Design and application of composite RNA aptamers

Shengchun Wang

University at Albany, State University of New York, sw774864@albany.edu

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Biochemistry Commons, Biology Commons, and the Molecular Biology Commons

Recommended Citation
https://scholarsarchive.library.albany.edu/legacy-etd/475

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive. Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
DESIGN AND APPLICATION
OF COMPOSITE RNA APTAMERS

by

Shengchun Wang

A Dissertation
Submitted to the University at Albany, State University of New York
In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

College of Arts and Sciences
Department of Biological Sciences

2011
ABSTRACT

RNA aptamers are being developed as an essential tool in many fields of biological research. Their utility is not limited to being protein inhibitors; a lot of novel functions can be realized. However, in vivo application of RNA aptamers still faces many challenges. The aim of this dissertation is to design and apply composite aptamers in multiple expression and delivery systems to address some critical issues, such as correct folding, high level production, degradation by nucleases, excessive consumption of cellular resource and potential toxic effect.

In one project, I used an intron homing process to mediate aptamer incorporation into the yeast rDNA, resulting in large amount accumulation of a HSF targeting aptamer without affecting normal cellular behavior. The aptamer expressed in this manner inhibited yeast heat shock response. In another project, I used a forward engineering approach to develop a multivalent RNA aptamer to assist the formation of transcription pre-initiation complex at promoter, making it an RNA based transcription activator (taRNA).

The design and implementation of these methods have demonstrated the potential to be adopted and modified for more precise gene regulation and have expanded the repertoire of tools in synthetic biology research using aptamer and other nucleic acid based materials.
ACKNOWLEDGEMENT

Throughout the past six years, there have been many people kindly provided intellectual and spiritual support to help me approach where I intended to go. First and foremost, I would like to express my gratitude to my advisor, Dr. Hua Shi, one of the smartest, most diligent and compassionate people I know. Starting from the time I first walked into his office on that Monday afternoon of Thanksgiving week, he has trusted me and has taken me under his wings. During these years, it felt like I learned something new from him every single day, not only in science, but also in other aspects of life. Having him as a mentor is a blessing as well as a privilege.

I also like to show my appreciation to all the faculty and staff members in the department, especially my committee members, Dr. Richard Zitomer, Dr. Douglas Conklin and Dr. Min-Ho Lee. They have been the best committee anyone can ask for. When I discussed my career plan with them, not only were they mentally supportive, they even lend an active hand, helping me take critical steps along the process. More importantly, I am lucky to know a different yet equally energetic and enthusiastic side of the biology professors outside the laboratory. The connection we developed has moved beyond science and has definitely enlightened my years in graduate school and will be a great lifetime memory.

I am extremely grateful to my scientific collaborators. Drs. Robert Suran and Volker Vogt helped with the design of the intron homing project, sharing reagents as
well as advices. Drs. Xiaoching Zhao and John Lis communicated unpublished results. Dr. Jason Shepard performed the living cell array assay. They and others helped me go through some of the difficult struggles through experiments.

I would never imagine teaching can be so fun and fulfilling if it weren’t for the two outstanding supervisors I had, Dr. Alice Jacklet and Dr. Christine Gervasi. It was a great experience working with both of them along with other teaching assistants. I would not have been able to make those experiments work without the help of troubleshooting from my labmates, for this I would especially thank Dr. Probhat Mallik, Dr. Vamsee Chatakonda, Dr. Daiying Xu, and all other past and current members.

To the friends I have met through these years, I could not have come this far without their companionship and camaraderie. For this I would like to thank Gautham Sarathy, Jung-Hoon Doh, Boris Shkolnik, Dr. Qinglu Zeng, Dr. Tao He, Yuanyuan Liu, Kuangnan Xiong, Dr. Jingjing Xie, Mingliang Wan, Dr. Zhen Huang and Mohammad Qneibi, the list goes on and please forgive me for just naming a few.

Last but not the least, I cannot be more thankful to my family members. To my parents, Lanzhi Liu and Rui Wang, for spiritually supporting their only child from across the globe; to my uncle Dr. Beiqing Liu and aunt Ming Chen for being the family I have here in Albany; to my wife Yan Han, for believing in me, for her encouragement and unconditional support. I would not have accomplished anything
without her being there for and with me through thick and thin; to my cat Meredith
who simply paused all my problems, even for just a second; and to the game of
basketball, which I consider a part of my life, for reminding me who I am, keeping me
passionate and scrappy for all I treasure and all that matters.
TABLE OF CONTENTS

CHAPTER I: INTRODUCTION
1.1. Nucleic acids in synthetic biology
1.2. RNA aptamers generated through SELEX
   1.2.1. Scheme of SELEX
   1.2.2. Targets of aptamer
1.3. Utility of aptamers
   1.3.1. Aptamers in clinical diagnostics
   1.3.2. From basic science towards therapeutics
   1.3.3. Application of aptamers as biosensors
   1.3.4. Aptamer nanotechnology
1.4. Challenges to using aptamers in living cells
   1.4.1. Integrity and stability of aptamers
   1.4.2. Folding pattern of aptamers
   1.4.3. Intracellular accumulation of aptamers
   1.4.4. Excessive consumption of cellular resource
   1.4.5. Global vs. localized effect
1.5. Dissertation objectives and outlines
   1.5.1. Chapter II: Materials and Methods
   1.5.2. Chapter III: In vivo effect of an aptamer expressed as part of an intron
   1.5.3. Chapter IV: Engineer an aptamer based transcription activator (taRNA)
   1.5.4. Chapter V: Functional analysis of AptB4 for designing potential regulatory mechanisms

CHAPTER II: MATERIAL AND METHODS
2.1. Yeast strains and media
   2.1.1. Yeast strain for intron homing project
   2.1.2. Yeast strain and media for taRNA project
2.2. Plasmids
   2.2.1. Plasmids for the intron homing project
   2.2.2. Plasmids for the taRNA project
2.3. Gene synthesis and cloning
   2.3.1. DNA template synthesis
   2.3.2. Gel purification of the digested inserts
   2.3.3. Sub-cloning of inserts into designated vectors
2.4. RNA isolation
2.5. Reverse transcription (RT) and PCR
2.6. Northern blot analysis
2.7. Yeast transformation
2.8. Yeast growth under normal and stressful conditions
2.9. Reporter gene analysis
2.9.1. Beta-galactosidase filter lift assay
2.9.2. Beta-galactosidase quantitative assay
2.9.3. HIS3 reporter gene analysis
2.10. Single cell analysis
2.11. Secondary structure prediction and confirmation
2.12. In vitro transcription and radioactive labeling
2.13. Protein production and purification
2.14. Nitrocellulose filter-binding assay
2.15. Electrophoresis Mobility Shift Assay (EMSA)

CHAPTER III: IN VIVO EFFECT OF AN APTAMER EXPRESSED AS PART OF AN INTRON
3.1. Introduction
3.2. Conceptual design of an indirect aptamer delivery system
3.3. Construction of yeast strains and confirmation of aptamer expression
3.4. Cellular phenotypes caused by the HSF aptamer
3.5. Further increasing the aptamer dosage by two configurations of aptamer dimers
3.6. Molecular phenotypes caused by the HSF aptamer
3.7. Discussion
3.8. Figures

CHAPTER IV: ENGINEER AN APTAMER BASED TRANSCRIPTION ACTIVATOR (taRNA)
4.1. Introduction
4.2. Mechanism-driven choice of an aptamer
4.3. Refinement of AptTFIIB-4
4.4. Molecular and genetic design
4.5. RNA expression and ribozyme function in yeast
4.6. Activity of the transcription activator RNA
4.7. Specificity of the transcription activator RNA
4.8. Discussion
4.9. Figures

CHAPTER V: FUNCTIONAL ANALYSIS OF AptB4 FOR DESIGNING POTENTIAL REGULATORY MECHANISMS
5.1. Introduction
5.2. Minimization of the original AptB4
5.3. Sequence modifications on MinB4
5.4. Replacing the original 3-way junction with other variants
5.5. Inhibiting AptB4 binding by specific antidotes
5.6. Perspectives on development of RNA based regulatory elements
   5.6.1. Dynamic Range
   5.6.2. Amplification of signal
5.6.3. Naturally existing RNA switches
5.6.4. Designing of RNA based-regulatory components
5.6.5. Directed evolution
5.6.6. *In trans* control of RNA activity
5.7. Potential application of controllable aptamer
5.8. Figures

BIBLIOGRAPHY
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>VEGF</td>
<td>vascular endothelial growth factor</td>
</tr>
<tr>
<td>HIV</td>
<td>human immunodeficiency virus</td>
</tr>
<tr>
<td>HSF</td>
<td>heat shock factor</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>reverse transcription polymerase chain reaction</td>
</tr>
<tr>
<td>O.D.</td>
<td>optical density</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotide</td>
</tr>
<tr>
<td>ml</td>
<td>milliliter</td>
</tr>
<tr>
<td>APS</td>
<td>ammonium persulfate</td>
</tr>
<tr>
<td>TEMED</td>
<td>tetramethylethylenediamine</td>
</tr>
<tr>
<td>EB</td>
<td>ethidium bromide</td>
</tr>
<tr>
<td>ETOH</td>
<td>ethanol</td>
</tr>
<tr>
<td>NEB</td>
<td>New England Biolab</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-(N-morpholino)propanesulfonic acid</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate-polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>DEPC</td>
<td>diethylpyrocarbonate</td>
</tr>
<tr>
<td>NC</td>
<td>nitrocellulose</td>
</tr>
<tr>
<td>ONPG</td>
<td>O-nitrophenyl-b-Dgalactopyranoside</td>
</tr>
<tr>
<td>3-AT</td>
<td>3-aminotriazole</td>
</tr>
<tr>
<td>C12FDG</td>
<td>5-dodecanoylaminofluorescein di-b-D-galactopyranoside</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>X-gal</td>
<td>bromo-chloro-indolyl-galactopyranoside</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethanesulfonylfluoride</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>Pol I</td>
<td>polymerase I</td>
</tr>
<tr>
<td>Pol III</td>
<td>polymerase III</td>
</tr>
<tr>
<td>GTFs</td>
<td>general transcription factors</td>
</tr>
<tr>
<td>TBP</td>
<td>tata binding protein</td>
</tr>
<tr>
<td>PIC</td>
<td>pre-initiation complex</td>
</tr>
<tr>
<td>IRP1</td>
<td>iron regulatory protein 1</td>
</tr>
<tr>
<td>IRE</td>
<td>iron response element</td>
</tr>
<tr>
<td>cIIB</td>
<td>c-terminal core domain</td>
</tr>
<tr>
<td>EMSA</td>
<td>electrophoretic mobility shift assay</td>
</tr>
<tr>
<td>μm</td>
<td>micro meter</td>
</tr>
<tr>
<td>Tris</td>
<td>tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>PEG</td>
<td>polyethylene glycol</td>
</tr>
<tr>
<td>°C</td>
<td>degree Celsius</td>
</tr>
<tr>
<td>%</td>
<td>percent</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1.</td>
<td>The indirect aptamer delivery system</td>
</tr>
<tr>
<td>3.2.</td>
<td>Functionality of the aptamer delivery system</td>
</tr>
<tr>
<td>3.3. A</td>
<td>Predicted secondary structure of the monomer and the dimer</td>
</tr>
<tr>
<td>3.3. B</td>
<td>Cellular phenotype of the inhibitory HSF aptamer</td>
</tr>
<tr>
<td>3.4.</td>
<td>Molecular phenotype of the inhibitory HSF aptamer monoer</td>
</tr>
<tr>
<td>3.5.</td>
<td>Molecular phenotype of the inhibitory HSF aptamer dimer</td>
</tr>
<tr>
<td>4.1.</td>
<td>Refinement of AptTFIIB-4</td>
</tr>
<tr>
<td>4.2.</td>
<td>AptTFIIB-4 and its derivatives</td>
</tr>
<tr>
<td>4.3. AB</td>
<td>Molecular and genetic design of taRNA</td>
</tr>
<tr>
<td>4.3. C</td>
<td>Molecular and genetic design of taRNA</td>
</tr>
<tr>
<td>4.4.</td>
<td>RNA expression and ribozyme function in yeast</td>
</tr>
<tr>
<td>4.5.</td>
<td>Activity of the taRNA as measured by the HIS3 reporter gene</td>
</tr>
<tr>
<td>4.6. AB</td>
<td>Activity of the taRNA as measured by ß-galactosidase activity</td>
</tr>
<tr>
<td>4.6. CD</td>
<td>Optical array analysis of living single cells using C12FDG</td>
</tr>
<tr>
<td>4.7.</td>
<td>Specificity of the taRNA</td>
</tr>
<tr>
<td>5.1.</td>
<td>AptB4 deletion constructs binding activity</td>
</tr>
<tr>
<td>5.2.</td>
<td>MinB4 deletion constructs binding activity</td>
</tr>
<tr>
<td>5.3.</td>
<td>MinB4 mutation constructs binding activity</td>
</tr>
<tr>
<td>5.4.</td>
<td>MinB4-au-flip-S modification</td>
</tr>
<tr>
<td>5.5.</td>
<td>Test specific sequence requirement on short stem</td>
</tr>
<tr>
<td>5.6.</td>
<td>Test the overall structural requirement of MinB4</td>
</tr>
<tr>
<td>5.7.</td>
<td>Antidote treatment of MinB4</td>
</tr>
</tbody>
</table>
CHAPTER I: INTRODUCTION

1.1. Nucleic acid in synthetic biology

It has been acknowledged more than a quarter of a century ago, that molecular biology had been in the phase of descriptive study, and that the real challenge would start when it enters the era of synthetic biology, in which devised control elements are added to existing genomes to establish novel behaviors or even construct original “synthetic” organisms. Nevertheless, this term was largely unused or abandoned until earlier this century, when it is referred to mostly as a combination of science and engineering in order to design and build novel biological functions and systems. The pressing need in modern synthetic biology is to establish and enrich the repertoire of fundamental technologies and tools, so that further constructing and engineering can be done by combining or optimizing the existing methodologies, as well as assembling or decoupling existing parts. The task of contemporary synthetic biology, hence, lays partially in the development of standardized methods and exchangeable biological parts, which would accelerate the development of both biology and biotechnology.

Nucleic acid molecules have long been the fundamental elements for building synthetic biological systems. Increasing functions and applications of nucleic acids have been emerging beyond its traditional recognition as the storage form of genetic information. The fact that single stranded nucleic acids form complex structures with surprisingly sophisticated function, of which natural DNA and RNA only make
occasional use, brings out an intriguing challenge, that is to take nucleic acids beyond their proven use and realize their true potential. Inspired by the natural versatility of RNA molecules, there has been an increasing number of engineered RNA molecules with new biological functions. Efforts spent in synthetic biology have produced novel, synthetic RNA components capable of regulating gene expression, sensing environmental stimuli and controlling cell behaviors (Isaacs, Dwyer et al. 2006).

1.2. RNA aptamers generated through SELEX

1.2.1. Scheme of SELEX

RNA aptamers are single stranded RNA ligands selected for different targets from a huge library of molecules containing randomized sequences. The selection process is called Systematic Evolution of Ligands by EXponential enrichment (SELEX), first reported in 1990 (Ellington and Szostak 1990; Gold 1995). The SELEX process involves iterative cycles of selection and amplification starting from a large library of oligonucleotides with different sequences, generally $10^{15}$ or more. After the incubation with the specific target and the partitioning of the binding from the non-binding molecules, the oligonucleotides that are selected are amplified to create a new mixture enriched in those nucleic acid molecules having a higher affinity for the target. After several cycles of the selection process, the pool is enriched in the high affinity sequences at the expense of the low affinity binders. The number of cycles required depends on the stringency of selection, but, once obtained and once the sequence is known, unlimited amount of the aptamer can be produced by chemical
synthesis (Ngundi and Taitt 2006). Moreover, the aptamer selection process can be manipulated to obtain aptamers that bind a specific region of the target and with specific binding properties in different binding conditions.

1.2.2. Targets of aptamer

1.2.2.1. Proteins targeting aptamers

RNA aptamers can be selected towards specific binding to a target protein. The selection scheme can be designed in such a way that the aptamers enriched during selection binds to a certain functional or nonfunctional domain of its target protein, leaving the other portions on the target untouched. Some aptamers are selected for their activity of blocking the action of a functional domain such as the DNA binding domain of a DNA binding protein, or a site on the protein that carries enzymatic activity, in which cases the aptamer functions in a similar fashion as a protein antibody, giving aptamers potential as protein function inhibitors that have proven to be valuable for therapeutic applications (Spiel, Mayr et al. 2009). In other cases, the functional domain on a certain protein is already blocked by another reagent to exclude the enrichment of RNA species that bears binding activity to it. When selection is carried out in this manner, aptamers generated from the pool will possess binding activity to other parts on the target protein while leaving the functional domain untouched. By this design of selection, aptamers will have the potential to bring about other interesting activities through predetermined effects on protein function. I will elaborate on this issue in later chapters by showing a design for
composite aptamer construction that establishes connection between two proteins that normally do not interact with each other. This design led to a predicted modification of protein-protein interaction, resulting in predicted activity through a molecular by-pass pathway.

