Chapter II: A synthetic library of RNA control modules for predictable tuning of gene expression in yeast

Abstract

Advances in synthetic biology have resulted in the development of genetic tools that support the design of complex biological systems encoding desired functions. The majority of efforts have focused on the development of regulatory tools in bacteria, whereas fewer tools exist for the tuning of expression levels in eukaryotic organisms. Here, we describe a novel class of RNA-based control modules that provide predictable tuning of expression levels in the yeast *Saccharomyces cerevisiae*. A library of synthetic control modules that act through posttranscriptional RNase cleavage mechanisms was generated through an *in vivo* screen, where structural engineering methods were applied to enhance the insulation and modularity of the resulting components. The library of sixteen synthetic RNase substrates exhibit a wide range of gene regulatory activities (spanning 8% and 85% at the protein level). This new class of control elements can be combined with any promoter to support titration of regulatory strategies encoded in transcriptional regulators and thus more sophisticated control schemes. We applied these synthetic controllers to the systematic titration of flux through the ergosterol biosynthesis pathway, where feedback regulation was observed to maintain production of ergosterol and thus cellular growth rates. This work provides insight into endogenous control strategies and highlights the utility of this control module library for manipulating and probing biological systems.

2.1. Introduction

Synthetic biology is advancing capabilities for engineering biological systems exhibiting desired functions. The proper functioning of synthetic genetic circuits often relies on precise control and tuning of the expression levels of key protein components. For example, the proper functioning of synthetic gene networks exhibiting complex dynamic behaviors has been shown to depend on the appropriate matching of levels of protein components in the engineered networks[1-3]. The tuning of protein levels to obtain functioning circuits has been commonly achieved by screening randomized gene expression control elements for those sequences that provide the desired regulatory strength[3-5]. As another example, the optimization of engineered metabolic networks has been shown to depend on the precise control of enzyme levels and activities[4, 6-7]. The tuning of enzyme levels is critical for reducing metabolic burden due to enzyme overexpression[8-9], decreasing accumulation of toxic intermediates by balancing pathway flux[4, 10], and redirecting cellular resources from native pathways without negatively impacting the health and viability of the engineered host by knocking out required enzymes[11-12]. As such, the development of well-characterized gene expression control modules that can be used to predictably tune the levels of proteins are key to the design of robust genetic systems.

While many gene regulatory tools have been developed for use in *Escherichia coli*[13-16] fewer such tools exist for the precise tuning of expression levels in the budding yeast, *Saccharomyces cerevisiae*. However, *S. cerevisiae* is a relevant organism in industrial processes, including biosynthesis and biomanufacturing strategies[17-22], such that as more complex genetic networks are engineered into yeast it becomes critical to

have tools that allow for the facile programming of gene expression levels. The existing methods for tuning gene expression levels in *S. cerevisiae* rely on transcriptional control mechanisms in the form of inducible and constitutive promoter systems. Many inducible promoters do not provide tunable control systems due to their on/off switch-like behavior, where the amount of inducer molecule controls the likelihood that a given cell is repressed or fully expressing the desired protein[23]. While engineered variants have been constructed that offer more tunable responses[24-25], these systems can exhibit other undesirable properties, due to nonspecific or pleotropic effects associated with the inducing molecule or limitations associated with costs in using the inducing molecule in large-scale processes. As an alternative strategy, a promoter library was recently developed based on mutating the constitutive TEF1 promoter[26]. The resulting library of promoter parts comprised 11 promoter variants that spanned expression levels from 8% to 120%, providing a useful tool for controlling expression levels in yeast. However, control modules based on transcriptional mechanisms require the use of a particular promoter, which may be limiting to certain applications. For example, the use of a specific or native promoter may be desired to retain cellular control mechanisms associated with the given promoter. RNA-based control modules based on posttranscriptional mechanisms may offer an advantage by allowing these control elements to be coupled to any promoter of choice, providing for enhanced control strategies and finer resolution tuning of expression levels.

