Development of RNA-based Genetic Control Elements for Predictable Tuning of Protein Expression in Yeast

Thesis by

Andrew Harris Babiskin

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2011

(defended November 29, 2010)

© 2011

Andrew Harris Babiskin

All Rights Reserved

Acknowledgements

First, I would like to express my appreciation to my research advisor, Christina Smolke. Her enthusiasm for her students, her hands-on approach, and her accessibility were major factors in my development as an independent researcher. Her guidance has been immeasurable in the development and the progress of my research projects.

Over the years, I developed close relationships with several colleagues in lab and in the community at the California Institute of Technology. My fellow Smolke lab members have provided a great atmosphere for the facilitation of the sharing of information (and vectors) and for social interaction. Maung Win, Joe Liang, Leo d'Espaux and I spent numerous enjoyable hours together -- whether lunch at Chandler or coffee break at Red Door -- discussing research, politics, pop culture, and whatever else came to mind. I would like to thank Drew Kennedy for spending many half-price burger and chicken sandwich dinners with me during my time at Stanford University. Outside of lab, I developed close friendships with Armin Sorooshian (my roommate of 4 years), Marc Woodka, Ubaldo Córdova-Figueroa, Edgardo García-Berríos, and Muang Win. We enjoyed many visits to all-you-can-eat buffets, our sometimes successful seasons of intramural basketball, our unsuccessful attempt at intramural football, and assembling the greatest Gradiators team of all time. I am lucky to have great colleagues who are also such great friends. Outside of Caltech, I would like to thank Justin Aefsky, Mark Gaylord, and Peter Schmidt. My many trips to San Diego to visit them always gave me a nice, quick weekend away from graduate school. I express my appreciation to my friends back on the east coast who I always had a great time with whenever I was visiting.

For their constant support and love, I would like to thank my parents, Carol and Robert Babiskin, and my sisters, Rachel and Jenny. They have made me everything I am today and my success can be attributed to their involvement. To my deceased grandfather, Julius Babiskin, I would like to acknowledge his doctorate in physics as my source of inspiration for achieving my own doctorate. I would also like to thank my grandparents, Melvin and Lee Rosen, for their constant love and support. Being so far away was tough, but I enjoyed visiting home and going out to lunches with them. Even though they may not understand what I am actually doing, I was always glad to provide them with additional bragging material.

Last but not least, I would like to thank my wife, Rosario Babiskin, whom I love deeply. She has been with me for almost my entire graduate career. Because of her, I had a family away from home with whom I could spend numerous holidays with and eat ridiculous amounts of amazing food. We share together the cutest dog in the world, Cleo. When I moved away to Stanford, it was tough on the both of us being apart and not knowing when we would live together again, but she stayed strong. Though that period was difficult, we now have memories that we will cherish forever, including our wedding. Having her back living with me after our marriage has been amazing and her support and understanding has been invaluable in these final months as I finished my doctoral research and wrote my thesis.

Development of RNA-based Genetic Control Elements for Predictable Tuning of Protein Expression in Yeast

Andrew Harris Babiskin

B.S., University of Maryland

M.S., California Institute of Technology

Ph.D., California Institute of Technology

Abstract

The proper functioning of many biological processes and synthetic genetic networks depends on the precise tuning of expression levels of key protein components. With growing interests in eukaryotic hosts and the increasing complexity of networks in synthetic biology, there is a need for the expansion of the genetic toolbox, particularly for the bioprocessing and biosynthesis applications in the yeast *Saccharomyces cerevisiae*. The available control elements in yeast generally focus on the regulation of transcription through alternative promoter systems. Synthetic RNA-based control elements placed in the untranslated regions (UTRs) of transcripts have the ability to regulate the posttranscriptional mechanisms of translation initiation and transcript stability. Such posttranscriptional elements have the added advantage of being coupled to any promoter for enhanced control strategies.