1.2.2.2. Membrane protein and lipid targets

In the case of membrane proteins, the selections were performed against soluble protein fragments, detergent-membrane protein mixed micelles, whole cells, vesicles derived from cellular membranes, enveloped viruses, etc. This strategy is especially useful in selecting aptamers against channels or receptors, where the subsequent aptamer-target interaction is analyzed by physiological studies of such channels (Park, Wang et al. 2011). Although most aptamers selected through this scheme function as channel blockers, some of them seem to potentiate channel function. Liposomes were used as an experimental system for the selection of aptamers against membrane lipids. RNA aptamer binding for rafts in lipid vesicles was reported (Edwards and Baeumner 2007). There have also been some studies where the whole cell is used as the target. In one example, aptamer generated were able to bind to and distinguish leukemia cells from any other cell types, giving cell-based aptamer selection great promise for the development of specific molecular probes for cancer diagnosis and cancer biomarker discovery (Meng, Sefah et al. 2010).

1.2.2.3. Small molecule targeting aptamers
Based on the ever expanding data of genome research, various gene regulatory elements have been discovered in the 5'-untranslated region (UTR) of primary transcripts. Some of these elements are riboswitches which carry aptamer domains and gene regulatory domains (Lescoute and Westhof 2005). These naturally existing aptamer domains mainly function by sensing a small molecule metabolite, a nucleotide or an amino acid. Binding of these small molecules to the aptamer is usually confined to a pocket, partly because intercalation of the small molecule help reinforce the overall architecture and stability of the aptamer. Naturally existing small molecule binding aptamers also function in regulating gene expression by undergoing conformational changes upon binding, which often leads to steric hindrance of the downstream gene to repress translational or transcriptional events (Huang, Kim et al. 2009). Alternatively, aptamers can be selected in vitro for binding to small molecular weight compound drugs like theophylline and fluorescent probes like malachite green. Because of the similar binding properties of these in vitro selected aptamers to their naturally existing counterparts, these aptamers have been widely used in constructing aptamer based biosensing platforms (Ferapontova, Olsen et al. 2008).

1.3. Utility of aptamers

1.3.1. Aptamers in clinical diagnostics

To detect diseases at the cellular or molecular level, it is desirable to isolate unique biomarkers for each disease or stages of diseases. When biomarkers are identified, an efficient technique is needed for a rapid and accurate measurement of
the molecular recognition of such marker. Nucleic acid aptamers have recently attracted significant attention in the field of clinical diagnosis. As an alternative to antibodies, aptamers have various promising attributes such as high affinity, high specificity, small size, little immunogenicity, stable structures, and ease of synthesis, which enable it to be useful in such recognition based applications. There have been increasing studies using aptamers as probes for cell detection and molecular imaging. To detect protein biomarkers at picomolar concentrations, RNA aptamer microarrays has been developed, where adsorption of proteins onto the RNA microarray is detected by the formation of a surface aptamer-protein-antibody complex followed by an amplification of the signal (Li, Lee et al. 2007). In other studies, RNA aptamers were used to label cell surface markers or used in combination with antibodies for immunophenotyping assays such as flow cytometry and immunohistochemical staining of fixed tissue sections (Zhang, Zhao et al. 2009).

1.3.2. From basic science towards therapeutics

Aptamer isolations have been extended towards a wide range of proteins of importance for therapy and diagnostics, such as growth factors and cell surface antigens. As designer drugs, they exhibit high specificity, high affinity and modifiable bioavailability. Inhibitory aptamers have been generated against complex targets, such as membrane-bound receptors and even entire cells. When expressed in living cells as intramers, aptamers can be used to dissect intracellular signal transduction pathways. Because the interaction of aptamers with protein targets is confined to single
functional domains, aptamers have been widely used for drug target validation. Small molecular weight organic compounds can be screened to determine whether special aptamers block binding.

There have been various attempts that utilize aptamer’s protein inhibitory activity as the principle for aptamer based therapeutic agent. In this case, each aptamer molecule is a functional unit of the drug. To deliver them, most aptamer based drugs are given through intravitreal or intravenous injection. As a result, the target and disease range of these aptamer based drugs are largely limited to accessible tissues and organs, such as the anti-VEGF aptamer for ocular vascular diseases (Burmeister, Lewis et al. 2005), or aptamers targeting factors in coagulation pathways for coagulation related conditions (Gopinath, Shikamoto et al. 2006). Aptamers targeting cell surface proteins are being explored as promising delivery vehicles to target a distinct disease or tissue in a cell-type-specific manner. This concept is adapted to various anti-cancer and anti-HIV therapeutics. One example involves targeting a chemotherapy agent quantum dot conjugated Doxorubicin to ovarian tumor tissue that has mutated mucin overexpression (Lu, Shahzad et al. 2010). Such nanoparticle conjugated approaches are also widely developed for cell imaging technology. There are also studies using microfluidic device to sort and capture cells with the help of cell surface molecule specific aptamers (Martin, Phillips et al. 2011). Aptamers can also be used as part of the combinatorial approach involving other nucleic acid based therapies. In some studies, aptamers were applied in combination
with an RNAi agent, in which an aptamer-shRNA or aptamer-siRNA chimera, which helps to deliver shRNA to specific tissue or cells expressing the aptamer target, sensitizes the cells for the improved treatment efficacy and minimizes side effect (Zhou, Li et al. 2008; Zhou and Rossi 2011).

To increase the bioavailability of aptamers, chemical modifications have been added to aptamers to generate nuclease-resistant inhibitory RNA aptamers, decreasing the chances of degradation by body fluids. Most modifications involve replacement of the 2’-OH group with other chemical groups to avoid nuclease recognition. Such modifications can be carried out either on the entire SELEX pool before the selection, or on isolated aptamers (Burmeister, Lewis et al. 2005).

1.3.3. Application of aptamers as biosensors

RNA aptamers have been adapted to analytical applications not only as alternatives to antibodies, but as unique reagents in their own right. As an emerging class of recognition elements, RNA aptamers offer remarkable convenience in the design and modification of their structures. Aptamers can also be site-specifically modified during chemical or enzymatic synthesis to incorporate particular linkers, or other moieties (Merino and Weeks 2005). Being nucleic acids, aptamers are readily adapted to amplification methods (Zhou, Battig et al. 2010). These attributes of RNA aptamer have been used to generate a great variety of aptamer sensors that exhibit high sensitivity as well as specificity. In addition, their structures can be controlled or
modulated by intermolecular interactions when they make contact with small molecule ligands or their complementary sequences (Cho, Lee et al. 2009). These features have been useful for incorporating molecular reporters for developing biosensors and probes.

Since the essential role of a biosensor is to capture analytes and produce a detectable signal, to achieve more sensitive analysis, utilization of molecular reporters has been studied. As an example, Optical reporters have been chemically conjugated to aptamers for detection of analytes (Iliuk, Hu et al. 2011), such as in the case of a molecular beacon, where a fluorophore and a quencher are attached to the 3’ and 5’ ends of the aptamer. In the absence of analytes, aptamers assume a random, flexible format. In the presence of analytes, the aptamer and the analyte form a compact nanostructure so that the 3’ and 5’ ends of the aptamer will be brought close to each other. Thus, the analytes binding event quenched the fluorescent signal.

1.3.4. Aptamer nanotechnology

The development of nanoscience and nanotechnology has generated nanomaterials with novel properties. By integrating aptamers with novel nanomaterials, nanomaterial-assisted aptasensors are made and have been considered as an excellent sensing platform with wide applications. In some colorimetric assays, gold nanoparticles were coated with aptamers targeting specific cell surface molecules, giving sensitive detection for particular cell types (Zhang, Chen et al. 2010).
1.4. Challenges to using aptamer in living cells

1.4.1. Integrity and stability of aptamers

The primary limitation on the use of aptamers in bioanalytical methods has been their nuclease sensitivity, especially in the cases of RNA based aptamers (Famulok, Mayer et al. 2000). However, there has been various attempts to increase the stability of such molecules. One of the ways to make the improvement was by chemical modification of the ribose ring at the 2′-position (Pieken, Olsen et al. 1991). A different approach to stabilize aptamers comes from selection of RNA aptamers binding to stereoisomers of an intended target molecule, followed by chemical synthesis of the mirror-image of the selected sequences. As a consequence of molecular symmetry, the mirror-image aptamer (L-ribose) binds to the natural target molecule. Because of the substitution of the natural D-ribose with L-ribose, the mirror-image aptamer is stable (Klussmann, Nolte et al. 1996). There are other measures such as incorporating non-natural nucleic acids into a nucleic acid strand to convey resistance to nucleases and promote expression.

1.4.2. Folding pattern of aptamers

The activity of aptamers to bind to their targets exists in its tertiary structure. The folding pattern of aptamers must be maintained to the expected biological function. When aptamers were selected in vitro, the condition was often established to imitate the condition in vivo or even in the living cell, where the aptamer is supposed to
function. However, intracellular environmental conditions fluctuate and various unpredictable factors might play an important role in affecting aptamer folding, such as temperature, pH, ion concentration, etc. Moreover, multiple structures of one sequence can co-exist at equilibrium, when only one or a few of these structures have target binding activity. Some extreme examples exist where two structures of one RNA aptamer are assumed in roughly equal proportion, acting as a collaborating pair to initiate its biological function (Huang, Pei et al. 2009).

1.4.3. Intracellular accumulation of aptamers

To realize the potential of aptamers as protein inhibitors that function intracellularly, the amount of accumulation inside the cell needs to reach a certain level to cause an observable difference. One of the major difficulties in aptamer’s in vivo application is that most of aptamers, unlike small molecules, cannot be directly taken up by cells without external assistance. There are a few examples showing that aptamers can be internalized into living cells for functional studies (Xiao, Shangguan et al. 2008). But these applications are mainly limited to aptamers targeting specifically to cell surface molecules. Instead of being delivered into the cell, aptamers can also be produced from an endogenous resource. In these cases, the strength of the production, rate of degradation and the relative size of the aptamer product become a major concern. While chemical modifications of aptamers would significantly decrease their recognition by nucleases and increase their half-life, the rate of production and the relative large size of aptamers are yet other barriers to
overcome. Compared to small molecule organic compound drugs, which have molecular weights of several hundred as single molecules, aptamers have a much larger molecular weight. Each nucleotide is on average 300 Daltons, and a given aptamer is as many as 50-100 nucleotides in length, constituting to a much larger molecule. There are studies that put an aptamer that binds to its target with high affinity under the control of a strong promoter and expressed it from a vector. However, no phenotypic effect was elicited by the aptamer unless its target level was artificially decreased (Thomas, Chedin et al. 1997).

1.4.4. Excessive consumption of cellular resource

When aptamers are produced endogenously instead of delivered from outside, the cell may spend a large proportion of its resources to reach the required level of aptamer expression. In this case, the host cellular resource, especially transcriptional initiation and elongation factors, may have the tendency to become limited. Normal cellular behaviors might be affected if there is not a compensatory mechanism. Alternatively, the host cell might recognize this excessive consumption of cellular resource and selectively down regulate or even silence aptamer expression in order to maintain normal cellular behavior and function.

There are also possible situations where the aptamer expressed in the host cell targets and inhibits one of the major factors that are responsible for its own production, such as a general transcription factor or even the RNA polymerase itself. In such cases,
the more aptamer produced, the stronger the inhibition of that important factor, and the greater repression on aptamer expression, thereby establishing a negative feedback loop. As a result, it would be convenient to have a method of intracellular aptamer expression that avoids these problems, producing aptamers in high amount without depleting too much cellular resource or generating a negative inhibition on the rate of its own production.

1.4.5. Global vs. localized effect

To achieve the designated function of an aptamer, simply reaching high intracellular accumulation is not enough and will sometimes cause undesired consequences. This is especially true when the aptamer is targeting a crucial factor for cell survival. When the function of that factor is compromised by high amount aptamer elicited inhibition, it will greatly sicken the cell or cause cell death. As a result, localized effect of such an aptamer will be of benefit for the cell under certain conditions. To meet this challenge, an efficient technique is needed to localize the aptamer to the cellular location where it is expected to meet its target, thereby increase the local concentration of the target at designated site without greatly increase the global intracellular concentrations of the aptamer.

1.5. Dissertation objectives and outlines

The objective of my dissertation work is to design and implement strategies for aptamer expression and delivery to tackle some of the issues related to intracellular
application of aptamers discussed earlier. Using aptamers as fundamental building blocks and with the help of other functional RNA units such as the group I intron, ribozymes and other non-functional structural elements, I established strategies of in vivo application of aptamers that not only solved some of the problems but also provided methods that could be adapted by other researcher in synthetic biology and biotechnology using nucleic acid. The model organism used in my work is yeast *Saccharomyces cerevisiae*, a lower eukaryote. The strategies discussed in this dissertation may still need optimization and modifications before they can be used in higher eukaryote. Nevertheless, these strategies established principles whereby some of the challenges can be addressed and thereby enlarged the tool box of nucleic acid synthetic biology.

1.5.1. Chapter II: Materials and Methods

The second chapter of this dissertation contains the materials and methods used in Chapter III through Chapter V.

1.5.2. Chapter III: In vivo effect of an aptamer expressed as part of an intron

Heat Shock Factor 1 (HSF1) is a highly conserved transcription factor responsible for mobilizing specific genomic expression programs in response to stressful conditions such as elevated temperature. In multicellular organisms, this mechanism for enhancing cellular survival appears to enable and promote malignant growth and metastasis. In search of an effective HSF antagonist that would decouple
the transcriptional response from the heat shock stimulus, an RNA aptamer was
previously identified. This aptamer interferes with HSF binding to DNA and inhibits
the transcription of heat shock genes in vitro. In this study, I developed a powerful
expression system to produce this aptamer in yeast and examine its systemic effects.
This indirect delivery system utilized the intron homing process to ferry
aptamer-coding sequences to the middle of hundreds of rRNA genes, so that the
aptamers can be transcribed in large quantity by RNA polymerase I without any
additional promoter being introduced into the cell. When the HSF aptamer was
expressed this way, we observed a temperature sensitive growth phenotype and
specific decrease of heat shock gene expression. These results attest to the potential of
this aptamer as a specific and effective inhibitor of HSF1 activity. The intron-aptamer
expression system can be used to deliver other aptamers as well as other types of
functional RNA.

To solve the aptamer expression problem, I took advantage of the group I intron
homing process. Besides the self splicing activity, one group I intron encodes a
homing endonuclease that recognizes a specific sequence on rDNA, the genes for
ribosomal RNA, and cleaves to make double stranded break. The intron gene then is
used as template for double stranded break repair, resulting in incorporation of the
intron sequence into rDNA. Since there are 150 copies of rDNA gene in the yeast
genome, and ribosomal RNAs constitute more than 90% of yeast total RNA, the
homing process results in large amount of intron expression.
1.5.3. Chapter IV: Engineer an aptamer based transcription activator (taRNA)

The aim of this project is to utilize the protein binding activity of an aptamer to generate novel molecular connection among proteins. The resulting “molecular by-pass” functions to increase the local concentration of a certain functional domain of a protein, short circuiting transcriptional activation. In this case, an aptamer is used to recruit a general transcription factor to promoter. According to the recruitment model of transcriptional activation, an activator helps initiate transcription by bringing the RNA polymerase to a specific location on the DNA through interaction with components of the transcriptional machinery. However, it is difficult to isolate and define the activities of specific activator–target pairs experimentally through rearranging existing protein parts.

Here I demonstrated design and construction of an RNA-based transcriptional activator to not only study the specificity from both sides of the activator–target interface, but also test the principle of an aptamer localization method. Utilizing a well-characterized site-specific RNA aptamer for TFIIB, I was able to delineate some key features of this process. By rationally converting an inhibitory aptamer into the activation domain of the activator, I also introduced a new source of submolecular building blocks to synthetic biology.

1.5.4. Chapter V: Functional analysis of AptB4 for designing potential regulatory
mechanisms

The aim of this project is to dissect the functional requirement of the TFIIB aptamer, AptTFIIB-4, used in the taRNA in order to explore the possible ways of adding a regulatory mechanism to it, so that the activity of transcriptional activation can be temporally controlled.

I created and tested a series of deletion and mutation construct of the original TFIIB aptamer, AptTFIIB-4 (AptB4), to find out the minimal requirement for its binding activity. The result of this study implied that both of the apical-loops were required, although the putative stems connecting to the apical-loops could be slightly altered. It also appeared that the 3-way junction holding the two stem-loops was not required at sequence level, since replacement with different 3-way junctions with similar structures and higher stability seemed to maintain the overall target binding activity.