Endoribonucleases play key roles in RNA processing across diverse cellular systems[27]. In eukaryotic cells, endoribonuclease cleavage in the untranslated regions (UTRs) or coding regions of a transcript can result in rapid degradation of that transcript

by exoribonucleases. The RNase III family is a class of enzymes that cleaves double-stranded RNA (dsRNA)[28]. The *S. cerevisiae* RNase III enzyme, Rnt1p, recognizes RNA hairpins that contain a consensus AGNN tetraloop and cleaves its substrates 14 nucleotides (nt) upstream and 16 nt downstream of the tetraloop[29]. Rnt1p harbors an RNase III domain and a dsRNA-binding domain (dsRBD)[28], where the AGNN tetraloop of an Rnt1p substrate forms a predetermined fold that is recognized by the dsRBD[30]. Rnt1p is localized to the nucleus, where it has been shown to cleave cellular ribosomal RNA (rRNA) precursors, small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), and messenger RNAs (mRNA)[31-34]. However, despite extensive characterization of this RNA processing enzyme, neither natural nor synthetic Rnt1p substrates have been used to control gene expression levels in yeast.

We have demonstrated that Rnt1p substrates can be utilized as effective posttranscriptional gene control modules when placed in the 3' UTR of a target transcript. We utilized this Rnt1p regulatory construct with a cell-based screening strategy to develop a library of synthetic Rnt1p substrates that exhibit a wide range of gene regulatory activities (spanning 8% and 85%) to tune Rnt1p processing efficiency. *In vivo* and *in vitro* assays demonstrate that the library of control elements modulate transcript and protein levels through variations of the Rnt1p processing efficiency. The library of Rnt1p elements was applied to predictably modulate flux through an endogenous ergosterol biosynthesis network through the direct integration of the synthetic components with an endogenous gene target, highlighting the broader utility of these synthetic control modules. The described Rnt1p substrate library provides a new set of

control modules that can be used to predictably tune gene expression in yeast with any desired promoter.

2.2. Results

2.2.1. Implementing Rnt1p hairpins as RNA-based gene regulatory components

Rnt1p is an RNase III enzyme that cleaves consensus hairpin structures in *S. cerevisiae*. In order for a hairpin to be effectively recognized and cleaved by Rnt1p it must have the following consensus elements: an AGNN tetraloop and four base-pairs immediately below the tetraloop (Figure 2.1A). An Rnt1p substrate can be divided into three critical regions: the initial binding and position box (IBPB), comprising the tetraloop; the binding stability box (BSB), comprising the base-paired region immediately adjacent to the tetraloop; and the cleavage efficiency box (CEB), comprising the region containing and surrounding the cleavage site[29]. The CEB has no reported sequence or structural requirements. Rnt1p will initially position itself and bind to the tetraloop and cleave the hairpin at two locations within the CEB: between the 14th and 15th nts upstream of the tetraloop and the 16th and 17th nts downstream of the tetraloop. Most naturally-occurring Rnt1p hairpins have been identified in noncoding RNAs (ncRNAs), where Rnt1p plays a critical role in ncRNA processing[31-33]. Synthetic *trans*-acting RNA guide strands were recently utilized to direct Rnt1p processing of a target ncRNA[35]. Rnt1p hairpins have also been identified within the coding region of at least one endogenous yeast gene, *MIG2*, where Rnt1p was shown to play a role in controlling expression levels of that gene[36]. However, the ability of Rnt1p hairpins to function as

genetic control modules in regulating the expression of heterologous genes has not been previously examined.

