GAGGTTGTGGTCATTCTCTGG	*hDHFR*	This work
hDHFR (reverse)	AGAACATGGGCATCGGC	*hDHFR*	This work
y*dhodh* (forward)	TCCACCTGTACCGATAACTTTG	*yDHODH*	This work
y*dhodh* (reverse)	GATGTGGAGAAGGAGAGTGTAG	*yDHODH*	This work
Pf-ß-actin (forward)	AAAGAAGCAAGCAGGAATCCA	*P. falciparum* β-actin (PF3D7_1246200)	Augagneur *et al.*[137]
Pf-ß-actin (reverse)	TGATGGTGCAAGGGTTGTAA	*P. falciparum* β-actin (PF3D7_1246200)	Augagneur *et al.*[137]

CHAPTER 4. VERSATILE CONTROL OF *PLASMODIUM FALCIPARUM* GENE EXPRESSION WITH AN INDUCIBLE PROTEIN-RNA INTERACTION

4.1. Note

This chapter is intended to be submitted, in a modified form, for peer review and publication.

4.2. Abstract

The available tools for conditional gene expression in *Plasmodium falciparum* are limited in their utility and applicability to a given gene of interest. To enable researchers to reliably control the expression of target genes, we built a tool that modulates mRNA translation in response to a tetracycline analog. Here, we show the broad utility of the system for controlling the expression of reporter and endogenous parasite proteins trafficked to a variety of subcellular compartments. The system allows efficient post-transcriptional control of gene expression when implemented in several native and engineered promoter contexts. Induction and repression of gene expression are rapid and homogeneous across the cell population. By placing a drug resistance determinant under inducible control, we modulate *P. falciparum* response to antimalarial drug exposure, demonstrating the usefulness of the system for controlling relevant parasite biology.

4.3. Introduction

Malaria continues to impose an enormous burden of death and disability. *Plasmodium falciparum,* the parasite responsible for the vast majority of malaria deaths, can be cultured in the laboratory, but a comprehensive suite of molecular tools is not yet available[138]. A generally useful system for conditional gene expression in *P. falciparum* remains highly sought after. Tools based on inducible control of transcription[61,139] have proven unreliable[140] and yield a high degree of expression heterogeneity; construction of a robust transcriptional tool will likely require a better understanding of the molecular details of transcription in *P. falciparum*. A tool

based on inducible control of fusion protein stability has now been used effectively by several research groups[59,141]. However, it is not generalizable to many genes of interest, as the target protein must 1) tolerate a fusion to a ligand-sensitive degradation domain and 2) remain accessible to the proteasomal degradation apparatus, precluding the system's use for many important proteins trafficked to subcellular or host cell compartments. RNA interference does not appear to function in the parasite, likely due to the lack of necessary molecular machinery[142].

We previously engineered a system for inducible control of eukaryotic translation that uses a protein-RNA interaction controlled by a tetracycline analog[2]. As the molecular components of cytosolic translation in *P. falciparum* are thought to be well conserved with other eukaryotes[143], we reasoned that we should be able to implement our system for control of translation in the parasite. Such a system should enable experimenter control of target gene expression with regard to neither the mechanism of mRNA transcription nor the post-translational behavior of the target protein.

4.4. Results and discussion

We used a plasmid-based transient screening system to verify the basic functionality of the system in *P. falciparum.* After constructing a parasite strain that stably expresses TetR from a single-copy chromosomal locus (Fig. 4.S1 and Table 4.S1), we transfected the strain with dual-luciferase reporter plasmids expressing firefly luciferase (FLuc) under the control of the TetR-binding aptamer 5-1.17 (figs. 4.1A and 4.S2). Addition of anhydrotetracycline (aTc) to transiently transfected parasites yielded a ~5-fold change in FLuc expression (Fig. 4.1B), similar to what we observed in yeast[2]. Also similar to what we previously observed, behavior of the system was consistent across different promoter/5′ untranslated region (5′UTR) contexts, with a similar dynamic range produced by the Hsp86 or calmodulin 5′UTRs. The mutant sequence 5-

1.17m2 yielded no aTc-dependent regulation, as expected. To perform further characterization of system behavior, we built single-plasmid systems expressing aptamer-regulated FLuc, constitutive reference *Renilla* luciferase, TetR, and a G-418-selectable marker (fig. 4.1C). These constructs express polycistronic messages that use 2A-like peptides to yield multiple distinct proteins[3]. We transfected single-plasmid constructs into parasite strain Dd2 and measured aTc-dependent luciferase expression after G-418 selection. In this context, the system produced a similar, ~5-fold change on addition of aTc (Fig. 4.1D). Lower concentrations of aTc produced smaller increases in FLuc expression (fig 4.S3). In all experiments aTc concentrations remained well below the minimum concentration required to inhibit proliferation *in vitro*[144]. aTc-dependent regulation persisted after >12 months of continuous growth in the presence of G-418. To establish that regulation occurs at the point of translation, and not via a change in steady-state mRNA concentration, we measured FLuc mRNA by quantitative PCR (Fig. 4.S4). We also verified by anti-TetR western blot that addition of aTc does not alter the amount of TetR present in cells (Fig. 4.S5). Comparing aptamer-regulated and non-aptamer FLuc expression, we observed that inclusion of the aptamer sequence (in the absence of TetR binding) decreased expression by ~50% in one example (Fig. 4.1E), although this will likely vary based on the exact sequence context surrounding the aptamer.

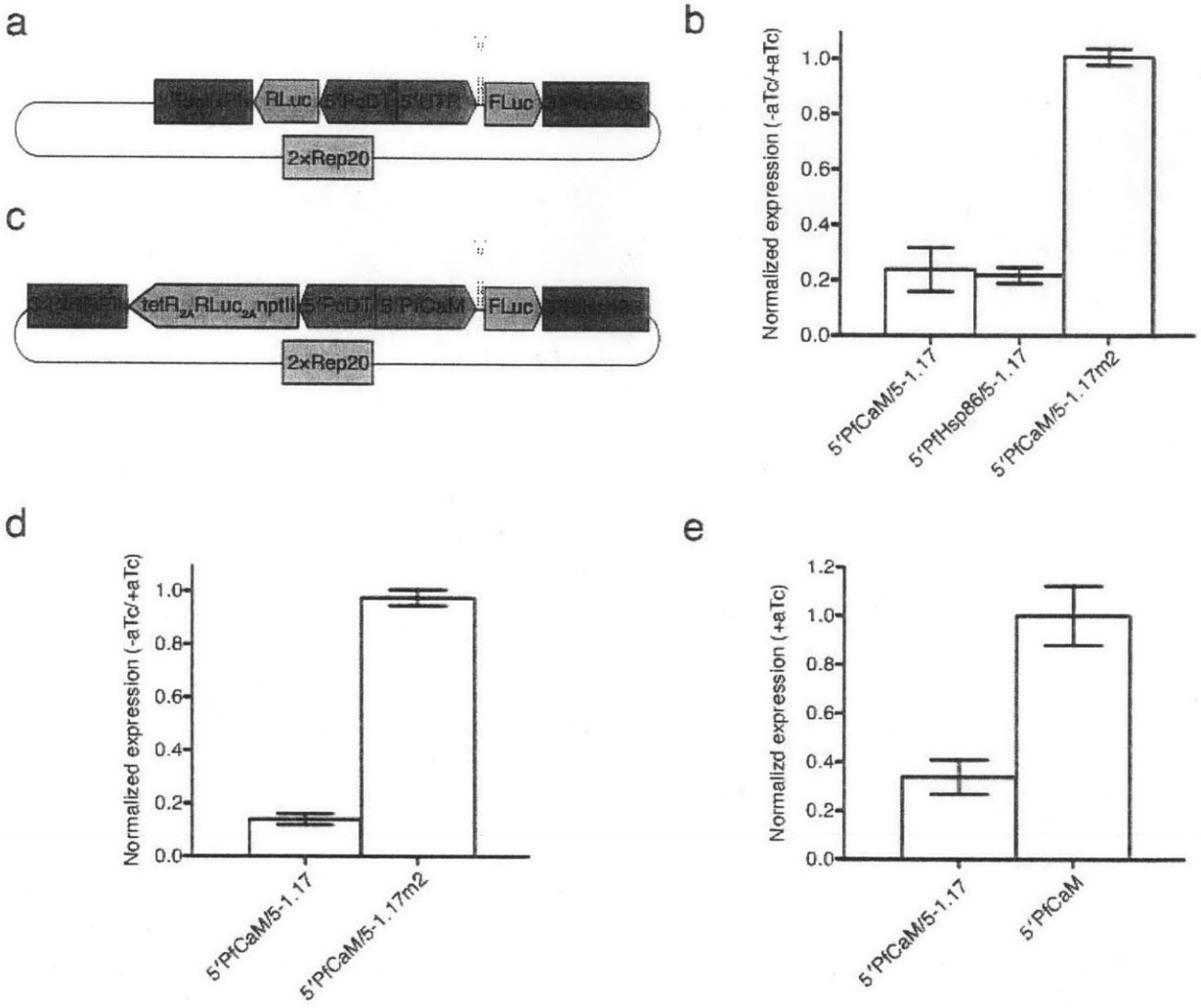

Figure 4.1. Aptamer 5-1.17 **enables aTc-dependent control of reporter gene expression.** (**A**) Transient screening vector: aptamer or control sequence (stem-loop) mediates FLuc expression in the context of a 5′UTR (PfCaM or PfHsp86) when transfected into a *tetR*$^{+}$ strain. *Renilla* luciferase (RLuc) is expressed constitutively. (**B**) Upon transient screening in the TetR-expressing strain B10, **5-1.17**, but not **5-1.17m2**, yields aTc-inducible repression of FLuc. Values are expressed as the FLuc:RLuc ratio in the -aTc condition normalized to the FLuc:RLuc ratio in the +aTc condition. (**C**) Single-vector stable episome: aptamer or control sequence mediates FLuc expression in the context of the PfCaM 5′UTR. TetR, RLuc, and a selectable marker are constitutively expressed from the same plasmid. (**D**) Upon stable expression of a single-vector construct, **5-1.17**, but not **5-1.17m2**, yields aTc-inducible repression of FLuc. (**D**)

Addition of **5-1.17** to 5′PfCaM minimally inhibits FLuc expression. All data are expressed as the mean ± s.d. of 2–4 independent experiments.

P. falciparum obeys a ~48-hour intraerythrocytic development cycle (IDC) during which individual genes are expressed in a highly dynamic, temporally variant manner[145]. To best study the IDC, the TetR-aptamer system should allow a rapid increase or decrease of target gene expression, with characteristic induction and washout times significantly less than the length of an IDC. We characterized the kinetics of aTc induction and washout by in alanine-synchronized parasites expressing a stable 5-1.17/FLuc or 5-1.17m2/FLuc reporter (Fig. 4.2). In one experiment, cultures were induced or repressed at late schizont stage: parasites maintained ± aTc were washed at ~40 hr post invasion, then grown subsequently ± aTc for another IDC (Fig. 4.2A,B). 12 hr after aTc addition, FLuc expression was indistinguishable from a culture continuously exposed to aTc. Repression after aTc washout was similarly rapid, but FLuc levels remained slightly elevated for most of the IDC. In a second experiment, induction/washout was performed at the time of invasion (very early ring stage) and FLuc expression was monitored continuously through the subsequent IDC (Fig. 4.2C,D). The changes in expression were nearly complete at the first time point (5 hr), although the magnitude of FLuc repression after aTc washout was not as great under these conditions.

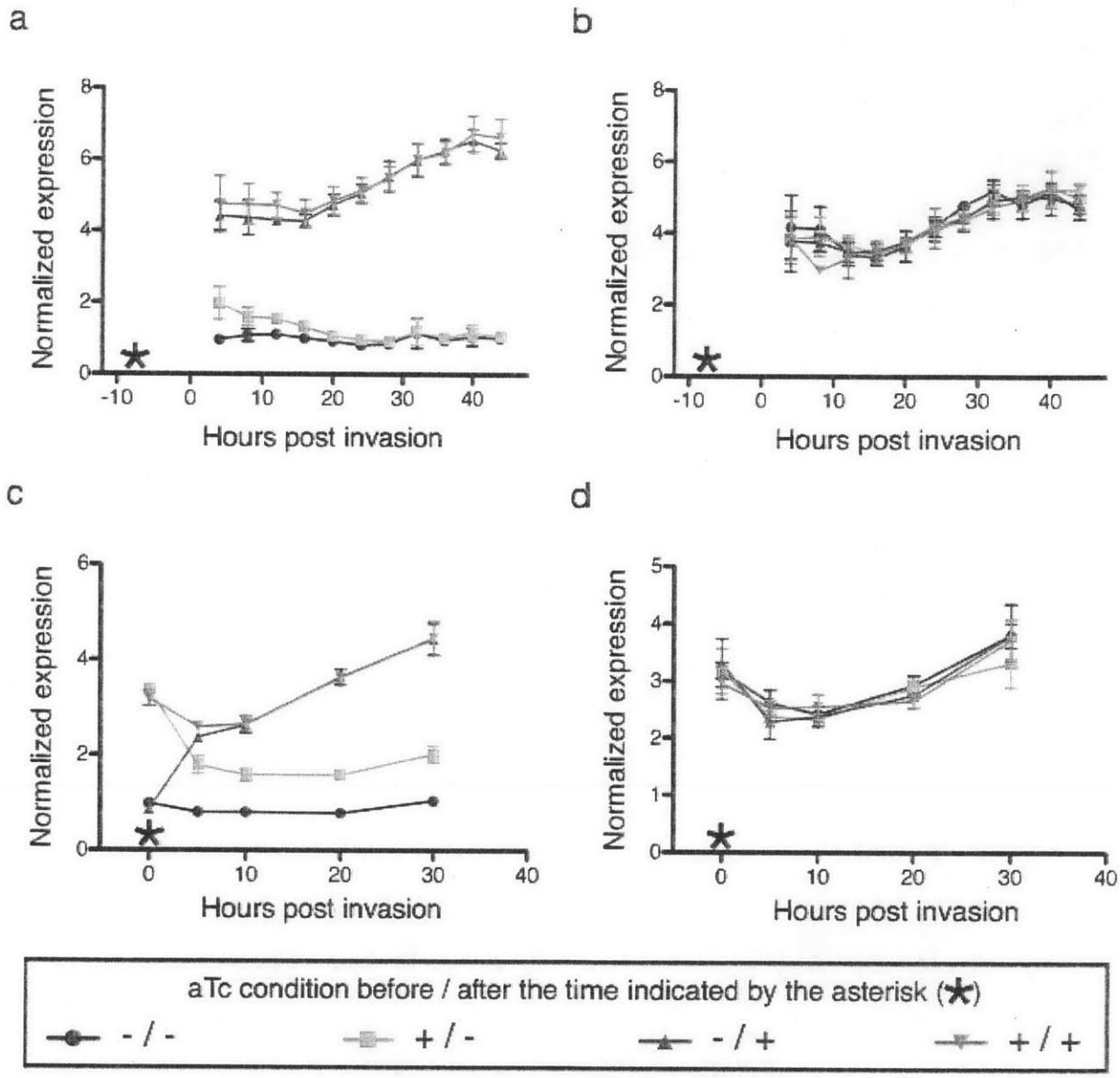

Figure 4.2. Kinetic characterization of induction and washout performed before or at the time of parasite egress and reinvasion. Dd2 derivatives harboring pSG231 (**5-1.17**/FLuc) or pSG252 (**5-1.17m2**/FLuc) were tightly synchronized by repeated sorbitol treatment. aTc (1 μM) was added, removed, or maintained at the time indicated by the asterisk. Dual luciferase measurements were taken at each time point and relative FLuc:RLuc ratios were calculated. Values represent the mean ± s.d. of 3 biological replicates.