Two types of posttranscriptional elements were examined in this thesis. The first type is a class of RNA hairpins baring AGNN tetraloops that are cleaved by the *S. cerevisiae* RNase III enzyme Rnt1p. By locating these hairpins in the 3’ UTR of a

transcript, the endonucleolytic cleavage due to Rnt1p activity resulted in the rapid degradation of the transcript. We developed two libraries of RNA hairpins based on the randomization of critical regions in Rnt1p substrates that affect the enzyme's ability to associate and/or cleave the hairpin. The modulation of the strength of binding and cleavage by Rnt1p resulted in changes in the steady-state transcript levels and thus protein levels. Through integration of an aptamer into the stem of an Rnt1p hairpin, we were able to develop a riboswitch based upon the direction inhibition of Rnt1p cleavage through association of the ligand in the sites of cleavage. The second type of posttranscriptional elements examined is the placement of internal ribosome entry sites (IRESes) in the 5' UTR that initiate translation independent of the 5' cap through direct interaction with the ribosomal machinery. We propose that the activity of small sequential IRESes can be tuned through varying the complementarity with the 18S ribosomal RNA (rRNA) to advance the creation of yeast multicistronic vectors. The application of Rnt1p hairpins and IRESes provide a key tool in synthetic biology for the construction of complex genetic networks in yeast where the predictable tuning of gene expression is necessitated.

Table of Contents

Acknowledgements iii
Abstract v
Table of Contents vii
List of Tables xii
List of Figures xiii

Chapter I: Introduction 1
1.1. Synthetic biology and metabolic engineering 1
1.2. Common pathways of transcriptional decay and translation in yeast 4
1.2.1. Deadenylation-dependent decapping pathway of transcript degradation 6
1.2.2. Cap-dependent translation initiation 7
1.3. Posttranscriptional regulation through transcript stability and translation 8
1.3.1. Control of transcript decay 8
1.3.2. Control of the initiation of translation 11
1.3.3. RNA processing by the RNase III enzyme Rnt1p 13
1.3.4. Translation initiation mediated through internal ribosome entry sites 14
1.4. Interrelationship among the thesis projects 15
References 17

Chapter II: Synthetic RNA control modules that tune gene expression in yeast through directed processing by Rnt1p 33
Abstract 33
2.1. Introduction 34
2.2. Results 37

2.2.1. Implementing Rnt1p hairpins as RNA-based gene regulatory components 37
2.2.2. Design and selection of an Rnt1p cleavage library to achieve tunable gene regulatory control....41
2.2.3. A synthetic Rnt1p hairpin library exhibits a range of gene regulatory activities in vivo....43
2.2.4. Rnt1p library hairpins maintain regulatory activity in a different genetic context 45
2.2.5. In vitro *characterization demonstrates that Rnt1p library members achieve differential activity through alterations in Rnt1p cleavage rates*....47
2.2.6. Control of endogenous ERG9 expression by 3' UTR replacement with Rnt1p library members....52

2.3. Discussion....58

2.4. Materials and Methods....63
2.4.1. Plasmid construction 63
2.4.2. 3' UTR replacement cassette and integration 67
2.4.3. Rnt1p substrate characterization assays 67
2.4.4. Fluorescence quantification....68
2.4.5. Quantification of cellular transcript levels....69
2.4.6. Cell growth rate determination....70
2.4.7. Cellular ergosterol quantification 70
2.4.8. In vitro *transcription of Rnt1p substrates*....71
2.4.9. Rnt1p expression and purification....72
2.4.10. In vitro *Rnt1p substrate cleavage assay* 73
2.4.11. In vitro *Rnt1p substrate mobility shift assay*....73

2.5. Supplementary Information 74
Supplementary Figures and Tables....74

Acknowledgements 79

References....80

Chapter III: Synthetic RNA modules for precise control of expression levels in yeast by tuning RNase III activity 87

Abstract 87

3.1. Introduction 88

3.2. Results 91

3.2.1. Design and selection of an Rnt1p binding library to achieve tunable gene regulatory control 91

3.2.2. A synthetic Rnt1p binding library exhibits a range of gene regulatory activities in vivo 97

3.2.3. Synthetic BSBs exhibit modular activity with different CEBs in vivo 100

3.2.4. In vitro *characterization demonstrates that Rnt1p binding library members achieve differential activity through alterations in Rnt1p cleavage rates and affinity* 102