I also implemented a different scheme of regulation, by adding complementary oligonucleotide to block the correct binding pattern. Although the level of such antisense oligonucleotide required for a sufficient inhibition is quite high, it opens a door for controlling taRNA activity in a different way. The result of this project provides preliminary information and can serve as a starting point of designing RNA regulatory elements.
CHAPTER II: MATERIALS AND METHODS

2. Yeast strains and media

2.1. Yeast strains for the intron homing project

The parental or host strains for intron homing were W303-1A (MATa; leu2-3,112; trp1-1; can1-100; ura3-1; ade2-1; HIS3-11,15), YPH500 (MATa; ura3-52; lys2-801_amber; ade2-101_ochre; trp1-Δ63; HIS3-Δ200; leu2-Δ1), and BY4741 (MATa; HIS3Δ1; leu2Δ0; met15Δ0 ura3Δ0). The two plasmids for trans-integration were co-transformed into a parental strain using the LiAc/ssDNA/PEG method (Gietz and Woods 2002). Transformants carrying both plasmids were identified on plates with SD-Ura-His media. To induce intron homing, these transformants were cultured on SGal-Ura-His plates and passed for several times. Samples were collected at several stages for colony PCR to monitor the progress of homing until the process was complete. The primers used in this assay have the following sequences. E1 (a.k.a. JL83, #1803-1824 of yeast 25S rDNA, towards downstream): 5’-TCACCCCGGAATGTTTATCC-3’. E2 (a.k.a. JL84, #2772-2753 of yeast 25S rDNA, towards upstream): 5’ CGAATGGGACCTTGAATGC-3’. I (R1): 5’-CAATTGACGTTCTTGCC-3’.

2.1.2. Yeast strains and media for tRNA project

The S. cerevisiae strain YBZ-1 (MATa, ura3-52, leu2-3, 112, HIS3-200, trp1-1,

* The methods in this chapter are presented as protocols I followed in the experiments described in later chapters.
ade2, LYS2::(LexA op)-HIS3, ura3::(LexA op)-LacZ, LexA-MS2-MS2 coat (N55K)) was a gift from Professor Marvin Wickens (University of Wisconsin, Madison).

Media consisted of yeast nitrogen base (USBiological), 2% glucose, and synthetic drop-out supplements lacking histidine or histidine and uracil (USBiological).

Transformation was performed according to standard protocol using lithium acetate.

Yeast cells were cultured either on agar plates or in liquid medium at 30°C if not otherwise indicated. Growth rate in liquid media was measured by cell density through turbidity at O.D. 600.

2.2. Plasmids

2.2.1. Plasmids for the intron homing project

The group I intron trans-integration system is composed of two plasmids, pCPIPpo, which carries the Physarum homing endonuclease I-PpoI, and pRSTtLSU1-ClaI, which carries the Tetrahymena intron TtLSU1 flanked by yeast rDNA sequences. pRSTtLSU1-ClaI has a unique ClaI site resulting from the mutation of 4 nucleotides in the P1 loop, which is used for the insertion of the aptamer AptHSF-RA1 in either monomeric or dimeric form. The monomer insert was generated through bi-directional extension of an overlapping pair of oligonucleotide primers purchased from Integrated DNA Technologies. The dimeric insert was synthesized by GenScript. The sequences of these ClaI inserts are the following.

RA-1(M): 5’-ATCGATGCGGCCG
CGAAATTCAACTGCCCATTGGCATCGCGATACAATAATTTTAGTTGAACGCAG
TTCGCGGCGCATCGAT-3’.

RA-1(D):
The ligation product was transformed into DH5 E. coli cells. Single colonies with an insert were confirmed using PCR and verified by sequencing.

2.2.2. Plasmids for the taRNA project

The plasmids pIIIA/IRE-MS2 and pAD-IRP, were gifts from Professor Wickens. The plasmid pDB-sansA was derived from pIIIA/MS2-1 by means of the following manipulations. First, the unique NotI site was destroyed by digesting with NotI, then the sticky ends were filled in using the Klenow fragment of DNA polymerase I, and the blunt ends were re-ligated. Second, the EcoRI fragment was removed and replaced with the following sequence containing a NotI site:

50'-ACTTGAGGGTCTGGGCTAAGCCCACTGATGAGTGCCTGAAATGCGACGA
AACCTCGAGTGACTGCTCCTACGCGCCGCGAGGCGCGCACTATTCCGGTTCGCGCA
GAAACATGAGGATCACCACCTGTCTGACCGCAACGCGTGAACATGAGGAGAATCCTCGAGTTCGAGGACGAGTACGTCTAGCGA
GATGTGGTTTCGCTACTGTAGTGAGTCCCGTGGAGGACGAAACGTCGAC-30.

The plasmids encoding taRNA and its derivatives were constructed by inserting a NotI fragment into the pDB-sansA vector. Each plasmid and the RNA it encoded were
named after the standardized aptamer or aptamer derivative being engrafted to the ‘DB-sansA’scaffold’ (e.g. the taRNA is ‘DB-B4’ encoded in the plasmid ‘pDB-B4’).

The positive control was derived from the RNA-based transcription activator m26-29 with the insert sequence

50-CGACTCTAGAGGATCGCTTCGGCGGCTAGAAACTAGTGGAATCCCCCGGGC
GCGGAAGATTGTTCCCCCAAGTGGATGCCTAAACCTCATGCAT-30.

The sequences of other inserts are each listed below after the name of plasmid.

pDB-B4:
50-AGCTAATGTAGGATGCTGGGGTAGTCCAGCCCTAGAATAAGCGCTAGTA
CTACAAGCT-30.

pDBB4mutS:
50-AGCTAATGTAGGATGCTGGCTTCGGCCAGCCCTAGAATAAGCGCTAGTA
CTACAAGCT-30.

pDB-B4mutL:
50-AGCTAATGTAGGATGCTGGGGTAGTCCAGCCCTAGCTTCGGCTAGTACTA
CTACAAGCT-30.

pDB-B4rev:
50-AGCTTGTAGTACTAGCGCTTATTCTAGGGCTGGACTACCCCAGCATCCTA
CATAGCT-30.

pDB-B60: 50-GGGAGAAATCTAGC
CATCTAGGCGGTGATCGCAGACACAGGGCAGAGATGCGGCTCCC-30.

pDB-TBP12:
50-GCCGTGCCCGGTGGATAGGCACATAAGAC-30.

pDBTBP101:
50-AGAAATTCAAACCTCTTCGGAGCCAAGGTAACAATTCAGTTAGTGGAATGAAACTG-30.

pDB-FC:
50-TCGCTCACGATAACGACTGATTGCGATGTTGATGCGTGATTGGCCACGCGCGA-30. pDB-RA1:
50-GAATTCATAACTGCTTTGCCCATTGCATACGATAAAAATTAAAGTTGAAACGCAGGTTC-30.

Inserts were prepared by bi-directional extension of overlapping oligonucleotides. All synthetic oligonucleotides were purchased from Integrated DNA Technologies, Inc.

2.3. Gene synthesis and cloning

2.3.1. DNA template synthesis

A bi-directional primer extension reaction was performed with two oligos that have a 25-30nt complementary region. The double stranded template is synthesized by one cycle of a typical PCR reaction. For a higher concentration of the template, more cycle can be added with a pair of additional primers. Product volume is decided by the initial input of partially complementary oligos. The PCR products were analyzed on native 8% polyacrylamide gel (acrylamide:bis-acrylamide=29:1). Correct template is identified by corresponding mobility and extracted by gel purification.
2.3.2. **Gel purification of the digested inserts**

Since the length of most inserts are short, it is necessary to run the digestion mixture on a large gel. Undigested inserts should run on the next lane for size comparison. 8% polyacrylamide gel with 10% glycerol is needed. Shown below is the recipe for 40ml gel.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>dH2O</td>
<td>13.5ml</td>
</tr>
<tr>
<td>30% Gel (37.5:1)</td>
<td>10.5ml</td>
</tr>
<tr>
<td>50% glycerol</td>
<td>8ml</td>
</tr>
<tr>
<td>5x TBE</td>
<td>8ml</td>
</tr>
<tr>
<td>APS</td>
<td>200ul</td>
</tr>
<tr>
<td>TEMED</td>
<td>20ul</td>
</tr>
</tbody>
</table>

Run the gel at 200V until the purple dye runs about 10cm to the bottom, and stain the gel for 5min in EB buffer. Observe the inserts under UV lamp, it should be easy to identify the cut ones by comparing to the uncut control. Slice the gel out with a new razor and put it into a fresh tube. Add 50μl dH2O. To crush the gel, it is better to burn the end of a 1ml pipette tip to melt it, hold it tip up until the melting part forms a sealed sphere. Wash the sphere with dH2O to cool it off and use it as tool to crush the gel. Glycerol gel is easy to crush. Make sure it contains no visible solid gel parts. Add dH2O to it to make the final volume about 5 time of the gel, usually 300μl in total is enough. Vortex briefly and incubate in a shaker at 37°C overnight.
Spin the gel purification mixture down for 3 minutes. Transfer the aqueous phase to fresh tube, spin it again, remove as much residual gel as possible. Extract with phenol, chloroform, and precipitate with ethanol. Re-dissolve the pellet in 20µl dH₂O, measure concentration by Nanodrop. Quality of the digested inserts can be seen on a 8% polyacrylamide gel.

2.3.3. Sub-cloning of inserts into designated vectors

When the insert is purified from a polyacrylamide gel, store it in TE buffer at -20°C or use it directly for cloning. Digest vector and inserts separately with designated restriction enzyme for no more than 2 hours. Longer digestion may result in star activity of some enzyme. Treat the vector with Antarctic Phosphatase (NEB) to remove the 5’ phosphate group to minimize vector self-ligation for symmetrical cloning. Start a ligation mixture by adding a vector and inserts to a ratio of 1:3. If ligation efficiency is low, try a ratio up to 1:10. Ligation is carried out by T4 ligase (NEB) at room temperature for 3 hours at room temperature or at 16°C for 8 hours to overnight. Inactivate the ligase by heating at 65°C for 10 minutes. Add 1/10 of the reaction mixture, no more than 3µl, to every 50µl of competent cell. Heat shock at 42°C for 30-45 seconds, move to ice for 2 minutes. Add 750µl of LB media, grow at 37°C for 30 minutes to allow antibiotic gene to be expressed. Inoculate 10-100µl of the culture on a plate that contains the designated antibiotic. Put the plate upside down in 37°C. Pick up single colonies the next day. Start another cell culture and isolate the plasmid according to the protocol from Qiagen.
2.4. RNA isolation

Yeast cells are grown in 30°C overnight and diluted to an OD600 value of 0.1–0.2, then grow until the OD600 value is between 0.6–1.0. Collect 3ml of yeast culture in a screw-cap microfuge tube containing 0.5 ml of 25% glycerol pre-cooled to -20°C. Spin in microfuge for 30 seconds at room temperature, discard supernatant. Cell pellets may be frozen at -70°C at this point. Resuspend pellet in 400 μl of AE buffer (50mM NaAc, 10mM, pH 5.0). Add 40 μl 10% SDS. Vortex briefly for 5–10 second, immediately add 500 μl hot phenol (preheated at 65°C). Vortex for 15 seconds, incubate at 65°C for 5 minutes. Vortex the mixture every half minute, each time for 5 seconds. The mixture is then moved to -70°C for 5 minutes or until frozen.

Centrifuge for 10 minutes in microfuge at 6000 rpm at 4°C. Remove ~450 μl of lower phenol layer. Leave behind cell pellet, interphase layer, and aqueous phase. Add fresh hot phenol into the tube and repeat the heat-freeze cycle for two additional times.

Transfer aqueous solution to a fresh 1.5 ml microfuge tube. Estimate volume. Extract with equal volume of chloroform. Vortex for 1 minute at RT. Spin for 3 minutes in the microfuge at top speed. Transfer the supernatant to a fresh 1.5 ml microfuge tube. Volume should be ~400 μl. Add 1/10 volume of 3 M NaAc. Vortex. Add 3 volumes of ethanol. Vortex. Hold at -70°C for 30 minutes or more.
Spin in the microfuge in 4°C at top speed for 30 minutes. Pellet is usually visible. Wash the pellet with 75% ethanol. Air dry using tubes covered with parafilm with holes. Resuspend the pellet in 20 μl ddH2O. If necessary, dissolve the RNA at 37°C. Measure the concentration by Nanodrop, OD260/280 ratio between 1.8-2.0 indicates good quality. Store RNAs at -70°C.

2.5. Reverse transcription (RT) and PCR

Liquid yeast culture grown at 22°C in log phase (O.D. 600 = 0.5~1.0.) is split into two halves. The first half was heat shocked (HS) at 37-39°C for 20 minutes, while the other half remained at 22°C as the non-heat shock (NHS) control. Both HS and NHS yeast cells were immediately spun down and total RNA was isolated. For reverse transcription, 4 pmole of each primer was hybridized with 100 ng of yeast total RNA for each 10 μl reaction. The RNA was then reverse-transcribed with SuperScript III reverse transcriptase (Invitrogen) at 37°C for 0.5 hour and 42°C for 0.5 hour. To increase the specificity of first-strand synthesis, use up to 45°C for reverse transcription. Treat the reaction mixture with RNase to eliminate the remaining RNA. The product is diluted 200 fold, and 2 μl of the diluted products is taken for further qPCR. The qPCR was performed with a SYBRgreen qPCR kit (Finnzyme) according to a protocol described in the manual. Serial dilutions of the input RNAs were used to plot the standard curves for each gene tested. In the experiment for heat shock genes, non-heat shock genes such as U6 and ADH1 were used as internal controls to
normalize the RNA levels of the heat shock genes and the aptamer being tested. The primers used are the following.

SSA3F: 5’-AGGGAGGCAGAACGAGTTCAGG-3’.
SSA3R: 5’-CTCCAGGACCTGCGCGGCACC-3’.
HSP82F: 5’-TCTGGGAATCCACGCTGGTG-3’.
HSP82R: 5’-CAGAATGTCTCTTGATAACTTCC-3’.
SSE2F: 5’-CAATAACTCAGTACTAGCAGTTGCC-3’.
SSE2R: 5’-ACCGCCGAACTCCACTCCACAC-3’.
HSP12F: 5’-TGAAGCCAGACTCTCAAAAGTC-3’.
HSP12R: 5’-CTTCTTGGTGGTGCTTTCTTCAAC-3’.
ADH1DNF: 5’-GGACAATTTGTTCGGTGCTTCTAAAGG-3’.
ADH1DNR: 5’-ATTCTTGGCAAGGTAGACAAGCCG-3’.
U6F: 5’-GTTCGCGAAGTAACCTTG-3’.
U6R: 5’-GAAATAAACCTCTTTGTAAACGG-3’.

2.6. Northern blot analysis

Suspend up to 20µg of total RNA in formaldehyde gel loading buffer. 1x loading buffer should contain—

5% glycerol

1mM EDTA (pH 8.0)

0.025% bromophenol blue
0.025% xylene cyanol FF.

Some Ethidium Bromide can be added for visualization.

Equal amounts of RNA are heated at 70°C for 10 minutes and resolved on Agarose gel (1.67%) containing 2.2M formaldehyde. This denaturing gel was made by adding 2.5g agarose to 108ml H₂O, dissolved in a microwave oven, and let cool to 55°C. Then add 15ml 10x MOPS buffer and 27ml deionized formaldehyde.

10x MOPS Running buffer should contain:

0.1 M MOPS (pH 7.0)

2mM sodium acetate

1mM EDTA (pH 8.0)

Observe the RNA on a UV light, if total RNA is loaded, two major band will be present indicating the small and large subunit of ribosomal RNA, a slight smear in between is mRNA. Blotting the RNA to a nitrocellulose membrane is done under the condition of 1x SSC media, overnight. First cut a corner on the gel to mark the orientation. Put the gel upside down on a 3mm Whatmann blotting paper, with the edge of the paper merged under 1x SSC media. Put a nitrocellulose or nylon membrane directly on top of the gel, mark orientation the same way as the gel. Put two other pieces of Whatmann paper on top of the membrane, the size of which should be slightly larger. Then put some paper towel on top to blot. Add something
heavy on top of the paper towel to keep balance. If the RNA is pre-intercalated with Ethidium Bromide, it should be visible under UV light at 302nm wavelength.

Crosslink the RNA to the membrane by a UV crosslinker, one minute for each side.

Hybridization at 55C overnight in solution containing:

Sodium phosphate (pH 7.2) 0.5M
7% (w/v) SDS
1mM EDTA (pH 7.0)

If internal labeling is preferred, probe can be labeled by in vitro transcription using radioactively labeled nucleotide (α-NTP). Alternatively, if end labeling is selected, DNA oligo can be labeled by T4 polynucleotide kinase, using radioactively labeled nucleotide (γ-NTP). See section 2.12 for more details.