To investigate the performance of the TetR-aptamer system in regulating endogenous parasite genes from a chromosomal context, we first made single-copy chromosomal integrants from our single-plasmid constructs. Using Bxb1 site-specific recombinase[146], we inserted *attP*$^{+}$ constructs at the *cg6* locus of strain NF54-attB[147] (Fig. 4.S6). A dual-luciferase reporter behaved similarly from this context (Fig. 4.3A), as did a green fluorescent protein (GFP) reporter, which displayed notably homogenous population behavior in both the repressed and induced state (figs. 4.3B,C). We further modified the GFP reporter by adding N-terminal targeting peptides for secretion and trafficking to the apicoplast, mitochondrion, or host erythrocyte cytoplasm[3]. As translation likely initiates within the aptamer[2], it is important to ensure that the aptamer-encoded leader peptide does not interfere with trafficking of the downstream protein. Fluorescence microscopy of live or antibody-probed cells revealed faithful protein trafficking unaffected by the inclusion of the aptamer sequence in the presence and absence of aTc (figs. 4.3D and 4.S7A). Flow cytometry analysis confirmed that aTc-dependent induction of GFP expression is homogeneous in these strains (Fig. 4.S8A), but N-terminal targeting peptides yielded low signal, making it difficult to accurately quantify the change in GFP intensity on aTc addition. Therefore, we quantified GFP expression by ELISA (Fig. 4.S8B), confirming that aTc-dependent regulation of apicoplast- and mitochondrial-targeted GFP was similar to that of the cytosolic protein.

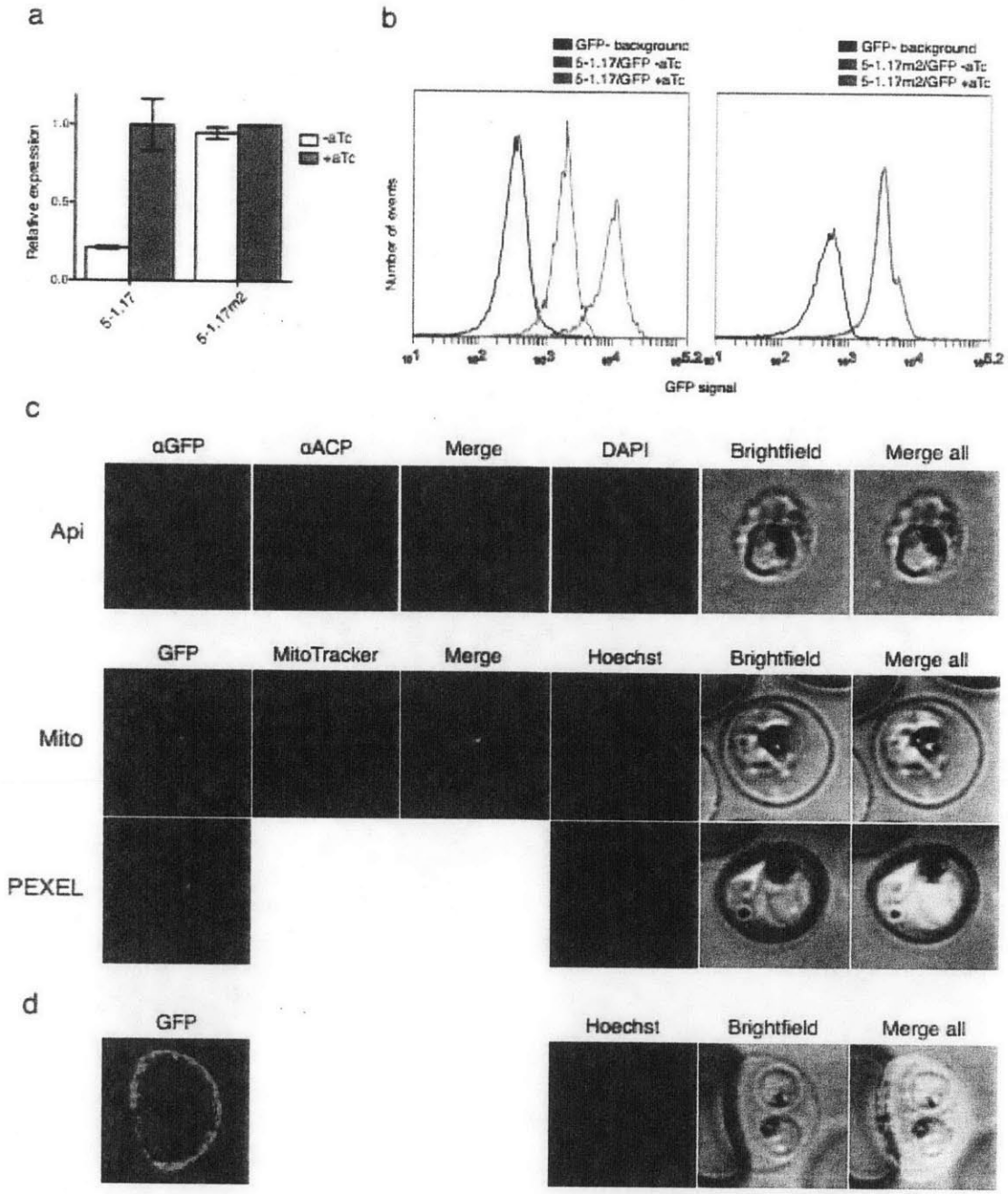

Figure 4.3. Aptamer control of a variety of protein targets. (**A**) **5-1.17**, but not **5-1.17m2**, mediates aTc-dependent control of a chromosomally integrated dual luciferase reporter. (**B**) Flow cytometry measurement of aptamer-regulated, chromosomally-integrated GFP reporter constructs. Note that the blue and red curves (± aTc) in the right panel very closely overlap. (**C**) Immunofluorescence and live fluorescence imaging of strains expressing chromosomally-integrated GFP reporters bearing an N-terminal targeting peptide (Api = apicoplast-targeted; Mito = mitochondrion-targeted; PEXEL =

protein export to erythrocyte cytoplasm). ACP = acyl carrier protein, used as an apicoplast marker. (**D**) Live fluorescence imaging of aptamer-regulated KAHRP-GFP fusion expressed from an episomal construct bearing the *P. falciparum* centromere-derived element *pfCEN5-1.5*.

We used the TetR-aptamer system to control two endogenous genes in their native 5′UTR context. First, we built an episome expressing a GFP fusion to knob-associated histidine-rich protein (KAHRP-GFP) driven by 2 kb of native KAHRP upstream sequence and controlled by 5-1.17. This episome also expresses TetR, a constitutive luciferase, and a selectable marker, and its maintenance is regulated by a *P. falciparum* centromeric DNA element[58] (Fig. 4.S9). Export of aptamer-regulated KAHRP-GFP to knob-like foci on the surface of infected erythrocytes is apparent by microscopy (Figs. 4.3E and 4.S7B), and addition of aTc yields an apparent 14-fold increase in KAHRP-GFP expression (Fig. 4.S10A). Although flow cytometry analysis showed the presence of a significant non- or low-expressing population, this is likely due to expression heterogeneity associated with the episome, and not related to the presence of the aptamer, as we only observed a single peak (with homoegenous response to aTc) in any of the chromosomally-integrated reporter strains. We confirmed expression of full-length fusion protein and aTc-dependent regulation by anti-KAHRP western blot (Fig. 4.S10B).

Second, we chose to control the *P. falciparum* chloroquine resistance transporter (PfCRT). PfCRT is an integral protein of the digestive vacuole membrane, and is thought to be both essential and inaccessible to regulation with the degradation domain system[148]. We made use of a validated zinc finger nuclease targeting the PfCRT locus[108] to efficiently insert an aptamer-regulated, chloroquine-resistant allele of PfCRT into the chloroquine-sensitive strain 3D7 (figs.

4.4A, 4.S11). The resulting strains grew at approximately half the rate of 3D7, likely due to an aptamer-mediated decrease in maximum PfCRT expression. Slow growth in the presence of decreased PfCRT expression has been observed previously[100]. We used a constitutively expressed, novel luciferase reporter (Fig. 4.S12) to measure parasite growth sensitivity to the potent chloroquine metabolite desethylchloroquine (dCq, figs. 4.4B,C). When PfCRT was controlled by an aptamer, addition of aTc increased the dCq IC_{50} from 135 to 348 nM; aTc had no effect on dCq sensitivity in the absence of a functional aptamer. Addition of verapamil, a calcium channel blocking agent with the putative ability to inhibit drug-resistant forms of PfCRT, uniformly increased dCq sensitivity, preserving the sensitivity-enhancing effect of aTc. To check whether aptamer-mediated control of dCq resistance was associated with a measurable change in PfCRT protein concentration, we measured PfCRT protein by western blot (Fig. 4.S13).

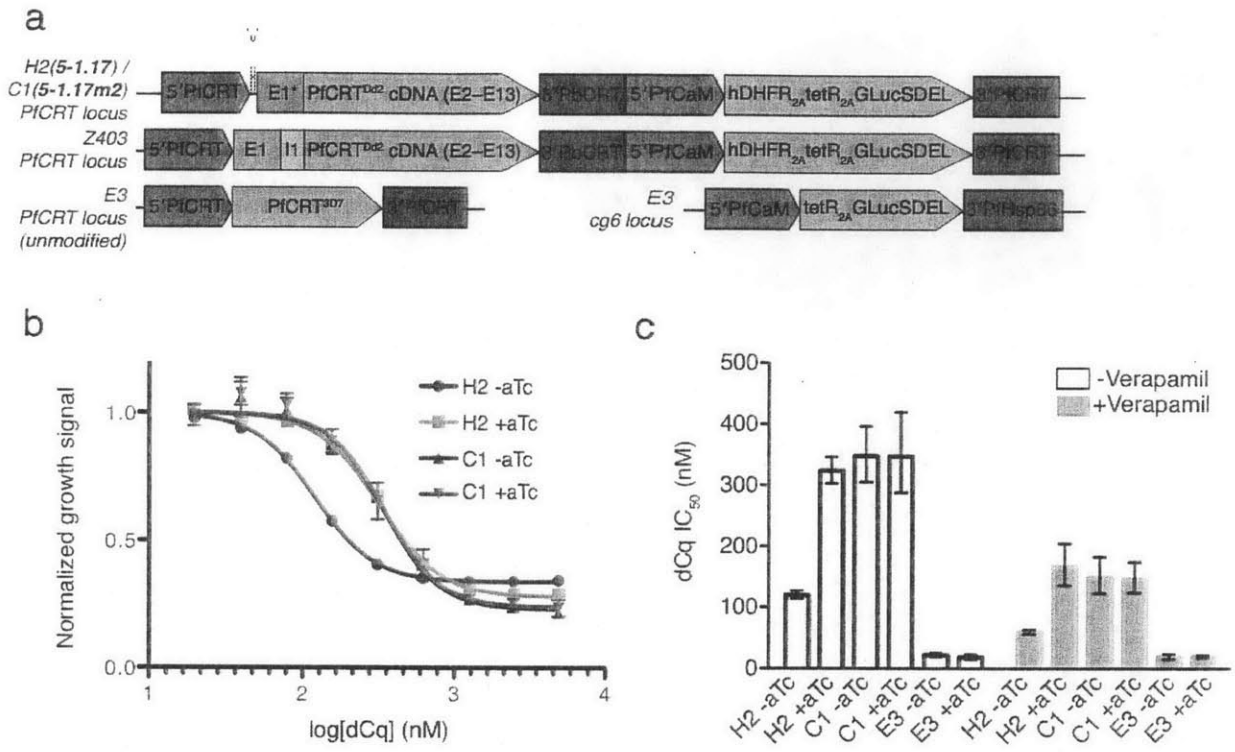

Figure 4.4. Aptamer-mediated control of *PfCRT* expression from its native chromosomal context. (**A**) Schematic of modified *PfCRT* loci generated in this study. The clone E3 has a wild-type (3D7) locus with a constitutive GLucSDEL reporter integrated at the *cg6* locus. (**B**) Aptamer-mediated repression with aTc induction of dCq sensitivity. (**C**) aTc modulates dCq sensitivity in the presence of the dCq-sensitizing agent verapamil.

In a previous study, Waller *et al.* placed PfCRT under the control of a truncated 3′UTR, generating strains with a constitutively low level of PfCRT expression, which exhibited increased sensitivity to chloroquine and dCq[100]. Here, we used the TetR-aptamer system to conditionally activate PfCRT expression in an isogenic system, establishing a direct causal relationship between PfCRT abundance and antimalarial resistance. This example demonstrates

the power of the TetR-aptamer system to generate conditional mutants of previously inaccessible genes.

We have built a tool for routine control of gene expression in *P. falciparum.* Because the system functions reliably without dependence on parameters such as 5′UTR length or sequence, downstream coding sequence, or trafficking of the regulated protein, it represents an important addition to the malaria research toolset. Furthermore, because the system was initially developed in yeast and then readily adapted for use in *P. falciparum* with minimal alteration, it may be readily adapted for use in other organisms that currently lack a robust system for inducible gene expression.

