Chapter VI: Conclusions

For many applications of synthetic biology, such as metabolic engineering, there is a need for the development of genetic tools that allow for the fine-tuning of gene expression in *Saccharomyces cerevisiae*. The majority of tools currently employed are limited to the use of native yeast promoter systems with or without the regulation of transcription factors and the use of synthetic promoters engineered rationally or through the screening of libraries[1]. Given constant culture conditions, output from constitutive promoters cannot be altered unless used in plasmid systems with different copy numbers. However, the difference in expression between low- and high-copy plasmids is considerable, with high-copy plasmids often resulting in overexpression of a desired plamid. Constitutive promoters can be used interchangeably, but they are limited in number and thus limited in what gene expression levels can be achieved. The regulatory range of a constitutive promoter can be modulated through the library of a mutatnt promoter library, such as the TEF1 promoter library created through random mutagenesis of the wild-type promoter[2]. Inducible promoter systems have an advantage in that expression output from the promoter can be controlled by setting the concentration of the small molecule inducers. However, many inducible promoter systems do not exhibit substantial titratable regimes, such that fine tuning expression levels can be difficult from such systems. Often times, inducible promoter systems exhibit heterogenous behavior, where a given cell in the population is either fully expressing or repressed in these systems and a change in the ratio of these two states across the population is observed by varying the inducer concentration[3]. Inducible promoters are also not desirable for industrial applications due to the additional cost and to nonspecific or pleotropic effects

associated with the inducing molecule. Synthetic promoter systems are also limiting in applications where expression of an endogenous gene is to be modulated, but it is desired to retain the cellular regulatory mechanisms associated with the native promoter.

RNA-based postranscriptional elements have the advantage of being coupled to any promoter of choice, allowing the utilization of enhanced control strategies. RNA is a diverse molecule with structural, enzymatic, and ligand-binding properties[4]. Once a coding RNA is transcribed, it goes through two processes: translation and degradation. Previously described RNA devices have controlled these process through antisense or ribozyme technology, respectively, in response to a small molecule effector[5-7]. In my thesis research, we developed novel genetic control modules to precisely tune gene expression in *S. cerevisiae* through posttranscriptional regulation. Two libraries of synthetic RNA hairpins placed in the 3' untranslated region (UTR) of transcripts mediated endonucleolytic cleavage by the RNase III enzyme Rnt1p. Processing by the enzyme leads to a reduction in transcript levels that ultimately results in decreased protein levels. Various intermediate levels of gene expression were achieved through the randomization of nucleotides associated with the cleavage efficiency box (CEB) and binding stability box (BSB) of Rnt1p substrates (Chapter II and Chapter III, respectively). These CEB and BSB modules were integrated combinatorially to further extend the range of the overall Rnt1p hairpin library (Chapter III). In addition, the ligand-sensing properties of RNA aptamers were incorporated into Rnt1p hairpins to develop an Rnt1p switch (Chapter IV). Here, the binding of ligand directly interfered with Rnt1p cleavage, resulting in increased gene expression. Switch dynamics were rationally modulated and improved through incorporating aptamers with differing affinities, BSB modules, and

switch modules. Lastly, we attempted to screen for a library of small sequential internal ribosomal entry sites (IRESes) at physiological conditions that differ in their ability to initiate translation (Chapter V). Such synthetic IRES elements will allow for the construction of yeast 'operons', where appropriate gene ratios can be achieved through the usage of IRESes of various strengths.

The Rnt1p hairpin library and the Rnt1p switches provides an efficient way to control gene expression levels in *S. cerevisiae*. The RNA modules take advantage of an unique property of the Rnt1p enzyme - its specific recognition of hairpins containing an AGNN tetraloop[8]. Other RNase III enzymes characterized to date have not demonstrated this property. As such, these hairpins will only function in *S. cerevisiae*, where Rnt1p is a critical enzyme involved in the processing of non-coding structural and functional RNAs[9-12]. It is possible that these hairpins can be extended as orthogonal control elements in additional yeast species, other eukaryotes, and prokaryotes by heterologously-expressing Rnt1p. The transfer of RNase-based regulatory strategies has been recently demonstrated in studies where a yeast variant of the RNase III Dicer enyzme from *S. castellii* and other factors were expressed in *S. cerevisiae* to introduce the mechanism of RNA interference (RNAi)[13].

