LIGHT-CONTROLLED REGULATION OF EUKARYOTIC GENE EXPRESSION USING A DIFFUSIBLE EFFECTOR

by

Nazanin Azrahimi
B.Sc. (Honors), Simon Fraser University, 2006

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In the
Department of Molecular Biology & Biochemistry

© Nazanin Azrahimi 2010
SIMON FRASER UNIVERSITY

Summer 2010

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.
Approval

Name: Nazanin Azrahimi
Degree: Master of Science
Title of Thesis: Light-controlled regulation of eukaryotic gene expression using a diffusible effector

Examinining Committee:
Chair: Dr. J. K. Scott
Professor of Molecular Biology and Biochemistry, SFU

__________________________
Dr. Dipankar Sen
Senior Supervisor
Professor of Molecular Biology and Biochemistry, SFU

__________________________
Dr. Nancy Hawkins
Committee Member
Assistant Professor of Molecular Biology and Biochemistry, SFU

__________________________
Dr. Peter Unrau
Committee Member
Associate Professor of Molecular Biology and Biochemistry, SFU

__________________________
Dr. N. Harden
Internal Examiner
Associate Professor of Molecular Biology and Biochemistry, SFU

Date Defended/Approved: July 28th, 2010
Declaration of Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the “Institutional Repository” link of the SFU Library website <www.lib.sfu.ca> at: <http://ir.lib.sfu.ca/handle/1892/112>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author’s written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada
Abstract

This project focuses on the design of a small synthetic, trans-acting noncoding RNA to directly regulate gene expression by light irradiation in a eukaryotic system at the mRNA level. This trans-acting noncoding RNA is designed to incorporate an \textit{in vitro}-selected small ligand-binding domain, called an aptamer domain. This aptamer domain selectively binds to one but not the other of the two isomers of a dihydropyrene photo-switch compound. Such an aptamer sequence allowed us to design an allosteric riboregulator, a RNA molecule with a regulatory function whose mode of activity depends on its photo-switch ligand binding to the aptamer domain, which in turn is influenced by light irradiation at an appropriate wavelength. This light-controlled riboregulator is referred to as an antiswitch. Such an antiswitch construct could offer a flexible control strategy by adapting active or inactive forms in response to a specific isomer of the photo-switch compound binding to the aptamer domain. This antiswitch is specifically designed to control the expression of Green Fluorescent protein (GFP) \textit{in vivo} in \textit{Saccharomyces cerevisiae} as a model organism. This light-controlled antiswitch could present a platform for programming cellular behavior in the developmental biology of model organisms with respect to cellular states and environmental stimuli in a non-invasive manner.
Dedication

To my parents, who supported me unconditionally, and for their love and all the sacrifices they made in their lives to make it possible for me to attain my educational achievements to this day. To my brother, who encouraged me throughout these years to excel in life, and supported me with dedication and his positive spirit. To my husband, who believes in me and encourages me constantly, and whose love and support gives me the will to thrive.
Acknowledgements

I would like to express my gratitude to Dr. Dipankar Sen for his support, guidance, and inspiration during the years of my graduate studies. I thank my supervisory committee members, Dr. Nancy Hawkins and Dr. Peter Unrau for their advice and valuable suggestions regarding my thesis project. I thank Dr. Christopher Beh for his collaboration, support and scientific ideas to plan and execute my project.

I thank Dr. Nicholas Harden for being the internal examiner of my defense.

I would like to thank present members of the Sen lab for their help and assistance during my graduate research: Kaushiki Sen, Lester Poon, Gurpreet Sekhon, Paul Cernak, and Janet Huang. Special thanks to Kaushiki Sen for her great friendship and support.

I would like to thank past members of the Sen lab for their guidance and discussion: Ed Leung, Yong Liu, Hyun-Wu Lee. With their guidance on experimental techniques I was able to establish good laboratory skills and habits.

Special thanks to Gabrielle Alfaro from Beh lab for his help, suggestions and guidance on experimental techniques involved with living cells.

I would like to thank Dr. Jonathan Choy and Winnie Enns for their support with flow cytometry.

I would like to thank friends and staff in the MBB department for their help and support throughout the years.
I also thank the MBB department, the university, NSERC Canada, and CIHR Canada for their financial support throughout the years.

Special gratitude to my parents, my brother, and my husband for their encouragement, support, dedication and love.
# Table of Contents

Approval ........................................................................................................ ii  
Abstract ......................................................................................................... iii  
Dedication ..................................................................................................... iv  
Acknowledgements ........................................................................................ v  
Table of Contents ......................................................................................... vii  
List of Figures ............................................................................................... viii  
List of Abbreviations .................................................................................... ix  

## Chapter 1: Introduction to noncoding RNA elements and their role in regulating gene expression

1.1 Introduction: .............................................................................................. 1  
1.2 *In Vitro* Selection and Nucleic Acid Aptamers ........................................... 3  
  1.2.1 *In vitro* Selection .................................................................................. 3  
  1.2.2 DNA and RNA Aptamers ....................................................................... 6  
1.3 Riboregulators: .......................................................................................... 9  
  1.3.1 Antisense RNA ....................................................................................... 9  
  1.3.2 MicroRNAs (miRNAs) .......................................................................... 14  
  1.3.3 Small interfering RNA (siRNA) .............................................................. 18  
  1.3.4 Piwi-interacting RNAs (piRNAs): ......................................................... 22  
  1.3.5 Riboswitches ........................................................................................ 24  
1.4 Thesis Overview: ...................................................................................... 36  

## Chapter 2: Design of light responsive, allosteric C8-antiswitch and its *in vitro* functional analysis

2.1 Introduction: .............................................................................................. 38  
  2.1.1 Photochromic Molecules ...................................................................... 40  
  2.1.2 *In Vitro* Selection of an RNA aptamer against a photochromic compound .... 43  
  2.1.3 A ligand-controlled riboregulator of eukaryotic gene expression ............... 49  
2.2 The design and of a light-sensitive antiswitch and its *in vitro* functional analysis .... 52  
2.3 Material and Methods ............................................................................. 56  
  2.3.1 Materials: ............................................................................................... 56  
  2.3.2 Uptake of BDHP-PEG compound into *S. cerevisiae* .............................. 56  
  2.3.3 Investigating the fluorescent property of BDHP-PEG .............................. 57  
  2.3.4 Antiswitch and EGFP *in vitro* transcription: ........................................ 58  
  2.3.5 *In-vitro* target binding affinity and ligand-induced structural switching of the C8-antiswitch ................................................................. 59  
  2.3.6 Generating C8-antiswitch with defined end-sequences by *cis*-cleaving dual-hammerhead ribozyme construct .................................................. 60  
2.4 Results and discussion: ........................................................................... 62  
  2.4.1 Uptake of BDHP-PEG photochromic compound into *S. cerevisiae* cells ...... 62
List of Figures

Figure 1.1 The cycles of in vitro selection of RNA .......................................................... 4
Figure 1.2 Transcriptional attenuation by antisense RNA in pT181 plasmid ....................... 12
Figure 1.3 microRNA biogenesis ..................................................................................... 16
Figure 1.4 siRNA biogenesis ........................................................................................... 20
Figure 1.5 The ping-pong model of piRNA action for transposon mRNA cleavage .......... 24
Figure 1.6 Common mechanism of riboswitch gene expression control ......................... 26
Figure 1.7 Structural and sequence representation of adenine and glmS riboswitches .... 28
Figure 1.8 Synthetic aptamer-based riboswitch mechanism to control gene expression .... 30
Figure 1.9 Antisense-based riboswitch function in a prokaryotic system ......................... 33
Figure 1.10 Antiswitch design and function ...................................................................... 35
Figure 2.1 photo-induced isomerization of spiropyrans and spirooxazines ...................... 41
Figure 2.2 Photo-induced isomerization of azobenzenes ................................................ 42
Figure 2.3 Photo-induced isomerization of dithienylethenes ......................................... 42
Figure 2.4 Reversible photo- and thermoinduced isomerization of the para quinone structure to the ana quinone structure ................................................................. 43
Figure 2.5 Structure and absorbance spectra of the two photochoromic isomers .......... 45
Figure 2.6 C8 aptamer and its incorporation to design an allosteric hammerhead ribozyme ......................................................................................................................... 46
Figure 2.7 RNA cleaving properties of the UGLOOP-hammerhead ribozyme construct .... 48
Figure 2.8 General illustration of the antiswitch molecule mechanism ............................ 50
Figure 2.9 C8-antiswitch structure and function to regulation EGFP target mRNA translation ............................................................... 54
Figure 2.10 BDHP-PEG compound uptake into yeast cells ............................................. 63
Figure 2.11 Preliminary BDHP-PEG emission spectrum ................................................ 66
Figure 2.12 Excitation and emission spectra of BDHP-PEG and BCPD-PEG samples ...... 67
Figure 2.13 In vitro target binding affinity assay of C8-antiswitch in response to BDHP-PEG ligand ..................................................................................................................... 69
Figure 2.14 Sequences and structures of the C8-Antiswitch stems as predicted by MFold ............................................................................................................................. 72
Figure 2.15  Mechanism of cis-cleaved C8-antiswitch by dual hammerhead ribozymes............. 74
Figure 2.16  Cis-cleaving activity of dual hammerhead ribozymes to generate C8-
antiswitch with defined sequence ends................................................................. 76
Figure 2.17  Cis-cleaving activity of the dual-hammerhead construct as a function of
time.................................................................................................................................. 78
Figure 3.1  Saccharomyces cerevisiae reproduction....................................................... 83
Figure 3.2  Fluorescent activated cell sorting (FACS) ..................................................... 87
Figure 3.3  Map of pJET1.2 plasmid ............................................................................. 104
Figure 3.4  Schematic representation of yeast plasmid constructs, pSwitch and pTarget ...... 106
Figure 3.5  Propidium iodide staining of compound-treated yeast cells.......................... 108
Figure 3.6  Histogram of FACS analysis showing EGFP fluorescence measurements........ 110
Figure 3.7  Flow cytometry and EGFP fluorescent histogram of compound treated
pSwitch+pTarget yeast cells......................................................................................... 113
Figure 3.8  Northern blot analysis for endogenous C8-Antiswitch detection....................... 117
Figure 3.9  Northern blot analysis of endogenous yeast RNA probed for Phenylalanine
detection ..................................................................................................................... 120
Figure 4.1  Sequences and structures of the alternative C8-Antiswitch stems as predicted
by Mfold...................................................................................................................... 129
Figure 4.2  Proposed light-responsive cis-acting C8-aptamer switch............................... 132
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS</td>
<td>Antiswitch</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>BCPD</td>
<td>Benzo-cyclophanediene</td>
</tr>
<tr>
<td>BCPD-PEG</td>
<td>Polyethylene glycol derivative of benzo-cyclophanediene</td>
</tr>
<tr>
<td>BDHP</td>
<td>Benzo-dimethylidihydropyrene</td>
</tr>
<tr>
<td>BDHP-PEG</td>
<td>Polyethylene glycol derivative of benzo-dimethylidihydropyrene</td>
</tr>
<tr>
<td>CTP</td>
<td>Cytidine triphosphate</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleotide triphosphate</td>
</tr>
<tr>
<td>ds DNA</td>
<td>Double-stranded DNA</td>
</tr>
<tr>
<td><em>E. Coli</em></td>
<td>Escherichia Coli</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EGFP</td>
<td>Enhanced green fluorescent protein</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescent activated cell sorter</td>
</tr>
<tr>
<td>FMN</td>
<td>Flavin mononucleotide</td>
</tr>
<tr>
<td>GlcN6P</td>
<td>Glucosamine-6-phosphate</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanosine triphosphate</td>
</tr>
<tr>
<td>Hm</td>
<td>Hammerhead ribozyme</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Hm-As-Hm</td>
<td>Hammerhead-Antiswitch-Hammerhead</td>
</tr>
<tr>
<td>Kb</td>
<td>Kilo bases</td>
</tr>
<tr>
<td>miRNA</td>
<td>MicroRNA</td>
</tr>
<tr>
<td>NTP</td>
<td>Nucleotide triphosphate</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>pGAL</td>
<td>Galactokinase promoter</td>
</tr>
<tr>
<td>pGPD</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase promoter</td>
</tr>
<tr>
<td>Phe</td>
<td>Phenylalanine</td>
</tr>
<tr>
<td>piRNA</td>
<td>Piwi-interacting RNA</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>Saccharomyces cerevisiae</td>
</tr>
<tr>
<td>SD</td>
<td>Shine Dalgarno</td>
</tr>
<tr>
<td>SELEX</td>
<td>Systematic evolution of ligands by exponential enrichment</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering RNA</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer ribonucleic acid</td>
</tr>
<tr>
<td>TPP</td>
<td>Thiamine pyrophosphate</td>
</tr>
<tr>
<td>UTP</td>
<td>Uridine triphosphate</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated region</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
</tbody>
</table>
Chapter 1: 
Introduction to noncoding RNA elements and their role in regulating gene expression

1.1 Introduction:

The ability of a cell to successfully regulate its cellular processes leads to its viability. In virtually all organisms critical cellular events are controlled by modulating gene expression. Jacob and Monod (Jacob et al. 1961) were the first to suggest that a molecular switch could enhance or repress gene expression and regulate cellular processes. Modulating gene expression constitutes a promising therapeutic approach in a variety of diseases, such as inhibition of tumor growth (Tseng et al. 1994, Calabretta et al. 91), treatment of cardiovascular diseases, and inhibition of viral infection. Over the past 20 years, there has been a growing interest in regulating gene expression using synthetic nucleic acids.

Nucleic acids are DNA and RNA molecules that comprise the genetic material of a living organism. DNA molecules exist in vivo as double-stranded helices, composed of a sense and an anti-sense strand. The sense DNA strand contains genetic sequences that are transcribed by RNA polymerases and transcription factors into an RNA sequence, which in turn is translated by ribosomes into a protein polypeptide. This sequential transfer of genetic information is referred to as the central dogma of molecular biology. In addition to coding genes, other non-coding DNA and RNA molecules are found to act as regulatory agents of various cellular processes. There have been studies in which...
RNAs are discovered to act as agents to inhibit gene expression in a sequence-specific manner. More specifically, *cis* and *trans* acting RNA elements have been studied as potential regulators of gene expression. These noncoding regulatory RNA elements are classified as antisense RNA (Good *et. al.* 2003, Vacek *et. al.* 2003).

Antisense RNAs are small, *trans*-acting RNA molecules that bind to target mRNA using Watson-Crick base-pairing. There are different categories of antisense RNAs: microRNAs (miRNAs) (Bartel *et. al.* 2004) are small, single-stranded RNAs that interact with their complementary mRNA sequence. Small interfering RNAs (siRNA) (Scherer *et. al.* 2004) are short double-stranded RNA molecules involved in the RNA interference (RNAi) pathway. Riboswitches (Mandal *et. al.* 2004, Winkler *et. al.* 2002, 2004) are *cis*-acting metabolite binding mRNA structures that control gene expression through modulating translation initiation, or transcriptional termination or mRNA cleavage by exhibiting a catalytic activity.

Recent studies have discovered widespread regulatory RNA elements, also referred to as riboregulators, in natural systems such as prokaryotes (Barrick *et. al.* 2004) as well as eukaryotes and even in humans (Yelin *et. al.* 2003). In addition, researchers have been able to synthetically engineer RNA elements with regulatory functions, also called riboregulators, to modulate gene expression. The discovery of aptamers, artificial nucleic acid molecules capable of binding to a specific ligand, has allowed scientists to design riboregulators with an allosteric control. This chapter will describe nucleic acid aptamers, and different classes of riboregulators along with examples of their structure and function.
1.2 In Vitro Selection and Nucleic Acid Aptamers

1.2.1 In vitro Selection

In vitro selection procedures can isolate nucleic acid molecules, known as aptamers, which bind ligands with a high affinity and selectivity comparable to highly evolved protein molecules (Ellington et al. 1990). An in vitro selection approach, referred to as SELEX, has led to the discovery of nucleic acid aptamers capable of selectively binding to a wide variety of target ligand molecules. The target ligands used for aptamer selection have included a wide range of sizes including small ion (Ciesiolka et al. 1996), small molecules (Ellington et al. 1990), peptides (Nieuwlandt et al. 1995) and single proteins (Tuerk et al. 1990), up to macroscopic entities like organelles (Ringquist et al. 1995), viruses (Pan et al. 1995), and even entire cells (Morris et al. 1998).

In vitro selection or SELEX (Systematic Evolution of Ligands by Exponential enrichment) is an approach to select a specific nucleic acid molecule with a specific property from large, random library of molecules. SELEX has been used to generate DNA and RNA aptamer with various functions (Ellington et al. 1990; Tuerk et al. 1990). This in vitro selection process is composed of three important processes. First, it requires introduction of genetic variations or mutations to create a random library of sequences. Second, it requires selection of variant molecules best suited for a particular process and lastly, the selected molecules are amplified.
Figure 1.1  The cycles of *in vitro* selection of RNA

The general procedure for *in vitro* selection of RNA aptamers is shown in Figure 1.1. The process starts with a large population of DNA template molecules. The starting library of DNA templates is chemically synthesized with a region of random or mutagenized sequences flanked on each end by constant sequences for primer binding and PCR amplification, and with a T7 RNA polymerase promoter sequence at the 5’-end of the DNA template strand. The DNA templates are transcribed into RNAs by T7 RNA polymerase. From a pool of RNA molecules, a fraction of RNA molecules are selected based on their ability to carry out specific function. The selected RNA molecules are copied into cDNAs by reverse transcriptase. PCR is used to amplify the cDNAs, and then the entire cycle is repeated. After successful rounds of selection, the resulting population of RNAs must be able to carry out a specific function more efficiently than the starting pool of RNAs. This resulting population of selected RNAs are cloned and
sequenced. In the end, representative number of sequences are tested for its desired function. DNA aptamers are selected in a similar manner except the DNA templates are not transcribed into RNA and the step involving reverse transcription of the selected molecules is omitted.

The diversity of the initial random library of sequences is an important factor for a successful in vitro selection (Osborne and Ellington 1997). The number of nucleotides (n) in the region of random sequences determines the total number of possible unique nucleic acid molecules from a random library. The theoretical number of the unique sequences is calculated by $4^n$. The size of a random sequence region for a given selection depends on desired functional diversity. A large random sequence region, for example $4^{80}$, would result in the selection of only a very small fraction of the total possible sequences. On the other hand, a small random sequence region (e.g. $4^{20}$) will limit the complexity of each unique sequence. The length of the oligonucleotides is another feature to be considered when designing a random library of oligonucleotides since synthesis of oligos longer than 150 nucleotides is not practical. Based on the above-mentioned criteria, an in vitro selection is typically carried out using a 40-nucleotide random sequence region. The total possible number of unique sequences from this random region is about $10^{24}$, and a library of $10^{14}$~$10^{15}$ unique sequences are typically generated (Osborne and Ellington 1997).

Another important factor to select the active oligos during in vitro selection procedures is the selective replication and amplification of these oligos. Depending on the activity of the selected molecule, there are different approaches to selectively replicate these molecules. If the oligos being selected have catalytic cleaving activity, gel
electrophoresis is used to physically separate the smaller self-cleaved molecules from the larger inactive molecules. In other cases if the active molecules are able to catalyze a reaction by covalently adding a sequence required for replication, then only the active molecules will be replicated and there is no need for physical separation of inactive molecules. In the third case, if the active molecule being selected for has a ligand binding activity, immobilized ligands are used to separate the active molecules from those that do not bind by a procedure similar to affinity chromatography. Molecules with binding affinity to the immobilized ligands are eluted and then replicated.

1.2.2 DNA and RNA Aptamers

DNA and RNA molecules with specific ligand binding affinity are known as aptamers. The functional characteristics of aptamers and their ability to selectively bind to target ligands make them similar to antibodies (Jayasena 1999). Aptamers are able to recognize and bind to their cognate ligand at target molecule concentrations ranging from micromolar to picomolar. Aptamers undergo structural changes upon ligand binding, which could be translated into and induce other functional activities.

These aptamers can be generated by in vitro selection and manipulated for clinical or diagnostic purposes (Breaker 2004). In fact the potential applications of aptamers in research studies are enormous. Immobilized aptamers on solid matrix could be used as chromatographic agents of in vitro studies (Romig et. al. 1999). Fluorescent tagging of such immobilized aptamers can be used to design biosensors for detecting specific proteins (Hamaguchi et. al. 2001) or small molecules (Jhaveri et. al. 2000) by monitoring changes in the fluorescent intensity of the reporter tag, which is induced by the structural rearrangement of the aptamer after target binding.
Aptamers could also have in vivo applications since they can be delivered into organisms or expressed in cells. They could be used for designing molecular switches by blocking protein synthesis (Suess et. al. 2003), or they can be incorporated in the design of allosteric ribozymes, catalytic RNA enzymes, to control their on and off activity (Mandal and Breaker 2004).

As previously described, a variety of aptamers have been discovered by in vitro selection to have strong affinity and specificity to a vast diversity of target ligands ranging in size, from small ions (Ciesiolka et. al. 1996) to proteins (Tuerk et. al. 1990), and even entire cells (Morris et. al. 1998). For example, RNA aptamers were selected with moderate binding affinities of 100-400 μM to free Zn\(^{2+}\) (Ciesiolka et. al. 1996). In a different study Bock and co-workers isolated single stranded DNA aptamers that bind, with high binding affinities ranging between 25-400 nM, and inhibit thrombin, an essential protein for blood clotting (Bock et. al. 1992). Binding of the 15-nucleotide consensus sequence of the thrombin aptamer to this protein increased the clotting time by seven-fold. Proteins can be used as targets for aptamer selection because the extensive grooves, ridges, projections and depressions on the surface of proteins could offer numerous sites for hydrogen bond donors and acceptors. Single-stranded RNA aptamers have been isolated for binding to proteins as well. Gold and co-workers (Tuerk et. al. 1992) used in vitro selection to isolated RNA aptamers capable of selectively binding to HIV reverse transcriptase (HIV-RT), with dissociation constants at about 50 nM, and inhibit its reverse transcriptase activity. Morris and co-workers (Morris et. al. 1998) also used In vitro selection to select multiple single stranded DNA aptamers with high affinities, in the nanomolar range, for binding to human red blood cell membrane.
The dissociation constant, $K_d$, is an equilibrium constant that represents the tendency of reversible dissociation of the aptamer-ligand complex into an aptamer and a ligand molecule (Mathews et. al. 2000). In other words, the dissociation constant is used to describe the affinity between a ligand and an aptamer in terms of how easy it is to separate the aptamer-ligand complex into two individual components. The dissociation constant is defined as $K_d = [\text{ligand}][\text{aptamer}]/[\text{ligand.aptamer}]$. From this equation, a high $K_d$ value is indicative of a weak binding strength between the ligand and the aptamer. On the other hand, a low $K_d$ represents a stronger binding affinity between the aptamer and the ligand to form a complex.