4.5. Supplementary figures

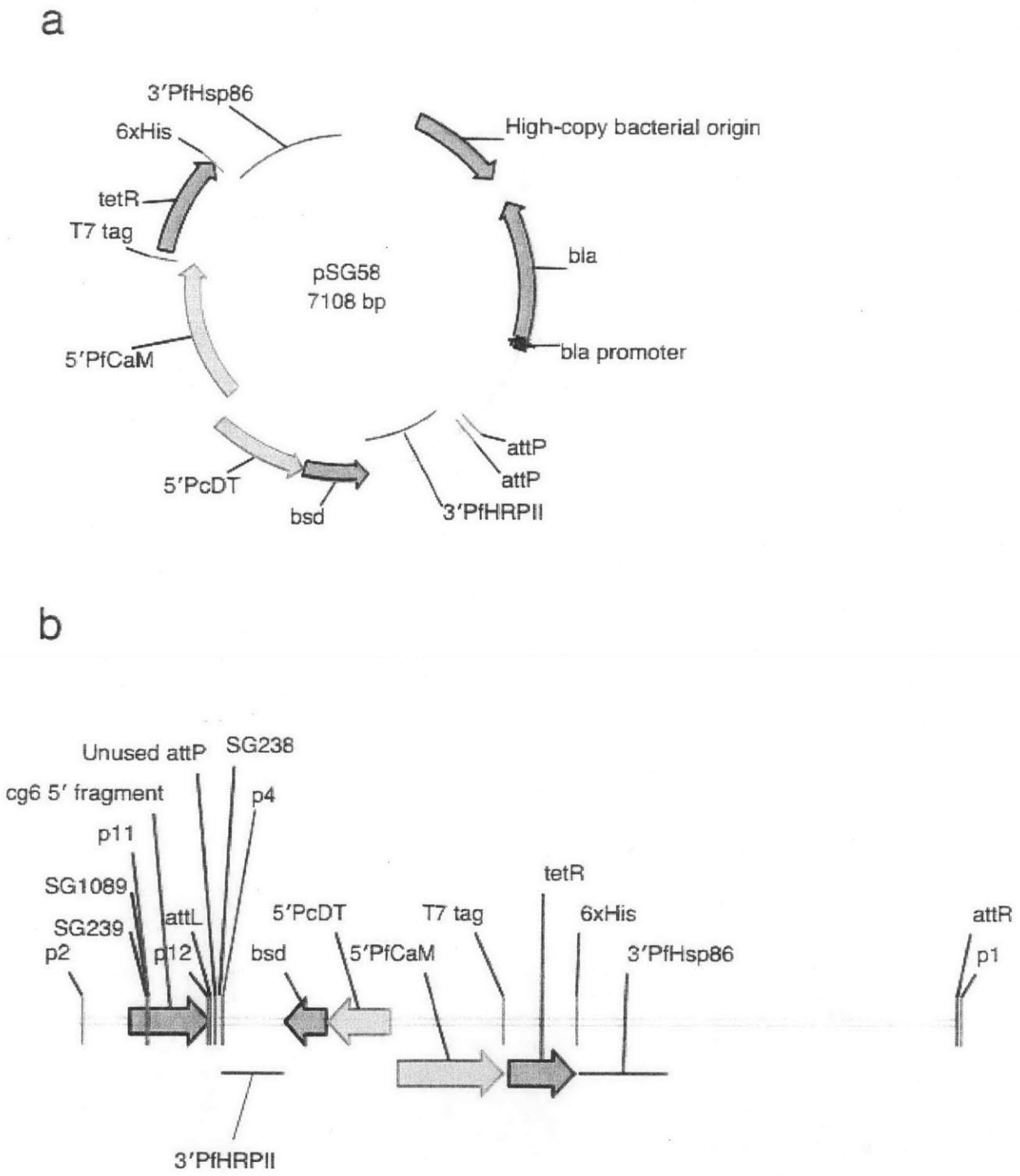

Figure 4.S1. Bxb1-mediated integration of *tetR*. (**A**) Plasmid construct carrying *tetR* flanked by an N-terminal T7 epitope tag and C-terminal 6×His epitope tag. (**B**) Schematic of integrated locus after Bxb1 recombination. p1, p2, p4, p11 and p12 are primer-binding

sites as published[55]. SG238 and SG239 are primer-binding sites used to confirm integration.

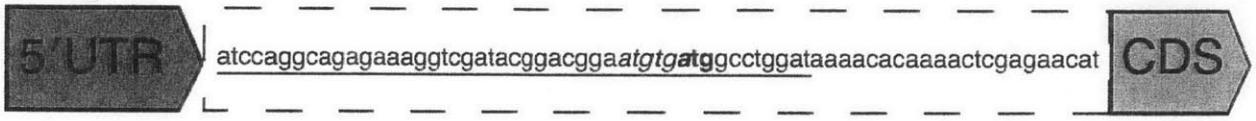

Figure 4.S2. Placement of aptamer-encoding DNA in the 5′UTR. 5′UTR = any upstream sequence. Dashed box = aptamer encoding sequence. Red **a** bases are mutated to **t** to yield the **5-1.17m2** control sequence. *Italic* bases indicate a short upstream start/stop found in many TetR-binding aptamers[1,2]. **Bold** bases indicate the productive start codon that is kept in frame with the downstream coding sequence. CDS = any coding sequence of interest. Note that CDS may carry its own 5′ start codon, or not, with no effect on expression or regulation by aTc.

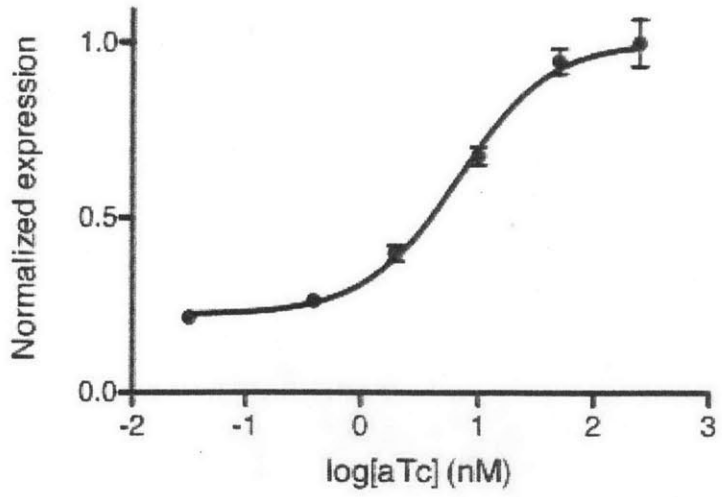

Figure 4.S3. Dose-dependent response to induction with aTc. Dd2 parasites were stably transfected with a dual-luciferase reporter and incubated for 48 hr with in the indicated concentration of aTc. Values represent aptamer-regulated FLuc signal

normalized to constitutive RLuc signal. Data represent the mean ± s.d. of three biological replicates.

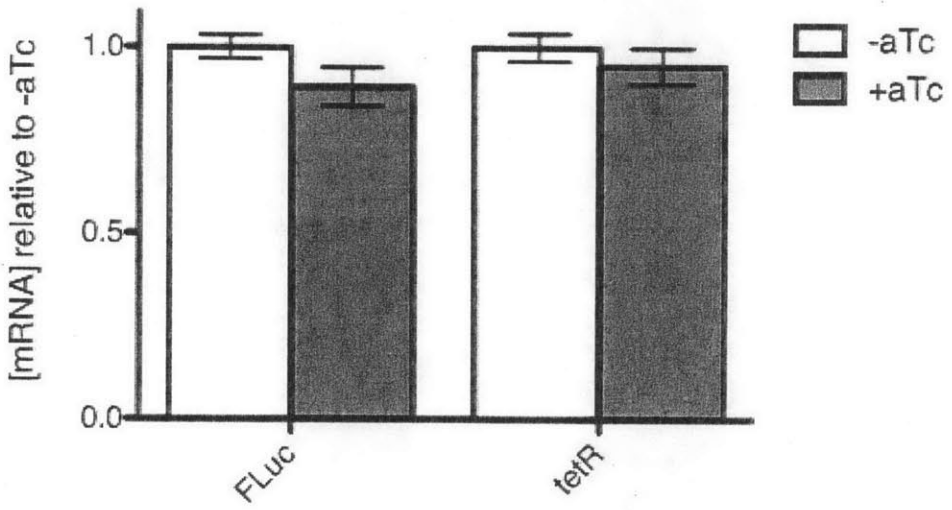

Figure 4.S4. Induction with aTc does not affect steady-state mRNA concentration. [mRNA] was calculated by TaqMan quantitative PCR of aptamer-regulated *FLuc* and *tetR* mRNAs in strain Dd2/pSG231 (**5-1.17**/FLuc), in the presence or absence of aTc.

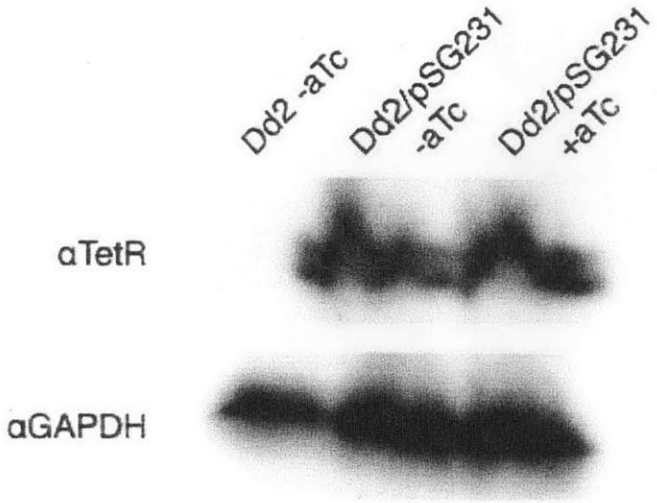

Figure 4.S5. Induction with aTc does not affect TetR concentration. TetR was assayed by western blot in Dd2 or Dd2/pSG231, in the repressed or induced state. GAPDH was used as a loading control.

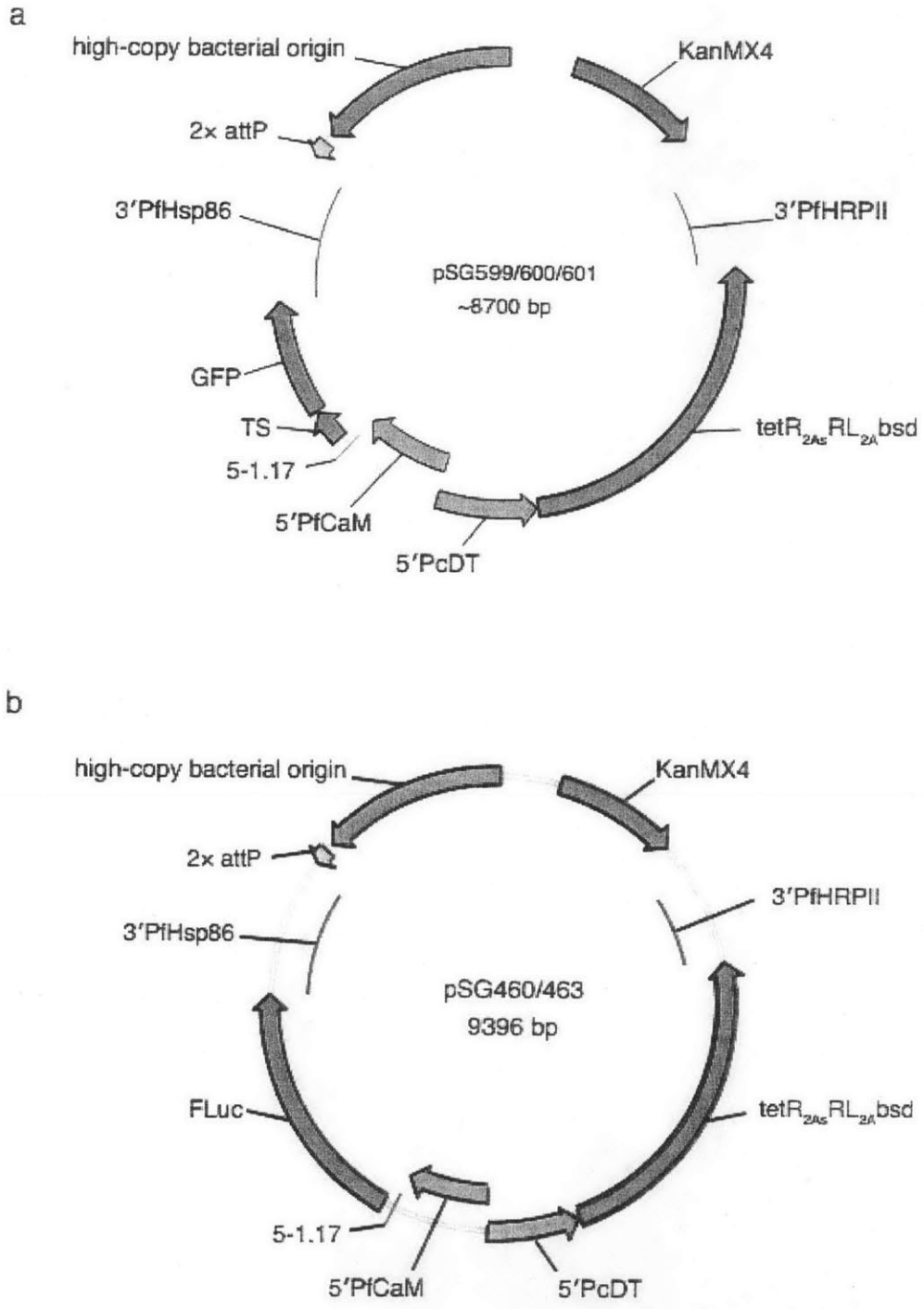

Figure 4.S6. Plasmid constructs used for Bxb1 integration. (**A**) Constructs for expression of aptamer-regulated GFP with N-terminal targeting sequences. TS = N-terminal targeting sequence: pSG599 = apicoplast targeting peptide; pSG600 = mitochondrial targeting peptide; pSG601 = PEXEL protein export element. (**B**)

Constructs for expression of aptamer-regulated FLuc. pSG460 is shown; in pSG463, **5-1.17** is replaced by **5-1.17m2.**

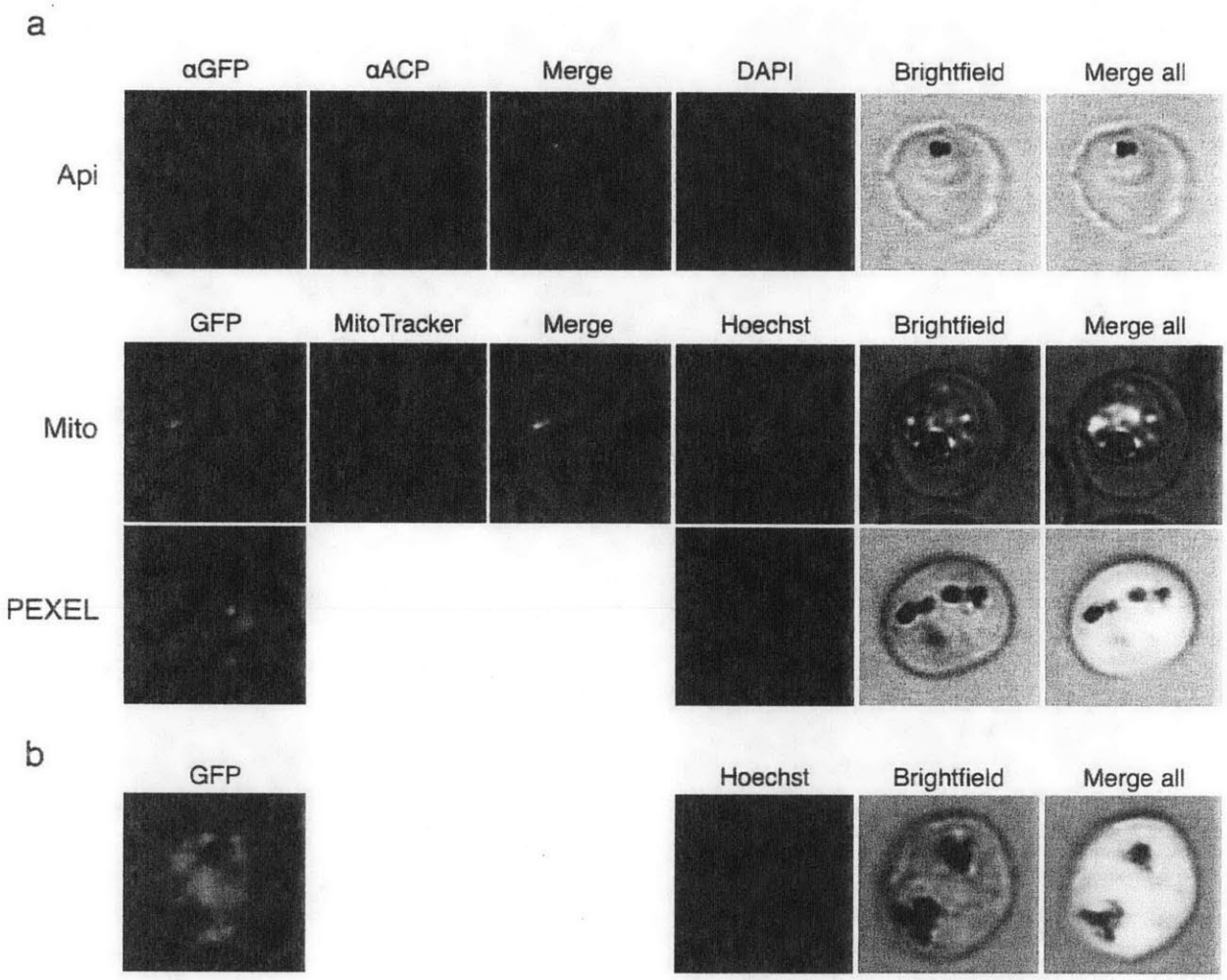

Figure 4.S7. Accurate subcellular localization of aptamer-regulated fluorescent proteins in the repressed (-aTc) state. Fluorescence microscopy indicates that in the repressed state, as in the induced state (Fig. S4.3), aptamer regulation does not interfere with proper subcellular localization of fluorescent reporters (**A**) or KAHRP-GFP fusion (**B**).