Rinse the membrane in 2xSSC at 23°C and then successively wash in 2xSSC, 0.5xSSC with 0.1% SDS, and 0.1xSSC with 0.1% SDS for 15 minutes each at 23°C. A final wash containing 0.1xSSC and 1% SDS is carried out at 50°C. Duration of washing should be increased to remove intense background. Dry the membrane on blotting paper and expose to a phosphoscreen. The image is scanned by a TyphoonTM imager and analyzed by ImageQuant software (GE Healthcare).

2.7. Yeast transformation
Grow yeast at 30°C, until the culture reached an OD600 reading of 0.5-1.0. Use 0.5-1ml cell culture, spin down for 3 minutes. Re-suspend in 50µl dH2O and add solutions in the following order:

- 50% PEG: 240µl
- LiAc (2M): 18µl
- ssDNA (2mg/ml): 50µl
- Plasmid (200ng/µl): 2µl

Salmon sperm ssDNA carrier DNA should be boiled for 5 minutes to denature and chill on ice before adding. Vortex briefly after adding each solution and incubate in a 42°C water bath for 50min to 1hr. Spin it down for 1min, remove solution and re-suspend the cell in 500µl dH2O (without DEPC treatment). Plate 100µl of cells on selective media. Put it in a 30°C incubator, upside down. Colony should form in 3-4 days. Pick up robust colonies for reporter gene expression analysis.

2.8. Yeast growth under normal and stress conditions

Yeast strains are grown continuously at 30°C. Collect same amount of cells from each strain and dilute into same volume of minimum media or ddH2O. Perform a serial dilution for each strain; with each concentration 10 times lower than the previous one. The serial dilution can be done in ddH2O to cease the cell growth during the time of operation. Five different concentrations of each strain is plated on SGal-Ura-His media to examine their phenotype under different conditions:
SGal-Ura-His media at 22°C, SGal-Ura-His media at 30°C, SGal-Ura-His media at 37°C, SGal-Ura-His media supplemented with 5 mM CuSO$_4$ at 30°C, SGal-Ura-His media with 0.3 mM H$_2$O$_2$ at 30°C, and SGal-Ura-His media with 1 M sorbitol at 30°C. The growth of the cell patches was compared and recorded after 3-4 days of incubation.

2.9. Reporter gene analysis

2.9.1. Beta-galactosidase filter lift assay

Streak a robust colony from transformation onto selective media; put all testing strains on one plate. Cut a piece of Nitrocellulose filter and place it on top of the yeast colonies, get rid of bubbles. Make sure the filter is in close contact with the cell, carefully label at the edge of the filter with pencil. Cut a slightly larger piece of Whatman filter paper, but small enough to fit in a Petri dish, and merge into 5-10ml of Z buffer containing X-gal. Let it sit until reaching equilibrium. Buffer recipe is shown below.

Z buffer with X-gal: 60 mM Na$_2$HPO$_4$, 60 mM NaH$_2$PO$_4$; 10 mM KCl; 1 mM MgSO$_4$, X-gal 1mg/ml. Note: X-gal stock solution: 100mg/ml in Dimethylformamide.

Remove excess buffer from the Petri dish. Pick up the NC filter with a pair of forceps and dip the filter into liquid nitrogen; let it stay for 10-15seconds to
permeabilize cell membrane. Take the filter out and dry briefly on a piece of aluminum foil. Place the filter on top of the Whatman filter paper, get rid of air bubbles. Close the lid of the Petri dish. Seal the dish with parafilm, and place at 30°C. Check for color change every 5 minutes. Depending on the level of activity, blue color will start to form between 20 minutes to 3 hours.

2.9.2. Beta-galactosidase quantitative assay

Quantitative analysis of beta-gal activity usually gives large standard deviation, even using miller unit. So it is better to normalize the activity to an internal standard with known activity. To obtain reliable results, measure at least four clones of each yeast transformation. Make a 3-ml overnight culture, dilute with fresh media to an OD600 value of about 0.1-0.2 (3 ml in total is sufficient). Incubate the cultures at 30°C to an OD600 value of 0.5 to 1.0. Determine the OD600 value for each culture; record the reading. Centrifuge 2 ml of the cultures for 3 minutes. Resuspend them in 665 µl Z buffer (refer to previous section for buffer recipe). Add 55 µl chloroform and 55 µl 0.1% SDS. Vortex at the highest setting for 1 min to permeabilize the cells. Add 125 µl 4mg/ml ONPG dissolved in Z buffer (it is hard to dissolve, but do not heat; vortex 1-3 min). Mix the solution vigorously. Shake the samples in a 37°C until a slight yellow coloring appears (usually this takes between 10 and 90 min), record the time. Stop the reaction by adding 400 µl 1M Na2CO3. Centrifuge for 5 minutes. Carefully transfer the supernatant into a photometric cuvette. Determine the OD at 420 nm.
Calculate Miller unit according to the following formula: Miller Units = \[ \frac{1000 \times OD_{420}}{V(\text{ml}) \times T(\text{min}) \times OD_{600}} \].

2.9.3. **HIS3 reporter gene analysis**

Streak out various strains on ura- his- media and grow the cells in 30°C. Various amount of HIS3 gene product inhibitor 3-Aminotriazole (3-AT) can be added to the media to increase the stringency. For an initial test, add 3-AT to a final concentration of 0, 0.2mM, 0.5mM, 1.0mM. To analyze HIS3 gene expression in liquid media, pick up more than three single colonies from each strain and grow in 30°C overnight and dilute each strain to OD600 value of 0.05. From the diluted culture, grow for another 2 hours and measure OD600 value again. Inoculate same amount of cells from each strain into equal volume of histidine dropout liquid media. Place into a rotator or shaker at 30°C, monitor the OD600 value for each strain every 2 hours. Plot the recorded OD value for each strain to determine the growth rate.

2.10. **Single cell analysis**

Collect an overnight culture for each strain and dilute to a OD600 value of 0.1-0.2. Grow the cells to a OD600 value of 0.5-0.8. Cells were kept alive in minimal media and localized onto the tip of an optical fiber that contains diamond shaped wells, with each well large enough to accommodate a single cell. Continuous optical measurements of individual localized cells are performed on a microfluidic gridded array, the LiveCell Array (Molecular Cytomics), with a standard microscope.
A 500-ml aliquot of cell suspension is combined with a 500-ml aliquot of 250 mg/ml concanavalin A-tetramethylrhodamine (ConATAMRA) conjugate, a labeling reagent that allowed the cells’ position to be registered. The cells are washed with media and loaded onto the array. The microfluidic chamber enables addition of the substrate 5-dodecanoylaminofluorescein di-b-D-galactopyranoside (C12FDG, Molecular Probes) in media to the localized cells. The measurements are taken with optical filter systems specific for fluorescein (C12FDG) and TAMRA.

2.11. Secondary structure prediction and confirmation

Secondary structures of RNA constructs were predicted using the mfold program (v. 3.2) (Mathews, Turner et al. 2007). To identify a stable stem as the point of integration, a series of derivatives were constructed based on the most thermodynamically stable predicted structure of an aptamer and tested for binding activity.

2.12. In vitro transcription and radioactive labeling

The [α-32P] CTP- incorporated RNA was prepared using the MAXIscript™ in vitro transcription kit (Ambion) according to the manufacturer’s instructions. Non-radioactive CTP was added to about 1/100 concentration of other non-radioactive nucleotide. Radioactively labeled CTP should be present at about 1/10 of the non-radioactive CTP, so it becomes the limiting factor. Reaction was carried out at 37°C for 3 hours, then treated with turboDNase for 15 minutes to remove the DNA.
template. Prior to their use, a small sample of each RNA preparation should be subjected to electrophoresis on an 8% polyacrylamide (acrylamide:bis-acrylamide=29:1), 8M urea gel and shown to be of the expected size.

2.13. Protein production and purification

Expression vector was amplified in DH5α competent cells and plasmids were isolated by miniprep (Qiagen). Transform the expression vector into BL21-AI (Invitrogen) competent cells. Cells are grown at 37°C overnight. Dilute the cell culture to 0.1-0.2 of OD600 value. Allow the cells to grown until the OD600 reading reaches a value between 0.4-0.6 before induction by adding L-arabinose 0.2% and 1mM IPTG for 4 hours.

Cells should be harvested and re-suspended in 1x resin binding buffer (20mM Sodium Phosphate, 500mM NaCl, pH 7.8), 5ml for a 500ml culture. Break down the cells by sonication (Sonabox II, Artek). The sonication parameters should be set at 50% amplitude, with pulse on for 0.3 second, pulse off for 1 second. To avoid over-heating the sample, embed the sample in crushed ice and chill the sample after 10 cycles of pulses. Repeat it for 1-2 minutes to thoroughly break down the cells. Resuspend each 100ml culture of cells in 5ml resin binding buffer and loaded on Nickel affinity gel (His-select, Sigma). 1 ml of affinity resin is need for each 100ml of cell culture. This column and sample mixture can be incubated or shaked at appropriate temperature to allow sufficient contact.
Wash the column with wash buffer (20mM Sodium Phosphate, 500mM NaCl, pH 6.0) containing increasing concentration of imidazole. Start with wash buffer containing no imidazole, wash with at least 10 column volume. To determine the appropriate elution condition, monitor the OD 260 value of the eluted product, the peak should come off the resin between 50-200nM of imidazole. A pilot experiment with lower volume of cell culture can be used to determine the elution condition before a larger size culture is used.

Dialysis of the purified product is done to change the buffer into a storage buffer containing: HEPES 20mM, NaCl 250mM, Glycerol 15%, EDTA 0.2mM, DTT 2mM, PMSF 1mM.

Protein is resolved on a 10% SDS gel with a BSA standard to estimated concentration and quality, and is aliquoted and stored in -80°C. If a higher concentration is needed, use Microcon (Millipore) centrifugation tube for appropriate size of protein. Other similar devices can also be used.

2.14. Nitrocellulose filter-binding assay

The protein-RNA mixture was incubated at a designated condition before loading on a nitrocellulose membrane (MilliPore, pore size 0.22μm) through a Microfiltration Apparatus (Bio-rad). Some random RNA and protein such as yeast tRNA and BSA
should be added into the mixture as non-specific competitors to rule out non-specific binding. The nitrocellulose membrane is pre-incubated in 1x binding buffer (25 mM phosphate, pH 7.2, 2 mM Mg2Cl2, 3% glycerol, 75 mM KCl. The binding buffer can be prepared to contain bovine serum albumin) before being used for binding assay. Cut the membrane into the size that fit exactly on the apparatus. Mark the orientation by making a diagonal cut on one edge. Check for leakage of the apparatus by loading same amount of binding buffer into each well, turn on vacuum to let it drain. The wells should have similar draining speed.

When the apparatus is ready, carefully load the samples onto the membrane without touching the membrane with the pipette tip. Wash each well with 500-100µl of 1x binding buffer after sample loading. Make sure that the sample and wash buffer go through the wells at similar speed. The vacuum force applied to drain the washing buffer should not be too strong. The nitrocellulose membrane is then air dried briefly until no liquid is visible. Wrap it around with plastic wrap and expose it to a storage phosphor screen (GE Healthcare), for appropriate time, roughly the time it takes to rinse and dry the apparatus (10-15 minutes). The image is scanned by a TyphoonTMimager with ImageQuant software (GE Healthcare). Semi-quantification of the signal can be done with the Imagequant software. Select the negative control lane as background and the comparative strength of the signal in experimental lanes is calculated as a number. Plot the number in a bar graph using a software such as GraphPad.
2.15. Electrophoresis Mobility Shift Assay (EMSA)

The protein-RNA mixture is prepared in the same way as for the nitrocellulose filter-binding assay. For convenience, half of the reaction mixture can be loaded on the filter and the other half saved for EMSA. For EMSA, add 2-4µl of 50% glycerol per 10µl reaction to allow the sample be loaded into the well, pipette up and down to mix thoroughly. Some dyes can be added to the mixture for loading convenience, 6-8% polyacrylamide gel (35.5: 1) containing 0.5x TGB buffer and 2.5mM MgCl2 is usually used for EMSA, Mg²⁺ is used to help maintain RNA stability. Alternatively, 2.5% agarose gel containing 1/4 TBE buffer can be used. (The resolution is better on polyacrylamide gel. Agarose gel is used when the protein size is big or aggregation is observed.) Label the wells before loading to keep track if no dye is added to the reaction mix. Run the gel in 4°C at 150V. If there is an extra well, load 5µl of DNA loading buffer to estimate how far the samples have moved. If there is no extra well, some dye can be added to the mixture to keep track, usually 3 hours of running at 150V will give a good resolution on 6-8% of polyacrylamide gel.

Stop the running when the purple dye reached the bottom of the gel. For polyacrylamide gel, use normal cycle to dry the gel at 80°C for 1 hour. To dry the gel, cut a DEAE filter paper (DE81) to the size slight larger than the gel, put it onto the reverse side (bottom) of the gel so the front side is exposed. Cover the top of the gel with a piece of plastic wrap so that the gel drier will not be contaminated. For agarose
gel, dry it using the gradient cycle, 50°C for 2 hours, expose to a storage phosphor screen (GE Healthcare) for appropriate time. Usually 1 hour is good enough for freshly labeled RNAs; prolonged exposure might give a strong background. The image is scanned by a Typhoon imager and ImageQuant software (GE Healthcare). Semi-quantification of the signals can be done with the Imagequant software. Select the negative control lane as background and the comparative strength of the signal in experimental lanes is calculated as a number. Plot the number in a bar graph using a software such as GraphPad.
CHAPTER III: IN VIVO EFFECT OF AN APTAMER EXPRESSED AS PART OF AN INTRON*

3.1. Introduction

Responsive to environmental changes, the cell adjusts its genomic expression program in a coordinated manner. This process is often governed by a master regulator in the form of a DNA-binding transcriptional activator (Gasch, Spellman et al. 2000). The Heat Shock Factor 1 (HSF1) is the major transcription regulator in stress response, and it is highly conserved among eukaryotes (Wu and Hampsey 1999; Birch-Machin, Gao et al. 2005). Under unfavorable conditions, especially elevated temperatures, HSF1 activates the expression of heat shock genes, producing heat shock proteins with chaperonin activity to repair cellular damages and boost cellular resistance to further injury (Lindquist and Craig 1988; Kregel 2002). Intriguingly, whereas it evolved to enhance cellular survival, this mechanism can impede organismal survival under certain circumstances. In particular, HSF1 can act in a multifaceted way as a powerful supporter of malignant transformation (Dai, Chen et al. 2007). Therefore, blocking the HSF activity may be a promising therapeutic modality to treat cancer.

A productive strategy in biomedical research is to identify fundamental questions that underlie the molecular mechanism of diseases and approach them in a facile

*Part of his chapter has been published in the following paper:
experimental vehicle like yeast. In mammals, while HSF1 is the dominant heat shock regulator and is responsible for the maintenance of cancerous phenotype, the multiple HSF isoforms, which appear to have somewhat specialized functions, complicate the analysis of HSF1 function (Rabindran, Giorgi et al. 1991; Sarge and Morimoto 1991; Schuetz, Gallo et al. 1991). Furthermore, the diversity of genetic background among various cancers makes it difficult to elucidate the mode of action of a potential HSF-targeting therapeutic agent. In contrast, in the fruit fly Drosophila melanogaster and the baker’s yeast Saccharomyces cerevisiae, the HSF activity is encoded by a single HSF1 gene; and in the yeast HSF1 is essential for its viability and vegetative growth, making it a convenient and inexpensive means to screen for reagents that would inhibit the HSF1 activity in a way potent enough to generate a systemic effect (Sorger and Pelham 1988; Wiederrecht, Seto et al. 1988).

Like many other transcription factors, HSF possesses a DNA binding domain and an activation domain. In addition, it also has a trimerization domain and a flexible linker that collaborates with flanking domains to allow the homotrimer of HSF to bind avidly to a Heat Shock Element (HSE) (Wu 1995). Based on this molecular configuration, an effective strategy of uncoupling the heat shock response is to mask the DNA binding and adjacent domains, thereby preventing HSF from binding to the HSEs. To this end, an RNA aptamer was generated for HSF, which competes with the HSE for binding to HSF1 (Zhao, Shi et al. 2006). This aptamer, named AptHSF-RA1, binds yeast and Drosophila HSF1 equally well through interaction with the
DNA-binding domain and a flanking linker region. An in vitro transcription assay demonstrated the capability of this aptamer to inhibit heat shock gene transcription (Zhao, Shi et al. 2006).

Due to their high affinity and specificity, RNA aptamers offer advantages over other types of reagents to manipulate and control protein function (Gold 1995; Wilson and Szostak 1999). However, their intracellular delivery is not as straightforward. A tactic employed to circumvent this problem is to deliver RNA aptamers as synthetic genes (Shi, Hoffman et al. 1999). In mammalian cells, RNA polymerase III (Pol III) promoters are often used to express aptamers and observe their effects (Mi, Zhang et al. 2006). But the same type of promoter when used in yeast was less successful (Thomas, Chedin et al. 1997). To address this issue, we have developed an expression system to achieve high level of aptamer accumulation. This system utilizes the process of group I intron homing to insert an aptamer coding sequence into multiple copies of yeast rDNA, so the aptamers can be produced through transcription by RNA polymerase I (Pol I) without the requirement of additional promoters. When the HSF aptamer was expressed using this system, I observed growth retardation of the yeast cells under heat shock condition. Consistent with this systemic phenotypic change, I also observed specific decrease of expression of genes activated by the HSF1.