Single-stranded DNA and RNA aptamers capable of binding to small molecules have been isolated by in vitro selection as well. Ellington and Szostak (Ellington and Szostak 1990) isolated RNA and DNA aptamers (Ellington and Szostak 1992) for selective binding to a variety of organic dye molecules with hydrogen bond donor and acceptor groups, and an overall negative charge. Examples of other RNA aptamers isolated for binding to small biomolecules with strong affinities are the ATP-binding RNA aptamer (Sassanfar and Szostak 1993), with $K_d$ of about 1\(\mu\)M and ATP-binding DNA aptamer (Huizenga and Szostak 1995) with $K_d$ value of about 6\(\mu\)M.

There are certain features that make some ligands more optimal for binding to an aptamer during the selection process. Generally planar ligand molecules, and those with hydrogen bond donor and acceptor sites and positively charged groups are ideal for recognition and binding to aptamers (Wilson and Szostak 1999). On the other hand, ligands with a non-planar structure and an overall negative or neutral charge are difficult to be recognized by aptamers. In addition, purine-rich loops in simple aptamer structures
tend to be involved with ligand recognition because these bases can associate in noncanonical base-pairing interactions to properly fold into a structure capable of ligand interaction. If DNA aptamers are converted into RNA and vice versa, in most cases they no longer bind to their cognate ligands. This is contributed by the presence of the 2’-hydroxyl groups in RNA and its role in helix stability and the overall tertiary structure stabilization of aptamers.

The enormous application of aptamers for *in vitro* and *in vivo* studies makes them an ideal tool to be used in diagnostic and therapeutic research. Initially, aptamers targeted to VEGF found utility in the treatment of wet macular degeneration (Ng E.W. *et al.* 2006). Aptamers that have properties more suitable for systemic administration are now being developed (Keef A.D. *et al.* 2010), primarily against targets in the bloodstream, such as thrombin, factor IXa, or on cell surfaces such as epidermal growth factor receptor (EGFR).

### 1.3 Riboregulators:

#### 1.3.1 Antisense RNA

Antisense RNA elements are single-stranded RNA molecules, ~50-300 nucleotides long, that specifically bind to their target RNA (sense RNAs) generally in *trans* based on Watson-Crick base pairing between complementary sequences. They are diffusible molecules, mostly untranslated and highly structured with one to four stem-loops, that upon binding to their target RNA could affect gene expression regulation (Brantl 2007). The first naturally occurring antisense RNA was discovered in plasmid ColE1 in 1981 (Tomizawa *et al.* 1981). Ever since they have been found in both
prokaryotes and eukaryotes. In general antisense RNAs are classified into two classes based on their origin of expression: first, cis-encoded RNAs are located in the same DNA region and are fully complementary to their target RNA over a long sequence stretch. Second, trans-encoded RNAs are located in another chromosomal location than their complementary target RNA and are only partially complementary to each other (Brantl 2007). Unlike cis-encoded RNAs, trans-encoded RNAs are diffusible elements. The studies on the mechanism of action of these two classes of antisense RNAs in eukaryotic and bacterial systems point to a surprising number of similarities.

**Cis-encoded Antisense RNAs**

The first evidence of cis-encoded antisense RNAs was discovered in bacterial plasmids, where they were found to control the expression of genes required for plasmid replication and inheritance (Storz et al. 2005). They were later found in transposons and phages as well. In addition to plasmid cis-encoded antisense RNAs, chromosomal cis-encoded antisense RNAs function in bacterial and eukaryotic cells to modulate expression of protein-coding genes.

The cis-encoded antisense RNAs achieve transcriptional inhibition through transcription attenuation. In prokaryotes, binding of antisense RNA to the target site induces a terminator stem-loop structure formation in the nascent mRNA and consequently, transcription is terminated prematurely upstream of the Shine-Dalgarno (SD) sequence (Brantl 2007). Antisense RNAs could also regulate gene expression through translation inhibition. In this mechanism, the antisense RNA directly blocks the ribosome binding site (RBS) and ribosome loading. Antisense RNAs could also influence mRNA stability to promote or inhibit mRNA degradation.
Plasmid-Encoded Antisense RNAs: The plasmid-encoded antisense RNAs function by base-pairing with their target mRNA and by doing so they influence transcription, stability or translation of the target mRNA and consequently regulate expression of genes involved in replication or inheritance of the plasmid. The plasmid-encoded antisense RNAs are transcribed constitutively, however their level of synthesis could be dependent on the copy number and plasmid concentrations inside a cell (Storz et. al. 2005). An example of this metabolically unstable system exists in the case of RNAI (~85 nt) and RNAII (~150 nt) antisense RNAs transcribed from the pT181 plasmid. These RNAs regulate RepC protein expression, a protein necessary for replication initiation, by promoting transcriptional attenuation. RNAI and RNAII start at position 246 on pT181 plasmid and base pair upstream of the nascent repC-mRNA, at position 339 and 413 respectively (Kumar et. al 1985), and stabilize a secondary structure, a terminator stem-loop, associated with transcription termination (Brantl et al. 2000). In this system, a lower plasmid copy number leads to decreased levels of RNAI and RNAII, which allows for higher repC-mRNA transcription and higher RepC levels and restores plasmid replication (Storz et al. 2005). On the other hand, a high plasmid copy number would cause decreased RepC levels and plasmid replication.
Figure 1.2 Transcriptional attenuation by antisense RNA in pT181 plasmid

Complementary sequences are designated as A, B, a and b. (Figure reproduced from Brantl et al. 2000)

Another system in which plasmid-encoded antisense RNAs could exert an influence on the condition of the cell is referred to as the plasmid addiction system. In this system the survival of the cell relies on the presence of the antisense-containing plasmid, and the absence of the plasmid confers lethality. A model example of this toxin-antitoxin system is the hok-sok gene of bacterial plasmid R1 (Gerdes et al. 1997). The hok (host killing) gene encodes a protein responsible for bacterial cell membrane damage, which causes cell death. The sok RNA (suppressor of killer) is also encoded within the plasmid R1 sequence and upon expression the sok antisense RNA inhibits hok mRNA translation. The hok mRNA levels (a toxin) are more stable than the sok RNA (an antidote) levels, and when plasmid R1 is lost, this differential stability causes sok RNA
levels to drop faster than the hok mRNA levels (Gerdes et al. 1997). The loss of inhibitory influence of sok RNA leads to Hok protein expression and consequently cell death.

Prokaryotic chromosomally Encoded Antisense RNAs: These antisense RNAs function similarly to plasmid-encoded antisense RNAs, however they are expressed from bacterial genome. There are fewer known example of chromosomally, cis-encoded RNAs, and they seem to be homologous to the hok-sok loci found on plasmids as already described. In addition this toxin-antitoxin role of bacterial cis-encoded antisense RNAs, they have been implicated with increased gene expression through transcript stabilization. For example 3’ end of GadY noncoding mRNA (105nt) in E.coli base pairs with the 3’-untranslated region (3’ UTR) end of gadX mRNA. This base pairing results in increased stability of gadX mRNA and therefore its expression level increases (Opdyke et al. 2004).

**Trans-encoded Antisense RNAs**

*Trans*-encoded antisense RNAs been found to influence gene expression levels in both bacteria and eukaryotes. They exert their function to destabilize mRNAs or to inhibit or activate translation. They are encoded at a distinct chromosomal site than their target RNA sequence. Unlike *cis*-encoded antisense RNA, they do not require perfect base pairing with their target sequence in order to regulate gene expression. Characterization studies of trans-encoded antisense RNAs in *E.coli* has led to the discovery of a RNA chaperon protein, called Hfq, which binds to these regulatory RNA elements and promotes base pairing between the antisense RNA and its target mRNA. Hfq also protects the antisense RNA fragments or the mRNA from RNases by inducing structural changes in these RNAs. Moreover, Hfq blocks RNaseE ribonuclease and by
doing so the antisense RNA and its target mRNA are protected against ribonuclease digestion. The eukaryotic homolog of Hfq is called the Sm and Sm-like proteins, which are a component of the RNA splicing and degradation complexes. Similar to Hfq, these eukaryotic proteins could also promote base pairing between antisense RNA and its target and induce changes in their structures.

1.3.2 MicroRNAs (miRNAs)

miRNAs are short, ~22 nucleotide regulatory endogenous, noncoding RNAs, first discovered in *Caenorhabditis elegans* (Lee *et al.* 1993). Further studies have identified miRNAs in plants and animals (Bartel 2004, Pillai *et al.* 2005) and more than 222 human microRNAs (Bentwich *et al.* 2005). Binding of miRNAs to their target complementary sequence in the 3’ untranslated region (3’UTR) of messenger RNAs (mRNAs) exerts the regulatory function of miRNAs in gene expression at post-transcriptional level (Obernosterer *et al.* 2006). This sequence specific binding of miRNAs leads to mRNA cleavage and mediates translational repression (Pillai 2005).

Figure 1.3 (Esquela-Kerscher A. and Slack F.J. 2006) shows miRNA biogenesis and its mode of function inside a cell to repress mRNA expression. miRNAs are initially transcribed as long, mono- or polycistronic precursor RNAs (pri-miRNAs) (Obernosterer *et al.* 2006). These long transcripts are further processed by a nuclear RNase III-like enzyme called Drosha, which cleaves pri-miRNAs into ~70 nucleotides-long precursor hairpins, known as pre-miRNAs (Lee *et al.* 2003). Further maturation of pre-miRNA hairpins takes place in the cytoplasm once they are exported out of the nucleus by Exportin-5 (Yi *et al.* 2003; Bohnsack *et al.* 2004; Lund *et al.* 2004). A second RNase III-like enzyme in the cytoplasm, termed Dicer, cleaves pre-miRNA hairpins into ~22-nt
duplexes (Bernstein et al. 2001; Hutvágner et al. 2001; Ketting et al. 2001; Knight & Bass 2001). Finally one of the two strands of the duplex accumulates asymmetrically as the mature miRNA in accordance with the thermodynamic asymmetry rule. The asymmetry rule is based on the thermodynamics of helices, and states that the strand in a duplex whose 5’-end lies at the end of the duplex with the lower thermodynamic energy is the mature miRNA strand (Tomari and Zamore 2005). The mature miRNA strand is then assembled into multiprotein effector complexes called miRNPs (miRNA-containing ribonucleo-protein particles), which are similar to multi-protein complexes in siRNA assembly called RNA-induced silencing complex (RISC) (Khvorova et al. 2003; Schwarz et al. 2003). After the miRNP assembly is complete, the miRNA directs the complex to its target by base pairing with its target mRNA and the target mRNA is either degraded or its expression into protein is inhibited (Pillai 205).
Figure 1.3 microRNA biogenesis

MicroRNA (miRNA) are transcribed in the cell nucleus by RNA Polymerase II (Pol II). The transcripts are large molecules called pri-miRNA and have a hairpin-like structure (the mature miRNA sequence is shown in red). These large pri-miRNA transcripts are processed into smaller ~70-nucleotide pre-miRNA precursor product by the RNase III enzyme Drosa and its co-factor, Pasha. The pre-miRNA is transported into the cytoplasm by RAN–GTP and exportin 5 proteins. The pre-miRNA is further processed in the cytoplasm by another RNase III enzyme, Dicer, to generate a transient ~22-nucleotide miRNA: miRNA* duplex (*marks the mature miRNA sequence). This duplex is then loaded into a multi-protein complex called the miRNA-associated multi-protein RNA-induced silencing complex (miRISC) (light blue). The miRISC complex includes the Argonaute proteins, and the mature single-stranded miRNA (red). The mature miRNA then targets and base pairs with its complementary mRNA sequence to inhibit gene expression. This inhibition is caused in one of two ways depending on perfect or imperfect complementary base pairing with the target mRNA. Imperfect base pairing between the miRNA and its target sequence inhibits protein translation and hence protein expression is blocked (lower left). Perfect complementarities between miRNA and its mRNA targets, mainly found in the coding sequence or open reading frame (ORF) of the target, induce target-mRNA cleavage (lower right). (Adapted with permission from Macmillan Publishers Ltd: Nature Reviews Cancer, Esquela-Kerscher A. and Slack F.J. copyright © 2006)
1.3.3 Small interfering RNA (siRNA)

Small interfering RNAs (siRNA) are another category of small, noncoding RNAs that regulate gene expression. They are known to be involved in double-strand RNA (dsRNA)-induced post-transcriptional gene silencing (PTGS) also known as RNA interference (RNAi) in plants, animals, and fungi (Kawasaki et al. 2005).

siRNAs are generated from long double-stranded RNAs that are processed by dicer enzyme into smaller, ~21-23 nucleotide, double-stranded RNAs with 2-nucleotide 3’-overhangs on either end (Figure 1.4). The dicer enzyme is one of the two groups of protein enzymes required by small RNAs to be able to exert their regulatory functions (Carthew R.W. and Sontheimer E. J. 2009). Dicers bind to RNA duplex ends with short (~2 nt) 3’ overhangs through their RNA-duplex binding domain, PAZ domain. A second domain, RNase III domain, then excises long double-stranded RNAs or small hairpin RNAs to generate new ends with ~2 nucleotide 3’-overhangs (Carthew R.W. and Sontheimer E. J. 2009).

A second group of protein enzymes that support siRNA silencing effector function are called Argonaute proteins (Tomari et al. 2005). Argonaut protein is one of the major component proteins of the RNA-induced silencing complex (RISC). The double-stranded processed siRNA is mounted into the RISC complex and is unwound. Only one of the two single strands, referred to as the guide strand, will associate with the Ago protein of RISC (Carthew R.W. and Sontheimer E. J. 2009), while the other strand is discarded. The Ago protein directs the guide strand to bind to its target mRNA sequence by Watson-Crick base pairing (Carthew R.W. and Sontheimer E. J. 2009). The Ago
protein through its RNase H-like activity catalyzes guide strand-dependent
endonucleolytic cleavage of the target strand, which consequently leads to RNA
SiRNAs can be generated by Dicer protein from several different transcript categories with double-stranded RNA structures. After processing by Dicer, the resulting siRNA molecule consists of two strands: a guide strand (red), which assembles into functional siRISC complex, containing Ago protein and additional protein factors, and a passenger strand (blue), which is ejected from siRISC and degraded. The guide strand within the siRISC then recognizes and base pairs with its target RNAs, and silencing ensues through one of several mechanisms. It is also possible for the siRNA duplexes to be amplified by RNA-dependent RNA polymerase (RdRP) enzymes. This amplification process further strengthens and preserves the gene silencing response. (Adapted from Cell, Vol. 136, Carthew R.W. and Sontheimer E. J. Copyright (2009) with permission from Elsevier)
1.3.4 Piwi-interacting RNAs (piRNAs):

A third category of small RNAs with gene regulatory function is referred to as Piwi-interacting RNAs (piRNAs). While siRNA and miRNA are found in various organisms and are generated from their double-stranded precursors, piRNAs have been mainly found in animals and are characterized by the single-stranded nature of their precursors. Presence of piRNAs has been detected in the germline, for example, in zebrafish and Drosophila ovaries, as well as in mammalian testes (Choudhuri S. 2009). Genes encoding piRNAs are found in clusters in the germline genome. Each cluster has a size between 1 to 100 kb and could encode 10 to 4500 piRNAs. piRNAs are encoded from one of the two DNA strands within a cluster. The processing enzymes involved in piRNA biogenesis from their precursor single-stranded RNA are yet to be identified, however it is known that piRNA biogenesis is Dicer independent since the precursor piRNAs are single-stranded RNAs (Choudhuri S. 2009). piRNAs are ~24-30 nucleotide single-stranded RNAs, with a 5’ uridine bias and a 2’-O-methylated 3’-end. In mammals piRNAs have been mapped to repeat sequences, and several classes of DNA transposons (O’Donnel K.A. and Boeke J.D. 2007).

piRNAs function in the germline by associating with the Piwi clade of Argonaute proteins (Malone C.D., Hannon G.J. 2009). Numerous studies have pointed that piRNAs, through binding to Piwi proteins, control the expression of transposon elements (Malone C.D., Hannon G.J. 2009) and this function has led to their recognition as the guardians of the germline genome. For example, mutations in piwi gene allows for gypsy retrotransposons mobilization in animals (Sarot et al. 2004). In Drosophila, three Piwi proteins found in the germline cells are required for piRNA function. They are called
Aub (Aubergine), Piwi, and Ago3. Homologs of these three proteins exist in other animals. In the germ cells, the action of Piwi proteins forms a cycle, which amplifies piRNAs that target active transposon transcripts (Figure 1.3). This amplification model is referred to as the “ping-pong model” (Choudhuri S. 2009). In this model, antisense piRNAs associate with Piwi/Aub to target and cleave the sense transcripts. The cleaved sense transcripts generate sense piRNAs that associate with Ago3 protein. Next, the sense piRNA-Ago3 complex targets antisense transcripts cleavage, creating antisense piRNAs. This cycle continues as antisense piRNA associates with Piwi/Aub again to cleave more sense target transcripts and further generate sense piRNAs. This ping-pong model serves to self-amplify the piRNAs (Choudhuri S. 2009).
Figure 1.5  The ping-pong model of piRNA action for transposon mRNA cleavage.

Sense piRNAs are generated from precursor sense transcripts cleaved by Piwi/Aub. The sense piRNA transcripts then associate with Ago3 to direct the cleavage of the antisense transcripts. This cleaving process generates the antisense piRNAs. These antisense piRNAs in turn bind to Piwi/Aub to form the Piwi/Aub-antisense piRNA complex. This complex then leads to the cleavage of more sense target transcripts and further generates sense piRNAs. Repeated cycles of these cleavage steps forms a self-amplifying loop to control the expression of transposon elements. (Adapted from Biochem Biophys. Res. Commun., Vol. 388, Choudhuri S., Copyright (2009) with permission from Elsevier).

1.3.5  Riboswitches

1.3.5.1  Natural riboswitches:

Riboswitches are regulatory elements of gene expression that are found at the 5’-untranslated region (5’-UTR) of messenger RNA (mRNA). Riboswitch structure is composed of an aptamer domain, a short nucleic acid sequence which binds to a specific metabolite or small ligand molecule, and an expression platform domain which undergoes secondary structure rearrangement in response to aptamer-ligand binding.
Riboswitches can sense diverse metabolites such as guanine, flavin mononucleotide (FMN), lysine, and thiamine pyrophosphate (TPP). The aptamer domain, which is RNA sequence domain, is a highly folded structure that is able to recognize and specifically bind to its cognate metabolite by hydrogen binding or base stacking mechanisms. These aptamers could be found naturally in the mRNA sequence or it could be synthetically selected by in vitro techniques and incorporated into the riboswitch structure. The RNA sequence of the expression platform domain of riboswitches undergoes a secondary structure rearrangement in response to the metabolite binding to the aptamer domain. The aptamer domain for a particular metabolite is conserved among divergent organisms, suggesting they have persisted through evolution. In contrast, the expression platform domain could have different sequences and structures among riboswitches of the same class, as well as among different riboswitch classes (Tucker B.J. and Breaker R.R 2005).

Riboswitches can be used as a form of feedback inhibition by their ability to control expression of proteins involved in the biosynthesis or transport of the metabolite they are able to sense. They exert their gene expression regulatory role by altering transcription termination to prevent full-length mRNA synthesis, or by inhibiting initiation of translation of mRNA (Figure 1.6). For transcription termination, metabolite binding to the aptamer promotes a stem structure formation in the expression platform domain that serves as an intrinsic transcription terminator. In the absence of the metabolite a competing anti-terminator structure is formed which allows transcription to proceed. In the case of translation initiation control, aptamer-metabolite binding prevents translation by blocking ribosome access to its binding site on the mRNA or the Shine-Dalgarno sequence (Tucker B.J. and Breaker R.R 2005). Riboswitches that function to
repress gene expression in response to metabolite binding are referred to as OFF switches, for example the theophylline riboswitch.