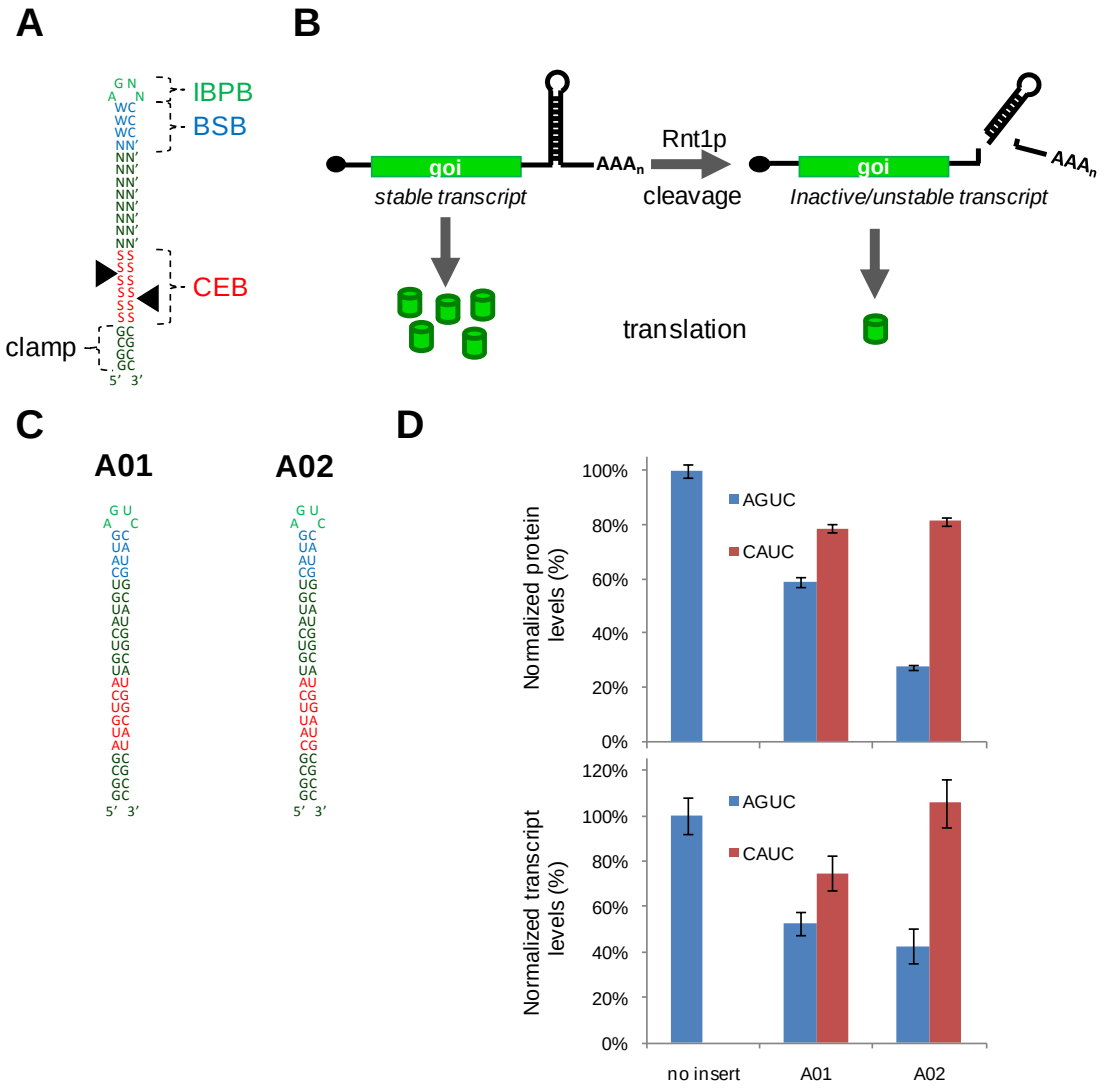


Figure 2.1. Genetic control elements based on Rnt1p hairpins. (A) Consensus elements of an Rnt1p hairpin. Color scheme is as follows: cleavage efficiency box (CEB), red; binding stability box (BSB), blue; initial binding and positioning box (IBPB), green. Black triangles represent location of cleavage sites. The clamp region is a synthetic sequence that acts to insulate and maintain the structure of the control element. (B) Schematic illustrating the mechanism by which Rnt1p hairpins act as gene control elements when placed in the 3' UTR of a gene of interest (goi). Barrels represent protein molecules. (C) Sequences and structures of Rnt1p hairpin controls. (D) The transcript and protein levels associated with Rnt1p hairpins and their corresponding mutated tetraloop (CAUC) controls support that the observed gene regulatory activity is due to Rnt1p processing. Normalized protein expression levels are determined by measuring the