3.2. Conceptual design of an indirect aptamer delivery system

When aptamers are used inside cells as “intramers” to modulate protein function,
it is important to balance the expression levels of the aptamer and the target protein to achieve the desired change in target function and cellular phenotype. Several additional issues also need to be addressed including rate of aptamer production, correctness of folding, stability, and subcellular localization. Our expression system aims primarily at generating high-level aptamer accumulation so that the effect of aptamers can be easily revealed and evaluated. It also ensures reasonable stability of correctly folded aptamers and keeps the aptamers within nuclei, the organelle that harbors the protein targets of our aptamers. This system comprises a set of molecular constructs that ferries the aptamer-coding DNA to a specified location in the middle of rRNA genes. The strategy exploits unique features of the nucleolar rRNA genes to achieve the goals set above: in yeast, rRNAs are transcribed at a high level by Pol I from about 150 copies of rDNA. The high capacity transcription of rRNA fortuitously generates a large amount of aptamers as byproducts.

Utilization of the power of Pol I transcription in this scheme requires the capability of inserting a heterologous DNA fragment into all copies of the rDNA and ensuring that normal rRNA production and maturation are not affected by this insertion. These problems were solved by using a group I intron trans-integration process. In our adaptation of this system, depicted in Figure 3.1A, we combined components from three different species: the yeast *Saccharomyces cerevisiae*, as the host, that harbors a group I intron, TiLSU1, from the ciliate *Tetrahymena thermophila* (Brehm and Cech 1983) and a homing endonuclease, I-PpoI, from the slime mold
Physarum polycephalum (Ellison and Vogt 1993). Homing of the group I intron into yeast rDNA was driven in trans by I-PpoI, which recognizes a 15-base pair site in each of the rDNA repeats on chromosome 12 and cleaves the DNA (Flick, Jurica et al. 1998). The *Tetrahymena* intron was chosen because its self-splicing ribozyme is far more active than that of the *Physarum* intron (Zhang, Jenkins et al. 1995), and it was found that it is able to accommodate an inserted sequence and retain the ability to splice quickly and accurately (Suran 2009). An aptamer-coding sequence receives a “piggyback ride” as an insert in the group I intron, which integrated into these rDNA repeats at the cleaved I-PpoI site. Using this system the intron RNA accumulation can reach 2% of the total RNA (Lin and Vogt 1998). As depicted in Figure 3.1B, we have designed a construct in which the aptamer is introduced as an extension of the P1 stem of the intron. At the interface, we used a “GC clamp” (Cassiday and Maher 2003) as a structural insulator between the intron and the aptamer, so that the splicing activity of the intron would not be compromised and the aptamer would be well-exposed.

### 3.3. Construction of yeast strains and confirmation of aptamer expression

I used three parental yeast strains, W303, YPH500, and BY4741, to test the aptamer delivery system. To form the aptamer-coding construct, a segment of DNA coding for a minimized version of AptHSF-RA1 was inserted into the sequence of the group I intron TtLSU1. The plasmid carrying the gene for this intron-aptamer, named pRSTtLSU1-RA1(M), was co-transformed into the parental yeast strains together.
with a plasmid carrying the homing endonuclease I-PpoI under Gal I promoter control, named pCPIPpo. In this arrangement, homing of the intron was controlled by the induced expression of the homing endonuclease. Because the insertion of the intron would abolish the I-PpoI site in rDNA, the homing process proceeded until all copies of rDNA had acquired an intron.

To observe this process and to assess our aptamer delivery system at the DNA level, we passed the transformants for several cultures and performed PCR with two flanking exon primers (see Figure 3.2A, lanes labeled EE) or an exon primer paired with an intron primer (lanes labeled EI). Intron homing would result in increased size of the EE fragment (from 870bp to 1425bp). The 745 bp EI fragment should only be produced by rDNA with homed intron. It should be pointed out that complete homing of the intron coding sequence into all copies of rDNA required several subcultures of the transformants. To monitor the homing process, I kept the induction of I-PpoI on by transferring the cells to new galactose containing media and performed colony PCR in several stages. In early cultures, no intron integration was observed. As the expression of the homing endonuclease continued and new colonies are allowed to form, some homing events were identified as the appearance of the EI fragment and partially up-shifted EE fragment. In later generations, intense EI fragment was seen with a completely up-shifted EE fragment. In Figure 3.2A, I also show the results of one completely homed clone compared to one parental clone. While these data were generated using the W303 parental strain, this experiment has been performed in all
three genetic backgrounds and the same profile was observed.

To confirm that this aptamer delivery system functioned properly, two more tests at the RNA level were required. First, the parasitic existence of aptamer should not affect the expression of rRNA and the splicing of the group I intron from rRNA transcripts should leave the mature rRNA intact. Second, the level of intron-aptamer accumulation should be sufficiently high as to be comparable to the rRNA level. For these purposes, we prepared RNA from the homed and parental strains, and measured the expression of the rRNA and the aptamer using Northern blot analysis. An oligonucleotide 41 bases in length covering the sequence of the downstream exon near the PCR primer-annealing site was used to detect the rRNA. To visualize the aptamer, we used a probe 56 bases in length complementary to the aptamer sequence. As shown in Figure 3.2B, the rRNA probe revealed a band with identical mobility and intensity in both parental (P) and homed (H) strains. The aptamer probe only detected a band in the homed strain, whose intensity is close to that of the rRNA band.

3.4. Cellular phenotypes caused by the HSF aptamer

The minimized “Core” of the AptHSF-RA1 comprises a three-way junction. Of the three stems eradiating from the junction, Stem 1 contains an apical loop and an internal loop, which are required for the aptamer’s activity. Stem 2 and 3 are short and both can serve as a point of integration with other structural or functional elements. As shown in the upper panel of Figure 3.3A, in the TtLSU1-RA1(M) construct, we
connected Stem 3 with the P1 stem of the intron and grafted a UUCG tetra-loop to the tip of Stem 2 to maintain strand continuity, as this arrangement was closer to the original full length version.

By preventing HSF1 from binding to its cognate DNA sites, its aptamer was anticipated to compromise yeast survival under elevated temperature. Prompted by this hypothesis, we examined the growth rate of yeast strains harboring the homed intron-aptamer at different temperatures. As shown in Figure 3.3B, there was insignificant difference in growth rate between the homed strain and the parental strain when they were cultured under 22 °C. As the temperature increased, I observed moderate growth retardation for the homed strain under 30 °C. The most severe phenotype was observed when the temperature was raised to 37 °C: the AptHSF(M) homed strain grew very slowly under this heat shock condition. As an additional control, I also tested homed strains with an intron carrying the antisense sequence of the aptamer (R1). Growth of this strain was normal under all three temperatures.

In addition to heat treatment, I also tested the growth of the aptamer expressing yeast under several other stressful conditions (Figure 3.3B). I found that growth of these strains was slightly compromised by CuSO₄. However, exposure to hydrogen peroxide and hyper-osmotic shock did not reveal any difference between the aptamer-expressing strains and the antisense and parental controls. These results strongly indicated a causal relationship between the aptamer expression and the
specifically compromised response to thermal stress.

3.5. Further increasing the aptamer dosage by two configurations of aptamer dimers

Having successfully tested the expression system and confirmed the aptamer induced inhibition of yeast heat shock response using the monomeric construct, we tried to further increase the dose of the aptamer and its inhibitory effect on heat response. To this end, we designed and constructed two dimeric constructs, so that each transcription event of the RNA construct will produce an extra copy of the aptamer in close proximity, further increase the binding avidity to the target protein. To ensure correct folding, an additional three-way junction was used to join the two aptamers and the intron together. This three-way junction had been shown previously to be stable among other variants (Diamond, Turner et al. 2001) and used to present multiple aptamers in molecular composites (Xu and Shi 2009). As shown in the lower panel of Figure 3.3A, we used either Stem 3 or Stem 2 to connect the aptamer to this additional three way junction to form the dimers RA1(D1 and D2).

With two stems available as points of integration, four different dimeric configuration could be arranged theoretically. However, among these four, two configurations would contain two near identical aptamer sequences in the same direction, which dramatically increase the likelihood of misfolding. Therefore, only the type of connection that forms two aptamers in opposite direction were constructed,
and named D1 and D2 respectively (Figure 3.3A). These two longer and more complex sequence also helped demonstrate the general utility of the aptamer expression system.

Yeast strains harboring each of the dimer configurations or their antisense were plated and placed in three different temperatures. This time, a growth retardation is observed for the RA1 D1 and D2 strains at both 22°C and 30°C, suggesting that HSF function is not limited to heat shock response. The cells harboring the dimers could not survive under 37°C, indicating a stronger inhibition of heat shock response generated by the dimers as compared to that by the monomer. The two configurations showed similar phenotypic effect, and the strains expressing the two antisense constructs did not have an observable difference in growth compared to wild-type.

3.6. Molecular phenotypes caused by the HSF aptamer

The HSF1 plays a predominant role in the induction of heat shock (HS) genes. The temperature sensitive phenotype of aptamer-expressing yeast strains described above led us to predict that the aptamer interfered with HSF activation of at least some HS genes. To investigate the mechanism underlying the systemic effect of the aptamer, we examined the mRNA level of three representative HS genes before and after heat treatment using the W303 stains expressing the AptHSF-RA1 in either monomeric (M) or dimeric forms (D1 or D2), and compared them with the strain harboring a control antisense version of each construct. These three HS genes, SSA3,
HSP82, and SSE2, all have HSEs in their promoter and are inducible by heat. As a control, we used the HSP12 gene, which is inducible by heat or other stresses but do not have typical HSEs in the promoter.

As shown in Figure 3.4A, using RT-PCR, we observed significant decrease of heat induced expression of all three typical HS genes in the strain expressing the AptHSF monomer. In addition, the basal level of SSA3 and SSE2 in this strain was also lower than that in the antisense control. In contrast, the expression level of HSP12 was not changed under both heat shock and non-heat shock conditions. To confirm this result, we used quantitative PCR (qPCR) to measure the mRNA level of the same three HS genes, shown in Figure 3.4B. In the strain expressing the AptHSF dimer, same HS genes are tested by RT-PCR and quantitative PCR. Figure 3.5 shows more severely decreased level of these mRNAs in the dimer expressing strains. In these experiments, we used the Pol II-driven ethanol inducible gene ADH1 and the Pol III-driven constitutive gene U6 to normalize the data. They were not induced by heat shock and not inhibited by the aptamer.

3.7. Discussion

HSF1 is a transcription factor involved in a wide variety of biological processes, in particular stress responses. It is also a master regulator of a cancer-enabling gene network and a potential target for therapeutic inhibition in mammalians. Interfering with DNA binding of a transcription factor is an efficient way of downregulating
genes controlled by this factor, a strategy demonstrated by an aptamer for NF-kB (Lebruska and Maher 1999; Huang and Szostak 2003; Mi, Zhang et al. 2006). The aptamer used in the present study was previously characterized to recognize the DNA binding domain and the adjacent linker region of HSF (Zhao, Shi et al. 2006). Here we showed that the growth of aptamer-expressing yeast was severely impeded by elevated temperature, but not by exposure to CuSO$_4$, hydrogen peroxide and hyper-osmotic shock. In yeast, there is a general response of gene expression to stress as well as specific genomic expression patterns for particular environmental conditions (Gasch, Spellman et al. 2000). Our analysis indicated that the specific downregulation of heat shock genes in these strains was responsible to the systemic effect. The congruence of cellular and molecular phenotypes demonstrated the specificity of the aptamer. Therefore, this type of mechanism-driven rational modulation of specific macromolecular interactions in the complex intracellular context can lead to changes in collective and emergent properties of gene networks and pathways, which in turn manifest as predictable and desirable changes in cellular phenotypes. Interestingly, the AptHSF-RA1 is also able to recognize mammalian HSF1 (Zhao, Shi et al. 2006), thus providing us with the exciting prospect of further developing it as a potential anticancer drug lead.

A construct similar to D1 described herein was tested in another lab (Zhao 2007). In that study, the intron donor plasmid used was pJLTtLSU1 (Lin and Vogt 1998), which is similar to pRSTtLSU1-ClaI and also has a HIS3 marker. However, for
unknown reasons, the transformants were selected on Uracil (U-) and Leucine (L-) dropout medium. Different from our observation, most of the cells died upon induction of I-PpoI. Nonetheless, some cells apparently having integrated introns were somehow isolated and further studied. The phenotypes of this strain were similar to those of my aptamer-expressing strains, with one exception: HSP12 expression was severely compromised. Because the promoter of this gene has no HSEs, there must have been some other uncontrolled events going on.

The forced expression of high levels of functional synthetic RNAs in vivo requires the use of powerful promoters to drive the transcription of genes encoding these RNAs. To avoid shunting the transcripts into a pre-mRNA pathway, Pol I- or Pol III-driven promoters are often used for this purpose. However, the high level transcription of a small number of housekeeping genes by these RNA polymerases may also suggest that these transcription systems are operating close to full capacity and leave less room to accommodate additional genes. This may explain the results previously published regarding poor aptamer expression using the promoter of the RNase P RNA gene (Thomas, Chedin et al. 1997). In this study, we created a parasitic expression system, in which the total number of Pol I transcription initiation events in the cell remained unchanged. This scheme ensured that there is no increased consumption of potentially limiting Pol I transcription factors by additional promoters. We successfully achieved a high level of aptamer accumulation by using a specific homing endonuclease to insert a self-splicing intron containing the aptamer sequence
into every copy of the 150 highly expressed rDNA genes. In general, homing endonucleases recognize sequences 14-40 base pairs in length (Stoddard 2005) and are being developed as powerful and precise tools to insert therapeutic genes into a chosen location to circumvent the hazards of random insertion technology (Chevalier, Kortemme et al. 2002; Budd, Reis et al. 2006; Rosen, Morrison et al. 2006). Our aptamer expression system provided an innovative demonstration of their utility and can be used to deliver other functional RNA.
3.8. Figures

Figure 3.1. The indirect aptamer delivery system. (A) Trans-integration of a group I intron. The three components, the host and the two plasmids, are depicted. The arrow shows the transition of rDNA to the intron homed state caused by Ppol cleavage of the rDNA and insertion of the intron containing the aptamer. (B) Schematic diagram of the intron-aptamer. P4-P6 and P3-P9 signify the two catalytic domains of the group I intron. The sequences that are illustrated as bases show the Clal sites and the GC clamp extending from the P1 stem.
**Figure 3.2. Functionality of the aptamer delivery system.** (A) Intron homing to rDNA as confirmed by PCR. The primer annealing sites on rDNA before and after homing is given at the top. The homing of the intron to all rDNA is shown for one clone in the W303-1A parental strain. EE designates bands produced by the two exonic primers. EI designates bands produced by the exonic primer E1 and the intronic primer I. “Marker” indicates molecular weight markers. (B) RNA expression as monitored by Northern blot analysis. “P” indicates parental strain. “H” indicates the homed strain. W303-1A was used for this experiment.
Figure 3.3A. Predicted secondary structure of the monomer (top panel) and the dimer (bottom panel). The sequence between the two ClaI sites is shown. Different stems are indicated for the monomer. The additional three-way junction used to present the dimer is encircled. I and II indicates two configurations of dimer constructs. S1, S2 and S3 indicates three stems of original AptHSF.
Figure 3.3B.  Cellular phenotype of the inhibitory HSF aptamer. (Top panel) The growth of the monomer (M)-expressing strain and control strains cultured on different solid media and under different temperatures. “Antisense (R1)” is a strain in which the antisense sequence of the monomer was expressed. (Bottom panel) The growth of the dimer (D)-expressing strains and control strains cultured under different temperatures. “Antisense” is a strain in which the antisense sequence of the dimer was expressed. In both (B) and (C) the parental strain was W303-1A.
Figure 3.4. Molecular phenotype of the inhibitory HSF aptamer monoer. (A) Effect of aptamer monomer (M) on the level of HS genes measured by conventional RT-PCR. AptHSF-Antisense is a strain expressing the antisense sequence of the aptamer monomer. HS = heat shock (20 minutes at 39°C), NHS = none heat shock. (B) Effect of aptamer monoer (M) on the level of HS genes measured by RT-qPCR. “R1” is a strain expressing the antisense sequence of the aptamer. The RNA level for each gene are presented as the ratio to the full HS induction level in the antisense control strain, which is set to 1. The expression level for each gene is normalized to that of ADH1. The error bars show the standard error from RT-qPCR experiments using three independently heat shocked yeast RNA preparations from the same strain. In both panel the parental strain used was W303-1A.
Figure 3.5.  Molecular phenotype of the inhibitory HSF aptamer dimer. (A) Effect of aptamer dimer (D1 and D2) on the level of HS genes measured by conventional RT-PCR. “D1-R and D2-R” are the strains expressing the antisense sequence of the aptamer dimer. HS = heat shock (20 minutes at 39°C), NHS = none heat shock. (B) Effect of aptamer dimer (D1 and D2) on the level of HS genes measured by RT-qPCR. The RNA level for each gene are presented as the ratio to the full HS induction level in the antisense control strain, which is set to 1. The expression level for each gene is normalized to that of ADH1. The error bars show the standard error from RT-qPCR experiments using three independently heat shocked yeast RNA preparations from the same strain. In both panel the parental strain used was W303-1A.
4.1. Introduction

In eukaryotic organisms, genes encoding messenger RNA are transcribed by RNA polymerase II (Pol II) with the help of general transcription factors (GTFs) (Fuda, Ardehali et al. 2009). To initiate transcription, the TATA-binding protein (TBP) first binds to DNA. Next, TFIIA and TFIIB bind to TBP and the core promoter, followed by TFIIF and Pol II. Finally, TFIIE and TFIIH join to complete the assembly of a Pre-Initiation Complex (PIC) (Woychik and Hampsey 2002; Kostrewa, Zeller et al. 2009). In addition, transcription of most genes requires activators, because the formation of chromatin makes the transcriptional ground state restrictive (Struhl 1999). There are two general mechanisms by which activators facilitate transcription: directly through interacting with members of the Pol II entourage or indirectly through altering chromatin structure (Ptashne and Gann 1997; Orphanides and Reinberg 2002). In either case, the location at which the activator binds to DNA determines which gene is activated. Therefore, a transcription activator requires a minimum of two domains, a DNA-binding domain and an activation domain. According to the recruitment model, the target of an activation domain is likely to be either a GTF or a subunit of the Pol II-mediator complex. Among the GTFs, TBP and

TFIIB are most strongly implicated as the targets of activators (Ptashne and Gann 1997).