![Figure 1.6 Common mechanism of riboswitch gene expression control.](image)

Riboswitches achieve transcription control by the stabilization of a specific conformation of the aptamer domain after metabolite binding. The stabilization of the aptamer domain in turn prevents the formation of a competing anti-terminator stem loop. As a result, a terminator stem is allowed to form which leads to transcription termination. Riboswitches could also inhibit translation by inducing a structural change in the mRNA due to metabolite binding to the aptamer domain. This secondary structure formation sequesters the ribosome-binding site (RBS), and therefore prevents ribosome binding to the mRNA and blocks translation. (Adapted from Curr. Opin. Struct. Biol., Vol. 15, Tucker B.J. and Breaker R.R., Copyright (2005) with permission from Elsevier)

As more studies of riboswitch structure and mechanism have been carried out, other riboswitches with differences to the prototypic riboswitch mechanisms described above have been discovered. For example the adenine riboswitch, as shown in Figure 1.7A (Mandal et al. 2004), promotes activation of gene expression in response to the presence of adenine. In this system the transcription terminator stem is allowed to form when the aptamer is not bound to its ligand. This type of riboswitch that activates gene expression in response to the presence of their cognate ligand is referred to as the ON
riboswitch. This is an opposite mechanism to the ‘OFF’ riboswitch, in which the presence of the ligand stabilizes the aptamer and induces the terminator stem to form and inhibits transcription. Bioinformatics searches have also identified additional riboswitches based on sequence similarities and conserved RNA domains in organisms intergenic regions. For example the glmS riboswitch (Figure 1.7B) is always found adjacent to the \textit{glmS} gene, which encodes glutamine-fructose-6-phosphate aminotransferase. This enzyme produces glucosamine-6-phosphate (GlcN6P), the metabolite that triggers the function of the glmS riboswitch. GlcN6P binding to the aptamer domain of the glmS riboswitch stabilizes the secondary structure of the expression platform domain of this riboswitch and allows it to carry out its self-cleaving ribozyme activity. The RNA is cleaved upstream of the \textit{glmS} gene, which leads to down-regulation of this gene. The decrease in gene expression is believed to be related to the cleaving activity of this riboswitch. The truncated mRNA products are eventually targeted for degradation by RNases.
1.3.5.2 Synthetic Riboswitches:

**Aptamer-based molecular switches**

Since riboswitches have shown to be able to control gene expression, scientists have been interested to generate synthetic riboswitches to develop artificial RNA regulators for conditional gene expression systems. To achieve this purpose engineered riboswitch modules are inserted into sensitive regions of the mRNA. Other than natural aptamers, the aptamer domain of the engineered riboswitch could be a synthetic, *in-vitro*
selected RNA molecule with high affinity and specificity for a respective ligand. Researchers have used an \textit{in vitro} selection technique, SELEX (Systemic Evolution of Ligands by Exponential Enrichment) to generate synthetic aptamers to incorporate into a riboswitch design. As described earlier, this aptamer selection technique involves many repeating rounds of \textit{in-vitro} selection and enrichment of RNA sequences that bind to a specific ligand in a column. This technique was first described in 1990 (Ellington, A.D. and Szostak, J.W. 1990) and researchers have used this approach to select different aptamers for differing molecules ranging from large peptide proteins, complex drug molecules, to small organic molecules or even metal ions (Stoltenburg R. \textit{et. al.} 2007).

To construct a synthetic riboswitch to control gene expression, aptamer sequences have been inserted into the 5'UTR of a reporter mRNA as shown in Figure 1.8. Ligand binding to the aptamer causes the aptamer to adapt a unique conformation in the 5'UTR, which can interfere with initiation of translation either by hindering successful scanning or binding of the ribosomal subunits (Bauer G. and Suess B. 2006). In such a system aptamers act as molecular switches, with the ligand as an integral part of the complex, to turn off the expression of a downstream gene (Bauer G. and Suess B. 2006).
Figure 1.8 Synthetic aptamer-based riboswitch mechanism to control gene expression.

The aptamer is inserted in the 5′-untranslation region (5′-UTR) of a eukaryotic mRNA. The addition of the ligand induces the formation of a ligand–aptamer complex, which interferes with translational initiation either by (A) hindering ribosomal scanning or (B) inhibiting binding of the ribosomal subunits. The mode of control depends on the position of the aptamer within the 5′-UTR. (Adapted from Journal of Biotechnology, Vol. 124, Bauer G. and Suess B., Copyright (2006) with permission from Elsevier)

The first example of engineered riboswitches was designed by Werstuck and Green (Werstuck G. and Green M.R., 1998), who selected an aptamer for Hoechst dye H33258, and inserted this aptamer into the 5′UTR of a mammalian β-galactosidase expression plasmid. Chinese hamster ovary cells were transfected with the β-galactosidase expression plasmid and analyzed for β-galactosidase activity in the presence or absence of the drug ligand. The presence of the Hoechst ligand in this system led to 10-fold decrease in reporter gene in the cells. Another example of this kind of riboswitch activity was reported in yeast cells (Grate D. and Wilson C. 2001). They inserted an in vitro selected malachite green aptamer motif into the 5′UTR of CLB2 cyclin gene, which encodes a B-type cyclin protein responsible for directing the cell cycle transition from G2 to mitosis in budding yeast (Grate D. and Wilson C. 2001). The aptamer-ligand complex formation in this system reduced CLB2 translation by 10-fold,
and therefore, cell cycle progression was slowed down and yeast cell morphology was affected. In this study Grate and Wilson’s results confirm the role of synthetic riboswitches in altering cellular processes and shed light onto the potentials of riboswitches as genetics tools for manipulation and understanding of cellular pathways (Bauer G. and Suess B. 2006).

In another study carried out by Suess group (Suess et. al. 2003) they looked more closely into the mechanism with which tetracycline-binding aptamer (Berens et. al. 2001) inserted into the 5’UTR of several reporter genes in yeast was inhibiting translation. In their following studies (Hanson et. al. 2003) they inserted the tetracycline aptamer motif immediately downstream of the cap structure in the 5’UTR and studied the mode of function of this riboswitch in controlling reporter gene expression in yeast cells. Their finding suggests the stem loop structure formed because of tetracycline binding to the aptamer motif blocks small ribosomal subunit binding to the cap motif (Hanson et. al. 2003). They also tested the effect of insertion of the aptamer motif distal to the cap motif and observed slightly higher level of inhibition in reporter gene expression (Hanson et. al. 2003). These findings suggested the formation of stem-loop structure due to aptamer-tetracycline bindings, regardless of its position relative to the cap structure, was perhaps blocking ribosome scanning and therefore inhibiting the formation of the 80S ribosome (Bauer G. and Suess B. 2006). Similar studies in higher eukaryotic systems (Harvey et. al 2002) involving different aptamer motifs, such theophylline-aptamer (Jenison et al. 1994) in Xenopus oocytes, produced contradictory results to those obtained in yeast. In higher eukaryotes cap proximal stem-loop motifs produced more efficient regulation in
gene expression (Kozak M. 1986), whereas in yeast riboswitch-controlled gene regulation was independent of the position of the aptamer motif (Bauer G. and Suess B. 2006).

**Antisense based riboswitches**

As described earlier, natural riboswitches are able to regulate gene expression at a post-transcriptional level by sequestering the ribosomal binding site within a stem loop structure of a riboswitch (Tucker B.J. and Breaker R.R 2005). In this system the riboswitch is positioned in *cis* relative to the target gene. By employing the mode of function of such natural riboswitches, scientists have integrated synthetic aptamer sequences within *cis*-acting riboswitch elements. This *cis*-element upstream of the Shine-Dalgarno (SD) sequence forms a stem-loop structure, which in turn interferes with translation initiation in a prokaryotic system and causes down-regulation of a reporter gene expression. To reverse the induced translational repression in this system, researchers (Issac *et. al.* 2004) designed and expressed a *trans*-acting antisense RNA that forms complementary base-pairing to one arm of the original sequestering helix and destabilizes the stem-loop that obstructs the ribosomal binding site (Figure 1.9). The activity of such antisense RNA would in turn lead to de-repression of the reporter gene.
Figure 1.9 Antisense-based riboswitch function in a prokaryotic system.

The *cis*-element upstream of the Shine Dalgarno sequence forms a stem-loop structure. The SD sequence is sequestered within this structure, preventing ribosome from binding to SD sequence and gene expression is inhibited. To reverse the translational repression a synthetic antisense RNA molecule forms complementary base pairing with the *cis*-element, which destabilizes the stem-loop structure formation. The resulting free SD sequence promotes translation. (Adapted from Journal of Biotechnology, Vol. 124, Bauer G. and Suess B., Copyright (2006) with permission from Elsevier)

In a study done by Smolke group (Bayer T.S., Smolke C.D. 2005), the antisense riboswitch design was further enhanced (Bauer G. and Suess B. 2006). They designed a trans-acting riboswitch by combining an aptamer domain that binds to a specific effector molecule with a sequence recognition domain (antisense domain) complementary to a target mRNA as shown in Figure 1.10 (Bauer G. and Suess B. 2006). They referred to this regulator RNA module as an antiswitch (Bayer T.S., Smolke C.D. 2005) which is defined as an antisense RNA capable of blocking ribosomal translation by base pairing with the target mRNA at the 5′ untranslated region. The antiswitch secondary structure stabilizes reversibly between two forms based on theophylline ligand sensing by the aptamer domain (Bayer T.S., Smolke C.D. 2005). In the absence of theophylline, the antisense sequence is sequestered in a stem-loop structure, allowing the target GFP mRNA to be translated. Upon theophylline binding to the aptamer, the aptamer stem-
loop becomes more energetically stable than the antisense stem-loop thereby allowing the antisense sequence to readily base-pair at the 5'UTR of the reporter GFP gene and block ribosomal access to the AUG start codon (Bayer T.S., Smolke C.D. 2005). This antiswitch can function as an “on” or “off” switch by sequestering or releasing the antisense domain in the absence or presence of the theophylline ligand, respectively. As shown in Figure 1.10A upon ligand binding to the aptamer domain, the antiswitch acts as an OFF-switch that, through complementary base pairing of the antisense domain with the target mRNA, is able to repress the target GFP mRNA translation. On the other hand, in the absence of ligand, the antiswitch serves as an ON-switch since the antisense sequence domain is sequestered in a thermodynamically stable stem loop structure and GFP expression is upregulated. Theophylline aptamer-incorporating antiswitch (containing an anti-GFP antisense sequence) was expressed efficiently in yeast cells by Smolke group and they were able to show GFP expression with exquisite response to theophylline levels in the growth media (Bayer T.S., Smolke C.D. 2005).
Antiswitch structure is composed of an antisense sequence domain and an aptamer domain. Aptamer-ligand binding causes the secondary structure of the antiswitch to alternate between two forms. (A) In the case of the “ON” switch and in the absence of the ligand, the antisense sequence cannot bind to the target mRNA since it is sequestered in a stem-loop structure. The mRNA is readily scanned by the ribosome and mRNA translation takes place. Ligand binding to the aptamer induces an aptamer stem-loop structure to form, allowing the antisense sequence to be free to base pair with the target mRNA and block ribosomal scanning and translation. (B) Alternatively for the “OFF” switch design, in the absence of the ligand, the antisense sequence is free to base pair with the target mRNA sequence and suppresses gene expression. In this model, ligand binding to the aptamer sequesters the antisense domain within a secondary structure, permitting target gene expression. (Adapted from Journal of Biotechnology, Vol. 124, Bauer G. and Suess B., Copyright (2006) with permission from Elsevier)

The antiswitch mechanism proposed by Smolke group could be redesigned to act as an “OFF” switch in the absence of the ligand to act as a block to ribosomal translation initiation. In such a system as shown in Figure 1.8B, presence of the ligand allows translation initiation to proceed without antiswitch interfering with the ribosomal scanning and the antiswitch act as an “ON” switch.
1.4 Thesis Overview:

The technique proposed by Smolke group for allosteric control of gene expression, by sensing and responding to intracellular metabolite levels, provides a tool for analyzing cellular pathways. However, controlling gene expression reversibly in temporal and tissue specific manner could provide even more control for a researcher to influence intracellular gene expression levels. Such a system provides researchers with a tool to study the effect of genetic manipulations in developmental biology as well as cancer research, disease development, and other areas of research in science. My research proposes a system in which gene expression can be controlled reversibly by using a photochromic compound as an effector molecule which is able to reversibly isomerize between two forms in response to light irradiation at different wavelengths. In other words, light could act as an inducer of antiswitch activity to reversibly control intracellular gene expression level in a temporal and tissue specific manner.

Chapter 2 describes the rational design of light-responsive antiswitch, referred to as C8-antiswitch. This molecular switch was designed to act in trans to control gene expression, similarly to an antisense oligonucleotide. The antiswitch structure was designed to incorporate an aptamer domain, the C8 aptamer, to dictate the secondary structure configuration of the antiswitch upon binding of a photochromic ligand molecule, BDHP-PEG. Therefore, light could be used as an inducer to reversibly control antiswitch secondary structure stabilization and consequently its activity to bind to the target mRNA.

Chapter 3 describes my approach for introducing the light-inducible C8-antiswitch into S. cerevisiae cells to control Green Fluorescent protein (GFP) expression
in vivo by light. GFP expression levels in response to antiswitch-ligand binding in *S. cerevisiae* cells was measured by flow cytometry using a fluorescent activated cell sorter (FACS). Chapter 4 summarizes the findings of my research, and proposes alternative approaches for controlling intracellular gene expression by C8-antiswitch.
Chapter 2:
Design of light responsive, allosteric C8-antiswitch and its in vitro functional analysis

2.1 Introduction:

As described in the previous chapter, riboswitches can be designed to act as antisense elements to regulate gene expression. The mode of function of such riboregulator can be controlled by the availability of a ligand molecule as the inducer of the riboswitch function. We are interested to use light as an inducer of molecular functions since it could provide a fast, non-invasive approach to control riboregulators’ function in a time and tissue specific manner.

To be able to achieve this goal, we took advantage of the properties of a photochromic switch compound, which isomerizes reversibly by light irradiation at visible and ultraviolet wavelengths. An RNA aptamer was previously selected by in vitro selection techniques to bind specifically to one isomer of this photochromic compound and not the other. This RNA aptamer sequence was used in the design of a riboswitch and was incorporated as the sensory domain of this riboswitch. The functional domain of this riboswitch is designed to act as an antisense RNA in response to light-induced ligand binding/unbinding at the aptamer domain. Our interest is to use this light-responsive riboregulator to control eukaryotic gene expression in vivo.
This chapter explains the design fundamentals of this light-responsive riboregulator, also referred to as an antiswitch. It will also describe the \textit{in vitro} studies performed to characterize the function of this light-responsive antiswitch \textit{in vitro}.
2.1.1 Photochromic Molecules

Photochromism is the reversible transformation of a chemical molecule between two distinct isomeric species in response to light absorptions at different wavelengths. This light-induced process leads to a colour change of the photochromic compound in a reversible manner, meaning the light-induced isomeric form must be able to revert back to the original molecular species photochemically by irradiation at a wavelength sufficiently different from that causing the initial change. Generally, visible light is employed to encourage de-coloration and ultraviolet radiation is used to promote coloration. The reversible isomerization process could also be induced thermally where a photochromic molecule isomerizes back to its more stable form at some rate, and heating accelerates this isomerization. An arbitrary requirement for using photochromic species is to have a photochromic molecule, which has two thermally stable isomeric states under ambient conditions for a reasonable time. Organic molecules with photochromic properties that have been identified and studied in detail rely on unimolecular reactions for their coloration and de-coloration, which generally involve ring closing and opening steps or cis-trans isomerizations. Photo-induced cycloaddition or electron transfer has also been employed to design photochromic systems.

Photochromic molecules are classified into 4 groups based on their structure rearrangement during the light-induced isomerization process.

Spiropyans and Spirooxazines

Spiropyans are the most studied family of photochromic materials. Spirooxazines are closely related to spiropyrans. The chemical composition of these
compounds consists of an oxazine, a heterocyclic group containing an oxygen and a nitrogen, arranged in a spiro form, a bicyclic organic group connected through just one atom. UV light irradiation causes the bond between the oxazine and the spiro-carbon to break, leading to ring opening and the molecule becomes planar. The new conjugated system is now able to absorb light of visible wavelength and hence appears as a coloured molecule. The coloured form isomerizes back to the original colourless molecular form once UV-light irradiation is removed. At this stage, the carbon-oxygen bond reforms and the molecule returns to its ground state. The coloured isomer of spirooxazine is thermodynamically less stable and reverts to the more stable colourless isomer in the dark unless cooled to low temperatures.

Figure 2.1 photo-induced isomerization of spiropyrans and spirooxazines

Azobenzenes

Azobenzenes are chemical compounds composed of two phenyl groups attached to an azo group (N=N double bond). An azobenzene molecule photo-isomerizes between its trans (orange-red colour) and cis (colourless) isomers. Trans-to-cis isomerization is induced by UV-light, and cis-to-trans isomerization is induced by visible light. The trans isomer is more stable than the cis isomer, thus cis-azobenzene will thermally back-isomerize to the trans-azobenzene.
Dithienylethenes

Dithienylethenes are chemical compounds composed of an alkene with a thiophene ring, a heterocyclic group with the formula C4H4S, attached to both sides. The 2-position of the thiophene is substituted with a methyl group, preventing oxidation of the closed ring form. Often the two free α-positions on the double bond are connected in a 5 or 6-membered ring in order to lock the double bond into the cis-form. This makes the dithienylethene undergo only open-closed ring isomerization.

Photochromic quinones

Quinones are cyclic compounds comprising of a six-member unsaturated ring to which carbonyl groups are attached. Photochromic properties of quinones and phenoxy napthacene quinone result from transfer of a phenyl group from one oxygen
group to another between the *ana* - and *para*-quinoid forms. Thermally stable quinones have been prepared. Quinones also have a redox activity, which is a feature that has led to construction of many-state molecular switches that operate by a mixture of photonic and electronic stimuli.

![Figure 2.4 Reversible photo- and thermoinduced isomerization of the *para* quinone structure to the *ana* quinone structure](image)

2.1.2 *In Vitro* Selection of an RNA aptamer against a photochromic compound

In a study published from our laboratory (H.W. Lee et. al. 2007), a light responsive hammerhead ribozyme was designed with its cleaving activity controlled by an aptamer domain, which sensed and selectively bound to an isomeric form of a photochromic molecule (Lee H.W. et. al. 2007). H.W. Lee successfully selected and characterized a RNA aptamer, C8 aptamer, by using SELEX, which recognizes and binds to one isomeric form of a photochromic compound but not the other isomer. This is a photo-isomerized dihydropyrene compound (Figure 2.5 A), which isomerizes reversibly between two forms, Benzo-dimethyl-dihydropyrene (BDHP) and Benzo-cyclophanediene (BCPD), by light irradiation at two different wavelengths (Mitchell R.H. et al. 2003). The pink benzo-dimethyl-dihydropyrene (BDHP) has a closed stable rigid planar extended $\pi$-system and was conjugated to 10-15 Polyethelyne Glycol (PEG) groups to
assist with the solubility of this compound in an aqueous solution (Lee H.W. et al. 2007). This isomer is also referred to as the ‘Closed’ isomer. The BDHP-PEG isomerizes to a second isomeric form of the compound by visible light irradiation. The second isomer is a colourless compound called benzo-cyclophanediene (BCPD). It has an open rigid step-shaped molecule and hence it is referred to as the ‘Open’ isomer (Lee H.W. et al. 2007). The open isomer can be converted reversibly back to the closed form isomer by ultraviolet (UV) light irradiation of wavelength less than 300 nm (Figure 2.5 A, B) (Mitchell R.H. et al. 2003). The selected C8 aptamer (Lee H.W. et al. 2007) specifically recognizes and binds to BDHP isomer and not the other isomer.
The C8 aptamer was then conjugated to a hammerhead ribozyme (Figure 2.6 A, B), an RNA cleaving RNA enzyme, to construct an allosteric hammerhead ribozyme (Lee H.W. *et al.* 2007). In this construct a distinctive RNA secondary structure element called a UG communication module was used to transmit ligand-induced structural stabilization of the aptamer motif to the catalytic motif of the allosteric ribozyme.
Figure 2.6  C8 aptamer and its incorporation to design an allosteric hammerhead ribozyme

(A) RNA sequence of C8 aptamer, specific for binding BDHP-PEG with Kd = 2.7 (+/-0.4) µM.

(B) Secondary structure and sequence of the C8-conjugated allosteric hammerhead ribozyme. In red: The G.U communication module connecting the C8 aptamer and the ribozyme. The 14-nucleotide RNA substrate forms complementary base pairs with the arms of the hammerhead ribozyme. (Adapted from Journal of Molecular Biology, Vol. 371, Lee H.W. et al., Copyright (2007) with permission from Elsevier)

The activity of this hammerhead ribozyme was under allosteric regulation, meaning its cleaving activity was only enabled by binding of the effector molecule. Hence this ribozyme was referred to as the UGLOOP allosteric hammerhead ribozyme. Its cleaving activity switched between active and inactive states in response to selectively binding of the BDHP-PEG isomeric form of the compound to the aptamer domain (Lee H.W. Lee et al. 2007). In the active state of the ribozyme the stable secondary structure of the catalytic domain allowed the target binding arms of the hammerhead ribozyme to base pair with the target RNA sequence and cleave the substrate into a smaller fragment (Figure 2.7 A). On the other hand, in the inactive state of the ribozyme the secondary
structure of the hammerhead ribozyme is destabilized and the catalytic core is unable to cleave the substrate. Ultimately, the cleaving activity of this allosteric hammerhead ribozyme was controlled by light, Figure 2.7 B, since the isomeric state of the effector compound, BDHP-PEG, was changed by visible light irradiation into a second isomeric form, BCPD-PEG, which did not bind to the C8 aptamer domain of the allosteric hammerhead ribozyme. Consequently, the ribozyme was inactive in the presence of the BCPD-PEG since its secondary structure was not stabilized (Lee H.W. et al. 2007). The cleaving activity was retained by UV light irradiation to isomerize the BCPD-PEG back to BDHP-PEG ligand (Figure 2.7 B).
Figure 2.7  RNA cleaving properties of the UGLOOP-hammerhead ribozyme construct

(A) The RNA-cleavage properties of the UGLOOP- hammerhead ribozyme construct in presence of 1mM BDHP-PEG or 1mM BCPD-PEG. The labels on the right-hand side indicate the presence or the absence of the various reaction mixtures.