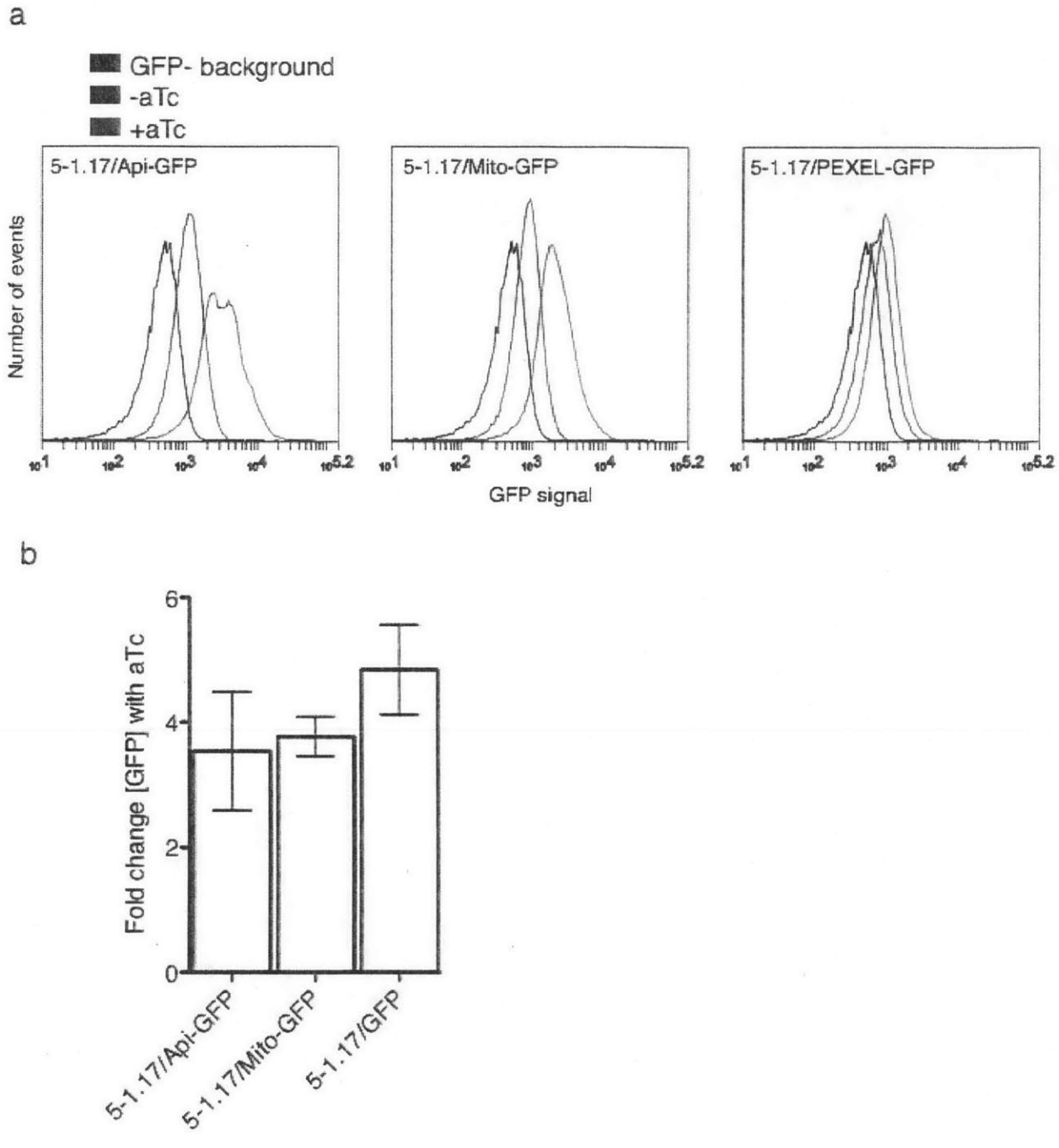

Figure 4.S8. aTc-regulated GFP expression from chromosomally-integrated GFPs with N-terminal targeting peptides. (**A**) Flow cytometry analysis of targeted GFP constructs grown ± aTc. (**B**) Quantification of aTc-regulated GFP expression by ELISA, with cytosolic GFP (no N-terminal targeting peptide) for comparison. Data represent the mean ± s.d. of four technical replicates.

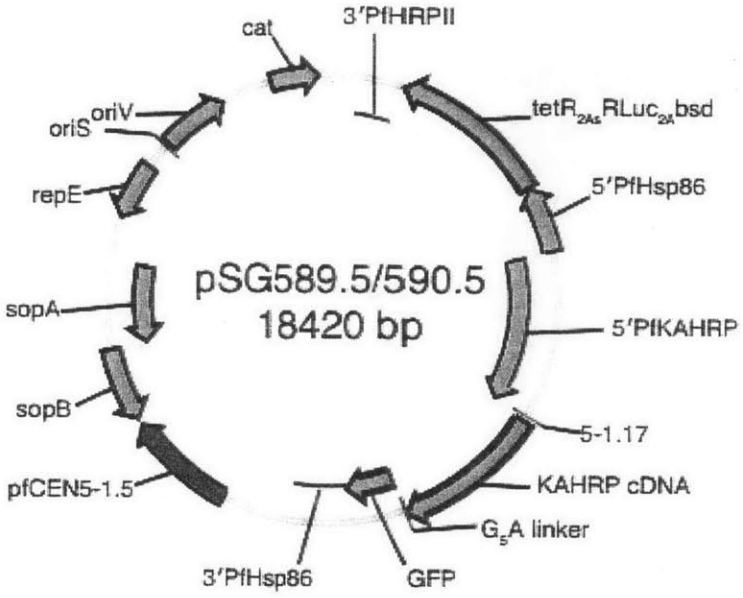

Figure 4.S9. Constructs for episomal expression of aptamer-regulated KAHRP-GFP protein fusion. pSG589.5 is shown. pSG590.5 is shown; pSG589.5 contains no aptamer sequence and the native 5′KAHRP-KAHRP CDS junction sequence is preserved.

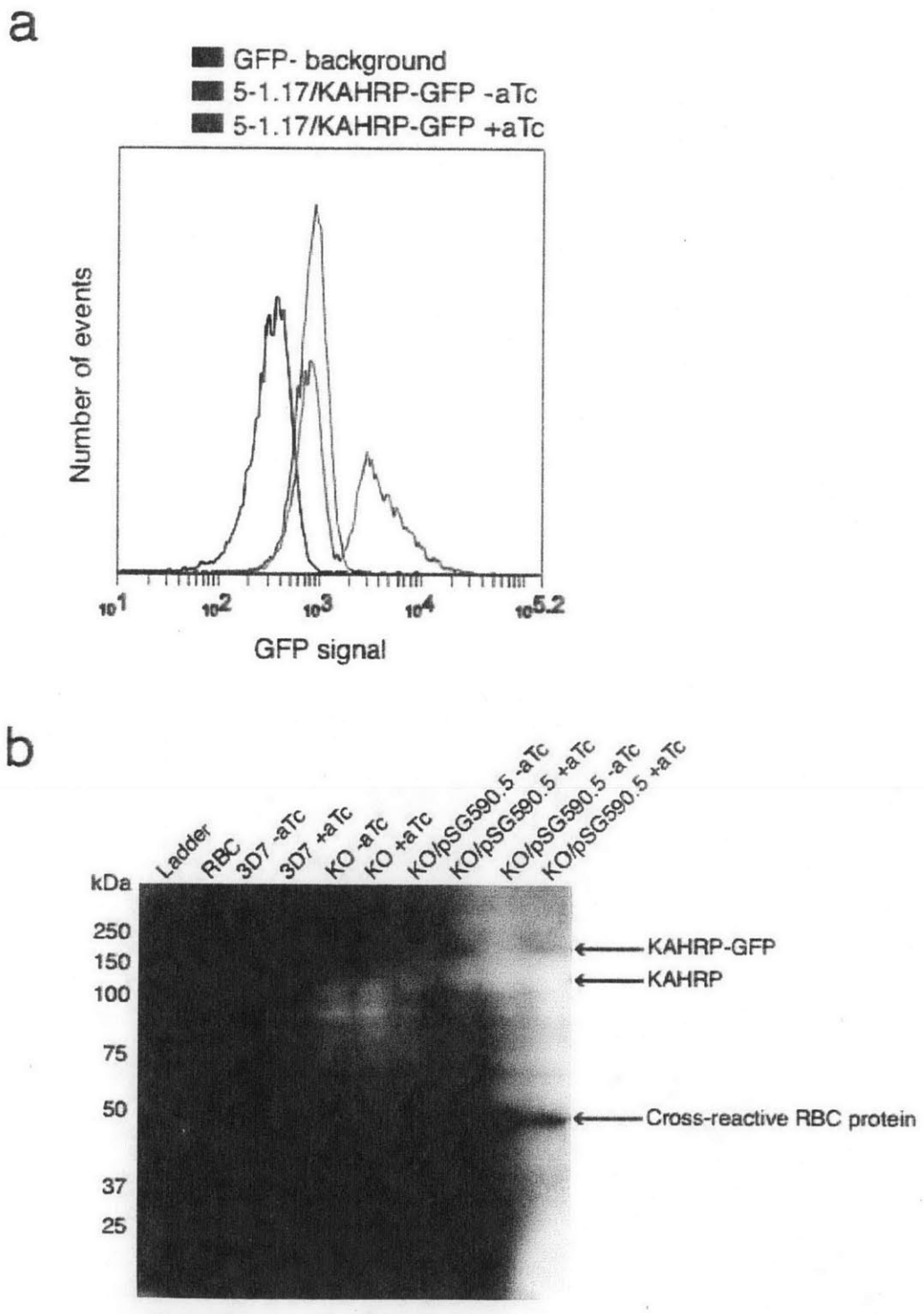

Figure 4.S10. aTc-dependent regulation of episomally-expressed KAHRP-GFP fusion protein by aptamer 5-1.17. (**A**) Flow cytometry analysis of *KAHRP*⁻ host transfected with episome expressing 5′KAHRP/5-1.17/KAHRP-GFP (pSG590.5). (**B**) Anti-KAHRP western blot of iRBC ghost fractions. *KO, KAHRP*⁻ knockout strain. Note

that KAHRP and KAHRP-GFP proteins run more slowly than predicted, as previously observed[149].

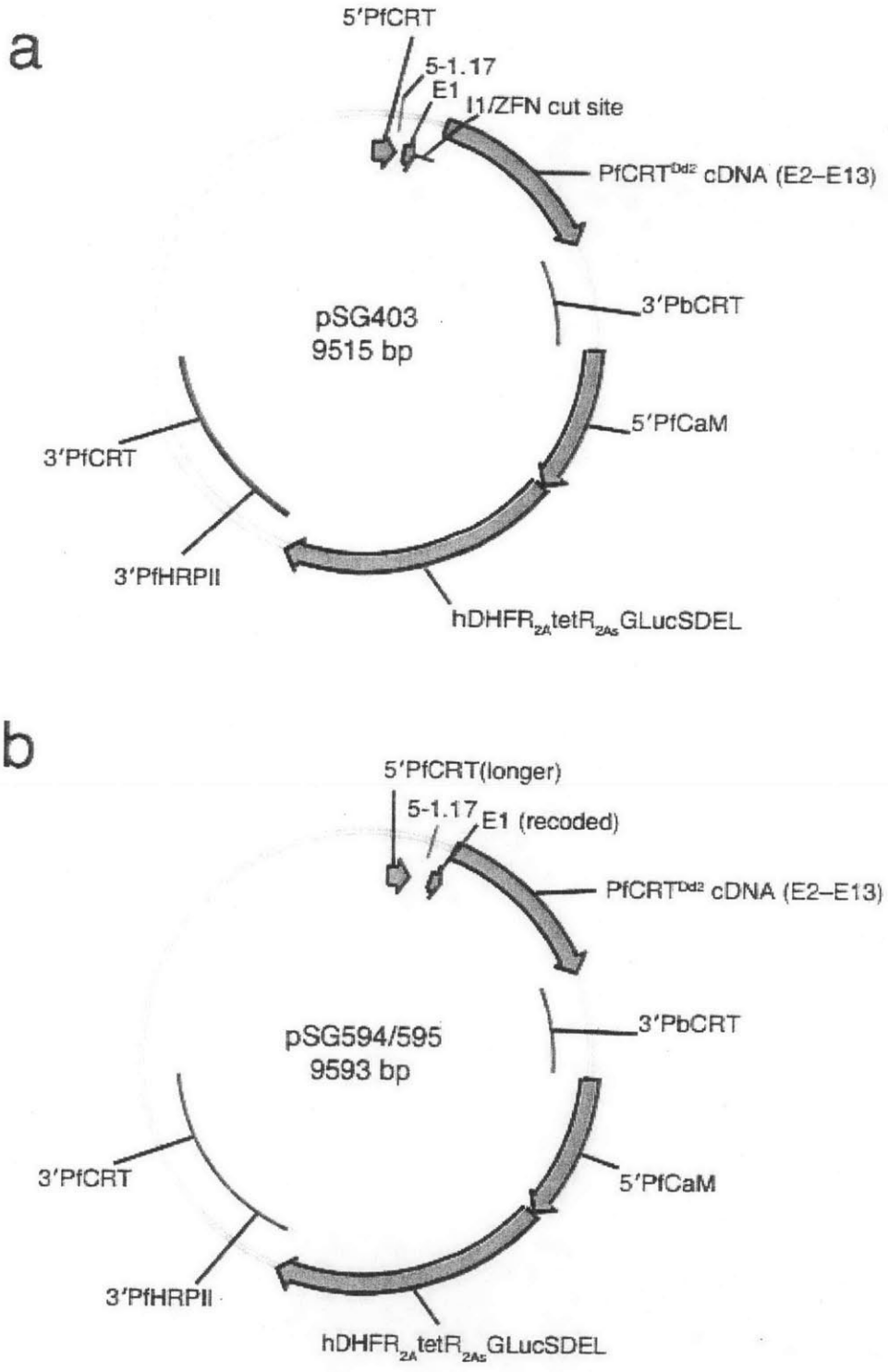

Figure 4.S11. Gene targeting vectors for ZFN-mediated replacement of *PfCRT*. (**A**) pSG403 is based on the donor vector from Fig. 3 of Straimer *et al.* The hDHFR resistance

marker of that vector has been replaced with a 2A-separated cassette additionally expressing tetR and GLucSDEL. DNA encoding **5-1.17** was inserted upstream of Exon 1. When co-transfected with the ZFN-expressing vector, sequence downstream of Exon 1 was replaced with vector-borne sequence, but no integration of the aptamer could be detected by PCR, indicating efficient repair downstream of the aptamer sequence. (**B**) pSG594 and pSG595 are based on pSG403, but Exon 1 has been recoded to change its DNA sequence without altering the polypeptide sequence of PfCRT. Additionally, Intron 1 (including the ZFN cut site) has been removed. This forces repair of the ZFN-induced break to occur upstream in the 5′UTR homology region (which has been expanded), including the aptamer in the repaired strain.