median GFP levels from a cell population harboring the appropriate construct through flow cytometry analysis and values are reported relative to that from an identical construct lacking a hairpin module (no insert). Reported values and their error are calculated from the mean and standard deviation from the three identically-grown samples, respectively. Transcript levels are determined by measuring transcript levels of *yEGFP3* and a house-keeping gene, *ACT1*, through qRT-PCR and normalizing the *yEGFP3* levels with their corresponding *ACT1* levels. Normalized transcript levels are reported relative to that from an identical construct lacking a hairpin module. Reported values and their error are calculated from the mean and standard deviation from three identically-prepared qRT-PCR reactions, respectively.

We designed a system that utilizes Rnt1p-mediated hairpin cleavage to regulate gene expression in yeast through the modular insertion of Rnt1p hairpins in UTRs of a gene (Figure 2.1B). Specifically, we inserted Rnt1p hairpins as gene control elements within the 3' UTR of a transcript to direct cleavage to that region, thereby inactivating the transcript and resulting in rapid transcript degradation. While directing cleavage to the 5' UTR of a transcript would be expected to similarly inactivate the transcript, insertion of secondary structures in the 5' UTR of eukaryotic transcripts has been shown to result in nonspecific translational inhibition due to affects of structural elements on ribosomal scanning[37], such that resulting gene regulatory effects would likely not be specific to the desired cleavage mechanism. We designed and built a low-copy Rnt1p characterization plasmid (pCS321) to quantify the gene regulatory properties of Rnt1p substrates in yeast. Unique restriction sites for inserting Rnt1p hairpins were located 2 nts downstream of the stop codon of a gene encoding a yeast enhanced green fluorescent protein (*yEGFP3*)[38].

We first examined the ability of Rnt1p hairpins to function as gene control elements when placed downstream of a heterologous reporter gene in yeast. We adapted two Rnt1p hairpins with different CEBs, A01 and A02, that had been previously characterized through *in vitro* assays (R31-27 and R31D, respectively, in [39]) (Figure 2.1C). The hairpins were modified by placing a G-C rich base-paired region, or clamp,

below the 18 nt stem of the Rnt1p substrates to ensure structural stability of the hairpins when placed within the context of our Rnt1p characterization construct and to provide a proper stem length for effective cleavage *in vivo*. The hairpins were inserted into the characterization plasmid, and regulatory efficiencies were determined by monitoring cellular fluorescence by flow cytometry and transcript levels by quantitative real-time PCR (qRT-PCR) (Table 2.1). Negative controls for Rnt1p hairpins were constructed by mutating the tetraloop sequence to CAUC to impede Rnt1p activity while maintaining the secondary structure of the hairpins. The fluorescence and transcript data for A01 and A02 show that nucleotide modifications in the CEB result in the attenuation of *in vivo* gene expression, and the mutated tetraloop controls support that the observed regulatory effects are due to Rnt1p processing (Figure 2.1D). Flow cytometry histograms of the control hairpins demonstrate that regulatory activity causes a population shift with reduced median levels (Supplementary Figure 2.1A).

Table 2.1. *In vivo* characterization data for the Rnt1p cleavage library. All normalized protein and transcript levels are determined as described in Figure 2.1D.