(B) The substrate RNA cleavage rates by the UGLOOP-hammerhead ribozyme, in the presence of BDHP-PEG, while being irradiated in real time with either ultraviolet (280–375 nm) or visible (>400 nm) light. For the first 60 min (■), under UV irradiation, the BDHP isomer is expected to be dominant and percentage of cleaved RNA substrate increases with time. With visible irradiation (▲), BDHP is converted to BCPD and the fraction of cleaved RNA substrate remains unchanged; upon UV irradiation (▼), BCPD is expected to re-convert to BDHP, and the fraction of the cleaved substrate continues to increase with time. (Adapted from Journal of Molecular Biology, Vol. 371, Lee H.W. et. al., Copyright (2007) with permission from Elsevier)
2.1.3 A ligand-controlled riboregulator of eukaryotic gene expression

As previously described in chapter one, an RNA-based regulator was designed (Bayer T.S., Smolke C.D. 2005) which could regulate gene expression in *S. cerevisiae* by base pairing in *trans* with its target transcript in response to presence of theophylline ligand, and block ribosomal scanning and translation initiation (Bayer T.S., Smolke C.D. 2005). This antiswitch, referred to as S1, was composed of an aptamer domain specific for theophylline recognition and binding (K$_d$ = 0.29 µM) and an antisense domain, which contained an antisense sequence complementary to the target mRNA encoding green fluorescent protein (GFP) to control GFP gene expression. Theophylline ligand binding to the aptamer domain stabilizes an aptamer stem formation, which induces a conformational change in the secondary structure of this molecular switch. At this stage the antisense portion of the switch is forced into a single stranded state to allow the antisense sequence to base pair with a 15-nucleotide region around the start codon of a GFP target mRNA to affect translation. In the absence of the ligand, the antiswitch adopts a second conformation, in which the antisense domain is sequestered within an antisense stem. The aptamer stem and the antisense stem were designed such that the antisense stem was thermodynamically slightly more stable than the aptamer stem (Figure 2.8 a and b)
Figure 2.8 General illustration of the antiswitch molecule mechanism

(a) General mechanism used by an antiswitch molecule to regulate gene expression in vivo. The antisense sequence is indicated in red; switching ‘aptamer stem’ is shown in blue. In the absence of effector, the antisense domain is bound in a double-stranded region of the RNA referred to as the ‘antisense stem’ and the antiswitch is in the ‘off’ state. In this state the antiswitch is unable to bind to its target transcript, which has a gfp coding region, and as a result, GFP production is on. In the presence of effector, the antiswitch binds the molecule, forcing the aptamer stem to form, switching its confirmation to the ‘on’ state. In this state the antisense domain of the antiswitch will bind to its target transcript and through an antisense mechanism turn the production of GFP off.

(b) Sequence and predicted structural switching of a theophylline-responsive antiswitch, s1, and its target mRNA. On s1, the antisense sequence is indicated in red; switching aptamer stem sequence is indicated in blue; the stability of each switching stem is indicated. On the target mRNA, the start codon is indicated in green. (c) In vivo temporal response of s1 inhibiting GFP expression upon addition of effector to cells that have accumulated steady-state levels of GFP and antiswitch s1. No theophylline, blue; 2 mM theophylline, red. (Adapted by permission from Macmillan Publishers Ltd: Nature Biotechnology, Vol. 23, Bayer T.S., Smolke C.D. Copyright © 2005)
By introducing this antiswitch into *S. cerevisiae*, Smolke group were able to show down-regulation of GFP expression in cells harbouring the antiswitch grown in the presence of theophylline, and relative to the appropriate controls, GFP expression level in these cells was unaltered in the absence of theophylline (Figure 2.8c). Quantitative real-time PCR (qRT-PCR) data obtained from these cells determined relative levels of target transcript did not change substantially between cells harbouring the antiswitch grown in the absence or presence of theophylline, indicating that antiswitch function through translational inhibition rather than reducing target RNA levels (Bayer T.S., Smolke C.D. 2005).

### 2.2 The design and of a light-sensitive antiswitch and its in vitro functional analysis

Using the C8 aptamer sequence (Lee H.W. *et. al.* 2007), the goal of my thesis is to design an antiswitch (Bayer T.S., Smolke C.D. 2005), which could control gene expression in response to light. Antiswitches should be excellent vehicles for adaptation to our photoswitch approach, by incorporating the C8 aptamer, in modular fashion, into antiswitches.

The design of my light responsive antiswitch was based on the rational design of the antiswitch model proposed by Bayer T.S. and Smolke C.D. (Bayer T.S., Smolke C. 2005) to control reporter gene expression in *Saccharomyces cerevisiae*. In principle, this new antiswitch model is designed to incorporate the C8-aptamer domain and be responsive to our photochromic ligand, BDHP-PEG to regulate the expression of a target transcript in response to light.
The antiswitch model I propose uses an antisense sequence domain to control gene expression and an aptamer domain, the C8-aptamer, to recognize and bind to our photochromic effector ligand BDHP (Figure 2.9). In the absence of the specific ligand, the antiswitch secondary structure stabilizes into a conformation which allows the antisense sequence domain to interact with the target mRNA to inhibit translation. At this stage the antiswitch functions as an ‘OFF’ switch. In the presence of the BDHP-PEG ligand, the C8-aptamer domain stabilization induces an alternative antiswitch structure, in which the antisense domain is sequestered within the double-stranded stem of the C8-aptamer and is no longer available to interact with the target transcript. The C8-antiswitch is functioning as an ‘ON’ switch in the presence of BDHP ligand.
In the absence of the ligand or in the presence of the open isomer of our photochromic ligand, which does not bind to the C8-aptamer (in blue), the antisense sequence (in red) is free to base pair with the target EGFP mRNA (in green) and negatively affect EGFP translation. In this stage the antisense stem (in black) is stabilized thermodynamically.

In the presence of BDHP ligand (represented as the purple circle), which could be isomerized from the BCPD (open isomer) form by ultraviolet light irradiation, the C8-aptamer is expected to bind to the BDHP ligand. The free energy associated with the aptamer-ligand binding is believed to stabilize the aptamer stem, sequestering the antisense sequence within this stem. As a result in the presence of BDHP ligand, the target EGFP mRNA is expected to remain free to be translated.

The EGFP translation could be reversed by visible light irradiation to isomerize the BDHP ligand (closed isomer) back to BCPD (open isomer) form, which does not bind to the C8-aptamer and the former antiswitch structure is stabilized to inhibit EGFP translation.
Specifically in my design, the 5’-end of the C8 aptamer sequence is conjugated to an antisense RNA sequence designed to target, in the absence of the BDHP-PEG ligand, and base pair with a 10-nucleotide region around the AUG start codon of the target mRNA encoding enhanced green fluorescent protein (EGFP). The 3’-end of the C8 aptamer is conjugated to a stem-loop domain, called the antisense stem domain, which is designed to be thermodynamically more favourable to form in the absence of the BDHP-PEG ligand. Upon BDHP-PEG ligand binding to the C8-aptamer domain, an alternative stem-loop structure is expected to form, called the aptamer stem domain, between the antisense sequence at the 5’-end of the aptamer domain and its complementary sequence at the 3’-end of the C8-aptamer domain.

The C8-antiswitch is based on a dual-stem and I expect aptamer-ligand binding and antisense binding to its target mRNA would lead to the structural changes in the C8-antiswitch molecule, allowing it to act as an allosteric riboregulator. *In vitro* characterization studies were conducted to examine C8 antiswitch-ligand affinity and conformational changes associated with antiswitch response.
2.3 Material and Methods

2.3.1 Materials:

DNA oligomers were purchased from Integrated DNA Technologies (IDT) and are listed in the Appendix section. All oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis (PAGE), visualized by UV shadowing, cut from the gel and eluted from the gel by crush-soaking in a buffer containing 10mM Tris-HCl (pH8.0 at 23°C) and 0.1 mM EDTA followed by ethanol precipitation. T7 RNA polymerase was purchased from Invitrogen. W303α yeast cells were donated by Beh lab, Simon Fraser University.

2.3.2 Uptake of BDHP-PEG compound into S. cerevisiae

A single colony of S. cerevisiae cells of strain W303α was inoculated overnight at 30°C in synthetic complete media (SC) and 2% glucose. The next day the cells were back diluted in fresh SC media and grown to OD₆₀₀ 0.8. 1mL cell culture aliquots were removed and the cell pellet was washed 3 times by PBS-Sorbitol (1.2M Sorbitol in PBS, pH 6.5). The cells were then re-suspended in 1ml of PBS-Sorbitol, 6% DMSO in presence or absence of 1mM BDHP-PEG compound. The cells were incubated overnight at 30°C, with the eppendorf tube containing the BDHP-PEG compound wrapped in aluminium foil to avoid incidental light irradiation and isomerization of BDHP-PEG by visible light. The next day, the cells were washed with fresh PBS-Sorbitol to remove BDHP-PEG compound. The cells were re-suspended in fresh PBS buffer, pH 7.4 and a small aliquot of cells was loaded onto microscopy slides treated with Concavalin A and viewed using a Leica fluorescent microscope.
The fluorescence microscope used was a Leica DMI 6000B model. The A4 filter cube from Leica Microsystems was used for DAPI detection, which includes a 400 nm longpass dichromic mirror, a 360-40 nm bandpass excitation filter, and a 470-40 nm bandpass emission filter. Images were captured using 510 millisecond exposure, 255 gain, and 55% FMI (Fluorescent Manager Intensity).

### 2.3.3 Investigating the fluorescent property of BDHP-PEG

To investigate the possible fluorescence characteristics of BDHP-PEG at a high concentration, 1 ml solution of 50 μM BDHP-PEG in 4% DMSO was transferred into a 10mm pathlength Quartz cuvette. This sample was excited at a range of wavelength between 340-380 nm, at 10 nm excitation wavelength intervals, and the fluorescent emission spectra were generated and analyzed for any detectable emission peaks between 450-600 nm. The buffer used as the blank sample was 4% DMSO. The spectrometer used in this experiment was a HORIBA Jobin Yvon, type FL-1039, with a 450W Xenon Arc lamp from USHIO Inc. DATAMAX software was used for data analysis.

The emission spectra of 50 μM BDHP-PEG as shown in Figure 2.12 was obtained by exciting the sample cuvette at different excitation wavelengths between 340 to 380 nm, at 10 nm excitation intervals. The spectrometer was set to detect the fluorescent emission spectra of BDHP-PEG between 450-600 nm, using 0.1 sec integration time, 2 nm excitation bandpass and 8 nm emission bandpass. The emission spectra as shown in Figure 2.12 was generated after blank sample correction using the DATAMAX software.

The emission spectrum of a 0.38 μM BDHP-PEG as shown in Figure 2.13A was generated using the same spectrometer setting except the integration time was increased.
to 1 second. The emission spectrum of BDHP-PEG was scanned between 400-600 nm by exciting the sample cuvette at $\lambda_{ex}=370$ nm. The excitation spectrum as shown in Figure 2.13A was generated by exciting the sample cuvette between 350-450 nm at 2nm increments and the spectrometer was set to detect emission at $\lambda_{em}=508$ nm. To determine the potential excitation and emission spectrum of BCPD-PEG, the same sample used for BDHP-PEG fluorescent measurement in Figure 2.13A was irradiated with visible light using a hand-held halogen lamp for 30 minutes to isomerize the 0.38 $\mu$M BDHP-PEG to BCPD-PEG. The excitation and emission spectra of 0.38 $\mu$M BCPD-PEG was then explored using the exact same spectrometer setting as explained for BDHP-PEG, to generate the spectra shown in Figure 2.13B. The spectra in Figure 2.13A and B were corrected for the blank solution of 0.4 % DMSO.

2.3.4 Antiswitch and EGFP in-vitro transcription:

*In vitro* Transcription reactions were performed using full-length purchased double-stranded DNA template. The sequences used are Antiswitch sense strand: 5’-CTA ATA CGA CTC ACT ATA GGTC ACC ATG GTG GTC ATC CTA CAC CAT GGT GAG CCC TCG CTC ACC ATG GT-3’ and the Antiswitch antisense strand: 5’-ACC ATG GTG AGC GAG GGC TCA CCA TGG TGT AGG ATG ACC ACC ATG GTG ACC TAT AGT GAG TCG TAT TAG-3’. The T7 promoter sequences are underlined at the 5’ and 3’-ends of these oligos. The EGFP sense sequence is: 5’-CTA ATA CGA CTC ACT ATA GGA TGG TGA GCA AGG GC-3’ and the EGFP antisense sequence is: 5’-GCC CTT GCT CAC CAT CCT ATA GTG AGT CGT ATT AG-3’. 
Transcription reactions (50 μL) containing 50 pmole of double-stranded DNA template were incubated in 40mM Tris-HCl (pH 7.9 at 23°C), 26 mM MgCl₂, 2.5 mM spermidine, 10 mM DTT, 0.01% triton X-100, 8mM GTP, 4mM ATP, 4mM CTP, 2mM UTP, 150 units of T7 RNA polymerase for 2 hrs at 37°C. Ethanol precipitation was performed following transcription, and the RNA were purified by a 12% denaturing PAGE, visualized by UV shadowing and eluted from the gel by crush-soaking in a buffer at 4°C containing 10mM Tris-HCl (pH8.0 at 23°C) and 0.1 mM EDTA followed by ethanol precipitation.

To be able to add a radioactive [γ-32P]-ATP at the 5’-end of the transcribed RNA, alkaline phosphatase enzyme from calf intestine kit (Roche) was used according to standard protocol to remove two pyrophosphates from the 5’-end of RNA, leaving one 5’-pyrophosphate on RNA. Following alkaline phosphatase treatment, the RNA was 5’-end labelled with [γ-32P]-ATP (Perkin Elmer) using standard phosphorylation protocols of T4 Kinase Kit (Invitrogen), and then re-purified by a 8% denaturing PAGE, visualized by autoradiography and eluted from the gel by crush-soaking in a buffer at 4°C containing 10mM Tris-HCl (pH8.0 at 23°C) and 0.1 mM EDTA followed by ethanol precipitation.

### 2.3.5 In-vitro target binding affinity and ligand-induced structural switching of the C8-antiswitch

*In-vitro* antiswitch affinity assays for all reaction conditions were conducted with 32-P-labeled antiswitch transcript (~300nM) with excess amount (~700nM) of unlabeled EGFP target transcript. Transcription of the antiswitch and the EGFP target sequence was explained in section 2.3.4. Antiswitch RNA was heated at 90°C for 1 min in a buffer
solution containing 50mM Tris-HCl (pH 7.5 at 23°C), 100mM NaCl, 5mM MgCl₂, 0.05% Triton X-100, and 1% DMSO. The antiswitch was allowed to cool down slowly in the buffer to room temperature in the presence or absence of BDHP-PEG or BCPD-PEG. Next, the EGFP RNA was heated in TE buffer (10mM Tris-HCl pH 8.0 at 23°C, 0.1 mM EDTA) for 30 seconds at 90°C and quickly added to the antiswitch solution. The antiswitch and EGFP in the presence or absence of BDHP-PEG or BCPD-PEG was mixed by pipetting, and this mixture was incubated for 15 min at room temperature. The reaction was terminated by addition of non-denaturing loading buffer (0.05% bromophenol blue, 0.05% xylene cyanol, 15% ficoll) and equal number of radioactive counts from each reaction condition was run on a 8% non-denaturing gel at 4°C for ~2.5 hrs. The reaction products from the gel were visualized and quantified using PhosphoImager and ImageQuant software (Molecular Dynamics).

2.3.6 Generating C8-antiswitch with defined end-sequences by cis-cleaving dual-hammerhead ribozyme construct

The dual-hammerhead ribozyme construct flanking the C8-antiswitch sequence, also referred to as the hammerhead-Antiswitch-Hammerhead (Hm-As-Hm), was transcribed in-vitro using double-stranded synthetic DNA oligos purchased from IDT (Integrated DNA Technologies). The Hm-As-Hm sense DNA sequence is: 5’-CTA ATA CGA CTC ACT ATA GGG TGA CTG ATG AGT CCG TGA GGA CGA AAC GGT AGG AAT TCC TAC CGT CTC ACC ATG GTG GTC ATC CTA CAC CAT GGT GAG CCC TCG CTC ACC ATG GTC AAC CGG AGT CGA CTC CGG TCT GAT GAG TCC GTG AGG ACG AAA CCA-3’ and the Hm-As-Hm anti-sense sequence is 5’-TCC TTT CGT CCT CAC GGA CTC ATC AGA CCG GAG TCG ACT CCG GTT
GAC CAT GGT GAG CGA GGG CTC ACC ATG GTG TAG GAT GAC CAC CAT
GGT GAG ACG GTA GGA ATT CCT ACC GTT TCG TCC TCA CGG ACT CAT
CAG TCA CCC TAT AGT GAG TCG TAT TAG-3’. These DNA oligos contained the
T7 promoter sequence (underlined) and were transcribed by T7 RNA polymerase and the
corresponding RNA transcript was internally radioactively labelled using α-UTP. In-
vitro transcription reaction (50 μL) containing 50 pmole of double-stranded DNA
template was incubated in 40mM Tris-HCl (pH 7.9 at 23°C), 26 mM MgCl₂, 2.5 mM
 spermidine, 10 mM DTT, 0.01% triton X-100, 8mM GTP, 4mM ATP, 4mM CTP, 2mM
UTP, and 1μL α-UTP and 150 units of T7 RNA polymerase, from Invitrogen, for 2 hrs at
37°C. Ethanol precipitation was performed following transcription, and the α–UTP
labelled RNA were purified by a 10% denaturing PAGE, visualized by autoradiography
and eluted from the gel by crush-soaking in a buffer at 4°C containing 10mM Tris-HCl
(pH8.0 at 23°C) and 0.1 mM EDTA followed by ethanol precipitation. The precipitated
RNA product was re-suspended and stored in TE buffer. The radioactive RNA product
was run on an 8% denaturing PAGE to visualize the cis-cleaving activity of the two
hammerhead ribozymes flanking both ends of the antiswitch RNA. To analyze the time-
dependent cis-cleaving activity of the dual-hammerhead construct, as shown in Figure
2.17, the transcription reaction was allowed to proceed for 4 hrs at 37°C, with aliquots
removed from the reaction at 30 minutes intervals for analysis on denaturing gel.
2.4 Results and discussion:

2.4.1 Uptake of BDHP-PEG photochromic compound into *S. cerevisiae* cells

The initial experiment performed prior to testing the functionality of C8-antiswitch design to control gene expression *in-vitro* and eventually *in-vivo*, was to confirm the BDHP-PEG ligand is able to permeate into *S. cerevisiae* cells. As mentioned earlier the Polyethylene glycol group (PEG) was attached to the BDHP molecule to increase the solubility of this compound in an aqueous incubation media.

A single colony of *S. cerevisiae* cells of strain W303α was inoculated overnight at 30°C in synthetic complete media (SC) and 2% glucose. 1 mL cell culture aliquots were removed and the cell pellet was washed by PBS-Sorbitol. The cells were then re-suspended in 1ml of PBS-Sorbitol, 6% DMSO in presence or absence of 1mM BDHP-PEG compound. The cells were incubated overnight at 30°C, with the eppendorf tube containing the BDHP-PEG compound wrapped in aluminium foil to avoid light irradiation and isomerization of BDHP-PEG by visible light. The next day, the cells were washed with fresh PBS-Sorbitol to remove BDHP-PEG compound. The cells were re-suspended in fresh PBS buffer and viewed using a Leica fluorescent microscope. The control cells were incubated overnight in the presence of 6% DMSO only.
Figure 2.10 BDHP-PEG compound uptake into yeast cells

A) Control, untreated yeast cell pellet, left, is a white pellet, whereas BDHP-PEG treated yeast cell pellet, right, appears as a pink pellet. Fluorescent microscopy images of B) control, untreated cells; C) BDHP-PEG treated yeast cells appearing as brighter, more fluorescent cells; D) BDHP-PEG treated cells after 20 minutes of visible light irradiation, with no apparent changes in fluorescence after light irradiation.
Figure 2.10 shows the images of control and BDHP-PEG treated cells as explained. As seen in Figure 2.10 A, the control cell pellet appears as a white colony whereas the cell pellet of BDHP-PEG treated cells appears to have a pink colour in appearance. Considering the BDHP-PEG compound is a pink colour compound, the pink colour of these cells is contributed by the overnight BDHP-PEG treatment. Uptake of BDHP-PEG compound into *Drosophila Melanogaster* and *E.coli* was also attempted using appropriate incubation conditions, however the presence of compound into these model organism could not be detected visually nor by microscopy.

To further investigate the uptake of the BDHP-PEG compound into the yeast cells, the cells were viewed by a fluorescent microscope, using a DAPI filter, with a 360/40nm bandpass excitation filter, and a 470/40nm bandpass emission filter. Figure 2.10 B shows the untreated yeast cell as controls. In comparison to the control cells, the BDHP-PEG treated cells, Figure 2.10 C, appear as weakly fluorescent cells under the mentioned microscope setting. The fluorescence is seen mainly as bright spots throughout the cytoplasm of the cells. We then irradiated with same slide and field of view of the BDHP-treated cells with visible light by the bright field microscopy light of the microscope to see whether the appearance of the cells changes by light irradiation. However, after 20 minutes of light irradiation, which is a sufficient length of time to isomerize the BDHP-PEG to BCPD-PEG form, no apparent changes are seen in the fluorescent appearance of the cells.

In a separate study by another group (Plant A.L. *et. al.* 1985) benzo(a)pyrene, a carcinogenic compound derived from the benzo-pyrene family of compounds, was spontaneously transferred into human fibroblasts and viewed using fluorescent
microscopy. The intracellular accumulation of benzo(a)pyrene resulted in an increased fluorescence at certain spots inside the fibroblast cells. After co-localization of benzo(a)pyrene and a fluorescent cholesterol analog it was determined that benzo(a)pyrene was partitioning into cytoplasmic lipid droplets (Plant A.L. et. al. 1985). Considering this finding, I suspected the bright spots seen in BDHP-PEG treated cells, is perhaps due to BDHP-PEG cytoplasmic accumulation of this molecule.

2.4.2 Investigating the fluorescent property of BDHP-PEG

It was still unclear whether the cytoplasmic fluorescence detected by the fluorescent microscope is produced by BDHP-PEG. If so, BDHP-PEG should emit fluorescence if excited at the same excitation wavelength of the DAPI filter of the fluorescent microscope, with a 360/40nm bandpass excitation, and a 470/40nm bandpass emission filter. To investigate this possibility, a high concentration of BDHP-PEG, 50 µM, was excited at a range of wavelength between 340-380 nm, at 10 nm excitation wavelength intervals, and the fluorescent emission spectra were generated and analyzed for any detectable emission peaks between 450-600 nm. Figure 2.11 shows the preliminary, background corrected emission spectra of BDHP-PEG, excited at 10 nm intervals between 340-380 nm. According to Figure 2.11, it appears that BDHP-PEG has a fluorescent emission peak at 508 nm, and an excitation wavelength at 370 nm induces the highest fluorescent intensity.
Figure 2.11 Preliminary BDHP-PEG emission spectrum

The fluorescent emission spectrum of 50 μM BDHP-PEG in 4% DMSO solution. The emission spectrum was scanned at excitation wavelengths between 350-380 nm at 10 nm intervals. Emission peak is observed at 508 nm by 370 nm excitation.