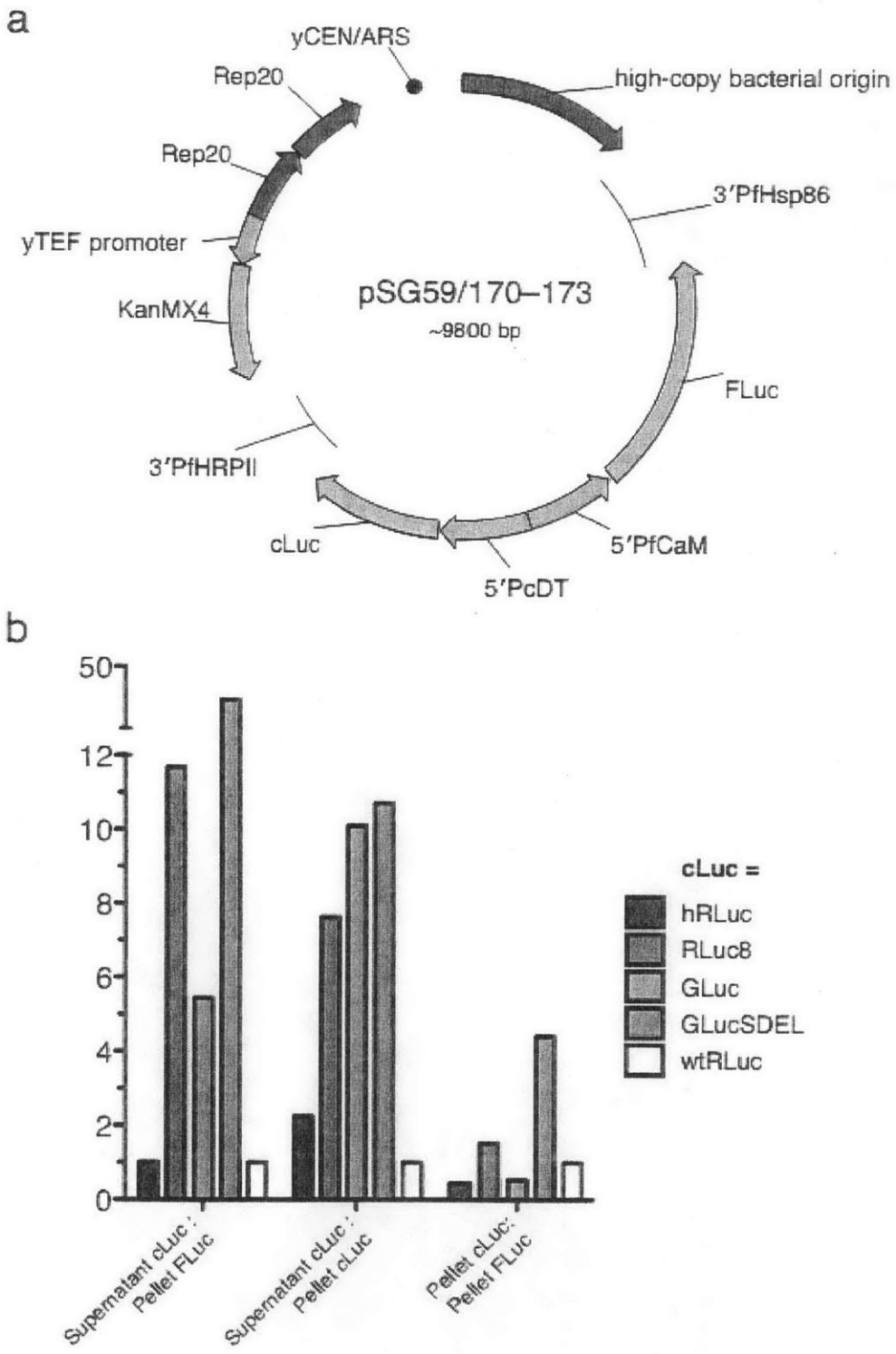

Figure 4.S12. Screening coelenterazine-utilizing luciferase genes for improved reporter activity in *P. falciparum*. (**A**) Plasmid vectors for transient transfection. Firefly luciferase (FLuc) is used for normalization. cLuc = coelenterazine-utilizing luciferase: in

pSG59, wild-type *Renilla* luciferase (wtRLuc); pSG170, human codon-optimized *Renilla* luciferase (hRLuc, Promega); pSG171, enhanced *Renilla* luciferase[150] (RLuc8); pSG172, human codon-optimized *Gaussia* luciferase bearing its native N-terminal secretion signal[151] (GLuc); pSG173, GLuc with the SDEL peptide fused to its C terminal (GLucSDEL). (**B**) Transient luciferase activity. *P. falciparum* 3D7 was transfected with pSG59, 170, 171, 172, or 173 and grown for 96 hr before dual luciferase measurement of iRBC and cell culture supernatant fractions. Luminescence values are expressed as ratios between FLuc or cLuc signal in various fractions, and normalized to pSG59 (wtRLuc). In mammalian cells, addition of a C-terminal ER retention signal to GLuc causes it to be retained intracellularly, instead of being secreted. In *P. falciparum,* that the ER retention signal SDEL does not affect the ratio of supernatant: pellet GLuc signal, presumably because the N-terminal GLuc secretion signal does not allow protein export from the *P. falciparum*-infected RBC. However, SDEL does increase GLuc activity ~10-fold, possibly due to enhanced formation of essential disulfide bonds in the ER. This greatly increased activity, together with the inherent stability of GLuc[151], makes GLucSDEL a convenient reporter for non-destructive monitoring of *P. falciparum* growth via monitoring of cell culture supernatant activity.

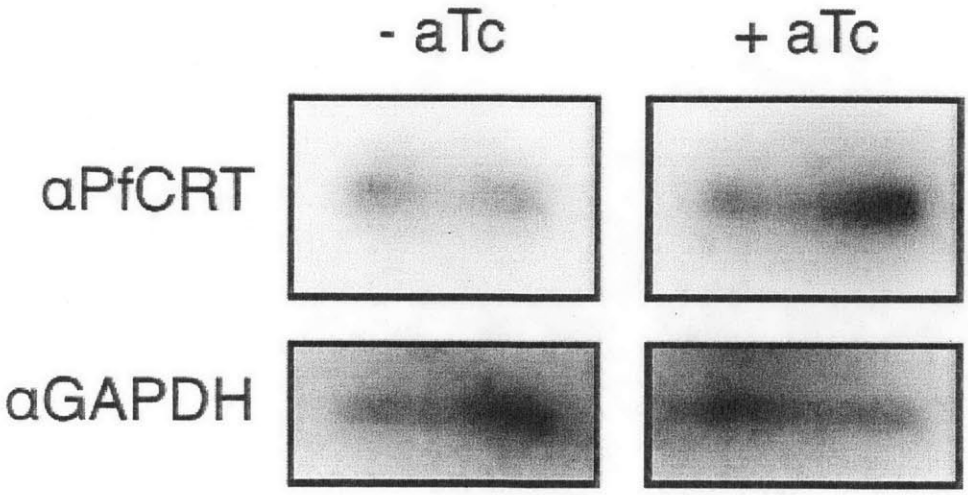

Figure 4.S13. Measurement of PfCRT protein by western blot. Cell line H2 (**5-1.17/*PfCRT***) was grown ± aTc for 48 hr and detergent-solublized extracts were prepared. Equal amounts of total protein were loaded in each lane.

4.6. Experimental procedures

Molecular cloning. Cloning was performed by standard restriction/ligation, yeast homologous recombination, or Gibson assembly methods as described[3]. Unless otherwise noted, enzymes were from New England Biolabs and reagents were from Sigma-Aldrich, Research Products International, GoldBio Technology, or Alfa Aesar. PCR of parasite DNA was performed on purified DNA template or directly on iRBCs (diluted 1:50 into the reaction) with Hemo KlenTaq mixed 15:1 (v:v) with PfuTurbo (Agilent). Plasmid DNA was prepared for transfection with Maxi Columns (Epoch Life Science) or the Xtra Midi Kit (Clontech).

Plasmid DNA sequences and sources. GenBank files for the constructs used in this study are provided as supplementary files. *FLuc* is wild-type firefly luciferase DNA carrying the K549E

mutation for cytosolic targeting in yeast[152]. RLuc8 DNA was provided by Andreas Loening and Sanjiv Gambhir (Stanford University). Green fluorescent protein is GFPmut2[153], cloned from pFGNr (Malaria Research and Reference Reagent Resource Center, MR4; www.mr4.org).

***P. falciparum* culture and transfection.** Parasites were grown in human erythrocytes at 2% hematocrit under 5% O_2 and 5% CO_2 in RPMI-1640 media supplemented with 5 g/L Albumax II (Life Technologies), 2 g/L $NaHCO_3$, 25 mM HEPES-K pH 7.4, 1 mM hypoxanthine and 50 mg/L gentamicin. Transfections used 25–75 μg of each plasmid per 200 μL packed RBC, adjusted to 50% hematocrit, and were performed by the spontaneous DNA uptake method[134]. Electroporation was performed on a Bio-Rad Gene Pulser with 8 square-wave pulses of 365 V for 1 ms, separated by 100 ms. Transfected parasites were selected with 2.5 mg/L Blasticidin S, 1.5 μM DSM-1, 5 nM WR9920 (Jacobus Pharmaceuticals) and/or 250 mg/L G-418 beginning 4 days after transfection. In the case of Bxb1-mediated integration, no selection was applied on the integrase-expressing plasmid. All chromosomally-modified strains were confirmed by diagnostic PCR and sequencing of PCR products.

Luciferase measurement. Dual measurement of FLuc and coelenterazine-utilizing luciferases was performed either with the Dual Luciferase Assay (Promega) or with homemade reagents as described[2]. Single measurement of GLucSDEL was performed by adding 1 volume of sample to 10 volumes of assay buffer (0.5× phosphate-buffered saline, 0.025% Nonidet P-40, 1 μg/mL coelenterazine). Luciferase intensity was measured on a GloMax 20/20 luminometer (Turner Biosystems) or a Spectramax L plate reader (Molecular Devices).

Western blot. For analysis of parasite fractions, ~10^6 late-stage parasites were harvested by lysis of infected RBCs with 0.5 g/L saponin in phosphate-buffered saline (PBS) and then lysed by

heating in Laemmli sample buffer at 95 °C for 10 min. For analysis of iRBC membrane-bound KAHRP, iRBC ghosts were prepared by lysis and repeated washing of iRBC membranes in 5 mM Tris-Cl, pH 8.0, supplemented with 10 μL/mL Protease Inhibitor Cocktail (Sigma). Ghosts were then solubilized by mixing 1:1 with 2× Laemmli sample buffer and heating at 55 °C for 15 min. Each lane was loaded with sample prepared from an equivalent amount of RBCs or iRBCs. For analysis of solubilized PfCRT extracts, equal numbers of parasites were harvested by saponin lysis and resuspended in 5 volumes of PBS. After 3 freeze-thaw cycles, n-decyl-β-D-maltopyranoside (Affymetrix) was added to a final concentration of 1%. Proteins were extracted by agitation for 1 hr at 4° C and insoluble material was removed by centrifugation at 18,500*g* for 10 min. After separation by SDS-PAGE, proteins were transferred to a PVDF membrane and probed with a mouse monoclonal antibody against TetR (Clontech), KAHRP (antibody provided by Diane Taylor), or PfCRT (antibody provided by David Fidock); or rabbit polyclonal antibody against GAPDH (Abcam ab9485). Blots were then imaged using a horseradish peroxidase-coupled secondary antibody and SuperSignal West Femto substrate (Thermo Scientific).

Flow cytometry. Late-stage iRBCs were enriched with a MACS LD column (Miltenyi Biotec) and stained for nucleic acid content with 1 μM SYTO 61 (Life Technologies) before analysis on an Accuri C6 flow cytometer (BD Biosciences). GFP was analyzed in FL1-H and SYTO 61 in FL4-H. Parasites were gated on high FL4 intensity to eliminate contaminating uninfected RBCs. Addition of aTc to nonfluorescent background iRBCs or uninfected RBCs had no effect on FL1-H intensity.

GFP ELISA. Late-stage iRBCs were enriched with a MACS LD column (Miltenyi Biotec) and RBC membranes were lysed with 0.5 g/L saponin in PBS. Parasites were washed repeatedly with

PBS to remove all trace of red color and stored at -80 °C. As hemoglobin-depleted parasite fractions were analyzed, KAHRP-targeted GFP was not analyzed by this method. Parasite lysates were prepared and GFP content was analyzed with the GFP ELISA kit (Abcam ab117992). Wells were loaded with equal amounts of total extracted protein as determined by BCA protein assay (Thermo Scientific). GFP$^-$ 3D7 parasites were used as a background control and to establish that the inclusion of aTc did not affect background signal intensity.

Sample preparation for immunofluorescent microscopy. Late-stage iRBCs were prepared by a method modified from Gallagher *et al.* Cells were pelleted and resuspended in fixation solution (PBS + 4% formaldehyde + 4% glutaraldehyde) at 0.2% hematocrit. Fixation solution was prepared fresh from reagents in methanol-free ampules. Cells were immediately spotted onto polyethyleneimine-coated glass slides and incubated at room temperature for 30 min. An Immunopen (EMD Biosciences) was used to contain the cell suspension in a small droplet. Attached cells were washed with PBS and permeablized by incubating in PBS + 1% Triton X-100 for 10 min. Cells were then washed with PBS and blocked/quenched with PBS + 3% bovine serum albumin + 150 mM glycine for 1 hr. Cells were probed with mouse monoclonal anti-GFP (12A6, 1:25) and rabbit anti-ACP (1:500) in PBS + 3% BSA at 4 °C overnight, washed with PBS, and probed with CF488A-conjugated donkey anti-mouse (Biotium 20014-1, 1:1000) and CF647-conjugated goat anti-rabbit (Biotium 20282-1, 1:1000) for 1 hr at room temperature in the dark. Slides were washed with PBS and a #1.5 coverslip was mounted with ProLong Gold + DAPI (Life Technologies) followed by incubation in the dark for at least 12 hr before sealing and imaging. The GFP monoclonal antibody 12A6, developed by P. J. Sanchez, K. J. Daniels, and D. R. Soll, was obtained from the Developmental Studies Hybridoma Bank, developed under the

auspices of the NICHD and maintained by The University of Iowa, Department of Biology, Iowa City, IA 52242.