3.3. Discussion 107

3.4. Materials and Methods 110

3.4.1. Plasmid construction 110

3.4.2. Library-scale yeast transformation 111

3.4.3. FACS and sorted library retransformation 113

3.4.4. Rnt1p substrate characterization assays 114

3.4.5. Fluorescence quantification 114

3.4.6. Quantification of cellular transcript levels 115

3.4.7. In vitro *transcription of Rnt1p substrates* 116

3.4.8. Rnt1p expression and purification 117

3.4.9. In vitro *Rnt1p substrate cleavage assay* 118

3.4.10. In vitro *Rnt1p substrate mobility shift assay* 118

3.5. Supplementary Information 120

Supplementary Figures and Tables 120

Acknowledgements 124

References 124

Chapter IV: Engineering ligand-responsive RNA controllers in yeast through the assembly of RNase III tuning modules 131

Abstract 131

4.1. Introduction 132

4.2. Results 134

4.2.1. Design of a ligand-responsive RNA switch based on Rnt1p processing 134

4.2.2. Replacement of the aptamer sequence modulates ligand responsiveness and Rnt1p processing 141

4.2.3. Incorporation of synthetic BSBs modulates ligand responsiveness and processing of the Rnt1p switch 143

4.2.4. The application of multiple switch modules decreases theophylline responsiveness and increases fold-change 145

4.2.5. Combined tuning strategies support the rational design of Rnt1p switch control systems with enhanced regulatory properties 147

4.3. Discussion 150

4.4. Materials and Methods 155

4.4.1. Plasmid construction 155

4.4.2. Rnt1p substrate characterization assays 157

4.4.3. Fluorescence quantification 158

4.4.4. Quantification of cellular transcript levels 158

4.4.5. In vitro *transcription of Rnt1p substrates* 159

4.4.6. Rnt1p expression and purification 160

4.4.7. In vitro *Rnt1p substrate cleavage assay* 161

4.5. Supplementary Information 162

Supplementary Figures and Tables 162

Acknowledgements 165

References 165

Chapter V: Synthetic IRESes promoting translation under normal physiological conditions in *S. cerevisiae* 170

Abstract 170

5.1. Introduction 171

5.2. Results 174

5.2.1. Implementing internal ribosome sites as RNA-based gene regulatory elements in dicistronic vectors 174

5.2.2. Development of a plate-based screen for IRES activity 176

5.2.3. Implementation of short sequential IRESes in tandem drives translation initiation of MEL1 177

5.2.4. Design of an IRES library to achieve tunable gene regulatory control 180

5.3. Discussion and Future Work 182

5.4. Materials and Methods 186

5.4.1. Plasmid and strain construction 186

5.4.2. MEL1 quantification 191

5.4.3. CyPET and YPET fluorescence distribution 192

5.5. Supplementary Information 193

Supplementary Figures and Tables 193

Acknowledgements 194

References 195

Chapter VI: Conclusions 199

References 206

Appendix: Table of plasmids and yeast strains and their relation to the laboratory database 209

List of Tables

Table 2.1 *In vivo* characterization data for the Rnt1p cleavage library 40

Table 2.2. *In vitro* characterization data for Rnt1p cleavage library 50

Table 2.3. Gene regulatory and phenotypic measures of the impact of Rnt1p hairpins on *ERG9* expression..56

Table 3.1. *In vivo* characterization data for the binding library..........................99

Table 3.2. *In vitro* characterization data for the binding library 105

Table 4.1. Relevant parameters for all RS-based switches and the Rnt1p and ligand binding controls, RSN and RSnt 140

Supplementary Table 2.1. Sequence and i*n vivo* characterization data for all screened Rnt1p hairpins. 77

Supplementary Table 2.2. *In vivo* characterization data for the Rnt1p cleavage library in the context of *ymCherry* (pCS1749). 79

Supplementary Table 3.1. Sequence and *in vivo* characterization data of all tested Rnt1p hairpins..123

Supplementary Table 4.1. The theoretical fold-change and dynamic range of all Rnt1p switches examined in this study as determined from experimentally measured baseline expression at 0 mM theophylline (*b*) and the theoretical maximal output (*M*) calculated by fitting the dose response data to the binding model..163

Supplementary Table 4.2. The previously reported gene-regulatory activity of the synthetic BSB modules selected for use in this study in the context of the Rnt1p hairpin genetic control element (A02) 164

Supplementary Table 4.3. Oligonucleotide template sequences for all switches built in this study..164

Supplementary Table 5.1. The oligonucleotide template sequences of all synthetic IRESes tested in this study 194