To have a more thorough fluorescence analysis, I also explored an excitation scan for BDHP-PEG and BCPD-PEG between a range of 350-450 nm and an emission scan between a range of 400-600 nm for both isomers. I also used a much lower concentration of each isomer to avoid potential inconsistency of the intensity of the excitation light reaching the sample solution. Figure 2.12 A and B shows the excitation and emission scan of BDHP-PEG and BCPD-PEG. Figure 2.12 A shows the excitation and emission spectra of 0.38 μM BDHP-PEG in 0.4% DMSO. It appears BDHP-PEG is excited at 370 nm, which results in a fluorescence emission peak at 508 nm. Figure 2.12 B shows the excitation and emission spectra of 0.38 μM BCPD-PEG in 0.4% DMSO, to see if there are any changes in the potential fluorescent properties of these two isomers.
Figure 2.12 Excitation and emission spectra of BDHP-PEG and BCPD-PEG samples

Excitation spectra (blue) and emission spectra (red) of 0.38 μM of A) BDHP-PEG, and B) BCPD-PEG.
The excitation spectrum scan between 350-450 nm was generated at $\lambda_{em}=508$ nm. The emission spectrum scan between 400-600 nm was generated at $\lambda_{ex}=370$ nm.

As shown in Figure 2.12 A and B, the BCPD-PEG isomer is also producing identical excitation and emission peaks as BDHP-PEG at 370 nm and 508 nm respectively. After observing identical excitation and emission peaks for these two isomers by the fluorometry experiment, we concluded the weak fluorescence that is observed is more likely caused by a contaminating precursor molecule in our stock BDHP-PEG sample.

Mitchell group at University of Victoria has synthesized the BDHP-PEG compound for us (Mitchell R. H. et. al. 2003). It is possible for traces of precursor fluorescent molecules to have remained as impurities in our BDHP-PEG stock, especially if pyrene, a fluorescent molecule in a solution with excitation and emission wavelength of 372 nm and 480 nm respectively (Sigman M.E. et. al. 1998), was used during BDHP-PEG synthesis. The excitation and emission wavelengths of pyrene correspond with the
excitation and emission bandpass of DAPI filter on our fluorescent microscope, 360/40nm excitation and a 470/40nm emission bandpass filter.

Considering this fluorescence analysis, BDHP-PEG and BCPD-PEG are non-fluorescent molecules, and what we are seeing as cytoplasmic fluorescent spots in BDHP-PEG treated yeast cells is most likely caused by fluorescent pyrene contaminant molecules that are perhaps partitioning into lipid droplets as suggested by a similar finding in the study done by Plant group (Plant A.L. et. al. 1985). The existence of a potential contaminating molecule can be tested in the future by HPLC analysis of the BDHP-PEG stock solution. Although cytoplasmic accumulation of BDHP-PEG could not be visualized by fluorescent microscopy we believe it is still possible that this compound has entered the cell cytoplasm and is co-localized with the fluorescent pyrene contaminant inside the cells.

2.4.3 *In vitro* target binding affinity and ligand-induced structural switching of the C8-antiswitch

An *in vitro* experiment was performed to examine the ligand and target binding affinity of the C8-antiswitch as shown in Figure 2.9. Gel-shift experiments were conducted using *In-vitro* transcribed, 5'-end radiolabelled C8-antiswitch, in the presence of unlabeled, equimolar amounts of an *In-vitro* transcribed, short (17 nucleotides) EGFP target transcript. The target RNA contained sequences of the regions upstream and downstream of the AUG start codon of EGFP mRNA. To examine antiswitch-target binding affinity in response to ligand sensing the reaction was performed in the presence of varying concentrations of BDHP-PEG (closed isomer) or BCPD-PEG ligands (open isomer). Figure 2.13 shows the gel shift as the result of the C8-antiswitch binding to its
target mRNA in the absence of BDHP ligand or in the presence of BCPD-PEG (open isomer) as expected.

Figure 2.13 In vitro target binding affinity assay of C8-antiswitch in response to BDHP-PEG ligand

Lane 1: 5’-end labelled C8-antiswitch (As), lanes 2-5: 5’-end labelled C8-antiswitch incubated in the presence of unlabelled target transcript in presence of 0 μM BDHP-PEG (lane 2), 0.4 μM BDHP-PEG denoted as C (lane 3), 40 μM BDHP-PEG (lane 4), 40 μM BCPD-PEG denoted as O (lane 5).

Lane 1 shows the 5’-end ³²P-labeled C8 antiswitch mobility in a non-denaturing polyacrylamide gel. Lane 2, shows a sharp shift in antiswitch mobility at 0 μm BDHP (denoted by 0 μM C) in the presence of unlabelled target transcript. This shift is indicative of the antisense sequence binding to the target EGFP transcript since the C8-antiswitch is acting as an “Off” switch to bind to the target mRNA in the absence of the effector ligand. Lanes 3, and 4, shows faster gel mobility of C8-antiswitch in presence of varying concentration of BDHP ligand (0.4 and 40 μM BDHP). This change in mobility is indicative of ligand binding to the aptamer domain of the C8-antiswitch. This
stabilizes the aptamer double-stranded stem, which sequesters the antisense sequence in the stem. As a result the antiswitch is in the “ON” state and cannot bind to the target EGFP sequence, hence it can migrate at a faster rate in the gel. Lane 5 shows another sharp shift in the mobility of the C8-antiswitch in the presence of the BCPD ligand. As previously discussed, this ligand is not able to bind to the aptamer domain of the C8-antiswitch, and therefore, the antiswitch is in the “Off” state again.

This experiment investigated the C8-antiswitch target binding affinity in response to the BDHP-ligand. The C8-antiswitch was able to sense the presence of this specific ligand in the reaction buffer, and binding of BDHP-PEG to the aptamer domain induced a secondary structure formation as predicted in Figure 2.9. This structural stabilization lead the antiswitch to act as an ‘ON’ switch and not bind to the target in-vitro transcribed mRNA fragment. On the other hand, the C8-antiswitch could sense the absence of the ligand and fold into an alternative secondary structure. In this conformation the antisense sequence domain of the antiswitch was able to base pair with the target mRNA and the C8-antiswitch acted as an ‘OFF’ switch.

In Figure 2.13, lanes 2 and 5, there are minor products with the same mobility on the gel as the C8-antiswitch in the absence of the target and the ligand, lane 1. The percentage of these minor products relative to the total C8-antiswitch was calculated using the ImageQuant Software and was determined to be 7%, whereas the C8-antiswitch experiencing a gel shift was determined to be 80%. Similarly in the presence of the BDHP ligand, lane 3 and 4, there are minor products with calculated average densitometry of 18% in comparison to the major product, which does not exhibit gel shift mobility, with densitometry of 69% relative to the total $^{32}$P-labelled C8-antiswitch.
These minor products formation is perhaps due to mis-folding of the C8-antiswitch in the absence or presence of the ligand.

The aptamer stem and antisense stem are designed such that the antisense stem is slightly more stable thermodynamically than the aptamer stem (ΔG= -24.2 kcal/mole and ΔG= -22.2 kcal/mole respectively). In retrospect, in the absence of ligand binding to the C8 aptamer, the formation of the antisense stem is expected to dominate over the aptamer stem formation. Alternatively, I anticipate the free energy associated with the binding of the BDHP-PEG ligand to the C8-aptamer domain could induce the formation of the aptamer stem and maintain it as the dominant stem structure. As a result, in the presence of BDHP-PEG ligand, the antisense sequence is expected to be sequestered within this aptamer stem and would not be available to interact with the target transcript. Figure 2.14 shows the sequences, secondary structure predictions, and the free energies associated with the formation of (a) the aptamer stem, (b) the antisense stem, and (c) the overall C8-antiswitch. This information is generated by an online secondary structure prediction tool, MFold (www.idtdna.com/Scitools/Applications/ mFold/).
Figure 2.14 Sequences and structures of the C8-Antiswitch stems as predicted by MFold.

a) Aptamer stem sequences and structure; red, antisense sequences; blue, C8-aptamer sequences. The free energy associated with the aptamer stem formation is estimated by MFold to be -22.2 kcal/mol.

b) Antisense stem sequences and structure. The free energy associated with the antisense stem formation is estimated by MFold to be -24.2 kcal/mol. The antisense stem is slightly more stable than the aptamer stem due to its slightly lower free energy of formation.

c) The overall structure of the C8-antiswitch expected to form in the absence of the BDHP ligand; red, antisense sequences; blue, C8-aptamer sequences; black, antisense stem.

2.4.4 Generating C8-antiswitch with defined end-sequences

To introduce the C8-antiswitch in vivo it is crucial for the antiswitch to have defined sequence ends to allow proper folding and structure stabilization. To achieve this purpose the C8-antiswitch sequence was designed to have a flanking hammerhead
ribozyme, known to self-cleave \textit{in vivo}, at each of its ends (Figure 2.15). Once this construct is transcribed, the flanking hammerhead ribozymes will stabilize into its catalytically active secondary structure, which enables the ribozyme to cis-cleave at specific sequences (Taira K. \textit{et al.} 1991, Bayer T.S. and Smolke C.D. 2005). As a result C8-antiswitch in between the two hammerhead ribozymes will be cut out with defined 5’ and 3’ sequences.
Figure 2.15 Mechanism of cis-cleaved C8-antiswitch by dual hammerhead ribozymes

Top: The C8-antiswitch (denoted as AS here) is attached to a hammerhead ribozyme (black) at each end. Once the hammerhead ribozymes stabilize into their catalytically active form, they are expected to cis-cleave at particular sequences shown by the red arrow. This results in the release of the C8-antiswitch sequence with defined 5’ and 3’ sequence ends.

Bottom: Sequence and cleavage mechanism of Hm-As-Hm expression construct. The expression construct enables cloning of general sequences between two hammerhead ribozyme sequences through unique restriction sites KpnI, EcoRI, SalI, and PstI (indicated in green). Predicted cleavage sites are indicated by red arrows, C8-Antiswitch insert is indicated by a blue line or lettering.

To confirm the cis-cleaving activity and the efficiency of the dual hammerhead construct, the C8-antiswitch with flanking hammerheads was transcribed in vitro from
synthetic double-stranded DNA template using T7 primer sequences at the 5’-end of the DNA sense strand. Transcription was performed by T7 polymerase enzyme and the transcripts were internally radiolabelled by addition of 1ul of α-UTP to the transcription reaction mixture. The magnesium concentration (2.5 mM) of the transcription buffer stabilizes the catalytically active secondary structure of the hammerhead ribozymes as transcription is taking place, which induces the ribozymes to cis-cleave and release the C8-antiswitch RNA fragment, which is 49 nucleotides in length. The cis cleaved products were visualized by running the transcripts onto a denaturing polyacrylamide gel and the percentages of the cleaved products were determined using the ImageQuant software (Figure 2.16).
Figure 2.16 Cis-cleaving activity of dual hammerhead ribozymes to generate C8-antiswitch with defined sequence ends.

Antiswitch lane: indicates the mobility of the internally α-UTP labelled C8-antiswitch, 49 nucleotides in length, transcribed from a synthetic DNA template by using T7 RNA polymerase. Hm-As-Hm lane: the hammerhead-antiswitch-hammerhead RNA (145 nucleotides in size) was transcribed from a synthetic DNA template and radiolabelled internally by α-UTP. Three different transcript lengths, 49 nt, 98 nt, and 145 nt are produced as the result of the cis-cleaving activity of the dual hammerhead ribozymes. The third lane from the left shows a 10 base pair $^{32}$P labelled DNA ladder as an approximate size marker.
As shown in Figure 2.16, the dual hammerhead construct flanking the C8-antiswitch, denoted as Hm-As-Hm, generates three transcripts of different sizes. The smallest fragment appears to have the same size as the control, *in vitro* transcribed, α-UTP labelled C8-antiswitch RNA of size 49 nucleotides. This is the C8-antiswitch RNA released from the dual hammerhead construct as the result of the cis-cleaving function of the two ribozymes. The efficiency of production of this fragment from the dual hammerhead construct is calculated by densitometry to be 61%. The second largest fragment, with 20% intensity, appears to be approximately 100 nucleotides in length by comparison with the ladder. This transcript is believed to be generated by cis-cleaving activity of only one of the hammerhead ribozymes, whereas the second ribozyme is not cleaving. In this case, a C8-antiswitch attached to a hammerhead ribozyme (denoted as Hm-As) is generated with a length of 98 nucleotides. Based on the Hm-As RNA fragment gel mobility in Figure 2.16, it is not possible to tell which of the two flanking hammerhead ribozymes (5’ or 3’ hammerhead to the C8-antiswitch) is still attached to the C8-antiswitch since both hammerheads are of equal size and incompletely cleaved construct will generate products of the same length. The largest fragment on the gel appears to be the intact, uncleaved, dual hammerhead ribozyme flanking the C8-antiswitch with a length of 145 nucleotides. The efficiency of C8-antiswitch generated from the dual hammerhead construct appears to be 61% of the total transcribed RNA after a 2 hr transcription reaction.

To monitor the progress of the cleaving activity of the dual hammerhead ribozymes as a function of time to generate the C8-antiswitch RNA fragment with defined sequence ends, aliquots of the transcription reaction were take every half-hour
and loaded into a denaturing gel. The product percentages were calculated by
densitometry using ImageQuant. Figure 2.17 shows the percentages of the three products
generated from the dual-hammerhead construct cis-cleaving activity.

Figure 2.17 Cis-cleaving activity of the dual-hammerhead construct as a function of time

The cis-cleaving activity of the dual-hammerhead ribozymes is monitored
over 4 hrs as transcription reaction progress. The percentage of each cleaved
product at each time point is shown in the lower box.

According to Figure 2.17, as the transcription reaction progressed, the percentage
of the C8-antiswitch (denoted as AS) generated from the dual hammerhead construct
increased with time. Concurrently, the percentage of the uncleaved (referred to as Hm-
As-Hm) products, and the transcript products cleaved at one-end only (referred to as Hm-
As) decreased with time. It appears that after three hours, the percentage of the C8-
antiswitch (63%) cleaved from the dual hammerhead construct has reached a plateau. At
this time it appears the Hm-As and Hm-As-Hm transcripts level have also decreased to a
plateau level (28% and 6% respectively). After 3 hours of in vitro transcription, the C8-antiswitch transcript appears to be the dominant fragment generated from the dual-hammerhead construct.

2.5 Conclusion:

The target binding affinity of the designed C8-antiswitch was tested. The results show the C8-antiswitch is able to fold into predicted secondary structures in the presence and absence of the BDHP-PEG photochromic compound. The stabilized secondary structure of the C8-antiswitch dictates whether the target binding domain, the antisense domain, of this antiswitch is able to form complementary base pairs with the target mRNA sequence.

It is crucial for the C8-antiswitch to have defined sequence ends to allow proper folding and structure stabilization in vivo. The dual hammerhead ribozyme flanking the C8-antiswitch is designed to achieve this purpose.

The data from the compound uptake experiment suggests the uptake of BDHP-PEG into S. cerevisiae. Although fluorescence microscopy could not directly confirm intracellular accumulation of BDHP-PEG, the presence of a fluorescent contaminant in the BDHP-PEG stock solution, could be co-localizing with the non-fluorescent BDHP-PEG inside the cells.

In Chapter 3, the C8-antiswitch flanked by two hammerhead ribozymes is introduced into S. cerevisiae, and the C8-antiswitch is expected to be generated from the cis-cleaving activity of the dual-hammerhead construct transcribed in vivo. The photochromic ligand-dependant target binding activity of the C8-antiswitch will be tested
in *S. cerevisiae* to monitor any regulatory effects of the C8-antiswitch on EGFP mRNA expression.
Chapter 3: Light responsive regulation of reporter gene expression by C8-antiswitch in *Saccharomyces cerevisiae*

3.1 Introduction:

The budding yeast *Saccharomyces cerevisiae* is a non-pathogenic, safe, biological model organism, and the first eukaryote to have its genome sequenced in 1996. The basic genetics, biochemistry and cellular biology of *S. cerevisiae* have placed this microorganism in the forefront of eukaryotic cellular and molecular biology research. Furthermore, yeast has been used as a model organism to study diseases in mammalian and higher eukaryotic systems. A large-scale molecular analysis in human diseases has become possible since at least 31% of yeast proteins are discovered to have human homologues, and 50% of human genes involved in heritable diseases are found to have homologous genes in yeast (Hartwell L.H. 2004, Conconi A. 2008).

Features such as autonomously replicating plasmids, whole cell transformation system, and rapid colony formation on simple media, makes this a tractable model organism (Conconi A. 2008). DNA transformation in yeast is extensively used for studies involving gene cloning and genetic engineering techniques. Scientists have identified and studied structural genes and their corresponding genetic traits by complementation from plasmid libraries. Plasmids are introduced into yeast either by genome integration through homologous recombination or as autonomously replicating molecules. Also unique in yeast, is the ability to transform synthetic oligonucleotides to
produce altered forms of genes and proteins. This approach has been exploited to study regulation of genes, structure-function relationships of proteins, analysis of chromosome structure, and to answer other general questions in cell biology (Sherman F. 1997). Genetic studies and transformations are performed with distinct varieties of yeast strains with different properties, carrying different mutations and genetic markers.

In *S. cerevisiae* vegetative cell division is the major method of reproduction and it involves mitosis followed by budding, hence they are referred to as budding yeast. A daughter cell is formed as an outgrowth or a bud from the mother cell. The budding process continues by mitotic division, followed by cell-wall formation of the daughter cell, and terminates by cell separation and the daughter cell budding-off of the mother cell. A mother cell can undergo 20-30 budding cycles, and each process leaves a bud scar on the mother’s cell wall (Sherman F. 1997). Yeast cells exist as either haploid or diploid cells, with 16 and 32 chromosomes respectively (Sherman F. 1997). Diploid cells are typically larger than haploid cells with 5-6 μm and 4 μm diameters respectively (Sherman F. 1997).

Sexual reproduction in *S. cerevisiae* is an alternative mode of reproduction, which usually occurs under conditions of nutrient deficiency (Sherman F. 1997). The mating process can only occur between haploid cells with two different mating types. Figure 3.1 shows reproduction in *S. cerevisiae*. A simple sexual differentiation is available with mating type ‘a’ and ‘α’ on a single locus, MAT. Haploid cells can mate with other haploid cells of the opposite mating type (MATa cell can mate with MATα cell and vice versa) and a diploid cell is produced. Diploid cells undergo meiosis to produce four haploid spores, two a and two α spores. These haploid spores could
germinate into haploid cells, which could then undergo vegetative cell division by mitosis as described previously. Haploid spores can be mechanically separated by micromanipulators and stably maintained as haploid strains (Sherman F. 1997).

Figure 3.1 Saccharomyces cerevisiae reproduction

S. cerevisiae, or budding yeast, can exist as haploid or diploid cells. Haploid cells have two mating types: MATα (shown here as brown cells), or MATα (shown as purple cells) mating types.

1- the vegetative cell division and budding process as a result of mitosis in both haploid and diploid cells.

2- mating or conjugation between a MATα and a MATα haploid cell to produce diploid cells (shown as light purple).

3- diploid cells undergoing meiosis to produce four haploid spores, two MATα and two MATα cells, which could germinate into haploid cells.
(Image obtained from http://en.wikipedia.org/wiki/Yeast as public domain)

Yeast culture growth involves three phases. The initial growth phase is called the lag phase during which the genes required for cell growth are turned on. The following phase is called the log phase, or the exponential growth phase, since the cells population number increases exponentially. The log phase consists of three stages: the early-log, mid-log, and late-log stage. Finally, the cells enter a stationary phase upon nutrient depletion in the media, and consequently cell growth terminates. Light scattering (Day
A. et. al. 2004) by a spectrophotometer at 600 nm is used to determine the phase and the number of cells in a culture. The doubling time for wild type haploid cells during their exponential growth phase in complete YPD medium (1% yeast extract, 2% peptone, and 2% glucose) is 90 minutes and approximately 140 minutes in synthetic media at 30°C. At optimal density of 600nm (OD$_{600}$), and 1mL of yeast culture media with an OD$_{600}$ measurement of 1.0 corresponds to approximately 2-3x10$^7$ cells in that volume of media (Day A. et. al. 2004). However, this number varies between different strains and different conditions (Day A. et. al. 2004).

Exogenous DNA can be introduced into yeast cells to be inherited and expressed in cells by the process of transformation. Similarly, synthetic DNA sequences can be cloned into shuttle vectors under the influence of promoters to express modified genes and proteins in yeast cells. Shuttle vectors are plasmid vectors that contain sequences required for plasmid propagation and selection in yeast as well as in E.coli. This characteristic is convenient for in vitro amplification and genetic alterations. A shuttle vector contains a sequence for an E.coli origin of replication (ori), for its replication and maintenance in E.coli, and an antibiotic resistant marker, conferring resistance to an antibiotic such as ampicillin (Amp$^R$) or tetracycline (Tet$^R$). The antibiotic resistant gene allows for selection of the transformed cells. A Shuttle vector also contains nutrient markers to complement specific auxotrophic mutations in yeast cells. For example URA3, HIS3, TRP1 nutrient markers on a shuttle vector allows for transformed yeast strains with mutations in ura3, his3, and trp1 to survive on a nutrient limiting media and hence be selected (Sherman F. 1997). Shuttle vectors also contain a yeast origin of replication. Depending on the type of origin, the frequency of replication, or the copy
number, of the shuttle plasmid varies in yeast. High copy-number plasmids are called YEp vectors with a 2 µ ori, with the frequency of replication between 10-40 copies per cell (Sherman F. 1997). YCp vectors are low copy plasmid vectors in yeast, producing 1-3 copies of the plasmid per cell. They contain centromere sequences, CEN, and an autonomously replicating sequence, ARS (Sherman F. 1997).