Sample preparation for live fluorescence microscopy. Late-stage iRBC cultures were resuspended in Albumax-free growth media and attached to polyethyleneimine-coated glass-bottom dishes (MatTek) before staining for 30 min with 2 μg/mL Hoechst 33342 (Life Technologies), washing, and resuspension in Opti-Klear imaging buffer (Marker Gene Technologies). For mitochondrial staining, cultures were incubated at 37 °C for 15 min in growth media + 25 nM MitoTracker Deep Red FM (Life Technologies) before attachment.

Fluorescent microscopy. Imaging was performed on a Zeiss Axio Observer microcope with a 100× objective. For each field, a 30-image Z-stack was deconvolved with Huygens Essential software (SVI). Maximum intensity Z-projections, false coloring, and image merging were performed with ImageJ software[136].

Quantitative PCR. qPCR was performed as described[2] with PrimeTime 5′ nuclease assays (Integrated DNA Technologies). *P. falciparum* RNA was extracted directly from late-stage iRBCs with Trizol (Life Technologies) and purified with an RNeasy kit (Qiagen). cDNA was prepared with SuperScript III reverse transcriptase (Life Technologies). PF13_0170 (*glutaminyl-tRNA synthetase*) was used as a housekeeping reference gene. Primer and probe sequences are listed in Table 4.S2.

ZFN-mediated editing of the *PfCRT* locus. Parasites were co-transfected with ZFN-expressing plasmid (expressing blasticidin resistance) and a donor vector (expressing WR99210 resistance). 4 days after transfection, cultures were selected with 2.5 μg/mL blasticidin and 2.5 nM

WR99210 for 2 weeks, at which point blasticidin selection was removed. Transfections were maintained on 0.4 μM aTc for the first 2 months after transfection, although it was later confirmed that growth in the presence or absence of aTc is similar for all strains generated. Proper repair of the *PfCRT* locus with aptamer integration was confirmed by PCR in bulk culture and after cloning by limiting dilution. The ZFN-expressing plasmid and donor vector were generously provided by David A. Fidock (Columbia University). We thank Andrew H. Lee (Columbia University) for technical advice.

4.7. Supplementary tables

Table 4.S1. *P. falciparum* cell lines and DNA constructs used in this study

Figure	Cell line name	Clone/other name	Functional designation
4.1B	3D7-tetR	B10	Constitutive expression of integrated *tetR*
4.1B	B10/pSG75 (transient)		5′PfCaM/5-1.17/FLuc; 5′PcDT/RLuc
4.1B	B10/pSG76 (transient)		5′PfCaM/5-1.17m2/FLuc; 5′PcDT/RLuc
4.1B	B10/pSG216 (transient)		5′PfHsp86/5-1.17/FLuc; 5′PcDT/RLuc
4.1D, 4.1E, 4.2,	Dd2/pSG231		5′PfCaM/5-1.17/FLuc; 5′PfHsp86/tetR$_{2As}$RLuc$_{2A}$neo

4.S3,			
4.S4,			
4.S5			
4.1D,	Dd2/pSG252		5′PfCaM/5-1.17m2/FLuc;
4.2,			5′PfHsp86/tetR$_{2As}$RLuc$_{2A}$neo
4.S5B			
4.1E	Dd2/pSG230		5′PfCaM/FLuc;
			5′PfHsp86/tetR$_{2As}$RLuc$_{2A}$neo
4.3A	NF54-attB	H6	5′PfCaM/5-1.17/FLuc;
	cg6::pSG460		5′PcDT/tetR$_{2As}$RLuc$_{2A}$bsd
4.3A	NF54-attB	G11	5′PfCaM/5-1.17m2/FLuc;
	cg6::pSG463		5′PcDT/bsd$_{2As}$tetR$_{2A}$GLucSDEL
4.3B	NF54-attB	F9	5′PfCaM/5-1.17/GFP;
	cg6::pSG584		5′PcDT/tetR$_{2As}$RLuc$_{2A}$bsd
4.3B	NF54-attB	D2	5′PfCaM/5-1.17m2/GFP;
	cg6::pSG586		5′PcDT/tetR$_{2As}$RLuc$_{2A}$bsd
4.3C,	NF54-attB		5′PfCaM/5-1.17/Api-GFP;
4.S7A,	*cg6*::pSG599		5′PcDT/tetR$_{2As}$RLuc$_{2A}$bsd

4.S8			
4.3C, 4.S7A, 4.S8	NF54-attB *cg6*::pSG600		5′PfCaM/5-1.17/Mito-GFP; 5′PcDT/tetR$_{2As}$RLuc$_{2A}$bsd
4.3C, 4.S7A, 4.S8A	NF54-attB *cg6*::pSG601		5′PfCaM/5-1.17/PEXEL-GFP; 5′PcDT/tetR$_{2As}$RLuc$_{2A}$bsd
4.3D, 4.S7B	MRA-554(*KAHRP*$^{-}$)/pSG590.5		5′PfKAHRP/5-1.17/KAHRP(cDNA); 5′PfHsp86/ tetR$_{2As}$RLuc$_{2A}$bsd
4.4	3D7 *PfCRT*::pSG594	H2	5′PfCRT/5-1.17/PfCRT-E1^{re}-PfCRT^{Dd2}(cDNA); 5′PfCaM/hDHFR2AtetR2AGLucSDEL
4.4A	3D7 *PfCRT*::pSG595	C1	5′PfCRT/5-1.17m2/PfCRT-E1^{re}-PfCRT^{Dd2}(cDNA); 5′PfCaM/hDHFR2AtetR2AGLucSDEL
4.4	3D7 *PfCRT*::pSG403		5′PfCRT/5-1.17/PfCRT-E1-I1-PfCRT^{Dd2}(cDNA); 5′PfCaM/hDHFR2AtetR2AGLucSDEL
4.4	NF54-attB	E3	5′PfCaM/5-1.17/FLuc; 5′PcDT/bsd;

	cg6::pSG479::pSG593	5′PfCaM/tetR$_{2As}$GLucSDEL
4.S12	3D7/pSG59 (transient)	5′PfCaM/FLuc; 5′PcDT/wtRLuc
4.S12	3D7/pSG170 (transient)	5′PfCaM/FLuc; 5′PcDT/hRLuc
4.S12	3D7/pSG171 (transient)	5′PfCaM/FLuc; 5′PcDT/RLuc8
4.S12	3D7/pSG172 (transient)	5′PfCaM/FLuc; 5′PcDT/GLuc
4.S12	3D7/pSG173 (transient)	5′PfCaM/FLuc; 5′PcDT/GLucSDEL

Table 4.S2. Primers and probes used for qPCR

Oligonucleotide	Sequence
FLuc fwd	5′ TCCTCTGACACATAATTCGCC
FLuc rev	5′ GCTATTCTGATTACACCCGAGG
FLuc probe	5′ /HEX/TCC AGA TCC/Zen/ACA ACC TTC GCT TCA AA/IABkFQ/
tetR fwd	5′ TGC TTA ATG AGG TCG GAA TCG
tetR rev	5′ ATC TCA ATG GCT AAG GCG TC
tetR probe	5′ /6-FAM/TTC TGG GCG/Zen/AGT TTA CGG GTT GT/IABkFQ/
PF13_0170 fwd	5′ GAG TTG AGA TGT ACC TAC GAC C

PF13_0170 rev	5′ GTT ATC CCC ATC ACC ACC TTG
PF13_0170 probe	5′ /Cy5/TGA TTG GTA CAG AGC CCA CTT TTA CTG C/IAbRQSp/

CHAPTER 5. FUTURE CHALLENGES AND OPPORTUNITIES

5.1. Unanswered questions about mechanism

The TetR-aptamer system for controlling translation was inspired by the natural IRE/IRP mechanism of translational repression (Section 1.4). As this mechanism is known to involve inhibition of an early step in translational initiation[25–27], we anticipated that the TetR-aptamer system would function similarly. However, the our polysome profiling studies (Section 1.10) indicated that translational repression mediated by our synthetic system, in fact, operates downstream of mature 80S ribosome assembly, putatively at some point in the elongation process prior to generation of the mature polypeptide. Unfortunately, this type of experiment can only report on the association of the target mRNA with ribosomal components of different average masses. It would likely be fruitful to perform a ribosome profiling experiment[154] and thereby obtain a higher-resolution picture of mature ribosome occupancy on the TetR-bound and free transcripts. Such an analysis could help define the precise mechanism of TetR-mediated repression. The details of this mechanism would surely accelerate the engineering of improved variants of the system.

Another notable difference between the IRP/IRE system and our TetR-aptamer system lies in their respective dependencies on proximity to the 5′ cap, or lack thereof. IRP/IRE-mediated repression, in its native mammalian context, requires that the interaction lay fairly close to the cap; cap-distal IRP-IRE interactions in the 5′UTR do have a slight translation-repressing effect in other systems (e.g. yeast), but the magnitude of repression is significantly reduced when compared with cap-proximal placement[155]. The apparent lack of position-dependence of the TetR-aptamer system again suggests that there is an important mechanistic difference between it

and the IRE/IRP system, though both are mediated by an apparently quite similar interaction between a noncatalytic protein and a short RNA stem-loop structure. Several conversations with researchers in this field indicate that it is not uncommon to identify strong mRNA-protein interactions in the 5′UTR that are entirely ineffective in repressing translation, though it is not known why. Simple factors such as the size or net charge of the protein do not seem to be the principal determinant of repression activity.

One further experiment that could be performed is an *in vitro* kinetic analysis of ribosome-mRNA association, in the presence of GMP-PNP to prevent productive AUG recognition mature 80S assembly[155]. Our polysome experiments were performed on exponentially-growing yeast, and it is implied in performing this experiment that the system of mRNA transcription, translational initiation, and translational elongation as measured by western blot (and the associated RNA and protein degradative processes) is at steady state. However, if this were not the case, it could be possible that the apparent lack of difference in polysome occupancy between induced (+aTc, high protein level) and repressed (-aTc, low protein level) samples merely reflected a "post-repression" phase of translation wherein a TetR-aptamer interaction in cells harvested at that particular growth phase no longer productively repress translation. This seems highly unlikely, however, as we have never observed a significant deviation from the ~5–10-fold difference in protein level between +aTc and -aTc states, whether making this measurement at very early exponential phase, or long into stationary phase.

5.2. Attractive future applications for the established tool

As mentioned in Section 4.4, the straightforward manner in which the TetR-aptamer system was developed into a robust tool in *P. falciparum* suggests that it may offer similar utility in other organisms. Several collaborators working in non-model organisms have begun investigating the

potential utility of the system in their research. Furthermore, the ability to control translation in the absence of transcriptional control – or in the total absence of transcription altogether – is attractive to many experimentalists in the fields of molecular biology and synthetic biology, and this has also become an area of collaboration.

5.3. Attractive future applications for the underlying technology

The TetR-aptamer system presented herein is, fundamentally, a ligand-sensitive protein-RNA interaction. It may be employed in a variety of different applications outside of "simple" translational repression at the 5′UTR. Most proximally to the work described herein, expression regulation mediated by 3′UTR interactions is an intriguing possibility that we have begun to explore. Although the 2–3-fold repression of maximum translation caused by 5′UTR aptamer placement is generally tolerable, clearly there are circumstances where it is not tolerable, and in generally it would be preferred if one could eliminate it altogether. Aptamer placement in the 3′UTR may offer this. Of course, the mechanism of gene expression control must differ in this case, and we have begun exploring approaches to control from the 3′UTR, primarily involving protein fusions to TetR. In this way, TetR becomes both a sensor for Tc and an adaptor for some RNA-related activity to which it is tethered. This activity may be that of a nuclease or other mRNA-metabolizing enzyme, or act as an adaptor to some other endogenous process in the cell. Non-switchable tethering experiments have repeatedly validated this concept[156–158]. Furthermore, such systems could offer improved dynamic range, particularly if used in combination with 5′UTR aptamers.

Outside of control of gene expression strength, a switchable RNA-protein interaction can form the basis of many ligand-controlled processes. As an adaptor between experimenter control and endogenous RNA-related processes, the possibilities are numerous. One example to have been

published by our laboratory so far is the use of the TetR-aptamer system to control subcellular mRNA localization in yeast[159].

Further abstracting the concept of ligand-switchable protein-RNA interactions, the concept is generalizable to other sensor-type proteins. The TetR protein family contains over 2,000 known members[160], many of which are known to respond to a specific small molecule inducer. The discovery of more inducer-repressor-aptamer cognates is an ongoing area of investigation in our laboratory. In principle, each of these sets should act orthogonally, allowing multidimensional control of post-transcriptional processes in a single cell that expresses these different systems simultaneously.

REFERENCES

1. Belmont BJ, Niles JC. Engineering a Direct and Inducible Protein–RNA Interaction To Regulate RNA Biology. *ACS Chem. Biol.* 2010; 5: 851–61.
2. Goldfless SJ, Belmont BJ, de Paz AM, Liu JF, Niles JC. Direct and specific chemical control of eukaryotic translation with a synthetic RNA–protein interaction. *Nucleic Acids Res.* 2012; 40: e64.
3. Wagner JC, Goldfless SJ, Ganesan SM, Lee MCS, Fidock DA, Niles JC. An integrated strategy for efficient vector construction and multi-gene expression in Plasmodium falciparum. *Submitted.* 2013;
4. Crick F. On protein synthesis. *Symp. Soc. Exp. Biol.* 1958; 12: 138–63.
5. Crick F. Ideas on protein synthesis. 1956; Available from http://profiles.nlm.nih.gov/ps/retrieve/ResourceMetadata/SCBBFT
6. Ahlquist P, Noueiry AO, Lee W-M, Kushner DB, Dye BT. Host Factors in Positive-Strand RNA Virus Genome Replication. *J. Virol.* 2003; 77: 8181–6.
7. Weake VM, Workman JL. Inducible gene expression: diverse regulatory mechanisms. *Nat. Rev. Genet.* 2010; 11: 426–37.
8. Reinstein E, Ciechanover A. Narrative review: protein degradation and human diseases: the ubiquitin connection. *Ann. Intern. Med.* 2006; 145: 676–84.
9. Wu X, Brewer G. The regulation of mRNA stability in mammalian cells: 2.0. *Gene.* 2012; 500: 10–21.
10. De Nadal E, Ammerer G, Posas F. Controlling gene expression in response to stress. *Nat. Rev. Genet.* 2011; 12: 833–45.
11. Müller-Hill B. *The lac operon: a short history of a genetic paradigm.* New York: Walter de Gruyter; 1996.
12. Bro C, Knudsen S, Regenberg B, Olsson L, Nielsen J. Improvement of Galactose Uptake in Saccharomyces cerevisiae through Overexpression of Phosphoglucomutase: Example of Transcript Analysis as a Tool in Inverse Metabolic Engineering. *Appl. Environ. Microbiol.* 2005; 71: 6465–72.
13. Hu MC-T, Davidson N. The inducible lac operator-repressor system is functional in mammalian cells. *Cell.* 1987; 48: 555–66.
14. Garí E, Piedrafita L, Aldea M, Herrero E. A Set of Vectors with a Tetracycline-Regulatable Promoter System for Modulated Gene Expression in Saccharomyces cerevisiae. *Yeast.* 1997; 13: 837–48.
15. Deuschle U, Meyer WK, Thiesen HJ. Tetracycline-reversible silencing of eukaryotic promoters. *Mol. Cell. Biol.* 1995; 15: 1907–14.
16. Labow MA, Baim SB, Shenk T, Levine AJ. Conversion of the lac repressor into an allosterically regulated transcriptional activator for mammalian cells. *Mol. Cell. Biol.* 1990; 10: 3343–56.
17. Gossen M, Bujard H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U. S. A.* 1992; 89: 5547–51.
18. Wishart JA, Hayes A, Wardleworth L, Zhang N, Oliver SG. Doxycycline, the drug used to control the tet-regulatable promoter system, has no effect on global gene expression in Saccharomyces cerevisiae. *Yeast Chichester Engl.* 2005; 22: 565–9.
19. Bashirullah A, Cooperstock RL, Lipshitz HD. RNA localization in development. *Annu. Rev. Biochem.* 1998; 67: 335–94.
20. Blower MD. Molecular insights into intracellular RNA localization. *Int. Rev. Cell Mol. Biol.* 2013; 302: 1–39.
21. Cogoni C, Macino G. Post-transcriptional gene silencing across kingdoms. *Curr. Opin. Genet. Dev.* 2000; 10: 638–43.
22. Mandal M, Breaker RR. Gene regulation by riboswitches. *Nat. Rev. Mol. Cell Biol.* 2004; 5: 451–63.