Substrate	Normalized protein levels (%)	Normalized transcript levels (%)
C01	84% ± 6%	68% ± 3%
C02	80% ± 3%	71% ± 7%
C03	55% ± 1%	60% ± 9%
C04	20% ± 1%	31% ± 4%
C05	55% ± 2%	51% ± 5%
C06	33% ± 2%	55% ± 5%
C07	41% ± 1%	67% ± 13%
C08	11% ± 0%	12% ± 2%
C09	25% ± 1%	28% ± 4%
C10	46% ± 1%	66% ± 8%
C11	11% ± 0%	56% ± 10%
C12	81% ± 6%	75% ± 12%
C13	8% ± 0%	12% ± 1%
C14	85% ± 3%	83% ± 6%
A01	59% ± 2%	53% ± 5%
A02	28% ± 1%	43% ± 8%
no hairpin	100% ± 3%	100% ± 8%

2.2.2. Design and selection of an Rnt1p cleavage library to achieve tunable gene regulatory control

Gene regulatory elements that allow the precise and predictable tuning of expression levels are important tools in synthetic biology for the control of gene circuits. The generation of well-characterized libraries of gene control elements that exhibit varying regulatory properties has resulted in useful tools in bacteria[13-15]. Similar strategies have been applied to develop libraries of transcriptional control elements, specifically constitutive promoters, in yeast[26, 40]. However, certain applications will require circuit designs where either native or inducible promoter systems will be desired, such that the ability to integrate posttranscriptional control elements that act downstream of desired promoter systems will be required. Libraries of tuned posttranscriptional stability control elements have not been developed to date in *S. cerevisiae*.

Based on the different gene regulatory activities observed from A01 and A02, we examined whether a larger library of synthetic Rnt1p hairpins could be engineered to develop a set of tuned posttranscriptional control elements. We developed an Rnt1p library based on randomizing the CEB (12 nt) to generate Rnt1p hairpins that exhibit different gene regulatory activities due to altered enzyme processing rates and identified synthetic Rnt1p substrates through a cell-based fluorescence screen (Figure 2.2A). The designed library has a diversity of ~1.7 x 10^7 different hairpin sequences. Due to the flexibility of the structural and sequence requirements for the CEB and the ability of each of the library members to bind Rnt1p through the maintained tetraloop structure, we anticipated that a large percentage of the library members would exhibit some cleavage activity. The goal of the functional screen of the library was to identify a set of Rnt1p

hairpins that will provide a range of different regulatory activities and be useful as modular gene control elements.