With all the properties and advantages offered by yeast as a model organism, we decided to use this versatile eukaryotic organism to explore the effects of the C8-antiswitch in regulating gene expression in yeast. This chapter discusses the construction of plasmids to introduce the C8-antiswitch into yeast cells as well as the construction of an expression plasmid to express a reporter protein, EGFP, in yeast cells. It also examines my approach to explore the possibility of the light-induced regulatory effect of this riboregulator on the expression of EGFP as a reporter protein.
3.2 Expression construct of the C8-antiswitch and the reporter protein in *S. cerevisiae*

As described in Chapter 2, I designed an antiswitch, called the C8-antiswitch, using a previously selected aptamer, the C8 aptamer, for specifically binding to one isomeric form (BDHP-PEG) of our photochromic compound with affinity of 2.7 (+/- 0.4) μM (Lee H.W. *et. al.* 2007). The stem sequence of the C8-antiswitch is designed to base pair with the mRNA sequence of EGFP, as a reporter protein. Bioinformatic sequence homology search using *S. cerevisiae* WU-BLAST2 Search (http://www.yeastgenome.org/cgi-bin/blast-sgd.pl) did not reveal significant homology between the C8-antiswitch stem sequence and other genomic DNA sequences in *S. cerevisiae*.

Fluorescence Activated Cell Sorting (FACS) was used to confirm EGFP expression from pTarget plasmid in yeast cells. FACS is a specialized flow cytometry instrument, which is practical for fast analysis of population of cells and it provides quantitative recording of light scattering and fluorescent characteristics of individual cells one cell at a time. A flow cytometer can be used to study cells between 1 and 15 micrometers in diameter. Therefore, it can accommodate *S. cerevisiae* cells that usually have a diameter of 5-10 micrometers. In addition to flow cytometry, this scientific instrument is able to sort a heterogeneous mixture of cells by physically collecting cells of particular interest into two or more containers for further analysis. This is achieved by breaking the fluidic stream of cells into droplets containing single cells, shown in Figure 3.2. The droplets will be electrically charged in a high-voltage field between deflection plates. The charged droplets are sorted by electrical deflection into collection containers.
Flow cytometry is extensively used for immuno-phenotyping. Immuno-phenotyping involves classification of cells based on identifying their specific antigens. This is achieved by conjugating fluorophore molecules to antibodies and adding these conjugated antibodies to a mixture of cells to identify particular cells based on antibody-antigen interactions on cell surface. Other applications of flow cytometers are in ploidy analysis, cell counting, and GFP expression analysis. The general basis of all these analytical studies using a flow cytometer is based on separating cells into a single-cell line in a sheath fluid stream and passing the cells through a laser beam one at a time to capture the scattered light and fluorescence emitted from each cell. The signals are converted into digital data, which is analyzed by flow cytometry software to generate
information about the size, complexity, phenotype and health of the cells within a population.

As shown in Figure 3.2, light scattered as each cell passes through the laser beam is collected at two angles: Forward Scatter (FSC) and Side Scatter (SSC). Forward scatter measures scattered light in the direction of the laser path (the forward direction). The magnitude of FSC is proportional to the size of the cell, with small cells producing a small amount of FSC and large cells producing a large amount of FSC as they pass through the laser. The intensity of the scattered light is converted into voltage pulse and is quantified by a detector in the cytometer. Side scatter (SSC) measures scattered light at 90 degrees to the laser path and is collected and quantified by a separate detector. The amount of side scatter light (SSC) corresponds to the degree of granularity and structural complexity inside of the cell. In immuno-phenotyping applications, as the cells pass through the laser beam of appropriate wavelength, the fluorophore molecules attached to them are excited and they fluoresce. The fluorescent light travels along the same path as the side scatter signal. At this stage filters direct the emitted fluorescent light to the appropriate detectors where the light is collected and converted into a voltage pulse. The magnitude of the voltage pulse is proportional to the amount of fluorescence emitted as each cell crosses the path of the laser. Similar events take place when the magnitude of GFP fluorescence is measured from the cells except in this case the fluorophore molecule is GFP. All the data collected by the cytometer is recorded and could be presented graphically as histograms by flow cytometry software, which also allow analysis of multiple parameters from a sample on two or three-dimensional plots.
3.3 Material and Methods:

3.3.1 Materials:

The yeast expression plasmids, YCplac22, and YCplac33 were generously donated by Chris Beh lab. EGFP was PCR-amplified from a template plasmid, pAGX2, also donated by Chris Beh lab. Hm-As-Hm template sense and anti-sense DNA templates were purchased from Integrated DNA Technologies (IDT). The GAL1 promoter was PCR amplified from *S. cerevisiae* genomic DNA (Invitrogen). The pGPD promoter sequence was subcloned from pCB74 donated by Chris Beh lab. PJET1.2 PCR cloning kit, and Fast Digest restriction endonucleases were purchased from Fermentas. T4 DNA ligase, and Taq Polymerase were purchased from Invitrogen. Deoxynucleotide triphosphates (dNTPs) were purchased from Pharmacia. QIAgen Gel extraction kit and plasmid miniprep kit were used for PCR gel extraction and plasmid miniprep and were used according to manufacturer instructions. All synthetic DNA sequences and primers were ordered from Integrated DNA Technologies (IDT) and are listed in the Appendix section.

3.3.2 pSwitch and pTarget plasmid construction:

pSwitch plasmid construction:

Polymerase chain reaction (PCR) was performed in 100 ul reaction volume using final concentrations of 20 mM Tris pH 8.4, 50 mM KCl, 2 mM MgCl$_2$, 0.2 mM of each dATP, dTTP, dCTP, dGTP, 2.5 units of Taq polymerase enzyme, 0.5 μM of each Hm-As-Hm forward (5’- CTA ACT AAG GTA CCG TGA CTG ATG AGT CCG-3’, KpnI sequence is underlined) and Hm-As-Hm reverse primers (5’- CTA ACT AAC TGC AG
TGG TTT CGT CCT CAC GGA-3’, PstI sequence is underlined), and 200 pmol of synthetic Hm-As-Hm DNA template. The Hm-As-Hm DNA template strand sequences are listed in the Appendix section. The Hm-As-Hm template (or anti-sense) DNA sequence is 5’-CTA ACT AAC TGC AGG GTC TTT CGT CCT CAC GGA CTC ATC AGA CCG GAG TCG ACT CCG GTT GAC CAT GGT GAG CGA GGG CTC ACC ATG GTG TAG GAT GAC CAC CAT GGT GAG ACG GTA GGA ATT CCT ACC GTT TCG TCC TCA CGG ACT CAT CAG TCA CGG TAC CTT AGT TAG-3’ (PstI and KpnI restriction sequences are underlined at the 5’ and 3’ ends respectively).

All the reaction components except for Taq enzyme were mixed in a 0.6 mL microcentrifuge tube by pipetting and incubated in a thermal cycler for 3 minutes at 94°C. The incubation temperature was then reduced to 80°C and Taq polymerase was added to the reaction and mixed by pipetting. 30 cycles of PCR reaction was carried out by heating the reaction for 45 seconds at 94°C, followed by 30 seconds at 55°C, and 90 seconds at 72°C. At the end of the 30 cycles the reaction was incubated at 72°C for an additional 10 min to allow for polymerization of short, incomplete DNA fragments to full-length products. The PCR products were gel purified using 0.8% agarose gel and QIAgen PCR gel extraction kit according to manufacturer protocols. The products were resuspended ddH2O and stored in -20°C. The PCR-amplified Hm-As-Hm DNA fragment was digested with KpnI and PstI restriction enzymes (Fermentas) and gel purified, using 0.8% agarose gel and QIAgen gel purification kit, to create sticky ends for subsequent cloning into pJET 1.2 vector.

A PJET 1.2-GAL1 vector (Fermentas) that already contained a KpnI restriction site from a previous GAL1 blunt-end cloning experiment (using EcoRI and KpnI
restriction sites on the 5’ and 3’ ends of GAL1 respectively), was digested with PstI and KpnI and gel purified to remove the previously subcloned fragment, and to create KpnI and PstI restriction site compatible with the PCR-amplified Hm-As-Hm DNA. Gel purification was performed using 0.8% agarose gel and QIAgen gel purification kit. The Hm-As-Hm DNA was ligated into pJET 1.2 cloning vector, using their compatible KpnI and PstI restriction sites.

The ligation reaction was performed in a 20 ul reaction volume containing 3:1 molar ratio of insert DNA : vector, and a final concentration of 50 mM Tris-HCl (pH 7.6), 10 mM MgCl2, 1 mM ATP, 1 mM DTT, 5% (W/V) polyethylene glycol-8000, and 1 unit of T4 DNA ligase enzyme. The reaction components were mixed in a 0.6 mL microcentrifuge tube by pipetting. The ligation reaction was carried out overnight at 15°C and stored at -20°C. The ligation reaction was transformed into E.coli cells and plated onto LB-Amp plates. The transformation protocol is described in section 3.3.3. Successful ligation into pJET1.2 plasmid disrupts the lethal gene eco47IR (Figure 3.3A), which enables positive selection of the recombinants. The ligated pJET1.2 plasmid was extracted from the recombinant colonies using QIAgen miniprep kit according to manufacturer’s protocols. Successful ligation of Hm-As-Hm into pJET 1.2 plasmid was confirmed by restriction endonuclease double digestion, and by plasmid sequencing using the Hm-As-Hm forward and reverse primers. This intermediate plasmid is referred to as pJET-As.

The pGPD promoter sequence was digested out of pCB74 plasmid, donated by Dr. Chris Beh lab, by a HindIII and KpnI double digestion. The pGPD promoter fragment was gel purified using 0.8% agarose gel and QIAgen gel purification kit. The
pJET-As plasmid was also digested with KpnI and HindIII and gel purified to create compatible sticky ends with pGPD promoter fragment. The pGPD promoter was ligated into pJET-As plasmid using their compatible restriction sites. This second intermediate plasmid is called pJET-As-pGPD. This plasmid was transformed into *E. coli* as described in section 3.3.3 and the pJET-As-pGPD plasmid was extracted from the recombinant cells using QIAgen miniprep kit. Successful insertion was confirmed by KpnI and HindIII restriction endonuclease double digestion and also by plasmid sequencing using the primer called pJET-As Forward: 5’-GTG CCA CCT GAC GTC GTG ACT GAT -3’.

In the next step the pJET-As-pGPD plasmid was digested with PstI enzyme to cut out the As-pGPD fragment from pJET 1.2 plasmid. I was able to use a single restriction endonuclease to digest As-pGPD fragment on both ends since pGPD contained a PstI restriction sequence immediately preceding the HindIII restriction sequence. The As-pGPD was gel purified using 0.8% agarose gel and QIAgen gel purification kit. The yeast expression plasmid, YCplac22 was also digested with PstI enzyme (Fermentas) to linearize this plasmid. 1 μL of Alkaline phosphatase enzyme was added directly into the PstI digestion reaction to remove the 5’-end phosphate from the linearized plasmid. This prevents re-ligation of the plasmid. The digestion reaction was purified using phenol-chloroform, followed by ethanol precipitation. The linearized plasmid was also gel purified prior to the next ligation reaction. The As-pGPD fragment was ligated into YCplac22 using their compatible PstI sticky ends by T4 DNA ligase. This plasmid is referred to as pSwitch plasmid. The pSwitch plasmid was transformed into *E.coli* cells and the cells were plated onto LB-Amp plates. The pSwitch plasmid was extracted from
E. coli cells using QIAgen miniprep kit and successful ligation reaction was confirmed by KpnI restriction endonuclease digestion. The pSwitch plasmid with the correct orientation of pGPD-As insert was transformed into W303α (MATα/ura3-1/leu2-3,11/his3-11/trp1-1/can1-100/ade2-1) yeast cells according to the protocols in section 3.3.4.

**pTarget plasmid construction:**

Polymerase chain reaction (PCR) was performed in 100 μL reaction volume for GAL1 promoter amplification from yeast genomic DNA template using the same reaction conditions as explained in pSwitch plasmid construction, except 0.3 μg of yeast genomic DNA was used as template DNA for GAL1 amplification. The primers used for GAL1 PCR amplification are GAL1 forward primer: 5’-CCT GAA TTC TTT ATA TTG AAT TTT CAA AAA TTC TTA CTT TTT TTT TGG-3’ (EcoRI restriction sequence is underlined), and GAL1 reverse primer: 5’-CCT GGT ACC TAT AGT TTT TTC TCC TTG ACG TTA AAG TAT AG-3’ (KpnI restriction sequence is underlined). The PCR-amplified GAL1 promoter product was gel purified using a 1% agarose gel and QIAgen gel purification kit and re-suspended in ddH2O and stored at -20°C.

The GAL1 PCR product was treated with a blunting enzyme from CloneJet PCR Cloning Kit (Fermentas) according to the manufacturer’s protocols to create GAL1 PCR products with blunt ends containing EcoRI and KpnI restriction sequences on the 5’ and the 3’-ends of the GAL1 DNA respectively. The pJET 1.2 cloning vector supplied with the CloneJet PCR Cloning Kit is a linearized vector with blunt end cloning sites. The blunt GAL1 PCR product (350 ng) was ligated into the blunt pJET 1.2 vector (50 ng) using 5 units T4 DNA ligase in a 20 μL reaction volume. The ligation reaction was
incubated at room temperature for 15 minutes, followed by phenol-chloroform extraction and ethanol precipitation. The reaction products were re-suspended in ddH2O and stored at -20°C. This intermediate ligation product was referred to as pJET-GAL. The pJET-GAL plasmid was transformed into E.coli cells using electroporation as explained in section 3.3.3 and plated onto LB-Amp plates and recombinant cells were cultured in LB-Amp media. The pJET-GAL plasmid was extracted from recombinant cells using QIAgen plasmid miniprep kit according to manufacturer’s protocols. Successful GAL1 ligation into pJET1.2 was confirmed by KpnI and EcoRI restriction double digestion, and also by sequencing using GAL1 forward primer.

The EGFP DNA fragment was PCR amplified from pAGX2 plasmid using EGFP forward primer: 5’-TAA GGT ACC ATG GTG AGC AAG GGC GAG-3’ (KpnI sequence is underlined) and EGFP reverse primer: 5’-TAA CTG CAG CTG CTA TTA CTT GTA CAG CTC GTC CA-3’ (PstI sequence is underlined). The PCR reaction was carried out using the same reaction conditions as already explained earlier in this section except 0.3 µg of pAGX2 plasmid was used as template DNA for EGFP amplification. The PCR-amplified EGFP DNA fragments were digested with KpnI and PstI to create sticky ends for cloning into pJET-GAL plasmid. The pJET-GAL plasmid was also digested with KpnI and PstI restriction endonucleases to create compatible sticky ends with EGFP DNA insert. The double-digested EGFP and pJET-GAL products were gel purified from 1% agarose gel and QIAgen gel purification kit. The ligation reaction was carried out in a 20 µL reaction volume using 70 ng double-digested pJET-GAL and 210 ng EGFP DNA insert and 5
units of T4 DNA ligase. The ligation reaction was incubated overnight at 15°C, followed by phenol-chloroform extraction and ethanol precipitation. The ligated products were resuspended in ddH2O and stored at -20°C. This second intermediate plasmid is referred to as pJET-EGFP-GAL1. This plasmid was transformed into E.coli and extracted using QIAgen plasmid miniprep. Successful ligation was confirmed by EcoRI and PstI restriction endonuclease double digestion and by sequencing using EGFP forward primer.

In the next cloning step, the EGFP-GAL1 fragment was cut out of pJET plasmid by EcoRI and PstI, and gel purified using 0.8% agarose gel and QIAgen gel purification kit. The yeast expression plasmid YCplac33 was also digested with the EcoRI and PstI to create compatible restriction sites for cloning EGFP-GAL1 DNA. Ligation reaction was carried out in a 20 µL reaction volume using 130 ng double-digested YCplac33 plasmid and 400 ng EGFP DNA insert using T4 DNA ligase. The ligation reaction was incubated overnight at 15°C, followed by phenol-chloroform extraction and ethanol precipitation. The ligated products were resuspended in ddH2O and stored at -20°C. This plasmid is referred to as pTarget plasmid. The pTarget plasmid was transformed into E.coli cells and the cells were plated onto LB-Amp plates. The pTarget plasmid was extracted from E.coli cell culture using QIAgen miniprep kit and EcoRI and PstI restriction endonuclease double digestion confirmed the successful ligation reactions. The Target plasmid was transformed into W303α yeast cells according to the protocols in section 3.3.4.
3.3.3 Transformation of competent \textit{E.coli} cells by electroporation

Before starting the reaction sterilized 1mm electroporation cuvettes were chilled on ice. Frozen electro-competent DH5\(\alpha\) \textit{E.coli} cells were thawed on ice and 50 ul of cells were added to 1 ul of desired ligated pJET1.2 plasmid. The cells were pulsed by CELL-PORATOR \textit{E.coli} System set to 1.8kV. Immediately after pulsing the cells 500 ul of LB (Luria–Bertani) liquid media was added into the cuvette and mixed with the cells by pipetting up and down. The LB media containing the pulsed \textit{E.coli} cells was transferred into 1.5 mL microcentrifuge tubes and the tubes were incubated at 37\(^{\circ}\)C water bath for 30 minutes. The transformed \textit{E.coli} cells were plated onto LB-Ampicillin plates and incubated overnight at 37\(^{\circ}\)C. The transformed \textit{E.coli} cells containing the pJET1.2 plasmid are ampicillin resistant. The lethal gene \textit{eco47IR} also allows for positive selection of the recombinant cells. The recombinant plasmid was extracted from bacterial culture using QIAgen miniprep kit according to manufacturer protocol and DNA insertion was confirmed by plasmid sequencing.

3.3.4 Yeast cell transformation

The ‘switch’ and ‘target’ plasmids are transformed into yeast cells of W303\(\alpha\) strain (MAT\(\alpha\)/ura3-1/leu2-3,11/his 3-11/trp1-1/can1-100/ade2-1), with mutant \textit{trp1} and \textit{ura3} mutation markers. The yeast cells were transformed by a chemical reaction referred to as the “Lazy Bones” plasmid transformation. For this transformation reaction W303\(\alpha\) cells were plated and grown for 1-2 days at 25\(^{\circ}\)C on Synthetic Complete (SC) media plates supplemented with all the commonly encountered auxotrophies prepared according to protocols (Kaiser C. \textit{et al.}, Methods in yeast Genetics, CSHL Press, 1994). Using a sterile toothpick a single yeast colony (~2 mm diameter) was placed in a 1.5 mL
microcentrifuge tube. 32 μg of salmon sperm carrier DNA (heated to 95°C in TE buffer for 5 minutes and rapidly cooled on ice) was added to the reaction tube, followed by the addition of 4 μL (total of 0.1-2 micrograms) of ligated plasmid DNA (YCplac 22 or YCplac 33). The content was mixed by pipetting. PLATE solution (40% Polyethylene Glycol (MW 3350), 0.1 M Lithium Acetate, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA) was added to the reaction tube. The transformation reaction was incubated at room temperature for 72 hours. The cells were pelleted by centrifuging for 3 minutes at 5000 rpm. The supernatant was carefully removed and discarded and the transformed yeast cells were resuspended in 200 μL of sterile ddH₂O. The transformed cells were spread on selective synthetic dropout media plates and grown at 25°C.

Five transformed cell lines were established by electroporation: 1) Cells with pSwitch1 plasmid, 2) cells with mutant pSwitch2 (containing mutant C8-antiswitch element), 3) cells with pTarget plasmid, 4) cells with pSwitch1 and pTarget, 5) cells with pSwitch2 and pTarget.

3.3.5  Cell culture preparation for flow cytometry

A single colony of each transformed strain was inoculated to OD₆₀₀~0.03 in 3mL of synthetic drop out media, with the appropriate amino acid drop out, and containing a 4% raffinose sugar content in the media. The synthetic drop out media supplemented with all the commonly encountered auxotrophies prepared according to protocols (Kaiser C. et al., Methods in yeast Genetics, CSHL Press, 1994) and 100 mg/L of adenine sulfate was added to the media to reduce the auto-fluorescence of W303α yeast cells. These cells were cultured overnight to an OD₆₀₀~1.0. At this point the cell cultures were back
diluted to an OD\textsubscript{600}~0.03 and each cultured strain was divided into three 1mL cultures and transferred into 1.6mL eppendorf tubes to be treated with either BDHP-PEG, BCPD-PEG or no compound. 1mM final concentration of the appropriate compound was added to the 1mL cell cultures, with extra synthetic drop out media added for samples with no compound treatment, and the cell samples were cultured again overnight at 30°C on an eppendorf rotator for aeration, in the presence or absence of compounds. Eppendorf tubes containing BDHP-PEG compound were wrapped with aluminium foil to block incidental visible light irradiation and isomerization of this compound. On the other hand, test tubes containing BCPD-PEG compound were exposed to light irradiation from a hand held lamp throughout the overnight incubation period to maintain this compound in this isomeric form. On the third day, each culture was back diluted to OD\textsubscript{600}~0.6 and 2% final concentration of galactose was added to each cell culture to induce EGFP expression in the strains with pTarget plasmid. After 3 hours galactose induction, cells were pelleted by 1 min centrifugation at 6000 rpm, the culture media containing the compounds was discarded, and the cell pellets were washed five more times with PBS buffer pH7.0. The cells were re-suspended in 1ml of fresh PBS buffer and transferred to FACS tubes for flow cytometry analysis.

3.3.6 Flow cytometry

To prepare the cells for FACS analysis the washed cells were resuspended in 1 mL fresh PBS buffer, pH 7.4, and 200 μL of this cell suspension, approximately 1.0x10\textsuperscript{6} cells, was transferred into UltidentBrand polystyrene culture tubes, product# 170-T3200, for FACS analysis. FACS analysis was performed with Becton Dickinson FACS ARIA I using an excitation wavelength of 488 nm and a bandpass filter of 530/30 nm for EGFP
fluorescence analysis. Data analysis was performed with the FlowJo Software version 6.3.2.