There are several challenges with implementation of the Rnt1p hairpin that should be addressed in future studies. The dependence of the hairpins' regulatory activity on Rnt1p may present disadvantages as gene expression is dependent on the cellular machinary of the host organism. For example, reductions in cellular Rnt1p levels may ultimately limit the regulatory range of the library and, in the case of Rnt1p switches, decrease the dynamic range of the switches, while increasing responsiveness to the

effector. On the other hand, it may be possible to counteract natural decreases of Rnt1p or to increase the silencing activity by Rnt1p by expressing more Rnt1p in the yeast strain. The Rnt1p substrates generated in this thesis generally act as modular units as demonstrated through the placement of the cleavage library in multiple genetic contexts (Chapter II) and the multiplicative baseline observed in the multiple module Rnt1p switches (Chapter IV). The mulitplicative baseline also suggests that the Rnt1p hairpins derived from the cleavage library can be placed in tandem to increase silencing activity. The initial attempts to place two hairpins in tandem in the 3' UTR were what led to the eventual removal of structurally weak structures from the cleavage library as hairpin structures appeared to be affecting each others' activity. The multiple module switches did not appear to have this issue as their structures were insulated by A-rich flanking sequences. The inclusion of such flanking sequences could help enhance the modularity of the cleavage library hairpins, but will likely also affect the knockdown observed from the hairpins leading to a recharacterization of the library or perhaps a screening of a new cleavage library with the flanking sequences included. Due to the small percentage of cleavage hairpins that lack modularity in a particular system, the described additional studies will only result in minor improvements. Instead, if multiple hairpin constructs are desired, it may be more prudent to perform research on linker sequences in the intervening regions between hairpins to ensure maintenance of activity.

The Rnt1p switches will serve as important tools in metabolic engineering by allowing the construction of synthetic feedback systems based on the sensing of metabolites and other molecules directly involved in the metabolic network and regulation of key enzyme activities. The development of tailored metabolite-responsive

Rnt1p switches is currently limited by the number of aptamers available that respond to ligands of interest and the subset of those aptamers that can be successfully integrated into the Rnt1p switch design due to structural constraints. In the development of the Rnt1p switch with the theophylline aptamer, the aptamer was positioned at different locations in the CEB (Supplementary Figure 4.1). While each design demonstrated the ability to be cleaved by Rnt1p, only one (RS) featured inhibition of cleavage upon the addition of ligand. This work highlights that the binding interactions of the ligand with the aptamer and its location in the CEB is critical for the ligand-dependent impedance of Rnt1p activity. The xanthine aptamer[14] and the tetracycline aptamer[15] were also built into the switch platform at multiple locations in the CEB, albeit unsuccessfully. Incorporation of both aptamers did not lead to any significant knockdown in the absence of ligand, i.e. the aptamer structure itself was inhibiting Rnt1p activity. These aptamers and the failed theophylline aptamers highlight the importance of the aptamer structure being processable by Rnt1p and the ability of the ligand binding interactions to interfere with Rnt1p activity.

The added structural requirements for the Rnt1p switch platform likely reduce the probability that an *in vitro* selected aptamer will respond in this platform. Therefore, future efforts to expand the Rnt1p switch platform to respond to alternative stimili should focus on the selection of aptamers with the switch platform *in vivo* in an appropriate reporter plasmid. For example, libraries can be designed with randomized aptamer/CEB region and functional sequences can be selected for decreased fluorescence in the absence of ligand (for ability to be cleaved by Rnt1p) and for an increase in fluorescence upon ligand addition (for selection of functional aptamers in which ligand binding interferes

with Rnt1p activity). Similar efforts are already under way in the Smolke laboratory with the selection of aptamers in the ribozyme-based riboswitch platform (J.C. Liang et. al., in preparation), and these strategies can be directly extended to the Rnt1p switch platform.

The Rnt1p hairpin library and the Rnt1p switches can be applied to a wide range of applications. We demonstrated the library's ability to examine the architecture of endogenous networks by pertubating metabolic flux around a major control point. Specifically, the expresson of ERG9, the first enzyme involved in the ergosterol biosynthesis, was modulated to reduce flux through the pathway (Chapter II). Native feedback control around this pathway was observed as ERG9 expression levels were maintained at a certain threshold value. In metabolic engineering applications, the Rnt1p-based modules can be employed to reduce metabolic burden, balance cytotoxic intermediate, and redirect cellular resources from native pathways.

We envision that the Rnt1p library can be employed by two methods. First, with the 16 CEB and 16 BSB modules described, 256 different Rnt1p hairpins can be constructed and integrated into the 3’ UTR of a gene of interest to screen for a desired phenotype. Ordering the oligonucleotides necessary to construct the 256 hairpins will likely be expensive. As an alternative option, since a wide range of gene expression was observed from the cleavage library, a library of randomized CEB sequences could be screened for a desired phenotype in a construct appropriate for the application of interest. Second, the two libraries can be applied rationally to limit the amount of constructs tested. The cleavage library was observed to have the largest range of gene regulatory activity and also the best distribution, while the binding library had a decreased range more suitable for the tuning of gene expression. The cleavage library can be employed

first to identify regulatory regions of interest. Then, more focused expression levels can be explored through the implementation of the binding library BSB modules with the appropriate cleavage library CEB modules.