23. Grskovic M, Hentze MW, Gebauer F. A co-repressor assembly nucleated by Sex-lethal in the 3[prime]UTR mediates translational control of Drosophila msl-2 mRNA. *EMBO J.* 2003; 22: 5571–81.
24. Nelson MR, Leidal AM, Smibert CA. Drosophila Cup is an eIF4E-binding protein that functions in Smaug-mediated translational repression. *EMBO J.* 2004; 23: 150–9.
25. Muckenthaler M, Gray NK, Hentze MW. IRP-1 Binding to Ferritin mRNA Prevents the Recruitment of the Small Ribosomal Subunit by the Cap-Binding Complex eIF4F. *Mol. Cell.* 1998; 2: 383–8.
26. Kim H-Y, Klausner RD, Rouault TA. Translational Repressor Activity Is Equivalent and Is Quantitatively Predicted by in Vitro RNA Binding for Two Iron-responsive Element-binding Proteins, IRP1 and IRP2. *J. Biol. Chem.* 1995; 270: 4983–6.
27. Preiss T, Hentze MW. Starting the protein synthesis machine: eukaryotic translation initiation. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 2003; 25: 1201–11.
28. Nie M, Htun H. Different modes and potencies of translational repression by sequence-specific RNA-protein interaction at the 5'-UTR. *Nucl Acids Res.* 2006; 34: 5528–40.
29. Kato J, Kobune M, Ohkubo S, Fujikawa K, Tanaka M, Takimoto R, et al. Iron/IRP-1-dependent regulation of mRNA expression for transferrin receptor, DMT1 and ferritin during human erythroid differentiation. *Exp. Hematol.* 2007; 35: 879–87.
30. Mowry KL. Complex formation between stage-specific oocyte factors and a Xenopus mRNA localization element. *Proc. Natl. Acad. Sci. U. S. A.* 1996; 93: 14608–13.
31. Schuldt AJ, Adams JHJ, Davidson CM, Micklem DR, Haseloff J, Johnston DS, et al. Miranda mediates asymmetric protein and RNA localization in the developing nervous system. *Genes Dev.* 1998; 12: 1847–57.
32. Proske D, Blank M, Buhmann R, Resch A. Aptamers--basic research, drug development, and clinical applications. *Appl. Microbiol. Biotechnol.* 2005; 69: 367–74.
33. Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science.* 1990; 249: 505–10.
34. Ellington AD, Szostak JW. In vitro selection of RNA molecules that bind specific ligands. *Nature.* 1990; 346: 818–22.
35. Suess B, Hanson S, Berens C, Fink B, Schroeder R, Hillen W. Conditional gene expression by controlling translation with tetracycline-binding aptamers. *Nucl Acids Res.* 2003; 31: 1853–8.
36. Suess B, Fink B, Berens C, Stentz R, Hillen W. A theophylline responsive riboswitch based on helix slipping controls gene expression in vivo. *Nucl Acids Res.* 2004; 32: 1610–4.
37. Fang X, Tan W. Aptamers Generated from Cell-SELEX for Molecular Medicine: A Chemical Biology Approach. *Accounts Chem. Res.* 2009; [Cited 2009 Nov 5] Available from http://www.ncbi.nlm.nih.gov/pubmed/19751057
38. Hanson S, Berthelot K, Fink B, McCarthy JEG, Suess B. Tetracycline-aptamer-mediated translational regulation in yeast. *Mol. Microbiol.* 2003; 49: 1627–37.
39. Kötter P, Weigand JE, Meyer B, Entian K-D, Suess B. A fast and efficient translational control system for conditional expression of yeast genes. *Nucl Acids Res.* 2009; 37: e120.
40. McCarthy JEG. Posttranscriptional Control of Gene Expression in Yeast. *Microbiol Mol Biol Rev.* 1998; 62: 1492–553.
41. Rich SM, Leendertz FH, Xu G, LeBreton M, Djoko CF, Aminake MN, et al. The origin of malignant malaria. *Proc. Natl. Acad. Sci.* 2009; 106: 14902–7.
42. Steenkamp ET, Wright J, Baldauf SL. The Protistan Origins of Animals and Fungi. *Mol Biol Evol.* 2006; 23: 93–106.
43. Prudencio M, Rodriguez A, Mota MM. The silent path to thousands of merozoites: the Plasmodium liver stage. *Nat Rev Micro.* 2006; 4: 849–56.

44. Wickramarachchi T, Devi YS, Mohmmed A, Chauhan VS. Identification and Characterization of a Novel Plasmodium falciparum Merozoite Apical Protein Involved in Erythrocyte Binding and Invasion. *PLoS ONE*. 2008; 3: e1732.
45. Talman A, Domarle O, McKenzie F, Ariey F, Robert V. Gametocytogenesis : the puberty of Plasmodium falciparum. *Malar. J.* 2004; 3: 24.
46. Hall N, Karras M, Raine JD, Carlton JM, Kooij TWA, Berriman M, et al. A Comprehensive Survey of the Plasmodium Life Cycle by Genomic, Transcriptomic, and Proteomic Analyses. *Science*. 2005; 307: 82–6.
47. Arnot DE, Gull K. The Plasmodium cell-cycle: facts and questions. *Ann. Trop. Med. Parasitol.* 1998; 92: 361–5.
48. Arnot DE, Ronander E, Bengtsson DC. The progression of the intra-erythrocytic cell cycle of Plasmodium falciparum and the role of the centriolar plaques in asynchronous mitotic division during schizogony. *Int. J. Parasitol.* 2010; [Cited 2010 Sep 13] Available from http://www.ncbi.nlm.nih.gov/pubmed/20816844
49. Trager W, Jensen J. Human malaria parasites in continuous culture. *Science*. 1976; 193: 673–5.
50. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, et al. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature*. 2002; 419: 498–511.
51. Aravind L, Iyer LM, Wellems TE, Miller LH. Plasmodium Biology: Genomic Gleanings. *Cell*. 2003; 115: 771–85.
52. Tuteja R, Ansari A, Chauhan VS. Emerging functions of transcription factors in malaria parasite. *J. Biomed. Biotechnol.* 2011; 2011: 461979.
53. Duffy MF, Selvarajah SA, Josling GA, Petter M. The role of chromatin in Plasmodium gene expression. *Cell. Microbiol.* 2012; 14: 819–28.
54. Wirth DF. Biological revelations. *Nature*. 2002; 419: 495–6.
55. Nkrumah LJ, Muhle RA, Moura PA, Ghosh P, Hatfull GF, Jacobs WR, et al. Efficient site-specific integration in Plasmodium falciparum chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nat. Methods*. 2006; 3: 615–21.
56. Duraisingh MT, Triglia T, Cowman AF. Negative selection of Plasmodium falciparum reveals targeted gene deletion by double crossover recombination. *Int. J. Parasitol.* 2002; 32: 81–9.
57. Crabb BS, Cowman AF. Characterization of promoters and stable transfection by homologous and nonhomologous recombination in Plasmodium falciparum. *Proc. Natl. Acad. Sci. U. S. A.* 1996; 93. Available from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC38976/
58. Iwanaga S, Kato T, Kaneko I, Yuda M. Centromere Plasmid: A New Genetic Tool for the Study of Plasmodium falciparum. *PLoS ONE*. 2012; 7: e33326.
59. Armstrong CM, Goldberg DE. An FKBP destabilization domain modulates protein levels in Plasmodium falciparum. *Nat Meth*. 2007; 4: 1007–9.
60. Epp C, Raskolnikov D, Deitsch K. A regulatable transgene expression system for cultured Plasmodium falciparum parasites. *Malar. J.* 2008; 7: 86.
61. Meissner M, Krejany E, Gilson PR, de Koning-Ward TF, Soldati D, Crabb BS. Tetracycline analogue-regulated transgene expression in Plasmodium falciparum blood stages using Toxoplasma gondii transactivators. *Proc. Natl. Acad. Sci. U. S. A.* 2005; 102: 2980–5.
62. Gebauer F, Hentze MW. Molecular mechanisms of translational control. *Nat Rev Mol Cell Biol*. 2004; 5: 827–35.
63. Besse F, Ephrussi A. Translational control of localized mRNAs: restricting protein synthesis in space and time. *Nat Rev Mol Cell Biol*. 2008; 9: 971–80.
64. Anderson P. Post-transcriptional regulons coordinate the initiation and resolution of inflammation. *Nat Rev Immunol*. 2010; 10: 24–35.
65. Hüttelmaier S, Zenklusen D, Lederer M, Dictenberg J, Lorenz M, Meng X, et al. Spatial regulation of beta-actin translation by Src-dependent phosphorylation of ZBP1. *Nature*. 2005; 438: 512–5.

66. Carthew RW, Sontheimer EJ. Origins and Mechanisms of miRNAs and siRNAs. *Cell.* 2009; 136: 642–55.
67. Topp S, Gallivan JP. Emerging Applications of Riboswitches in Chemical Biology. *ACS Chem. Biol.* 2010; 5: 139–48.
68. Jackson AL, Linsley PS. Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nat Rev Drug Discov.* 2010; 9: 57–67.
69. Matsukura S, Jones PA, Takai D. Establishment of conditional vectors for hairpin siRNA knockdowns. *Nucleic Acids Res.* 2003; 31: e77.
70. Condeelis J, Singer RH. How and why does beta-actin mRNA target? *Biol. Cell Auspices Eur. Cell Biol. Organ.* 2005; 97: 97–110.
71. Ausländer D, Wieland M, Ausländer S, Tigges M, Fussenegger M. Rational design of a small molecule-responsive intramer controlling transgene expression in mammalian cells. *Nucleic Acids Res.* 2011; [Cited 2011 Oct 22] Available from http://nar.oxfordjournals.org/content/early/2011/10/08/nar.gkr829.abstract
72. Harvey I, Garneau P, Pelletier J. Forced engagement of a RNA/protein complex by a chemical inducer of dimerization to modulate gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 2002; 99: 1882–7.
73. Macchi P, Hemraj I, Goetze B, Grunewald B, Mallardo M, Kiebler MA. A GFP-based System to Uncouple mRNA Transport from Translation in a Single Living Neuron. *Mol Biol Cell.* 2003; 14: 1570–82.
74. Paraskeva E, Atzberger A, Hentze MW. A translational repression assay procedure (TRAP) for RNA–protein interactions in vivo. *Proc. Natl. Acad. Sci. U. S. A.* 1998; 95: 951–6.
75. Plummer KA, Carothers JM, Yoshimura M, Szostak JW, Verdine GL. In vitro selection of RNA aptamers against a composite small molecule-protein surface. *Nucleic Acids Res.* 2005; 33: 5602–10.
76. Saito H, Kobayashi T, Hara T, Fujita Y, Hayashi K, Furushima R, et al. Synthetic translational regulation by an L7Ae–kink-turn RNP switch. *Nat Chem Biol.* 2010; 6: 71–8.
77. Oliveira CC, Goossen B, Zanchin NIT, McCarthy JEG, Hentze MW, Stripecke R. Translational repression by the human iron-regulatory factor (IRF) in Saccharomyces cerevisiae. *Nucl Acids Res.* 1993; 21: 5316–22.
78. Oliveira CC, Heuvel JJ, McCarthy JEG. Inhibition of translational initiation in Saccharomyces cerevisiae by secondary structure: the roles of the stability and position of stem-loops in the mRNA leader. *Mol. Microbiol.* 1993; 9: 521–32.
79. Win MN, Smolke CD. Higher-Order Cellular Information Processing with Synthetic RNA Devices. *Science.* 2008; 322: 456–60.
80. Urlinger S, Baron U, Thellmann M, Hasan MT, Bujard H, Hillen W. Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proc. Natl. Acad. Sci. U. S. A.* 2000; 97: 7963–8.
81. Scholz O, Henssler E-M, Bail J, Schubert P, Bogdanska-Urbaniak J, Sopp S, et al. Activity reversal of Tet repressor caused by single amino acid exchanges. *Mol. Microbiol.* 2004; 53: 777–89.
82. Theodorakis NG, Banerji SS, Morimoto RI. HSP70 mRNA translation in chicken reticulocytes is regulated at the level of elongation. *J. Biol. Chem.* 1988; 263: 14579–85.
83. Mootz D, Ho DM, Hunter CP. The STAR/Maxi-KH domain protein GLD-1 mediates a developmental switch in the translational control of C. elegans PAL-1. *Development.* 2004; 131: 3263–72.
84. Kaspar RL, Gehrke L. Peripheral blood mononuclear cells stimulated with C5a or lipopolysaccharide to synthesize equivalent levels of IL-1 beta mRNA show unequal IL-1 beta protein accumulation but similar polyribosome profiles. *J. Immunol.* 1994; 153: 277–86.