Although the general scheme of transcriptional activation by recruitment has been delineated in broad outline, certain important details remain elusive due to experimental difficulties. For example, an activator often interacts with multiple GTFs, and its effect on a single factor is therefore difficult to isolate; artificial recruitment of a single factor through fusion to a DNA-binding domain does not yield any information about the site or sites on the factor contacted by activators (Ptashne and Gann 1997). Many protein activators share a common amino-acid composition rather than exhibiting similarity in sequence or structure (Ma and Ptashne 1987); many RNA sequences have been isolated based on their capability to activate transcription, but the mechanistic basis for this activity is unknown (Buskirk, Kehayova et al. 2003; Saha, Ansari et al. 2003). Both observations raised questions regarding the specific features of surface topography that are essential for the function of an activation domain. An understanding of the mechanism underlying a phenomenon should enable the design and construction of different systems that are able to reproduce that phenomenon. Therefore, deliberate creation of novel molecules with explicitly and strictly defined biological function is a reliable way to test our current knowledge.

Following this principle, in the present study we implemented the mechanism of transcription activation by recruitment of a GTF using an RNA molecule assembled
from refined and standardized parts, especially those derived from aptamers. To
explore specificity inherent to both sides of the activator–target interface, we made
use of a well-characterized site-specific aptamer as the activation domain of a
synthetic activator. RNA aptamers are generated in an in vitro process emulating
Darwinian evolution (Ellington and Szostak 1990; Tuerk and Gold 1990). For many
proteins, aptamers with a dissociation constant in the nanomolar range have been
isolated. Because selection of an aptamer based on affinity for its target is performed
outside the cellular and organismal milieu, the aptamer often interferes with the
function of the protein when introduced into a living system (Shi, Hoffman et al.
1999). Consequently, aptamers are routinely used as inhibitors of protein activity.

Here we attempted to rationally convert this passive role of aptamers into an
active one by placing an aptamer in a designed molecular context, in which it
functions as one of several intentionally chosen interacting sites. In particular, I
constructed a ‘transcription activator RNA (taRNA)’ in the yeast *Saccharomyces
cerevisiae*, analogous to a protein-based activator. Using a set of modular parts in a
combinatorial manner, I specifically implemented the mechanism of transcriptional
activation by recruiting TFIIB to the promoter of reporter genes in the chromatin
environment. For this purpose, an RNA aptamer for TFIIB (Sevilimeedu, Shi et al.
2008), which is a potent inhibitor of transcription by default, was converted into the
activation domain of the taRNA by design. With the help of several other constructs
originally designed for the yeast three-hybrid system (Bernstein, Buter et al. 2002), I
was able to show that this synthetic RNA molecule activated transcription at a level comparable to a protein activator. Comparing the results obtained by creating new RNA-based factors with those obtained by reorganizing existing protein-based factors allowed us to highlight some critical features of this mechanism.

4.2. Mechanism-driven choice of an aptamer

Based on the recruitment model of PIC assembly, we surveyed published aptamers to find a candidate that would function as an activation domain when tethered to a promoter. Five aptamers were identified: one for Pol II, two for TBP, and two for TFIIB. All of these aptamers showed inhibitory effects on transcription either in vitro or in vivo. Four of them were deemed unfit to act as an activator based on the following mechanistic information. First, the aptamer for the Pol II, FC (Thomas, Chedin et al. 1997), binds in the Pol II active center cleft and prevents the DNA template from entering (Kettenberger, Eisenfuhr et al. 2006). Therefore, FC would not be able to activate transcription even if it were used to bring Pol II to a promoter. Second, the two aptamers for TBP, AptTBP-12 and AptTBP-101, recognize two discrete sites respectively (Fan, Shi et al. 2004; Shi, Fan et al. 2007). Both inhibit transcription by preventing PIC formation, although in mechanistically distinctive ways attributable to their specific binding sites (Shi, Fan et al. 2007). Third, one of the aptamers for TFIIB, AptTFIIB-60, inhibits transcription by preventing the incorporation of TFIIB into the PIC (Sevilimedu, Shi et al. 2008).
Intriguingly, the other TFIIB aptamer, AptTFIIB-4, inhibitory as it is, does not affect TFIIB occupancy at the PIC, nor does it affect TBP and TFIIA levels on a template (Sevilimedu, Shi et al. 2008). Therefore, AptTFIIB-4 met our design criteria and was chosen as the putative activation domain. Because it binds only TFIIB and not any other factor, its target in the Pol II machinery can be precisely assigned. Because it does not prevent TFIIB from being incorporated into the PIC, when tethered to a promoter I presumed AptTFIIB-4 would activate transcription either through recruitment of TFIIB to the PIC or through stabilization of a pre-bound complex on DNA. Because AptTFIIB-4 alone is inhibitory, its tethered form must activate transcription through the recruitment of TFIIB rather than by stabilizing the pre-formed PIC. Viewed retrospectively, the assay for the effect of the aptamer on in vitro transcription (Sevilimedu, Shi et al. 2008) was analogous to in vivo “squelching” experiments (Gill and Ptashne 1988; Farrell, Simkovich et al. 1996), thus excluding the possibility of activation through PIC stabilization. Taken together, this specific functional information allowed us to implement the predetermined mechanism from the bottom up.

4.3. Refinement of AptTFIIB-4

The first step of this implementation was to refine the “raw” aptamer into a standardized portable submolecular module. The method we used to stitch multiple elements together in a composite RNA molecule requires the identification of a double stranded stem in each element that can be used as the point of integration. For the
AptTFIIB-4, I identified such a stem through a series of deletions. Here I describe several critical and representative deletion constructs (Figure 4.1).

First, I deleted 28 nucleotides from the 5’ end to form the construct B4-Del5

\[5’-\text{GGGAGCUAAUGGAGUUGCUUGCCAGCCCUAGAAUAAGCGCUAGUACUACAAGCUUCUGGAGCUCGGU-3’}]. Second, I deleted 13 nucleotids from the 3’ end to form the construct B4-Del3 [5’-GGGAGAAUUCAACUGCCAU CUAGGCAAGAGCUAAUGGAGUUGCUGGGGUAGUCCAGCCCUAGAAU AAGCGCUAGUACUACAAGCU-3’]. Both deletions left a putative stem intact and, as shown in Figure 4.1, fully retained their binding activity. Next I deleted both the 5’ and the 3’ segment to form the MiniB4 construct [5’-GGGAGCUAAUGGAGGAUGCUUGCCAGCCCUAGAAU AAGCGCUAGUACUACAAGCU-3’]. This construct was still active as measured by the electrophoretic mobility shift assay (EMSA) shown in Figure 4.1.

The EMSA was performed using \(^{32}\)P labeled RNA probes prepared using the MAXIscript T7 in vitro transcription kit (Ambion) and [\(\alpha -^{32}\)P]CTP (GE Healthcare). The binding buffer contained 50 mM Tris-Cl [pH 7.6], 200 mM potassium acetate, 5 mM MgCl\(_2\), and 2.5 mM dithiothreitol (DTT). The binding mix was run on a 2.5% agarose gel in ¼ TBE buffer (22.25 mM Tris base, 22.25 mM borate, 0.5 mM EDTA) at 4ºC.
I used the program mfold to predict secondary structures for these three constructs and the full-length aptamer. As shown in Figure 4.1B, they all share the same three-way junction, which connects two stem-loops and the stem we identified as the point of integration with the rest of the taRNA molecule. This led to the proposed secondary structure depicted in Figure 4.2A, where the sequence and structure enclosed in the rectangular box is portable for TFIIB binding. This minimized version of the aptamer was composed of a three-way junction, exiting from which were two stem loops (“S” and “L”) and a stem with an open end (Figure 4.2B). To the open stem we added a “GC-clamp” (Cassiday and Maher 2003) by pairing the 5’ and 3’ termini, which served as a standardized interface with the rest of the composite molecule.

4.4. Molecular and genetic design

We converted the inhibitory AptTFIIB-4 into the activation domain of the taRNA by providing its minimized version (hereafter referred to as B4) with a designed molecular context (Figure 4.3A). To integrate it into the cellular regulatory network, the taRNA was produced by transcription from a synthetic gene. Both the taRNA and its expression system were designed using a modular approach, with each part previously tested individually in vivo. To design the composite taRNA molecule, I employed a method described elsewhere, which allows the use of a library of parts with a common junction at the submolecular level to simplify the prediction of function when the parts are combined (Xu and Shi 2009). Specifically, the taRNA
was composed of four functional elements— the B4 aptamer, two copies of the MS2 coat protein ligand, and a hybrid hammerhead ribozyme—connected in a single molecular entity through two three-way junctions (Figure 4.3B).

Protein activators achieve their sequence-specific association with DNA either directly through a DNA-binding domain [e.g. yeast Gal4 (Ma and Ptashne 1987)] or indirectly through an adaptor protein [e.g. VP16 of herpes simplex virus (Cress and Triezenberg 1991)]. For the taRNA we took the latter strategy: instead of a direct DNA–RNA interface, we employed a LexA-MS2 coat protein fusion construct that had been previously used to localize RNA to DNA through more specific DNA–protein and protein–RNA recognition (SenGupta, Zhang et al. 1996). As an indirect DNA association domain, we used a dimeric configuration of the coat protein ligand to improve its avidity to the MS2 coat protein.

To express the taRNA we designed a transcriptional unit driven by RNA polymerase III (Pol III), as TFIIB is not involved in Pol III transcription and its aptamer would not affect the production of the taRNA,. In addition to the coding region, the synthetic gene comprised the promoter and terminator of the RNase P RNA gene RPR1 (Lee, Rohlman et al. 1991). The same promoter–terminator system was used previously to express the aptamer FC, but failed to generate a detectable growth phenotype (Thomas, Chedin et al. 1997), suggesting that the taRNA thus produced would be kept at an appropriate level without causing systemic effects (e.g.
squelching). However, the RPR1 promoter is intragenic, and in the FC construct the leader sequence was not cleaved off (Thomas, Chedin et al. 1997). Also carried on the transcript is part of the terminator sequence including a stretch of uridines. To eliminate these flanking sequences, we used two cis-acting hammerhead ribozymes (Khvorova, Lescoute et al. 2003).

After self-cleavage, the 3’ half of the 5’ ribozyme and the 5’ half of the 3’ ribozyme would form a new ‘hybrid hammerhead ribozyme’, which is part of the taRNA shown in Figure 4.3B. Because it is not poly-adenylated, the taRNA would be retained inside the nuclei as demonstrated previously (Shi, Hoffman et al. 1999). Inside cells most RNA molecules are degraded by exonucleases that digest single-stranded RNA. Enhanced stability has been demonstrated by confining the termini of an RNA in a double-stranded stem (termed S35) (Thompson, Ayers et al. 1995). The Stem III of the hybrid ribozyme was such an S35 (Figure 4.3B). The hammerhead ribozyme cleaves the RNA backbone to yield a 5’ hydroxyl and a 2’, 3’ cyclic phosphodiester; under certain conditions it might ligate these products using the bond energy retained in the non-hydrolyzed cyclic product to re-form a phosphodiester (Buzayan, Feldstein et al. 1988). Sequences encoding all parts of the taRNA except the activation domain were embedded in the yeast/Escherichia coli shuttle vector pDB-sansA, which produced an RNA construct with DNA-binding (DB) activity conferred by the protein adaptor, but without an activation domain (hence ‘sansA’). Figure 4.3C depicts a section of the vector between the two EcoRI sites, in
alignment with the corresponding RNA transcript before and after ribozyme cleavage. The taRNA gene was generated by inserting the B4 sequence through subcloning into a NotI site to form a ‘DB-B4’ construct. The two resulting NotI sequences would form a GC-clamp to insulate the incoming aptamer from the rest of the molecule to ensure its correct folding. Alternative aptamers and positive or negative control units were also added to the ‘DB-sansA’ scaffold through this standardized procedure.

4.5. RNA expression and ribozyme function in yeast

After transformation, it is critical to confirm that the taRNA and the control constructs, both positive and negative, were expressed at a similar level, so their activity could be compared. It is also important to confirm the functionality of some designed features such as the ribosomal activity. For this type of conformation I used RT-PCR assays. In Figure 4.4A, I indicate the annealing sites of three primers on the RNA transcript. Their sequences are as follows: U1 [5’-CTGTGGTGCTCGCGGGAGGAACGAA-3’]; U2[5’-GCGACGAAACCTCGAGTCATACGTCG-3’]; D[5’-CGGCAGTACTCCGGAACGCTGGTGGA-3’]. Different aptamers or aptamer derivatives were inserted at a point between U2 and D, the cleavage site of the 5’ hammerhead ribozyme was located between U1 and U2.

For each RT-PCR reaction 100ng total RNA was used. Reverse transcription was done at 45°C, 20min, followed by 20 PCR cycles. Each RT-PCR was paired with a control reaction (NRT) in which the reverse transcriptase was not added. As shown in
Figure 4.4B, when the primer pair U2 and D was used, a similar level of expression was observed for taRNA and its control constructs, and no DNA contamination was detected in this assay. In contrast, when the primer pair U1 and D was used, no RT-PCR product was detected (Figure 4.4C), indicating the cleavage of the ribozyme.

4.6. Activity of the transcription activator RNA

The taRNA activity was measured through the transcription of reporter genes under the control of LexA operators, with the help of a LexA-MS2 coat protein fusion adaptor. The yeast strain YBZ-1, designed for the three-hybrid system, contains both the reporters and the adaptor. Importantly, the reporter genes HIS3 and lacZ are in the chromatin environment (Bernstein, Buter et al. 2002). To verify the predetermined function of the activation domain in the taRNA, I employed a positive control along with a battery of negative controls. For the positive control, I used the RNA-based transcription activator m26-29 (Buskirk, Kehayova et al. 2003). This RNA was selected for its capability of activating transcription, but its target and mechanism are unknown. My first negative control was the antisense sequence of B4 (B4rev)—a ‘DB-B4rev’ construct was a cloning byproduct of the taRNA. As shown in Figure 4.2B, B4rev was able to form a secondary structure similar to B4. We also generated two other negative controls, B4mutS and B4mutL, by replacing one of the two loops of B4 with a UUCG-tetraloop. The “empty” vector also served as a negative control as it would produce the “DB-sansA” RNA analogous to an activation domain deletion. After transformation, the expression of the RNA constructs in each yeast strain was
confirmed by RT–PCR (Figure 4.4 A and B), and the transformants expressing the
taRNA and the controls were plated at 30°C on media with or without histidine. As
shown in Figure 4.5A, all strains grew equally well on the plate with histidine,
indicating that, as predicted, taRNA at this level of expression had no significant
systemic effect. In contrast, on the plate without histidine, only strains expressing the
taRNA or the positive control DB-m26-29 were able to survive. To confirm this result,
I performed two more assays. First, I added increasing amounts of 3-aminotriazole
(3-AT), a competitive inhibitor of HIS3p activity, to the his- medium to test the
strength of the taRNA. As shown in Figure 4.5A, in media containing 5mM 3-AT,
some yeast colonies expressing the taRNA were still able to grow. Second, I
measured the growth rate of these strains in liquid media. As shown in Figure 4.5B,
the taRNA was able to sustain a growth rate comparable to DB-m26-29, although
none of the negative controls grew without histidine.