Propidium iodide staining was performed using propidium iodide (PI) added to a final concentration of 4 µg/mL to cell cultures resuspended in PBS. The cells were stained with PI for 20 minutes after which the red fluorescence of this dye staining the cells was measured using BD FACS ARIA I, 536 nm excitation wavelength, and 695/40 bandpass filter.

3.3.7 Endogenous C8-antiswitch RNA detection by Northern blotting

Endogenous small RNA was purified from 100 mL yeast cell cultures with pSwitch and pTarget plasmids. This was achieved by yeast total RNA extraction followed by small RNA enrichment by mirVANA miRNA isolation kit (Ambion). The cells were inoculated and grown overnight in synthetic drop out media, with the appropriate amino acid drop out, and containing 4% raffinose sugar content in the media. These cells were cultured overnight to an OD$_{600}$~1.0. Approximately 1x10$^9$ cells were spun down and resuspended in 1 mL H$_2$O and transferred to 15 mL Falcon tubes and H$_2$O was removed after a brief centrifuging. The cells were resuspended in 1 mL LETS-1% solution (0.1 M LiCl, 0.01 M EDTA, 0.01 M Tris-HCl, pH 7.6, 1% SDS) and lowered into liquid nitrogen to freeze the cells. 0.4 grams of glass beads was added to the frozen cell pellet, followed by addition of 0.2 mL acid Phenol:Chloroform (Ambion). The cells were vortexed at top speed for 3 minutes and centrifuged for 5 minutes. The aqueous phase was transferred into a new falcon tube and total yeast RNA was extracted by adding an additional 1 mL acid Phenol:Chloroform, followed by vortexing for 3 minutes
and centrifuging for 5 minutes. The last two steps were repeated until the interface was clear.

The aqueous phase containing the purified total yeast RNA enriched for small RNA fragments using the mirVANA miRNA Isolation Kit according to manufacturer protocols. 20 μg of the enriched small RNAs was run on a 10% denaturing polyacrylamide gel and northern blot was carried out using 5′-32P labelled DNA probes complementary to the full length of C8-antiswitch RNA. The northern blotting procedure was carried out according to Molecular Cloning Manual (Molecular Cloning, 3rd edition, by Joseph Sambrook and David W. Russell. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 2001).

The enriched small RNA purified from pSwitch+pTarget cells was also probed with 5′-32P labelled DNA probe complementary to antiswitch RNA, AS probe 5′-AC CAT GGT GAG CGA GGG CTC ACC ATG GTG TAG GAT GAC CAC CAT GGT GA-3′, for detecting the endogenously expressed antiswitch RNA. 18-nt DNA probe with the sequence 5′-GTC TGG CGC TCT CCC AAC -3′ was used for detecting phenylalanine tRNA from yeast RNA. The probes were 5′ end-labelled with [γ-32P]-ATP (Perkin Elmer) using standard phosphorylation protocols of T4 Kinase Kit (Invitrogen), and then re-purified by a 8% denaturing PAGE, visualized by autoradiography and eluted from the gel by crush-soaking in a buffer at 4°C containing 10mM Tris-HCl (pH8.0 at 23°C) and 0.1 mM EDTA followed by ethanol precipitation.

The *in vitro* transcribed Hm-As-Hm RNA used as detection control in Figure 3.8 and 3.9A was transcribed as explained in section 2.3.6.
3.4 Results and discussion:

To express the C8-antiswitches in *S. cerevisiae*, I planned to use a noncoding RNA expression construct similar to a previously described system shown in Figure 2.15. The expression construct enables cloning of the hammerhead-antiswitch-hammerhead construct through unique 5’ and 3’ restriction sites, KpnI and PstI respectively. Once this construct is transcribed *in vivo*, *cis*-cleavage activity of the flanking hammerhead ribozymes results in C8-antiswitch RNA release with defined 3’ and 5’ ends.

To express the C8-antiswitch in *S. cerevisiae*, I used a noncoding RNA expression construct, referred to as pSwitch, to transcribe the C8-antiswitch *in vivo* with defined sequence ends as previously described in Figure 2.15. The hammerhead-C8antiswitch-hammerhead (referred to as Hm-As-Hm) double-stranded DNA was PCR amplified and two different unique restriction sequences were added to the 5’ and 3’ ends of the amplified DNA by primer extension. Different restriction sites at each end determine the orientation of DNA insertion into the yeast shuttle plasmid. After transformation of this shuttle plasmid into *S. cerevisiae*, this construct is expected to be transcribed *in vivo* and the *cis*-cleaving activity of the flanking hammerhead ribozymes results in C8-antiswitch RNA release with defined 3’ and 5’ ends.

The PCR-amplified Hm-As-Hm DNA was inserted downstream of the pGPD (Glyceraldehyde-3-Phosphate Dehydrogenase) promoter (Bitter G.A. *et al.* 1984). The yeast shuttle plasmid used for cloning the promoter and the Hm-As-Hm construct is a low copy number plasmid, the CEN plasmid, expected to produce 1-3 copies of the plasmid per cell (Sherman F. 1997).
A second yeast expression plasmid was constructed by cloning the Enhanced Green Fluorescent Protein (EGFP) gene under the control of a galactose-inducible promoter (GAL1). An inducible promoter grants us more control to keep the expression levels of this fluorescent reporter protein at sub-maximal levels since over-expressed EGFP could conceal the regulatory effect of the C8-antswitch in vivo and small changes in the expression of the reporter protein will remain undetected. On the other hand, the pGPD is a relatively strong constitutively active promoter (Mumberg D. et. al. 1995), specifically chosen in this system to constitutively express the Hm-As-Hm construct which could lead to higher C8-antswitch mRNA levels in the yeast cells than the pGAL1 inducible promoter (Mumberg D. et. al. 1995; Piruzian E.S. et. al. 2002) selected for EGFP expression in the cells. The plasmid expressing the EGFP reporter protein is referred to as pTarget. It is also a low copy, CEN plasmid; however, the nutrition marker on pTarget is different from pSwitch. Yeast cells that are auxotrophic for both of these nutrition markers but have obtained both plasmids after transformation, are able to grow on a dropout media plate. EGFP sequence was PCR amplified using primers with two different restriction sequences to be incorporated at the 5' and 3' ends of the amplified EGFP DNA. Successful cloning of the pGPD-Hm-As-Hm and pGAL1-EGFP was confirmed by restriction enzyme digestions and sequencing of pSwitch and pTarget.

3.4.1 Yeast expression plasmid construction:

The pGPD (Glyceraldehyde-3-Phosphate Dehydrogenase) promoter was obtained from pCB74 plasmid donated by Beh lab by restriction endonuclease digestion. The pGPD promoter was subcloned into PCR cloning vector, pJET1.2, by HindIII and KpnI restriction site ligation. The PCR amplified hammerhead-Antiswitch-hammerhead (Hm-
As-Hm) DNA construct was cloned downstream of pGPD promoter into pJET1.2 by restriction site ligation through its KpnI and PstI restriction sites at the 5’ and 3’ ends respectively.

The ligated pJET-As-pGPD plasmids are transformed by electroporation into *E. coli* cells. PJET1.2 plasmid contains an *E. coli* origin of replication (pMB1) and selection marker for ampicillin resistance, which allows for colony growth on LB/Ampicillin plates. At the same time if there is no insertion into the pJET vector, re-circularized pJET1.2 vector expresses a lethal restriction enzyme (from eco47IR lethal gene as shown represented in Figure 3.3 A) after transformation and is not propagated. As a result, only recombinant clones containing the insert appear on culture plates. The cells were plated onto LB/Ampicillin plates overnight and positive colonies were picked and inoculated overnight. The ligated plasmids were purified by miniprep according to instruction manual from QIAGEN miniprep kit, and restriction double digestion was used to reconfirm correct orientation of insertion.
Figure 3.3  Map of pJET1.2 plasmid

A) pJET1.2 plasmid map from Fermentas.
B) Sticky-end cloning scheme for pSwitch plasmid. Hm-As-Hm is abbreviated to As in this figure. Note: pGPD in pJET plasmid contains HindIII followed by PstI restriction sites.
C) Sticky-end cloning scheme for pTarget plasmid.

The pGPD-HM-As-HM insert was then released from the PJET1.2 plasmid by restriction double digestion, using EcoRI and PstI restriction enzymes, and cloned into the multiple cloning site of the yeast expression plasmid, YCplac22 to generate pSwitch.
plasmid, as represented in Figure 3.3 B. YCplac22 was donated by Chris Beh lab and had a yeast ARS-CEN4 replication origin, which makes this plasmid a low copy, 1-2 copy number/cell. There is also a TRP1 selectable marker that allows for selection of transformed yeast cells that are auxotrophic for tryptophan synthesis. This plasmid construct allows Hm-As-Hm to be expressed from pSwitch constitutively and this allows for high levels of C8-antiswitch RNA to be available inside the cell before EGFP expression induction from a different plasmid, pTarget, containing the GAL1 promoter. In addition having the antiswitch expressed constitutively allows more time for the antiswitch RNA to be cleaved by the flanking hammerhead ribozymes and fold into its appropriate secondary structure in the presence or absence of the BDHP-PEG ligand.

A second YCplac type plasmid containing a subcloned yeast EGFP sequence was also constructed under the control of a subcloned GAL1 promoter. Chris Beh’s group donated the empty YCplac33 plasmid as well. After subcloning the EGFP and GAL1 promoter sequence this plasmid is referred to as pTarget as represented in Figure 3.4B. The GAL1 promoter sequence was PCR amplified using appropriate primers from yeast genomic DNA. The EGFP sequence was also PCR amplified from a plasmid, pAGX2, donated generously by Chris Beh lab. The cloning scheme of pTarget is similar to the scheme explained earlier in Figure 3.3B for pSwitch construction. pTarget has a different nutritional selectable marker than pSwitch. pTarget plasmid has a URA3 selectable marker. Having two different nutritional markers on pSwitch and pTarget allows for selection of transformed colonies that have acquired both plasmids after transformation into W303α cells with ura− and trp− mutations. The transformed cells will be able to survive on SC-Trp/Ura drop out media.
3.4.2 **Compound induced cell toxicity analysis**

Before measuring EGFP fluorescent level from compound-treated cells, these cells were stained with propidium iodide (PI) which is an intercalating fluorescent dye, excited by 488 nm laser light, and is used to assess for cell viability. Propidium iodide is able to permeate across the cell membrane of dead cells whereas it is not permeable into living cells. This test was performed to see whether compound treatment is toxic to cells or not. PI was added to a final concentration of 4 μg/mL to cell cultures resuspended in PBS. The cells were stained with PI for 20 minutes after which the red fluorescence of this dye staining the cells was measured using flow cytometry and 488 nm excitation laser. 10,000 cells were used from each sample and a PI histogram was generated for
each sample. PI staining was also compensated relative to negative, unstained cells, to create a minimum threshold for PI fluorescent signal from a population of cell.

The following propidium iodide histogram in Figure 3.5 was generated by 4% PI staining of pSwitch+pTarget cells incubated overnight in the absence and presence of both isomers of our photochromic compound individually. Figure 3.5A represents the PI staining histogram of cells incubated in the same media as sample cells in B) and C), but not treated with any compound, and this histogram is used as a negative control. Figure 3.5 B and C represent PI histograms of cells treated with BDHP-PEG isomer and BCPD-PEG isomer respectively. Relative comparison between our negative control histogram and lack of significant PI fluorescence level above $10^2$ indicates neither isomer of the compound is affecting cell viability after an overnight incubation.
Figure 3.5  Propidium iodide staining of compound-treated yeast cells

pSwitch+pTarget yeast cells were incubated overnight in the absence and presence of 1mM of BDHP-PEG or BCPD-PEG.
A) PI histogram of cells incubated in the absence of any compound. The majority of the cell population exhibit low PI fluorescence below $10^2$, which indicates PI is impermeable into these living cells.
B) PI histogram of cells incubated in the presence of BDHP-PEG.
C) PI histogram of cells incubated in the presence of BCPD-PEG.
No significant PI fluorescence histogram peak above $10^2$ in B) and C) is indicating the majority of the cell population is impermeable to PI and hence they are living cells.

3.4.3 Flow cytometry

To obtain preliminary EGFP induction data, and to confirm EGFP expression in yeast cells, yeast cell line with pTarget plasmid was used to induce EGFP expression from the GAL1 promoter. Yeast cultures were grown overnight at 30°C in SC media in the presence of 2% raffinose, a non-inducing sugar source that allows for culture growth however does not induce GAL1 promoter. Once the cell culture reached an OD$_{600}$~ 1.0 the cultures were back diluted to an OD$_{600}$~ 0.03 in a fresh SC/2% raffinose media and were transferred into 1.6 mL eppendorf tubes. 1mM final concentration of BDHP-PEG, BCPD-PEG, or no compound was added to this fresh media. The cells were incubated again overnight at 30°C on an eppendorf rotator, which allow for aeration. On the
following day each culture was back diluted to OD$_{600}$~ 0.6 and 2% final concentration of galactose was added to each cell culture to induce EGFP expression in the strains with pTarget plasmid. After 3 hours of EGFP induction by adding 2% galactose, EGFP levels were measured using FACS (Fluorescent Activated Cell Sorter). The Hm-As-Hm construct did not require induction since this element was under the control of the pGPD promoter, a constitutively active promoter.

To confirm EGFP expression from pTarget plasmid in yeast cells EGFP, the fluorescence levels of 10,000 cells per sample was measured. Untransformed W303α cells were used as a negative control for EGFP fluorescence. As positive controls, W303α cells were transformed with a plasmid that was already known to express EGFP under the influence of pGAL1 promoter in separate studies (Beh C. communications). This plasmid is referred to as pCB559, and it contains a 2μ, high copy number plasmid. Therefore, it is expected to express high levels of EGFP upon induction by galactose.

Figure 3.6 shows the EGFP fluorescence histograms generated by flow cytometry, using the FACS machine. The x-axis represents the fluorescence intensity level represented in a logarithmic value, and the y-axis represents the number of measured events. The histogram in red indicates the autofluorescence level of the negative control cells. The green histogram corresponds to the fluorescent intensity level generated from the positive control cells containing the pCB559 plasmid. The blue histogram in Figure 3.6 shown the fluorescence level generated by EGFP expression in pTarget cells. The highest percentage of negative control cells have the lowest (less than 10$^1$) fluorescent intensity level, while the positive control cells have a fluorescent intensity level greater than 10$^1$ and as much as 10$^4$. In comparison to the negative and positive control, pTarget
cells produce a low to intermediate EGFP fluorescence level between $10^1$ and $10^2$. Based on this experiment, relative to the negative and positive control EGFP histograms the pTarget cells containing the cloned EGFP sequence are able to express detectable levels of EGFP from the GAL1 promoter after 3 hours of galactose induction of cells in the culture media.

![Histogram of FACS analysis showing EGFP fluorescence measurements.](image)

(Red line) negative control W303α yeast cells. (Green) EGFP fluorescence from high copy# plasmid pCB559 containing GAL1-EGFP as my positive control. (Blue) EGFP fluorescence from my low copy# pTarget plasmid containing GAL1-EGFP.

After confirming EGFP expression from the transformed cells, I carried out an experiment to explore the effect of our photochromic compound on EGFP expression
level in cells containing pSwitch and pTarget plasmids. As described earlier in Figure 2.9, in the presence of BDHP ligand the C8-aptamer domain of the C8-antiswitch is expected to bind to the BDHP ligand. The free energy associated with the aptamer-ligand binding is believed to stabilize the aptamer stem, sequestering the antisense sequence within this stem. As a result in the presence of BDHP ligand, the target EGFP mRNA is expected to remain free to be translated and produce EGFP fluorescence in the cells. I expect the EGFP translation to be inhibited in the presence of the BCPD isomeric form of the photochromic compound. A similar outcome is expected in the absence of any ligand molecule in the culture media of cell transformed with both pSwitch and pTarget plasmids. In the latter two cases the C8-antiswitch is stabilized in a folded secondary structure in which the antisense domain of the antiswitch is free to bind to the target EGFP transcript at the start codon region and inhibit EGFP translation by rendering the ATG start codon inaccessible to the ribosome.

3.4.4 Compound-dependent flow cytometry and EGFP expression

Once it was established that the isomers of the photochromic compound were not affecting the cell viability, I moved on to analyze the potential effects of BDHP-PEG and BCPD-PEG on EGFP expression in pSwitch+pTarget yeast cells and the changes in fluorescence intensity levels of expressed EGFP in response to each isomer after compound uptake into the cells.

There are 3 strains of transformed W303α yeast cells for this experiment: 1) cells transformed with pSwitch plasmid, referred to as pSwitch cell 2) cells transformed with pTarget plasmid, referred to as pTarget cells 3) cells transformed with pSwitch and pTarget plasmids, referred to as pSwitch+pTarget cell. To conduct the compound-
dependant flow cytometry experiments, pSwitch cells and pTarget cells were used as negative and positive controls, respectively, for EGFP fluorescence. EGFP fluorescence in response to compound treatment from the control cells as well as from pSwitch+pTarget cell was analyzed by flow cytometry using FACS machine as described earlier.

Figure 3.7 shows the EGFP fluorescence histograms in the absence and presence of BDHP-PEG or BCPD-PEG. The X-axis represents EGFP fluorescent intensity and the Y-axis represents the percentage of cells from the total population of each sample. The red histogram is generated from flow cytometry of pSwitch cells used to represent negative control for EGFP expression. These control cells are void of pTarget plasmid and do not express EGFP protein. Therefore, the red histogram represents the background fluorescence or autofluorescence of the W303α cells. The blue histogram represents EGFP fluorescence from pTarget and pSwitch+pTarget cells. Fluorescent intensities were measured after 6 hours, 15 hours, and 24 hours after EGFP expression induction to monitor for maximum EGFP expression levels from the same sample of cells.
Figure 3.7  Flow cytometry and EGFP fluorescent histogram of compound treated pSwitch+pTarget yeast cells

A) EGFP histogram of pTarget (i) and pSwitch+pTarget (ii) cells not treated with compounds.
B) EGFP histogram of pTarget (i) and pSwitch+pTarget (ii) cells treated with BCPD-PEG compound.
C) EGFP histogram of pTarget (i) and pSwitch+pTarget (ii) cells treated with BDHP-PEG compound.
In each case, EGFP histograms were analyzed after 6 hours, 15 hours, and 24 hours of EGFP expression induction. Red histograms show fluorescent intensity of pSwitch cells, the negative control cells. Blue histograms show fluorescence measured from sample cells as noted in A), B), and C).
In histograms of Figure 3.7 A i) and ii) EGFP fluorescent is measured from pTarget and pSwitch+pTarget cells in the absence of compound. In this treatment condition, EGFP expression is expected to be reduced in pSwitch+pTarget cells relative to pTarget cells that are able to express EGFP without any inhibitory factors. However contrary to our expectation and regardless of the duration of EGFP induction, pSwitch+pTarget cells exhibit a similar intensity of EGFP fluorescence as the pTarget cells. This is suggestive of no inhibitory effect exerted by C8-antiswitch expressed from pSwitch plasmid in pSwitch+pTarget cells.

In histograms of Figure 3.7 B i) and ii), EGFP fluorescent intensities are measured by flow cytometry from pTarget cells and pSwitch+pTarget cells respectively in the presence of BCPD-PEG isomer. In the presence of this isomeric form of the compound, it is expected to see reduced EGFP expression levels due to the folding structure of the C8-antiswitch which allows its anti-sense domain to be free to target EGFP mRNA by base pairing and consequently blocking translation. However, EGFP histograms generated from pTarget cells as positive controls, with no inhibitory factors for EGFP expression, resemble the histogram of pSwitch+pTarget cells in the presence of BCPD-PEG. Again, contrary to our expectation, we are not able to conclude any inhibitory effect enforced by C8-antiswitch in the presence of BCPD-PEG.

In histograms of Figure 3.7 C i) and ii), EGFP fluorescent intensities are measured by flow cytometry from pTarget cells and pSwitch+pTarget cells respectively in the presence of BDPD-PEG isomer. This isomeric form of the compound is expected to have an opposite regulatory influence on EGFP expression level by allowing EGFP expression to proceed with no inhibitory effects from the antisense domain of the C8-
antiswitch. In the presence of BDHP-PEG, the aptamer domain of the C8-antiswitch is expected to bind to this ligand, which in turn induces an alternative secondary structure folding of the C8-antiswitch. In this alternative folding state, the antisense domain of the C8-antiswitch is sequestered in the aptamer domain and is no longer available to base pair with the EGFP mRNA sequence. As a result, EGFP mRNA is expected to be free to be translated and produce EGFP fluorescence. EGFP histograms generated from pTarget cells as positive controls with no inhibitory factors for EGFP expression resembles the histogram of pSwitch+pTarget cells in the presence of BDHP-PEG. This result suggests EGFP expression in the presence of BDHP-PEG is independent of C8-antiswitch activity.

3.4.5 Endogenous C8-antiswitch RNA detection

Since the C8-antiswitch seemed to have no regulatory effect on the EGFP expression level in both the absence and presence of BCPD-PEG, we decided to investigate whether this failure of function may be due to problems in C8-antiswitch expression from pSwitch plasmid inside the cells. The C8-antiswitch is flanked by a hammerhead ribozymes on both its 5’ and 3’ ends, each cleaving the C8-antiswitch at a specific nucleotide sequence to release the C8-antiswitch as a short 49 nucleotide long, single-stranded RNA molecule to be transported to the cytoplasm and fold in the presence or absence of the photochromic isomers. To detect the C8-antiswitch RNA inside the cell cytoplasm, total yeast RNA was extracted from the cells containing the pSwitch plasmid, and the total RNA was further purified to enrich the fraction of small RNA fragments according to standard miRVANA kit protocol. The enriched small RNAs was run on a 10% denaturing polyacrylamide gel and northern blot was carried out using 5’-32P labelled DNA probes complementary to the full length of C8-antiswitch
RNA. As a positive control for northern blotting, *in-vitro* transcribed Hm-As-Hm RNA was loaded in a separate lane on the same gel to be used as a size indicator.