A synthetic IRES library, once successfully developed, will allow the construction of multicistronic vectors in yeast where the relative ratios of the individual gene products can be altered through the integration of the library sequences. While a structured, native IRES in yeast provides similar levels of gene expression as cap-dependent translation (Figure 5.3), single copies of small IRES sequences with complementarity to the 18S rRNA are ineffective at producing substantial levels of gene product. Seven linked copies of these IRES modules were able to produce a visual phenotypic response due to MEL1 reponse (Figure 5.2), but levels were measured to be much lower compared to cap-dependent translation (Figure 5.3). Questions about the sensitivity of the MEL1 assay and issues with homologous recombination led to the attempt to construct a dual fluorescense reporter system with *CyPET* and *YPET*. However, homology between the gene pairs destabilized the resulting vectors. Recently, a new low-copy dual fluorescence reporter plasmid has been characterized in the Smolke laboratory that contains two transgenes, *ymCherry* and *yEGPF3* (J.C. Liang et al., in preparation). A future direction for the IRES project would be to rearrange this vector to produce an *ymCherry-yEGFP3* dicistronic vector. The ability of *yEGFP3* expression to report on IRES activity should first be tested through the placement of multiple copies of the IRES47 sequence in the intergenic region (IR). It is recommended that the method of gap repair be removed from this project since it depends on homologous recombination for successful cloning. It is also recommended that the entire IRES module be cloned from a single oligonucleotide template so that the

linker sequences are constant. Since the response from MEL1 and yEGFP3 may be different, there may need to be additional optimization around the number of IRES copies necessary for activity to be observed with yEGFP3. Once the dual-fluorescence reporter is optimized, the proposed library strategy can be employed to select active IRES modules through fluorescence-activated cell sorting (FACS).

In summary, we have developed synthetic genetic control modules for *S. cerevisiae* acting through posttranscriptional mechanisms that can be implemented with different genetic targets and promoters. The Rnt1p hairpin library provides a key tool for synthetic biology applications in yeast where the predictable tuning of gene expression is necessitated. Aptamer integration allows for the construction of ligand-responsive Rnt1p-based control devices. The development of synthetic IRESes will be useful in the creation of yeast 'operons' where various ratios of expression for several genes can be linked through a single transcriptional event. The synthetic control modules developed in this thesis will provide an important toolset for advancing yeast as a microbial host for bioprocessing and biosynthesis applications.

References

1. Nevoigt, E. Progress in metabolic engineering of Saccharomyces cerevisiae. *Microbiol Mol Biol Rev* **72**, 379-412 (2008).
2. Nevoigt, E. et al. Engineering of promoter replacement cassettes for fine-tuning of gene expression in Saccharomyces cerevisiae. *Appl Environ Microbiol* **72**, 5266-5273 (2006).

3. Louis, M. & Becskei, A. Binary and graded responses in gene networks. *Sci STKE* **2002**, pe33 (2002).

4. Win, M.N., Liang, J.C. & Smolke, C.D. Frameworks for programming biological function through RNA parts and devices. *Chem Biol* **16**, 298-310 (2009).

5. Bayer, T.S. & Smolke, C.D. Programmable ligand-controlled riboregulators of eukaryotic gene expression. *Nat Biotechnol* **23**, 337-343 (2005).

6. Win, M.N. & Smolke, C.D. A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. *Proc Natl Acad Sci U S A* **104**, 14283-14288 (2007).

7. Win, M.N. & Smolke, C.D. Higher-order cellular information processing with synthetic RNA devices. *Science* **322**, 456-460 (2008).

8. Lamontagne, B. et al. Sequence dependence of substrate recognition and cleavage by yeast RNase III. *J Mol Biol* **327**, 985-1000 (2003).

9. Catala, M., Lamontagne, B., Larose, S., Ghazal, G. & Elela, S.A. Cell cycle-dependent nuclear localization of yeast RNase III is required for efficient cell division. *Mol Biol Cell* **15**, 3015-3030 (2004).

10. Chanfreau, G., Elela, S.A., Ares, M., Jr. & Guthrie, C. Alternative 3'-end processing of U5 snRNA by RNase III. *Genes Dev* **11**, 2741-2751 (1997).

11. Chanfreau, G., Rotondo, G., Legrain, P. & Jacquier, A. Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. *EMBO J* **17**, 3726-3737 (1998).

12. Elela, S.A., Igel, H. & Ares, M., Jr. RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. *Cell* **85**, 115-124 (1996).

13. Drinnenberg, I.A. et al. RNAi in budding yeast. *Science* **326**, 544-550 (2009).

14. Kiga, D., Futamura, Y., Sakamoto, K. & Yokoyama, S. An RNA aptamer to the xanthine/guanine base with a distinctive mode of purine recognition. *Nucleic Acids Res* **26**, 1755-1760 (1998).

15. Berens, C., Thain, A. & Schroeder, R. A tetracycline-binding RNA aptamer. *Bioorg Med Chem* **9**, 2549-2556 (2001).