As an alternative and independent reporter gene I used lacZ, the activity of
whose product is easier to quantify. First, we measured the β-galactosidase activity
using a qualitative filter assay (Breeden and Nasmyth 1985). As shown in Figure 4.6A,
permeabilized cells in yeast colonies expressing the taRNA or the positive control
were able to convert X-gal to produce a blue color, but all the negative controls lacked
this capability. To quantify the specific activity of β-galactosidase, I next used a liquid
assay with cell lysates and ONPG (Ralser, Goehler et al. 2005). As a benchmark, I
compared the activity of taRNA to the positive control used in the three-hybrid
screening, in which a Gal4 activation domain is tethered to DNA using the binding of the iron regulatory protein 1 (IRP1) to the iron response element (IRE) (SenGupta, Zhang et al. 1996). As shown in Figure 4.6B, the taRNA has activity more than 2-fold greater than this control.

Furthermore, to gauge the population diversity, we determined the β-galactosidase activity in living cells using optical arrays of single cells (Biran and Walt 2002) with the chromogenic substrate C12FDG. This experiment was performed with the help of Dr. Jason Shepard (Department of chemistry, University at Albany, SUNY). The substrate contains a lipophilic tail enabling its entry into the cell. Once processed by the enzyme, this tail is lost, and a fluorescent product is trapped inside the cell. This assay allowed examination of cells individually under identical conditions. As shown in Figure 4.6C, the taRNA cells exhibited minimal heterogeneity. The standard deviations associated with the optical signal responses across the population were less than 2% of the measured signal intensity.

4.7. Specificity of the transcription activator RNA

Whereas B4 was the main focus of the present study, we used other aptamers for components of the Pol II machinery to help clarify and corroborate our mechanistic claims by comparing and contrasting their effects with those of B4’s. As shown in Figure 4.7A, we refined four other aptamers for TBP, TFIIB, or Pol II by trimming their sequences and converting them to the standardized form. Our modular design
allowed rapid and easy addition of these standardized aptamers to the DB-sansA scaffold. As predicted, none of these aptamers were able to function as an activation domain (Figure 4.7B). While the failure of FC, which binds the active cleft of Pol II, provides a straightforward illustration of the importance of the binding site on the target (Kettenberger, Eisenfuhr et al. 2006), the other three constructs afforded more subtle insights. Both TBP and TFIIB are primary targets of transcription activators, yet simply recruiting them to the promoter was not enough to activate transcription. In the case of TBP, fusion with a DNA-binding domain was able to activate (Chatterjee and Struhl 1995; Klages and Strubin 1995; Xiao, Friesen et al. 1995), but recruitment through a tethered aptamer was not. Presumably the fusion protein was incorporated into the holoenzyme, but aptamer binding prevented TBP from doing so, even when the binding site on TBP is the DNA-binding surface (Shi, Fan et al. 2007). A more interesting case is TFIIB. Both B4 and AptTFIIB-60 bind to TFIIB with similar affinity, and their contact sites overlap on the c-terminal core domain (cIIB) at or close to the linker region (Sevilimedu, Shi et al. 2008). Transcription activators such as Pho4 have also been shown to bind cIIB (Wu and Hampsey 1999). However, the two aptamers for TFIIB behaved differently both in vitro and in vivo. Taken together, this collection of constructs demonstrated that the specific site of contact and the mode of contact with the target were important for activation to occur.

Finally, we attempted to use these newly acquired mechanistic insights to build another aptamer-based transcription activator with a specified target. Previously, an
RNA aptamer for the heat-shock factor (HSF) was isolated (Zhao, Shi et al. 2006) and I have showed its inhibitory effect in vivo in Chapter 3. As mentioned in that chapter, HSF recognizes the heat-shock elements on DNA and upon heat shock activates transcription by recruiting other factors or complexes to the promoter. The aptamer, AptHSF-RA1, binds to the HSF DNA-binding domain and the linker region but not the domains involved in trimerization and activation. Therefore, we reasoned that AptHSF-RA1 might be used to recruit HSF to a non-heat-shock promoter. Not being occluded by the aptamer, the activation domain of the HSF would remain functional. Indeed, when we expressed a DB-RA1 construct (AptHSF-RA1 in the pDB-sansA vector) in yeast, we observed transcription activation of both reporter genes at a moderate level, lower than that produced by the taRNA but significantly above the background level (Figure 4.7B). However, this activity was indifferent to temperature change. Apparently, recruiting HSF this way was not able to fully recapitulate the natural molecular interactions and modification at the promoter (Sorger and Pelham 1988). Nonetheless, the activity we observed lends further support to the mechanism of recruitment illustrated by the B4-derived taRNA.

4.8. Discussion

In this study, we took a forward engineering approach to synthesize a transcription activator to implement the mechanism of transcription activation by recruitment of a GTF. By comparing the expected and observed behavior of the designed molecules, we were able not only to validate this mechanism, but also to
investigate the specificity between an activator and its target. In the future, the taRNA will be used as a model to study the events that occur in the process of transcription activation after the B4 associates with TFIIB.

Our work was greatly facilitated by existing constructs originally designed for the yeast three-hybrid screening (Bernstein, Buter et al. 2002). However, there is a fundamental difference between our work and the three-hybrid system. The general utility of the three-hybrid system lies in the fact that transcriptional activation in this system relies on the physical rather than the biological properties of the RNA (SenGupta, Zhang et al. 1996). Numerous types of RNA–protein interactions, including the binding specificity of aptamers, have been analyzed using this system, regardless of the normal function of the RNA molecule (Zhang, Kraemer et al. 2000; Konig, Julius et al. 2007). In contrast, here the taRNA was designed to provide the biological function and integrate itself into the functional context of PIC formation. This difference can be appreciated by comparing our work with a recently published study (Wurster, Bida et al. 2009). In this elegant piece of work, as well as similar works from the same group (Cassiday and Maher 2001; Cassiday and Maher 2003), an RNA aptamer was also involved and the target of the aptamer was also a transcription factor. However, the B4 aptamer in our construct is functionally not analogous to the anti-NF-kB aptamer, but instead analogous to the GAL4 activation domain in their system.
In order to regulate biological processes, proteins and other molecules associate with each other through a complex network of interactions. Modification of the network connectivity forms the basis of experimental perturbation and therapeutic intervention. Traditional methods modify such connectivity by blocking or abolishing molecular interactions. Following this strategy, RNA aptamers have been used as protein antagonists for more than a decade (Nimjee, Rusconi et al. 2005). An alternative and possibly more effective strategy is to introduce new links between non-interacting molecules. A recent study converted a monomeric aptamer into a non-covalent homodimer functioning as an agonist by inducing the targets to multimerize (Dollins, Nair et al. 2008). Differing from this approach, our implementation of the aptamer-based taRNA can be viewed as a molecular surgery that “rewired” the connectivity of an existing regulatory network to bypass a native activator. By transfiguring an inhibitory RNA aptamer into an activation domain and constructing covalent composites using a modular and combinatorial procedure, we have expanded the utility of aptamers and introduced a new type of building blocks, or “synthons,” for use in synthetic biology. This general method is also applicable to many real world situations, such as using aptamers targeting prostate specific membrane antigen and HIV envelop protein for tumor cell sensitization (Ni, Zhang et al. 2011) and HIV therapy (Zhou and Rossi 2011).
4.9. Figures

Figure 4.1. Refinement of AptTFIIB-4. (A) Activity of deletion constructs as measured by EMSA. (B) Predicted secondary structures for these constructs.
Figure 4.2. AptTFIIB-4 and its derivatives. (A) Predicted secondary structure of AptTFIIB-4. The structure was generated by mfold and supported by mutational analysis. The box indicates the ‘B4’ aptamer moiety. (B) Standardized aptamer B4 and its derivatives. The ‘S’ and ‘L’ loops are indicated in B4. The boxes indicate the identical GC-clamp in each structure.
Figure 4.3AB. Molecular and genetic design of taRNA. (A) Schematic diagram showing the function of the taRNA. (B) The most stable structure of the taRNA as predicted by mfold. Each structural and functional unit is indicated by a circle or a rectangle, with the name of the element next to it. The three-way junction connecting the two MS2 coat protein ligands was derived from the 5S rRNA of H. marismortui. The other three-way junction was a stable 5S rRNA variant named “System F”.
Figure 4.3C. Molecular and genetic design of taRNA. Schematic diagram depicting the EcoRI segment in the pDB series of plasmids before and after subcloning, and the RNA transcribed from the synthetic gene before and after further processing by hammerhead ribozyme cleavage. The DNA and RNA segments are aligned to show their relationship. The two ribozymes were taken from peach latent mosaic viroid (PLMVd) and tobacco ringspot virus satellite RNA (sTRSV), respectively (31). The cleavage sites of the ribozymes are indicated by arrowheads.
Figure 4.4. RNA expression and ribozyme function in yeast. (A) Schematic Diagram of the RNA transcript with primer annealing sites indicated. (B) RT-PCR Assay revealing similar level of expression for different constructs. NRT: no reverse transcriptase added. (C) Failure of U1-D primer pair to amplify the RNA transcript, indicating cleavage of the 5’ ribozyme. The taRNA transcript was used for the assay presented here. M: molecular weight marker.
Figure 4.5. Activity of the taRNA as measured by the HIS3 reporter gene transcription. (A) Growth on agar plates. Each strain is identified by the molecular construct it harbors and its position on the plates is indicated in the drawing (upper left). The upper right plate is a control plate (ura-, his+) demonstrating that all strains are viable and grow at a normal rate in the absence of selection for HIS3. The two lower plates lack histidine. The lower right plate also contains 5 mM 3-AT. (B) Growth curves measured in liquid media lacking histidine. Values represent the average of three independent cultures of each strain. Error bars show standard deviations.
Figure 4.6AB. Activity of the taRNA as measured by β-galactosidase activity. (A) Qualitative filter lift assay of permeabilized cells from a ura-, his- plate using X-gal. (B) Quantitative colorimetric assay of cell lysate using ONPG. The average activity of each strain is shown normalized relative to the Gal4 three-hybrid positive control (IRE-MS2+AD-IRP) (29). Error bars show standard deviations.
Figure 4.6C. Optical array analysis of living single cells using C12FDG. A total of 114 cells were measured one and two hours after adding the substrate. The inset shows a section of the array.
Figure 4.7. Specificity of the taRNA. (A) Secondary structures of five refined and standardized aptamers. Boxes indicate the identical GC-clamp in each structure. (B) Capability or lack thereof of each aptamer to function as an activation domain, as assayed by HIS3 reporter gene transcription and β-galactosidase activity.
CHAPTER V: FUNCTIONAL ANALYSIS OF APTB4 FOR DESIGNING POTENTIAL REGULATORY MECHNISMS

5.1. Introduction

The aim of the project discussed in this chapter is to further evaluate the basic requirement of AptB4 in order to generate a taRNA in which the activity is controlled in a time-specific manner.

In addition to the inherited function of molecular recognition, nucleic acid binding species, or aptamers, have emerged as components of powerful tools and widely adapted in biosensors, in drug-delivery systems, and as regulatory element that control gene expression (Bunka and Stockley 2006; Chu, Ebright et al. 2007). One of the major advantages of aptamers over their protein counterparts is that they have the potential to be easily coupled to other functional RNAs to realize sophisticated functions. Based on prediction of secondary and tertiary structures, aptamers and other functional nucleic acid molecules can be constructed in which they are extended, shortened, or modified in other manners, adding an extra layer of complexity to their original function. They have the potential to be generated through directed evolution or rational design to realize functions as biosensors (Zhou, Battig et al. 2010) and riboswitches (Block, Hammond et al. 2010), able to evaluate molecular information in the extracellular and intracellular environments and ultimately translate the information into observable response over varying time scales (Beisel and Smolke 2009).
In some cases, aptamers were engineered into allosterically controlled functional nucleic acids, the conformation of which is regulated by their cognate ligands. More than one conformations of the same RNA molecule may exist and possess different levels of activity. These conformations co-exist in equilibrium under certain conditions and might be interchangeable with the fluctuation of environmental conditions (Bayer and Smolke 2005; Barrick and Breaker 2007). Different conformations in a population also have their own distinct affinity for a ligand. Upon binding of the ligand, one state may be stabilized, and hence, favored in the population over the alternative conformations. If each conformation is associated with different activities, then increased ligand concentrations either increase or decrease the final readout of population activity, depending on which conformation is biased towards (Chen and Ellington 2009). As a result, specific and high affinity ligand binding induces a transition of total activity exhibited by the collection of RNA molecules. These interchangeable conformations have the potential to be regulatory components, utilized as part of genetic networks, tuning and controlling critical biological functions. Properties of these functions include sensing fluctuating environment (Acar, Mettetal et al. 2008), minimizing excessive expenditure in metabolism (Zaslaver, Mayo et al. 2004; Dekel, Mangan et al. 2005), determining developmental fate (Suel, Kulkarni et al. 2007) and ensuring correct transmission in signal transduction pathways (Yokobayashi, Weiss et al. 2002; Hao, Nayak et al. 2008).
In order to evaluate the potential of RNA aptamer to couple with other functional RNA unit to create a regulatory element, the first step is to extensively study the structural requirements of the aptamer. To this end, I made a series of modifications on the original AptB4 and tested them for binding activity to its target, TFIIB. My analysis revealed that both putative apical loops on the original AptB4 are required for its activity, although the putative stem connected to the two apical loops can be altered into other elements that hold the overall structure. Moreover, replacing the 3-way junction that holds the two stem-loops together with a structural junction from other source also seemed to maintain the binding activity. I tried to inhibit AptB4 activity using antisense oligonucleotides that were complementary to either the whole AptB4 sequence or only covers the two apical loops. These RNA antidote were able to repress AptB4 activity shown by in vitro binding assays, although the amount of antidote required was quite high, which can be another challenge if the AptB4 based taRNA activity is to be regulated in living cells.

5.2. Minimization of the original AptB4

The original AptB4 contains two apical stem loops, connected to a long stem by a three-way junction. The other end of the long stem is jointed with the original 5’ and 3’ primer sequences from SELEX, and this part was eliminated when AptB4 was incorporated into the taRNA based on the results described in Section 4.3. There I used EMSA to test a set of deletion constructs, including elimination of 5’ tail alone
(B4-Del5) or 3’ tail alone and a deletion (B4-Del3) of both 5’ and 3’ tail but leaving
the putative stem at its original length (MinB4). Both the 5’ tail and 3’ tail flanking the
putative stem can be eliminated without sacrificing the binding activity, as compared
to full length AptB4.

Here I made further deletions to probe the structural and sequence boundaries of
this aptamer, using a filter binding assay to examine aptamer binding to TFIIB. A very
high concentration of target protein is used in this assay to ensure that all RNA
molecules are saturated and that weak bindings can be observed. First, I made a
construct of 3’ deletion up to the three-way junction, B4-Del3’. As predicted by the
secondary structure generating software mfold, this elimination should result in
incorrect folding of AptB4 (Figure 5.1A), thus its binding affinity to target protein
TFIIB should be abolished or dramatically decreased. To my surprise, B4-Del 3’ up to
the three-way junction did not eliminate binding at high protein concentration (Figure
5.1B).

I hypothesized that the putative stem might add stability to the aptamer, but
might not be essential for binding. If that is true, then further minimization of AptB4
might still be possible. Based on the sequence of MinB4, I then tested constructs with
further shortened stem, including cutting above the A-A bulge (MinB4-Trunc1),
cutting below the three-way junction (MinB4-Trunc2), and elimination of the
three-way junction (MinB4-Trunc3) (Figure 5.2A). Cutting right below the three-way
junction did show binding, but slightly weaker possibly due to the lowered stability of the construct. No binding was observed for elimination of the three-way junction, even at 100nM TFIIB concentration (Figure 5.2B). This could be either due to instability of the construct or lower transcription efficiency caused by relatively short length (36nt).

5.3. Sequence modifications on MinB4

After testing these deletion constructs, I decided to investigate the sequence requirement for MinB4. I made several modified constructs that preserved the overall secondary structure of MinB4 but with changes in sequence. There were two major types of changes made to the MinB4 sequence. In one type, one of the two apical loops is replaced by a UUCG stable loop (MinB4-mut-L and MinB4-mut-S. Figure 5.3A). These constructs were used as control structures before in the phenotypic analysis of tRNA reporter gene expression assays (see last chapter for detailed description). Neither of the two constructs showed binding as compared to MinB4 (Figure 5.3B). In the other type, a base pair was flipped in one of the stems. When an A-U base pair was flipped on the long stem (MinB4-au-flip-L), binding was eliminated, meaning that this specific sequence is required for aptamer activity. When an A-U base pair at the bottom of the short stem was eliminated (MinB4-au-flip-S), however, binding to TFIIB was not affected, indicating that this sequence was not essential to maintain activity.
Based on the functionality of MinB4-au-flip-S, I then replaced the putative stem with a long G-C clamp (MinB4-au-flip-S-clamp. Figure 5.4A) to test the sequence requirement on the putative stem. As shown in Figure 5.4B, this replacement did not show an observable difference in binding activity compared to MinB4-au-flip-S. This means that the specific sequence on the putative stem is not essential.