In Figure 3.8 known amounts of *in vitro* transcribed Hm-As-Hm RNA were loaded in a serial dilution from 115 fmol to 0.05 fmol RNA to establish the lowest RNA detection threshold by Northern blotting. 20 μg of the cell purified small RNA from pSwitch+pTarget sample cells was loaded in a separate lane next to the controls. After transferring the RNA from the gel onto a membrane blot, the membrane was probed with 5′-32P labelled DNA probe complementary to C8-antiswitch RNA.
Northern blot analysis for endogenous C8-Antiswitch detection

Northern blot membrane was probed with 5’-32P labelled DNA oligos (49 nt) complementary to C8-antiswitch RNA. Lanes 1-12 show hybridization between DNA probe and the in-vitro transcribed Hm-As-Hm RNA controls. Lanes 1-12 contain 115fmol, 58, 29, 14, 7, 3.5, 1.8, 0.9, 0.45, 0.2, 0.1, and 0.055 fmol of in-vitro transcribed Hm-As-Hm RNA. Lane 13 contains 20 µg of purified endogenous yeast RNA; DNA probe hybridization is undetectable in this sample lane.

As shown in Figure 3.8, lanes 1-12 show hybridization between C8-antiswitch DNA probe and the in-vitro transcribed Hm-As-Hm RNA controls loaded in serial dilution into the wells. There are three hybridization bands in these wells. The highest band corresponds to the uncleaved 145 nucleotide long Hm-As-Hm RNA. The second
highest band corresponds to Hm-As RNA, 98 nt long RNA, with one hammerhead ribozyme cleaved off from one end only. The lowest band corresponds to the 49 nt long C8-antiswitch RNA cleaved by the two flanking hammerhead ribozymes at its 5’ and 3’ ends. According to densitometry calculations the percentage of the detectable products are 13.5%, 46%, and 3% for Hm-As-Hm, Hm-As, and As products. The level of detection of these RNA fragments by the northern blot procedure is as low as 0.05 fmol for the 145 and 98 nucleotide fragments. The 49 nucleotide C8-antiswitch RNA can be detected as low as 0.9 fmol, which corresponds to 5.4x10^{23} C8-antiswitch RNA transcripts. This is the lowest amount of endogenous RNA detected by Northern blotting by other researchers (Ebhardt H.A. and Unrau P.J. 2009), in a separate study. Lane 13 in Figure 3.8 contains 20 μg of small RNA purified from pSwitch+pTarget cells. This sample lane does not show a band corresponding to C8-antiswitch DNA probe hybridization to target endogenous RNA sample. This result suggests the lack of endogenous C8-antiswitch RNA in this sample, or it could also imply the amount of endogenous C8-antiswitch is below 5.4x10^{23} transcript detection limit.

To rule out the possibility of small RNA degradation during the endogenous RNA purification and northern blotting, and to have an internal control for small cellular RNA detection and degradation for the purified small, another northern blot gel was run. This time one lane of the gel contained 115 fmol of in vitro transcribed Hm-As-Hm RNA, as a positive control for small RNA detection by northern blotting. The other lanes of the gel were loaded with 20, 10, 5, 1, and 0.5 μg of purified endogenous small RNA from pSwitch+pTarget sample cells to be probed for detecting endogenous phenylalanine-
tRNA (Phe-tRNA) (Kelly N.J et al. 2003) as a control for the integrity of the purified small RNA from these cells.

After transferring the RNA onto a blot membrane, the membrane was cut into two pieces for probing with the appropriate DNA probes. The membrane piece containing the Hm-As-Hm RNA lane was probed with 5'-$^{32}$P labelled DNA probe complementary to C8-antiswitch RNA, and the other membrane piece containing the endogenous small RNA from pSwitch+pTarget cells was probed with 5'-$^{32}$P labelled 18-nt Phe-tRNA DNA probe. The latter membrane was not probed for C8-antiswitch since this was already done in the previous northern blot in Figure 3.8 lane 13, and there was no hybridization between the endogenous small RNA from pSwitch+pTarget cells and the C8-antiswitch DNA probe.
Figure 3.9  Northern blot analysis of endogenous yeast RNA probed for Phenylalanine detection

A) Lane 1 contains in-vitro transcribed Hm-As-Hm as a positive control and probed with 5'-32P labelled C8-antiswitch DNA oligos. 
B) Lanes 2-6 contains 20, 10, 5, 1, 0.5 μg of endogenous yeast RNA probed with 5'-32P labelled 18-nt Phe-tRNA DNA probe.

As shown in Figure 3.9A, lane 1, the C8-antiswitch DNA probe is able to base pair with its complementary in-vitro transcribed C8-antiswitch RNA similar to the same result found in Figure 3.8 lanes 1-12. Lanes 2-6 of Figure 3.9B shows hybridization between 18nt Phe-tRNA DNA probe and the endogenous small RNAs purified from pSwitch+pTarget cells. There are two hybridization bands in lanes 2-4, the higher bands
correspond to 95 bases, unspliced Phe-tRNA, and the lower bands correspond to the 76 bases, spliced Phe-tRNA. This result suggests that small (less than 100 nucleotides long) endogenous RNAs purified from yeast remain intact during the RNA purification and are detectable by northern blotting.

These finding confirm C8-antiswitch is not expressed from the pSwitch plasmid, or if transcribed, the transcript amount is less than the detectable limit. It is also possible to argue that the C8-antiswitch transcription does take place since the pSwitch plasmid sequencing had confirmed correct sequence for the pGPD promoter and the Hm-As-Hm cloning; but the transcripts are not transported from nucleus into the cytoplasm to be purified. In this case it is possible for the C8-antiswitch transcripts to have been degraded by the cellular small RNA degradation mechanism in the yeast cells prior to cellular RNA extraction and purification.
3.5 Conclusion:

The *in-vitro* functional analysis of the C8-answitch target binding affinity in respond to the specific isomer of photochromic compound was confirmed. The uptake of the photochromic compound into yeast cells was also confirmed. However, *in vivo* functional analysis of the C8-antiswitch to regulate EGFP expression, as a reporter protein, in response to the specific isomer of the photochromic compound in yeast cells, was inconclusive. Further northern blot experiments to detect endogenous C8-antiswitch RNA was also inconclusive, suggesting the EGFP expression control was unsuccessful in yeast cells since the level of C8-antiswitch RNA is too low to be able to exert any influence on the EGFP mRNA transcription. The level of C8-antiswitch transcripts must be at least equal to the level of the EGFP mRNA transcripts in a cell, for us to be able to expect any changes in the EGFP expression level by the C8-antiswitch. Since the C8-antiswitch regulatory influence in the presence of BDHP-PEG was inconclusive, additional experiments were not pursued to investigate light-controlled activity of the C8-antiswitch *in vivo*.

The design of the pSwitch and pTarget plasmid was based on this hypothesis and a YCplac plasmid, a low copy number plasmid, was used to construct both pSwitch and pTarget plasmids. However, to increase pSwitch transcript levels, the Hm-As-Hm domain expression was put under the control of pGPD promoter, a constitutively active promoter, as opposed to an inducible promoter for EGFP expression from the pTarget plasmid. However, since C8-antiswitch RNA levels was not detectable in extracted endogenous RNA from yeast cells containing the pSwitch plasmid, we conclude the lack
of inhibitory effects of C8-antiswitch on target gene expression is contributed to the lack or insufficient levels of intracellular C8-antiswitch mRNA levels.
Chapter 4: Conclusion and future directions

4.1 Thesis conclusion:

The focus of this thesis was to design a small, synthetic, trans-acting noncoding RNA to directly regulate gene expression by light irradiation in a eukaryotic system at the mRNA level. To achieve this goal we took advantage of light-induced, reversible isomerization of a photochromic compound, BDHP-PEG, to act as a light-responsive inducer of a riboregulator.

The design of a trans-acting riboregulator was proposed, which incorporated an in vitro-selected small ligand-binding domain, called an aptamer domain. This aptamer, called the C8 aptamer, was chosen since it was shown to be able to selectively bind to one but not the other of the two isomers of BDHP-PEG photochromic compound. The C8 aptamer was used to design an allosteric riboregulator, referred to as the C8-antiswitch. The design of this the C8-antiswitch was based on a proposed model of a RNA molecule with a regulatory function whose mode of activity depends on its photo-switch ligand binding to the C8 aptamer domain. It was also proposed that the ligand-binding event could be influenced by light irradiation at an appropriate wavelength. The functional domain of the C8-antiswitch was designed to act as an antisense domain. The antisense domain was incorporated in the overall structure of the C8-antiswitch and its target binding activity was influenced by the secondary structure stabilization induced by ligand-aptamer recognition and binding.
By using light as an inducer, the C8-antiswitch construct could offer a flexible control strategy by adapting active or inactive forms in response to a specific isomer of the photo-switch compound binding to the aptamer domain.

The ligand specific functional activity of the C8-antiswitch was tested in-vitro. The result showed the target binding affinity of the C8-antiswitch in response to one isomeric form of the ligand and not the other. This selective functionality of the C8-antiswitch could in turn be controlled in response to light-irradiation to reversibly isomerize the ligand molecule between the two forms.

After testing uptake of BDHP-PEG into S. cerevisiae cells, the design of the C8-antiswitch was taken a step forward by introducing it into S. cerevisiae to bind to EGFP reporter protein mRNA as the target transcript to interrupt ribosomal scanning and translation of this protein. After cloning the C8-antiswitch and EGFP into yeast expression plasmids, the fluorescent level of EGFP was analyzed by flow cytometry. Based on the proposed functional model of the C8-antiswitch we expected the C8-antiswitch to be able to exert its regulatory influence on EGFP expression in response to the appropriate isomeric form of the BDHP-PEG inducer, which reversibly isomerizes by light irradiation. However flow cytometry results did not show significant differences in EGFP fluorescent levels in response to BDHP-PEG treatment between controls and the sample cells believed to be expressing both EGFP and the C8-antiswitch.

Additional northern blot experiments were carried out to test the C8-antiswitch transcript levels in S. cerevisiae cells. These experiments did not detect endogenous C8-antiswitch mRNA and hence confirmed the lack of inhibitory influence of the C8-
antiswitch on EGFP expression is due to lack of or insufficient expression of C8-antiswitch in the cells.
4.2 Future directions:

To improve the *in vivo* functionality of the C8-antiswitch to control gene expression, I am proposing alternative approaches to the design of this system.

The first improvement to be made to this system is to use a stronger expression plasmid, such as a 2-micron plasmid, to make sure high levels of C8-antiswitch RNA is expressed in the cells. The level of C8-antiswitch transcripts must be at least equal to the level of the EGFP mRNA transcripts in a cell to be able to expect any changes in the EGFP expression level by the C8-antiswitch. To maintain a higher C8-antiswitch transcript level than the EGFP mRNA, I propose to keep the expression of EGFP protein as low as possible from a low-copy number plasmid with an inducible promoter, while the C8-antiswitch is constitutively expressed from a high-copy number plasmid.

It would also be ideal to be able to have a reporter system for intracellular expression of C8-antiswitch. A proposed system for this purpose is to clone the Hm-As-Hm domain within a toxic-lethal gene, under the control of an inducible promoter, in the pSwitch plasmid. The toxic-pSwitch plasmid and pTarget plasmid are transformed into wild type yeast cells with appropriate nutrition marker mutation. Intact toxic gene expression causes cells to die, but once the flanking hammerhead ribozymes of Hm-As-Hm are excising to release the C8-antiswitch domain, the toxic gene mRNA is cleaved and the cells will be able to survive. The viable yeast colonies will contain the C8-antiswitch RNA and these cells will be used for further FACS analysis and monitoring the EGFP expression level in response to BDHP-PEG treatment.

Improvements could also be made on the sequence design of the C8-antiswitch to further enhance the regulatory influence of the C8-antiswitch. The C8-antiswitch
sequence as proposed in this thesis contain a 10 nucleotide complementary region between its antisense domain and the target EGFP mRNA. If the length of the antisense domain is increased, perhaps the C8-antiswitch could bind to the target transcript with more affinity and specificity. A longer antisense domain could also act as a stronger inhibitor of ribosomal scanning and translation.

I am proposing an alternative C8-antiswitch design, expected to have the same secondary structure and functional properties in the presence or absence of BDHP-PEG as the existing model in this thesis. However, the sequence of the improved model is increased from 49 nt of the original C8-antiswitch to 73 nt: 5′-GCC CTT GCT CAC CAT GGT GGT CAT CCT ACA CCA TGG TGA GCA AGG GCC CTC GCC CTT GCT CAC CAT GGT GTA G-3′ (the bold letters represent the C8-aptamer sequence, and the underlined letters represent the antisense domain). This sequence is not yet prepared nor tested for its secondary structure folding in response to BDHP-PEG ligand.

Figure 4.1 shows the sequences, secondary structure predictions, and the free energies associated with the formation of (a) the aptamer stem, (b) the antisense stem, and (c) the overall alternative C8-antiswitch folded structure in the absence of BDHP-PEG. An online secondary structure prediction tool, MFold, generates these secondary structures (www.idtdna.com/Scitools/Applications/ mFold/). The aptamer stem of the alternative C8-antiswitch design and its antisense stem are designed such that the antisense stem is slightly more stable thermodynamically than the aptamer stem (ΔG= -28.9.2 kcal/mole and ΔG= -25.2 kcal/mole respectively).
Figure 4.1  Sequences and structures of the alternative C8-Antiswitch stems as predicted by MFold

a) Aptamer stem sequences and structure; red, antisense sequences; blue, C8-aptamer sequences. The free energy associated with the aptamer stem formation is estimated by MFold to be -25.2 kcal/mol.

b) Antisense stem sequences and structure. The free energy associated with the antisense stem formation is estimated by MFold to be -28.9 kcal/mol. The antisense stem is slightly more stable than the aptamer stem due to its slightly lower free energy of formation.

c) The overall structure of the alternative C8-antiswitch expected to form in the absence of the BDHP ligand; red, antisense sequences; blue, C8-aptamer sequences; black, antisense stem.
In retrospect, in the absence of ligand binding to the C8 aptamer, the formation of the antisense stem is expected to dominate over the aptamer stem formation. Alternatively, I anticipate the free energy associated with the binding of the BDHP-PEG ligand to the C8-aptamer domain could induce the formation of the aptamer stem and maintain it as the dominant stem structure. As a result, in the presence of BDHP-PEG ligand, the antisense sequence is expected to be sequestered within this aptamer stem and would not be available to interact with the target transcript.

This alternative C8-antiswitch model is redesigned to have a an enhanced and stronger target mRNA binding affinity by having a longer antisense sequence domain for complementary base pairing with the EGFP target mRNA sequence at its ATG start codon region. The antisense domain of the enhanced C8-antiswitch model is redesigned to base pair with an 18 nt-long region of EGFP mRNA around its start codon, as opposed to a 10 nt-long antisense domain in the original C8-antiswitch design. The free energy associated with target mRNA complementary base-pairing with the enhanced antisense domain is believed to be similar to the free energy of the antisense domain sequence of the S1 antiswitch (Bayer T.S., Smolke C.D. 2006), where successful ligand-induced gene expression regulation was achieved by the S1 antiswitch.

A separate direction to use light as an inducer of eukaryotic gene expression is to incorporate the C8-aptamer domain in a stem loop structure domain and insert this element in \textit{cis} immediately upstream of EGFP mRNA start codon. Modular \textit{in vivo} gene expression systems were constructed previously by another group (Vega Laso M.R. \textit{et. al.} 1993) to inhibit translational initiation in \textit{S. cerevisiae} by stem-loop structures inserted into the 5'-untranslated region of mRNA of \textit{cat} gene encoding chloramphenicol acetyl
transferase. In a different study (Suess B. et. al. 2003) conditional gene expression in S. cerevisiae was achieved by controlling gene translation initiation by inserting tetracycline-binding aptamers into the 5′-UTR of GFP encoding mRNA.

Based on these studies, I propose the design of a cis-acting light-responsive riboswitch, comprised of a stem-loop structure, with the C8-aptamer sequence defining the loop region. The secondary structure stability of this C8 aptamer stem-loop in response to BDHP-PEG ligand recognition could be mandated by light irradiation since visible light causes isomerization of BDHP-PEG to BCPD-PEG, which is not a ligand for C8-aptamer binding. As a result, the stem-loop structure could be stabilized only in the presence BDHP-PEG. The C8 aptamer stem-loop will be inserted into the 5′-UTR of EGFP encoding mRNA under the control of an inducible promoter, pGAL, in the pTarget plasmid by restriction site ligation.

In a previous work (Vega Laso M.R. et. al. 1993) a stem-loop with a predicted stability of ~14 kcal/mol and six G/C base pairs was inserted in the reporter gene mRNA, and the result showed inhibition of translation in vivo by at least 66%. I used these structural properties as a reference for designing a light-responsive molecular switch. This C8-aptamer sequence is incorporated in the loop region of the stem-loop. Figure 4.2 shows the sequence and secondary structure prediction of C8-aptamer switch, as well the schematic representation of the expression vector for functional characterization of this light-responsive cis-acting switch in S. cerevisiae.
Figure 4.2 Proposed light-responsive cis-acting C8-aptamer switch

a) The sequence and secondary structure of C8-aptamer hairpin. The C8-aptamer sequence is in blue.
b) Schematic view of the expression vector, pTarget, containing the cis-acting C8-aptamer hairpin element immediately upstream of the cloned EGFP, green box. The hairpin-stem sequence region, shown as the arrowheads, flanks the C8-aptamer region, represented by the blue line. The hairpin is inserted directly before the ATG start codon of egfp gene. The expression system is under the influence of galactose promoter, pGAL. The expression vector is a CEN4/ARS1 type plasmid, with URA3 nutrition selection marker. The arrow indicates the transcriptional start site.
c) Following transcription, the C8-aptamer hairpin forms in the 5’ UTR of the EGFP mRNA transcript.

Different designs of the stem loop, having varying lengths and sequence composition should be designed and tested in parallel in vivo, to assess the most effective design for regulatory effect of the stem loop in response to BDHP-PEG binding to C8-aptamer and stabilizing the stem loop structure. The position of the stem-loop insertion also affects the success of gene expression control (Suess B. et. al. 2003). It has been shown in S. cerevisiae that the tetracycline aptamer insertion and its secondary structure stabilization near the start codon has a more pronounced regulatory effect than cap-
proximal insertion of the stem-loop (Suess B. et al. 2003). Therefore, I propose to insert the C8 aptamer stem-loop immediately before the start codon of EGFP mRNA, as well as cap-proximal insertion of the stem-loop, to assess for light-induced regulation of gene expression by this light-responsive, cis-acting molecular switch based on the stability and position of the C8-aptamer stem-loop. Monitoring EGFP expression levels by analyzing intracellular EGFP fluorescent intensity by flow cytometry will be used to assess the regulatory influence of this light-responsive molecular switch in S. cerevisiae. This light-controlled molecular switch could present a platform for programming cellular behavior in the developmental biology of model organisms with respect to cellular state and environmental stimuli in a non-invasive manner.
Appendix:
DNA Sequences:

**T7 promoter:**
5’-CTA ATA CGA CTC ACT ATA GG-3’

**Antiswitch sense DNA:**
5’-CTA ATA CGA CTC ACT ATA GGTC ACC ATG GTG GTC ATC CTA CAC CAT GGT GAG CCC TCG CTC ACC ATG GT-3’

**Antiswitch anti-sense DNA:**
5’- ACC ATG GTG AGC GAG GGC TCA CCA TGG TGT AGG ATG ACC ACC ATG GTG ACC TAT AGT GAG TCG TAT TAG-3’

**GFP target fragment sense DNA:**
5’- CTA ATA CGA CTC ACT ATA GG-3’

**GFP target fragment anti-sense DNA:**
5’- GCC CTT GCT CAC CAT CCT ATA GTG AGT CGT ATT AG-3’

**Hm-As-Hm sense DNA for PCR:**

KpnI
5’-CTA ACT AAG GTA CC G TGA CTG ATG AGT CCG TGA GGA CGA AAG GGT AGG AAT TCC TAC CGT CTC ACC ATG GTG GTC ATC CTA CAC CAT GGT GAG CCC TCG CTC ACC ATG GTC AAG CGG AGT CGA CTC CGG TCT GAT GAG TCC GTG AGG ACG AAA GAC CCT GCA GTT AGT TAG-3’

PstI

**Hm-As-Hm anti-sense DNA for PCR:**
PstI
5’-CTA ACT AAC TGC AGG GTC TTT CGT CCT CAC GGA CTC ATC AGA CCG GAG TCG ACT CCG GTT GAC CAT GGT GAG CGA GGG CTC ACC ATG GTG TAG GAT GAC CAC CAT GGT GAG ACG GTA GGA ATT CCT ACC GTT TCG TCC TCA CGG ACT CAT CAG TCA CGG TAC CTT AGT TAG-3’

KpnI

**Hm-As-Hm Forward Primer:**

KpnI
5’- CTA ACT AAG GTA CCG TGA CTG AGT ATG CCG-3’

**Hm-As-Hm reverse Primer:**
PstI
5’- CTA ACT AAC TGC AGG GTC TTT CGT CCT CAC GGA-3’

**EGFP Forward primer:**
KpnI
5’-TAA GGT ACC ATG GTG AGC AAG GGC GAG-3’
EGFP Reverse primer:
\[ \text{PstI} \\
5'\text{-}TAA \text{ CTG CAG CTA CTT GTA CAG CTC GTC CA-3'} \]

GAL1 Forward primer:
\[ \text{EcoRI} \\
5'\text{-}\text{CCT GAA TTC TTT ATA TTG AAT TTT CAA AAA TTC TTA CTT TTT TTT TGG-3'} \]

GAL1 Reverse primer:
\[ \text{KpnI} \\
5'\text{-}\text{CCT GGT ACC TAT AGT TTT TTC TCC TTG ACG TTA AAG TAT AG-3'} \]

AS DNA probe:
\[ 5'\text{-}\text{ACC ATG GTG AGC GAG GGC TCA CCA TGG TGT AGG ATG ACC ACC ATG GTG A-3'} \]

Phenylalanine-tRNA DNA probe:
\[ 5'\text{-}\text{GTC TGG CGC TCT CCC AAC-3'} \]
Reference List