85. Ch'ng JL, Shoemaker DL, Schimmel P, Holmes EW. Reversal of creatine kinase translational repression by 3' untranslated sequences. *Science*. 1990; 248: 1003–6.
86. Berry JO, Carr JP, Klessig DF. mRNAs encoding ribulose-1,5-bisphosphate carboxylase remain bound to polysomes but are not translated in amaranth seedlings transferred to darkness. *Proc. Natl. Acad. Sci.* 1988; 85: 4190–4.
87. Fabian MR, Sonenberg N, Filipowicz W. Regulation of mRNA translation and stability by microRNAs. *Annu. Rev. Biochem.* 2010; 79: 351–79.
88. Maroney PA, Yu Y, Fisher J, Nilsen TW. Evidence that microRNAs are associated with translating messenger RNAs in human cells. *Nat. Struct. Mol. Biol.* 2006; 13: 1102–7.
89. Seggerson K, Tang L, Moss EG. Two genetic circuits repress the Caenorhabditis elegans heterochronic gene lin-28 after translation initiation. *Dev. Biol.* 2002; 243: 215–25.
90. Olsen PH, Ambros V. The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. *Dev. Biol.* 1999; 216: 671–80.
91. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucl Acids Res.* 2003; 31: 3406–15.
92. Koloteva N, Müller PP, McCarthy JEG. The Position Dependence of Translational Regulation via RNA-RNA and RNA-Protein Interactions in the 5′-Untranslated Region of Eukaryotic mRNA Is a Function of the Thermodynamic Competence of 40 S Ribosomes in Translational Initiation. *J. Biol. Chem.* 1997; 272: 16531 –16539.
93. Nagai T, Ibata K, Park ES, Kubota M, Mikoshiba K, Miyawaki A. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* 2002; 20: 87–90.
94. Ma H, Kunes S, Schatz PJ, Botstein D. Plasmid construction by homologous recombination in yeast. *Gene*. 1987; 58: 201–16.
95. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 2001; 29: e45.
96. WHO. World Malaria Report: 2012. Geneva: 2012.
97. Hyde JE. Drug-resistant malaria. *Trends Parasitol.* 2005; 21: 494–8.
98. Gardner MJ, Tettelin H, Carucci DJ, Cummings LM, Aravind L, Koonin EV, et al. Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum. *Science*. 1998; 282: 1126–32.
99. Bowman S, Lawson D, Basham D, Brown D, Chillingworth T, Churcher CM, et al. The complete nucleotide sequence of chromosome 3 of Plasmodium falciparum. *Nature*. 1999; 400: 532–8.
100. Waller KL, Muhle RA, Ursos LM, Horrocks P, Verdier-Pinard D, Sidhu ABS, et al. Chloroquine Resistance Modulated in Vitro by Expression Levels of the Plasmodium falciparum Chloroquine Resistance Transporter. *J. Biol. Chem.* 2003; 278: 33593–601.
101. Wong EH, Hasenkamp S, Horrocks P. Analysis of the molecular mechanisms governing the stage-specific expression of a prototypical housekeeping gene during intraerythrocytic development of P. falciparum. *J. Mol. Biol.* 2011; 408: 205–21.
102. Ben Mamoun C, Gluzman IY, Goyard S, Beverley SM, Goldberg DE. A set of independent selectable markers for transfection of the human malaria parasite Plasmodium falciparum. *Proc. Natl. Acad. Sci.* 1999; 96: 8716–20.
103. Goldberg DE, Janse CJ, Cowman AF, Waters AP. Has the time come for us to complement our malaria parasites? *Trends Parasitol.* 2010; In Press, Corrected Proof. [Cited 2010 Nov 3] Available from http://www.sciencedirect.com/science/article/B6W7G-50N2VF4-1/2/9f70b47e6cc06dc027096607d4cd3cf4
104. Balu B, Chauhan C, Maher SP, Shoue DA, Kissinger JC, Fraser MJ Jr, et al. piggyBac is an effective tool for functional analysis of the Plasmodium falciparum genome. *BMC Microbiol.* 2009; 9: 83.

105. Balu B, Shoue DA, Fraser MJ Jr, Adams JH. High-efficiency transformation of Plasmodium falciparum by the lepidopteran transposable element piggyBac. *Proc. Natl. Acad. Sci. U. S. A.* 2005; 102: 16391–6.
106. Crabb BS, Rug M, Gilberger T-W, Thompson JK, Triglia T, Maier AG, et al. Transfection of the human malaria parasite Plasmodium falciparum. *Methods Mol. Biol. Clifton NJ.* 2004; 270: 263–76.
107. De Felipe P, Luke GA, Hughes LE, Gani D, Halpin C, Ryan MD. E unum pluribus: multiple proteins from a self-processing polyprotein. *Trends Biotechnol.* 2006; 24: 68–75.
108. Straimer J, Lee MCS, Lee AH, Zeitler B, Williams AE, Pearl JR, et al. Site-specific genome editing in Plasmodium falciparum using engineered zinc-finger nucleases. *Nat. Methods.* 2012; 9: 993–8.
109. Oldenburg KR, Vo KT, Michaelis S, Paddon C. Recombination-mediated PCR-directed plasmid construction in vivo in yeast. *Nucleic Acids Res.* 1997; 25: 451–2.
110. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth.* 2009; 6: 343–5.
111. O'Donnell RA, Freitas-Junior LH, Preiser PR, Williamson DH, Duraisingh M, McElwain TF, et al. A genetic screen for improved plasmid segregation reveals a role for Rep20 in the interaction of Plasmodium falciparum chromosomes. *EMBO J.* 2002; 21: 1231–9.
112. De Koning-Ward TF, Fidock DA, Thathy V, Menard R, van Spaendonk RM, Waters AP, et al. The selectable marker human dihydrofolate reductase enables sequential genetic manipulation of the Plasmodium berghei genome. *Mol. Biochem. Parasitol.* 2000; 106: 199–212.
113. Ganesan SM, Morrisey JM, Ke H, Painter HJ, Laroiya K, Phillips MA, et al. Yeast dihydroorotate dehydrogenase as a new selectable marker for Plasmodium falciparum transfection. *Mol. Biochem. Parasitol.* 2011; 177: 29–34.
114. Gietz RD, Schiestl RH. Frozen competent yeast cells that can be transformed with high efficiency using the LiAc/SS carrier DNA/PEG method. *Nat Protoc.* 2007; 2: 1–4.
115. Tonkin CJ, van Dooren GG, Spurck TP, Struck NS, Good RT, Handman E, et al. Localization of organellar proteins in Plasmodium falciparum using a novel set of transfection vectors and a new immunofluorescence fixation method. *Mol. Biochem. Parasitol.* 2004; 137: 13–21.
116. Crabb B., Triglia T, Waterkeyn J., Cowman A. Stable transgene expression in Plasmodium falciparum. *Mol. Biochem. Parasitol.* 1997; 90: 131–44.
117. Duraisingh MT, Voss TS, Marty AJ, Duffy MF, Good RT, Thompson JK, et al. Heterochromatin Silencing and Locus Repositioning Linked to Regulation of Virulence Genes in Plasmodium falciparum. *Cell.* 2005; 121: 13–24.
118. Hasenkamp S, Merrick CJ, Horrocks P. A quantitative analysis of Plasmodium falciparum transfection using DNA-loaded erythrocytes. *Mol. Biochem. Parasitol.* 2013; 187: 117–20.
119. Skinner-Adams T, Lawrie P, Hawthorne P, Gardiner D, Trenholme K. Comparison of Plasmodium falciparum transfection methods. *Malar. J.* 2003; 2: 19.
120. Fernandez V, Hommel M, Chen Q, Hagblom P, Wahlgren M. Small, Clonally Variant Antigens Expressed on the Surface of the Plasmodium Falciparum–Infected Erythrocyte Are Encoded by the Rif Gene Family and Are the Target of Human Immune Responses. *J. Exp. Med.* 1999; 190: 1393–404.
121. Niang M, Yan Yam X, Preiser PR. The Plasmodium falciparum STEVOR multigene family mediates antigenic variation of the infected erythrocyte. *PLoS Pathog.* 2009; 5: e1000307.
122. Van Dooren GG, Stimmler LM, McFadden GI. Metabolic maps and functions of the Plasmodium mitochondrion. *FEMS Microbiol. Rev.* 2006; 30: 596–630.
123. Vaidya AB, Mather MW. A post-genomic view of the mitochondrion in malaria parasites. *Curr. Top. Microbiol. Immunol.* 2005; 295: 233–50.
124. Botté CY, Dubar F, McFadden GI, Maréchal E, Biot C. Plasmodium falciparum apicoplast drugs: targets or off-targets? *Chem. Rev.* 2012; 112: 1269–83.

125. Padmanaban G, Nagaraj VA, Rangarajan PN. An alternative model for heme biosynthesis in the malarial parasite. *Trends Biochem. Sci.* 2007; 32: 443–9.
126. Maier AG, Rug M, O'Neill MT, Brown M, Chakravorty S, Szestak T, et al. Exported Proteins Required for Virulence and Rigidity of Plasmodium falciparum-Infected Human Erythrocytes. *Cell.* 2008; 134: 48–61.
127. Szymczak AL, Workman CJ, Wang Y, Vignali KM, Dilioglou S, Vanin EF, et al. Correction of multi-gene deficiency in vivo using a single "self-cleaving" 2A peptide-based retroviral vector. *Nat Biotech.* 2004; 22: 589–94.
128. Waller RF, Reed MB, Cowman AF, McFadden GI. Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway. *EMBO J.* 2000; 19: 1794–802.
129. Sato S, Rangachari K, Wilson RJM. Targeting GFP to the malarial mitochondrion. *Mol. Biochem. Parasitol.* 2003; 130: 155–8.
130. Knuepfer E, Rug M, Cowman AF. Function of the plasmodium export element can be blocked by green fluorescent protein. *Mol. Biochem. Parasitol.* 2005; 142: 258–62.
131. De Felipe P, Luke GA, Brown JD, Ryan MD. Inhibition of 2A-mediated "cleavage" of certain artificial polyproteins bearing N-terminal signal sequences. *Biotechnol. J.* 2010; 5: 213–23.
132. Shanks RMQ, Kadouri DE, MacEachran DP, O'Toole GA. New yeast recombineering tools for bacteria. *Plasmid.* 2009; 62: 88–97.
133. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods.* 2009; 6: 343–5.
134. Deitsch KW, Driskill CL, Wellems TE. Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes. *Nucleic Acids Res.* 2001; 29: 850–3.
135. Fidock DA, Wellems TE. Transformation with human dihydrofolate reductase renders malaria parasites insensitive to WR99210 but does not affect the intrinsic activity of proguanil. *Proc. Natl. Acad. Sci. U. S. A.* 1997; 94: 10931–6.
136. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods.* 2012; 9: 671–5.
137. Augagneur Y, Wesolowski D, Tae HS, Altman S, Ben Mamoun C. Gene selective mRNA cleavage inhibits the development of Plasmodium falciparum. *Proc. Natl. Acad. Sci. U. S. A.* 2012; 109: 6235–40.
138. Maier AG. The malariologist's molecular toolbox. **In**: Malaria parasites: comparative genomics, evolution and molecular biology. Norfolk, UK: Caister Academic Press; 2013. p. 280.
139. Pino P, Sebastian S, Kim EA, Bush E, Brochet M, Volkmann K, et al. A Tetracycline-Repressible Transactivator System to Study Essential Genes in Malaria Parasites. *Cell Host Microbe.* 2012; 12: 824–34.
140. Ehlgen F, Pham JS, de Koning-Ward T, Cowman AF, Ralph SA. Investigation of the Plasmodium falciparum Food Vacuole through Inducible Expression of the Chloroquine Resistance Transporter (PfCRT). *PLoS ONE.* 2012; 7: e38781.
141. Dvorin JD, Martyn DC, Patel SD, Grimley JS, Collins CR, Hopp CS, et al. A Plant-Like Kinase in Plasmodium falciparum Regulates Parasite Egress from Erythrocytes. *Science.* 2010; 328: 910–2.
142. Baum J, Papenfuss AT, Mair GR, Janse CJ, Vlachou D, Waters AP, et al. Molecular genetics and comparative genomics reveal RNAi is not functional in malaria parasites. *Nucleic Acids Res.* 2009; 37: 3788–98.
143. Jackson KE, Habib S, Frugier M, Hoen R, Khan S, Pham JS, et al. Protein translation in Plasmodium parasites. *Trends Parasitol.* 2011; [Cited 2011 Jul 18] Available from http://www.ncbi.nlm.nih.gov/pubmed/21741312
144. Niles JC. Unpublished data.
145. Bozdech Z, Llinás M, Pulliam BL, Wong ED, Zhu J, DeRisi JL. The Transcriptome of the Intraerythrocytic Developmental Cycle of Plasmodium falciparum. *PLoS Biol.* 2003; 1: e5 EP –.

146. Nkrumah LJ, Muhle RA, Moura PA, Ghosh P, Hatfull GF, Jacobs WR, et al. Efficient site-specific integration in Plasmodium falciparum chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nat Meth.* 2006; 3: 615–21.
147. Adjalley SH, Johnston GL, Li T, Eastman RT, Ekland EH, Eappen AG, et al. Quantitative assessment of Plasmodium falciparum sexual development reveals potent transmission-blocking activity by methylene blue. *Proc. Natl. Acad. Sci. U. S. A.* 2011; 108: E1214–1223.
148. Fidock DA. Personal communication.
149. Rug M, Prescott SW, Fernandez KM, Cooke BM, Cowman AF. The role of KAHRP domains in knob formation and cytoadherence of P falciparum-infected human erythrocytes. *Blood.* 2006; 108: 370–8.
150. Loening AM, Fenn TD, Wu AM, Gambhir SS. Consensus guided mutagenesis of Renilla luciferase yields enhanced stability and light output. *Protein Eng. Des. Sel.* 2006; 19: 391–400.
151. Tannous BA, Kim D-E, Fernandez JL, Weissleder R, Breakefield XO. Codon-optimized Gaussia luciferase cDNA for mammalian gene expression in culture and in vivo. *Mol. Ther. J. Am. Soc. Gene Ther.* 2005; 11: 435–43.
152. Distel B, Gould SJ, Voorn-Brouwer T, van der Berg M, Tabak HF, Subramani S. The carboxyl-terminal tripeptide serine-lysine-leucine of firefly luciferase is necessary but not sufficient for peroxisomal import in yeast. *New Biol.* 1992; 4: 157–65.
153. Cormack BP, Valdivia RH, Falkow S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene.* 1996; 173: 33–8.
154. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science.* 2009; 324: 218–23.
155. Paraskeva E, Gray NK, Schläger B, Wehr K, Hentze MW. Ribosomal pausing and scanning arrest as mechanisms of translational regulation from cap-distal iron-responsive elements. *Mol. Cell. Biol.* 1999; 19: 807–16.
156. Daugeron M-C, Mauxion F, Séraphin B. The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation. *Nucleic Acids Res.* 2001; 29: 2448–55.
157. Unterholzner L, Izaurralde E. SMG7 Acts as a Molecular Link between mRNA Surveillance and mRNA Decay. *Mol. Cell.* 2004; 16: 587–96.
158. Coller J, Wickens M. Tethered Function Assays: An Adaptable Approach to Study RNA Regulatory Proteins. **In**: Translation Initiation: Extract Systems and Molecular Genetics. Academic Press; 2007 [Cited 2010 Aug 12]. p. 299–321. [Cited 2010 Aug 12] Available from http://www.sciencedirect.com/science/article/B7CV2-4PT0WXH-G/2/6d56fb16c6354060aaea99e402c1d2c9
159. Belmont BJ, Niles JC. Inducible Control of Subcellular RNA Localization Using a Synthetic Protein-RNA Aptamer Interaction. *PLoS ONE.* 2012; 7: e46868.
160. Ramos JL, Martinez-Bueno M, Molina-Henares AJ, Teran W, Watanabe K, Zhang X, et al. The TetR Family of Transcriptional Repressors. *Microbiol Mol Biol Rev.* 2005; 69: 326–56.