Since replacing either of the apical loops with UUCG stable loop eliminated the binding activity of AptB4, it can be concluded that both loops are required for the activity. However, flipping one base pair on each of the two stems had distinct effect on binding: loss of activity occurred if the flipping happens in the long stem but the activity was preserved when the flipping occurred in the short stem. I hypothesized that the specific sequence on the short stem is not required. To test this, I flipped two other base pairs on the short stem. One A-U based pair close to the apical loop is converted to U-A (MinB4-us-flip-S), and one C-G base pair at the bottom of the short stem, part of the original three-way junction, was switched to G-C (MinB4-gc-flip-S. Figure 5.5A). Both changes showed little decrease in binding activity compared to MinB4-au-flip-S-clamp (Figure 5B). Taken together, some sequences on the stems are required, but some other sequences can be altered without greatly affecting aptamer binding activity.

5.4. Replacing the original 3-way junction with other variants

Since flipping the C-G base pair at the bottom of the short stem essentially
changed the sequence of the original three-way junction, I further inferred that specific residues of the three way junction might be altered without affecting binding activity. In other words, the three way junction may function to hold the two stem loops but may not contribute to binding itself. If this were true, then I would be able to transplant the two stem loops onto other structures, such as another three-way junction. As long as the overall structure and the two apical loops were preserved, binding activity might be maintained.

To test this hypothesis, I selected a three-way junction derived from the E.coli 5S rRNA that was previously characterized to be extraordinarily stable (Diamond, Turner et al. 2001), grafted the two stem-loops on two adjacent open branches of the three-way junction, and stabilized the other branch with a G-C clamp. As a result of this combination, three different configurations were formed with the three-way junction rotating at three distinct orientations (MinB4-3wayjunc-5sv a, b, and c. Figure 5.6). The binding activity of these constructs was not significantly decreased compared to the positive control MinB4-au-flip-S-clamp.

A circular permutation of the MinB4 was also made, in which the original 5’ and 3’ ends are connected through a UUCG stable loop. New 5’ and 3’ ends were located between the two apical loops (MinB4-CirPermut-1. Figure 5.6). Reduced binding activity was seen in this circular permutation version, possibly due to accessible 5’ and 3’ end for nucleases access or unstable folding of the two stems.
5.5. Inhibiting AptB4 binding by specific antidotes

As an alternative to regulating aptamer activity by making sequence modifications to it, I attempted to add the regulatory element in trans using an aptamer antidote (Rusconi, Roberts et al. 2004). The antidote for MinB4 used in this case was antisense RNA molecules with sequence complementary to the original AptB4 sequence, the annealing of the antidote to MinB4 would prevent aptamer from folding correctly, thereby potentially down regulating tRNA activity. Since my earlier assays implied that the two apical loops are required for aptamer activity, but the putative stem is dispensable, I used two antidotes for this study, one covered the full length of MinB4 (full length), the other covered only the two apical loops (1/2 full length), with the hypothesis that the 1/2 should be enough to inhibit MinB4 activity. As shown in Figure 5.7, both antidotes repressed binding with similar strength. Although an inhibitory effect was observed for the antidote, the amount of antidote oligonucleotides needed is more than 100 times of the aptamer concentration, which would be difficult to achieve in vivo.

There are various reasons for high amount requirement of the antidotes. First, the antidote itself might be so long that it has a chance to fold on its own, thus limiting the available amount of denatured form for annealing to the aptamer. To this end, I also tested some shorter oligos, but found that an even higher titer is needed to inhibit aptamer function. Second, hybridization of the antidote to aptamer and aptamer
folding might prefer different conditions. To this end, I tested different time points for adding antidote, before, during and after aptamer-protein interaction and tried different temperatures that might favor hybridization, but did not see improvement of antidote mediated inhibition. Third, in vitro and in vivo conditions can be dramatically different, requirement of a high antidote to aptamer ratio in vitro might simply be due to the relatively large reaction volume. Such situations might be different in living cells, where molecules are confined to a small compartment, giving them higher chance for hybridization and competing with their own folding. To mimic an in vivo environment, I used several reagents to create a molecular crowding event, and found that antidote mediated inhibition was improved slightly in 5% PEG, but began to get worse in higher concentration of PEG, possibly due to the fact that PEG created a molecular crowding for the protein target as well, effectively increasing the concentration of aptamer target.

5. 6. Perspectives on development of RNA based regulatory elements

In designing and generating a regulatory mechanism on an existing functional aptamer, there are various issues to be addressed. Sometimes, a trade-off has to be negotiated when meeting some of requirements. For example, a high dynamic range of activity with sensitivity to ligand binding would compromise some other parameters. To adjust the design to maximum efficiency, attempts have been made in different ways, such as mimicking the naturally existed riboswitch, directed evolution with an artificially added selective pressure and mathematical models trying to balance the parameters for an optimal design.
5.6.1. Dynamic Range

A challenge that has been difficult to meet in generating nucleic acid based regulatory elements is a high dynamic range, which is defined by the difference between the highest and lowest activity. To increase the dynamic range, it often requires a low-activity conformation to lower the background noise. However, when such background activity is reduced to a low level by designed conformational changes, it is usually no longer possible to restore its original level of activity when activated, leaving the dynamic range smaller than the theoretically predicted maximum level. Besides, to trigger the conversion from the low-activity state to the high-activity state, the amount of ligand required may be too high to achieve, especially for in vivo purposes. This challenge makes most RNA based regulatory element merely a collection of molecules that possess two states of activity in different conditions, one with slightly higher activity than the other. As a result, the binary interpretation is often achieved through an amplification of downstream signals instead of immediate readout.

5.6.2. Amplification of signal

To overcome the difficulties caused by low dynamic range, it is required to use an efficient measurement method to distinguish the difference between the highest activity and the low background activity. The conformational changes in one RNA molecule upon ligand binding trigger the transition of one structure to the other. Such
transition in a population causes the RNA molecules either gain of function, in the case of ON switches, or loss of function, as in OFF switches (Beisel and Smolke 2009). The relative dominance of each conformation in a population is thus the primary parameter to monitor the efficiency of switches. However, strength of the switches was rarely presented as the immediate readout. More often than not, amplified signals in subsequent events were presented instead, such as strength of fluorescence activities (Wickiser, Winkler et al. 2005; Win and Smolke 2007), level of gene expression, or activity of an enzyme (Isaacs, Dwyer et al. 2006). This is due to the fact that the immediate readout is a change in proportion of a population of RNA molecules that exhibit certain conformation, and is hard to measure.

5.6.3. Naturally existing RNA switches

RNA based regulatory components exist in nature as riboswitches in all level of organisms (Wickiser, Winkler et al. 2005; Barrick and Breaker 2007). The term riboswitch was established to define RNAs that control gene expression by binding to metabolites without the need for protein factors. Riboswitches are primarily found in the 5’ untranslated regions of bacterial mRNAs (Mandal and Breaker 2004; Nudler and Mironov 2004; Winkler and Breaker 2005). Recently, more riboswitches have been found in eukaryotes such as fungi (Sudarsan, Barrick et al. 2003) and plants (Galagan, Calvo et al. 2005). These structured RNA motifs bind to specific ligands, which serves to regulate expression of the adjacent gene or set of genes, commonly by modulating premature transcription termination or translation initiation. Riboswitch
mediated gene control can function as a feedback mechanism, due to the fact that the targeted ligand is often synthesized or processed by the genes located downstream of the riboswitch. The discovery of riboswitches revealed that the ability of RNA to form precision binding pockets for small molecules is harnessed by modern cells to sense metabolites.

The structure of riboswitches generally consists of two regions. The first region usually consists of an aptamer or tandem of aptamers. The second region is an expression platform, which is responsible for modulating gene expression as a result of ligand binding. Since they fall into the larger category of RNAs that respond to changes in environment, the structure and regulatory mechanism can be used to design novel RNA based switches. Riboswitches translate changes in ligand concentration to changes in target gene expression through consecutive steps that begin with ligand binding by its aptamer sequence (Jensen, Atkinson et al. 1995; Osborne, Matsumura et al. 1997). High affinity and specificity binding transform the conformation biases partitioning toward the ligand-bound conformation. Subsequently, the biased conformation exploits diverse gene regulatory mechanisms at transcription, mRNA stability and translation level (Mandal and Breaker 2004).

5.6.4. Designing of RNA-based-regulatory components

In designing RNA based regulatory components, certain characteristics of the molecular reaction upon ligand binding are often assumed. One of the assumptions is
that the secondary structures correlate with the relative stability and thus enabling the prediction of functions. This assumption is questionable as discrepancies have been demonstrated that biosensor properties did not align with the stabilities based on secondary structure (Hall, Cater et al. 2009). Similarly, even greater discrepancies are shown when intracellular energetics are considered (Win and Smolke 2007). Another assumption is that the ligand binding initiated conformational change happens once as in the case of naturally existing riboswitch, when in fact multiple conformations are interchangeable and eventually reach an equilibrium. With this consideration, neither a pure kinetically or pure thermodynamically driven model is alone sufficient to predict switch performance.

Recently, there are studies using mathematical models of both kinetically and thermodynamically driven model to design riboswitches (Chen and Ellington 2009). To take into account various other parameters that would impact the design and performance of RNA switches, methods have been established for measuring rate of transcription, RNA degradation, and strength of ligand binding. But there are other parameters that are also critical in designing switches, which are difficult to measure or predict, such as RNA folding and conformational inter-conversion.

So far, most RNA-based regulatory components are still based largely on a standardized information transmission mechanism, where the transmission is realized by direct coupling of ligand binding domain, usually an aptamer, with an effector
domain, usually an RNA molecule with designated function (Win and Smolke 2007; Win and Smolke 2008). Mathematical optimization, on the other hand, is still largely theoretical and not yet incorporated into schemes of designing actual switches. To this end, selecting an effective regulatory mechanism is critical because many factors would have an impact on the performance of RNA switch (Chen and Ellington 2009).

5.6.5. Directed evolution

Earlier this decade, there has been a stream of studies using a combination of rational design and allosteric selection in generating RNA-based molecular switches. In their designs, the two elementary components of the switch, namely the recognition domain and the reporting domain, are connected through a structural bridge. The recognition domain functions as the ligand sensing element of the switch. As the name indicates, it usually consists of an aptamer with high affinity and specificity to its ligand, thus the ability to recognize input signal. Usually a small molecule is selected as the ligand since it is easier to reach a higher concentration and can readily diffuse into living cells, thus making the switch potentially more sensitive. The reporting domain serves to produce an output signal that can be observed and measured, such as ribozyme self-cleavage rate (Soukup and Breaker 1999), target gene expression level (Buskirk, Landrigan et al. 2004), or the change to a fluorophore (Stojanovic and Kolpashchikov 2004). The structural bridge plays the role of a communication module, to transmit the event of ligand binding to reporting domain activity. Such structural bridge is constructed by selection from a class of either pre-designed or
random stretch of RNA of various lengths. The scheme using randomized sequences was proven to be more successful for two reasons. Firstly, it allows selection for both ON and OFF type switches depending on expected outcome. Secondly, since the random pool contains a much greater diversity of sequences that might have the potential for desired function, it possess higher possibility to generate a switch with minimal background activity and smallest loss of maximum activity, resulting in a higher dynamic range. In contrast to pre-designed structural bridges, random bridges enable direct fitness competition among species of sequence, leaving the best few to survive multiple rounds of selection and stand out by outnumbering other species in the population (Breaker 2002).

5.6.6. In trans control of RNA activity

As stated above, RNA aptamers or other bioactive RNA molecules can be designed to possess a regulatory mechanism on their own, which is especially helpful when they are used to fulfill the functions as biosensors. However, there are other situations where other regulatory mechanism is needed to more efficiently control RNA activity, and the regulation is achieved with more than one RNA molecules. Such situations exist when aptamers are introduced as therapeutic agent. Therapeutic utility of RNA aptamers have been demonstrated by various studies, starting with the design of HIV trans-activating response region aptamer to inhibit HIV replication (Sullenger, Gallardo et al. 1990), and has continued to develop in other clinical fields such as antithrombotic and anticoagulation (Anderson, Li et al. 2007). In these cases,
rapid clearing of RNA based drugs is hard to achieve, but their activity can be controlled independently. By administering an antidote that does not eliminate RNA aptamer but forms aptamer-antidote complex, the aptamer activity is neutralized (Rusconi, Roberts et al. 2004). Due to the unique characteristics of oligonucleotides, antidotes can be generated using Watson-Crick base-pairing rules to sequester aptamer and reverse their activity (Dyke, Steinhubl et al. 2006). It proved that rationally designed aptamer-antidote pairs can be generated to provide control over drug activities.

A different approach may be taken with the AptB4, as characterized in this chapter. Since both apical loops are functionally required, either of the two apical loops can be separated from the other and synthesized from a different source. Re-integration of the two apical loops through some means would restore the aptamer activity, thus making an ON switch. This requires that each of the apical loops be carried on a single molecule, the stability of the molecule and a method of re-incorporation would be another challenge to overcome.

5.7. Potential application of controllable aptamer

With the continuous discovery of riboswitches (Meyer, Hammond et al. 2011) and artificial design of riboswitches based on experimental and mathematical models (Beisel and Smolke 2009), the idea of generating controllable aptamer is becoming an interesting field. There are many successful cases using small molecule aptamers to
regulate activity of a functional RNA, such as ribozyme, protein or fluorophore binding aptamer, and shRNA or other dicer substrate (Rossi 2006). If combined with the intracellular RNA delivery methods, RNAs could be produced in living cells and functionally regulated by small molecule ligand, creating RNA activities that are both temporally and spatially specific.
Figure 5.1. AptB4 deletion constructs binding activity. (A) Structures shown are predicted according to free energy minimization by software mfold. (B) From top to bottom are B4-Del5, B4-Del3, MinB4 and B4-Del3’ respectively.
Figure 5.2. MinB4 deletion construct binding activity. (A) Structures shown are predicted according to free energy minimization by software mfold. Dotted line indicates the position at which deletion is made, resulting deletion construct is shown. (B) From top to bottom are MinB4, MinB4-Trunc1, MinB4-Trunc2 and MinB4-Trunc3, respectively.
Figure 5.3. MinB4 mutation construct binding activity. (A) Structures shown are predicted according to free energy minimization by software mfold. Dotted circle and box indicates the location of mutation or flipped base-pair. Structure of MinB4 is shown here for comparison. (B) From top to bottom are MinB4, MinB4-mut-L, MinB4-mut-S, MinB4-au-flip-L and MinB4-au-flip-S, respectively. This filter binding assay was done on a different filtration apparatus from the one described in text.
Figure 5.4. MinB4-au-flip-S modification. (A) Binding activity of MinB4-au-flip-S, binding assay was done on the apparatus described in test, different from that of previous figure. (B) Original putative stem is replaced by G-C clamp, shown by arrow. Small dotted box indicates the flipped A-U base-pair on short stem. For MinB4-au-flip-S-clamp, double loading is applied for each protein concentration.
Figure 5.5. **Test specific sequence requirement on short stem.** (A) mfold predicted secondary structure of two other construct. Dotted box indicates the location of the flipped base-pair. These two addition construct are named MinB4-gc-flip-S and MinB4-ua-flip-S, respectively. MinB4-au-flip-S-clamp is shown for comparison. (B) From top to bottom are MinB4-gc-flip-S, MinB4-ua-flip-S, and MinB4-au-flip-S-clamp, respectively. Double loading is applied for each protein concentration.
Figure 5.6. Test the overall structural requirement of MinB4. (A) mfold predicted secondary structure new constructs. MinB4-ua-flip-S-clamp is used as positive control. MinB4-3wayjunc-5sv are three constructs with 5s RNA derived three-way junction replacing original three-way junction. There orientations of connection are shown. MinB4-CirPermut-1 is the circular permutation version of original. (B) From top to bottom are MinB4-ua-flip-S-clamp, three MinB4-3wayjunc-5sv variants as in (A), from left to right, and MinB4-CirPermut-1.
Figure 5.7. Antidote treatment of MinB4. (A) Filter binding assay for MinB4 treated with two antidote, full length and 1/2 full length. A decreased binding activity to the level of no target control was seen at antidote to taget ratio of 150:1. (B) Quantitative measurement of result from (A) by phosphoimager.
BIBLIOGRAPHY