List of Figures

Figure 1.1. Maturation of eukaryotic transcripts after transcription 5
Figure 1.2. The deadenylation-dependent decapping pathway of transcript degradation in eukaryotes 6
Figure 1.3. The mechanism of eukaryotic cap-dependent translation..................8
Figure 1.4. The processing of eukaryotic transcripts following endonucleolytic cleavage is independent of the 5’ cap and the 3’ poly(A) tail..11
Figure 1.5. Genetic elements that affect the initation of translation in eukaryotes 12
Figure 1.6. Simplified schematic of the interactions with structural and sequential IRESes with the translational machinery..........................15
Figure 2.1. Genetic control elements based on Rnt1p hairpins............................38
Figure 2.2. Design and *in vivo* screening of an Rnt1p cleavage library...............42
Figure 2.3. *In vivo* characterization of the selected Rnt1p cleavage library 44
Figure 2.4. Demonstration of functional modularity of the hairpin library in the context of a different genetic construct.......................................47
Figure 2.5. *In vitro* characterization of the Rnt1p library supports the tuning of gene regulatory activity through modulation of cleavage rates 48
Figure 2.6. Synthetic Rnt1p hairpins enable posttranscriptional control over endogenous *ERG9* expression levels 53
Figure 3.1. Implementation of Rnt1p hairpins as posttranscriptional genetic control elements and binding library design.....................................92
Figure 3.2. *In vivo* screening of an Rnt1p binding library 95
Figure 3.3. *In vivo* characterization of the selected Rnt1p binding library and demonstration of the modularity of the BSB sequences....................98
Figure 3.4. *In vitro* characterization of the binding library demonstrates that the observed tuning of gene regulatory activity is achieved through modulation of cleavage rates and binding affinities 104

Figure 4.1. Design and implementation of Rnt1p switches as posttranscriptional genetic control elements....................................135
Figure 4.2. Tuning the response curve of the Rnt1p switch through the integration of different theophylline aptamers..................................142
Figure 4.3. Tuning the response curve of the Rnt1p switch through the integration of different synthetic BSB modules................................143
Figure 4.4. Predictive tuning of the Rnt1p switch response curve through the integration of multiple copies of the switch module.........................146
Figure 4.5. Combinatorial implementation of multiple tuning modules results in predictive tuning of the Rnt1p switch regulatory response curve.....149
Figure 5.1. A yeast dicistronic vector based on insertion of an internal ribosome entry site (IRES) into the intercistronic region (IR) between the two genes of interest (goi1 and goi2) 175
Figure 5.2. Visualization of MEL1 activity on X-α-gal plates 177
Figure 5.3. Quantification of MEL1 activity of constructs baring IRES modules 178
Figure 5.4. Proposed design for selection of a 10-nt IRES library in a dicistronic vector..180
Figure 5.5. Expression profiles for the CyPET- and YPET-based constructs 182

Supplementary Figure 2.1. Flow cytometry histograms of pCS321-based constructs bearing control and library Rnt1p hairpins..74
Supplementary Figure 2.2. Sequences and structures of the final Rnt1p cleavage library and the positive control Rnt1p hairpins.............75
Supplementary Figure 2.3. Flow cytometry histograms of pCS1749-based constructs bearing control and library Rnt1p hairpins..75
Supplementary Figure 2.4. Correlation analysis of protein and transcript levels for all hairpins identified from the fluorescence-based *in vivo* screening assay..76
Supplementary Figure 2.5. Plasmid maps for key constructs used in this work......76

Supplementary Figure 3.1. FACS analysis and gating procedure for pCS1585 system on FACSAria 112
Supplementary Figure 3.2. FACS analysis and gating procedure for pCS1748 system on FACSAria II..113
Supplementary Figure 3.3. Sequences and structures of the selected Rnt1p binding library and control hairpins containing the 'parent' BSB 114
Supplementary Figure 3.4. Plasmid map of pCS321-based vectors.........................115
Supplementary Figure 4.1. Sequences illustrating the placement of the ΔTCT-4 aptamer within R31L-3B4Inv at multiple locations 162
Supplementary Figure 4.2. The dose response curves of RS, RS-B03, RS-B05, and RS-B06 indicate that these synthetic BSB modules increase baseline expression relative to the original Rnt1p switch (RS) 162
Supplementary Figure 4.3. Plasmid map of pCS321, the Rnt1p hairpin characterization plasmid 163
Supplementary Figure 5.1. Plasmid map of pRM, the dicistronic IRES characterization and screening plasmid 193