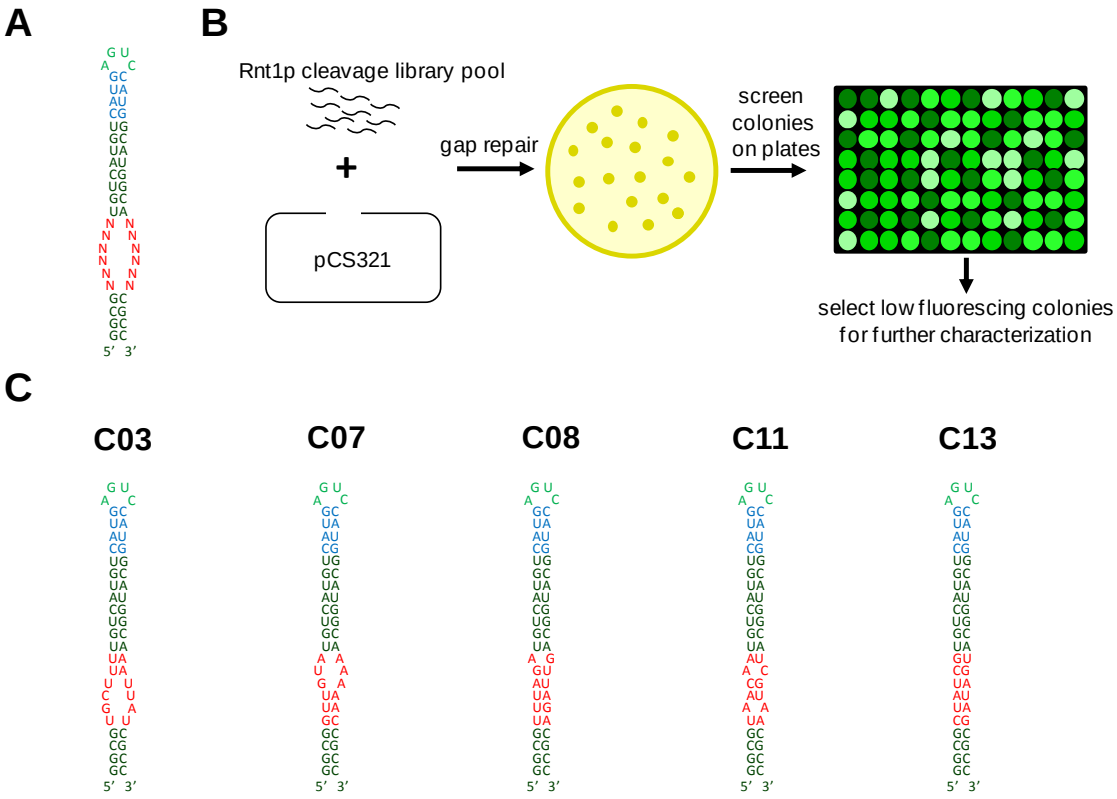


Figure 2.2. Design and *in vivo* screening of an Rnt1p cleavage library. (A) Sequence and structure of Rnt1p hairpin library containing the 12 randomized nucleotides in the CEB. (B) An *in vivo*, fluorescence-based screen of Rnt1p hairpin activity. The library pool is cloned through gap-repair into yeast, and clones are screened on a plate reader for sequences resulting in low fluorescence. (C) Sequences and structures of select library members highlight the diversity of the selected library sequences. The color scheme for hairpin sequences is described in Figure 2.1A.

The cleavage library was transformed into yeast through a gap-repair strategy and individual colonies were initially characterized for gene regulatory activity through assaying cellular fluorescence on a plate reader (Figure 2.2B). A total of 318 colonies were characterized and from this initial screening, constructs from 41 low-expressing colonies were sequenced (Supplementary Table 2.1). The sequences of the selected Rnt1p hairpins were analyzed by RNAstructure (http://rna.chem.rochester.edu/

RNAstructure.html) to determine the predicted secondary structure of the hairpins. No consensus secondary structure was identified from the 41 isolates due to the diversity of associated structures. The CEBs of the recovered hairpins are either completely base-paired or contain one or two bulges of different size and location (Figure 2.2C). To ensure modularity of the synthetic Rnt1p hairpins to other genetic constructs, we removed library candidates that were 'structurally weak'. Structurally-weak hairpins were identified based on two properties: 1) predicted ability of the hairpin sequence to fold into multiple secondary structure conformations with similar free energies and 2) interactions with flanking sequences. In total, 16 Rnt1p cleavage library substrates were identified as synthetic control modules (Table 2.1, Supplementary Figure 2.2).

2.2.3. A synthetic Rnt1p hairpin library exhibits a range of gene regulatory activities **in vivo**

The range of regulatory activities spanned by the cleavage library was measured at the protein expression and transcript levels. Flow cytometry analysis of the synthetic Rnt1p hairpins indicated that the selected set of hairpins spanned a large gene regulatory range – from 7.9% (C13) to 84.7% (C14) (Table 2.1, Figure 2.3A). The regulatory activities of the selected hairpins are fairly evenly distributed across this range allowing for precise tuning of expression levels based on insertion of different synthetic Rnt1p hairpins. Flow cytometry histograms of the library hairpins confirm that regulatory activity causes a population shift with reduced median GFP levels (Supplementary Figure 2.1B). The negative controls demonstrated that the majority of knockdown observed from each hairpin is due to Rnt1p processing (Figure 2.3B). The controls also indicate that the

hairpin structures can have slight effects on gene expression (compared to the construct with no hairpin insertion set at 100%), likely due to some effects of the inserted structures on normal translation or degradation processes.

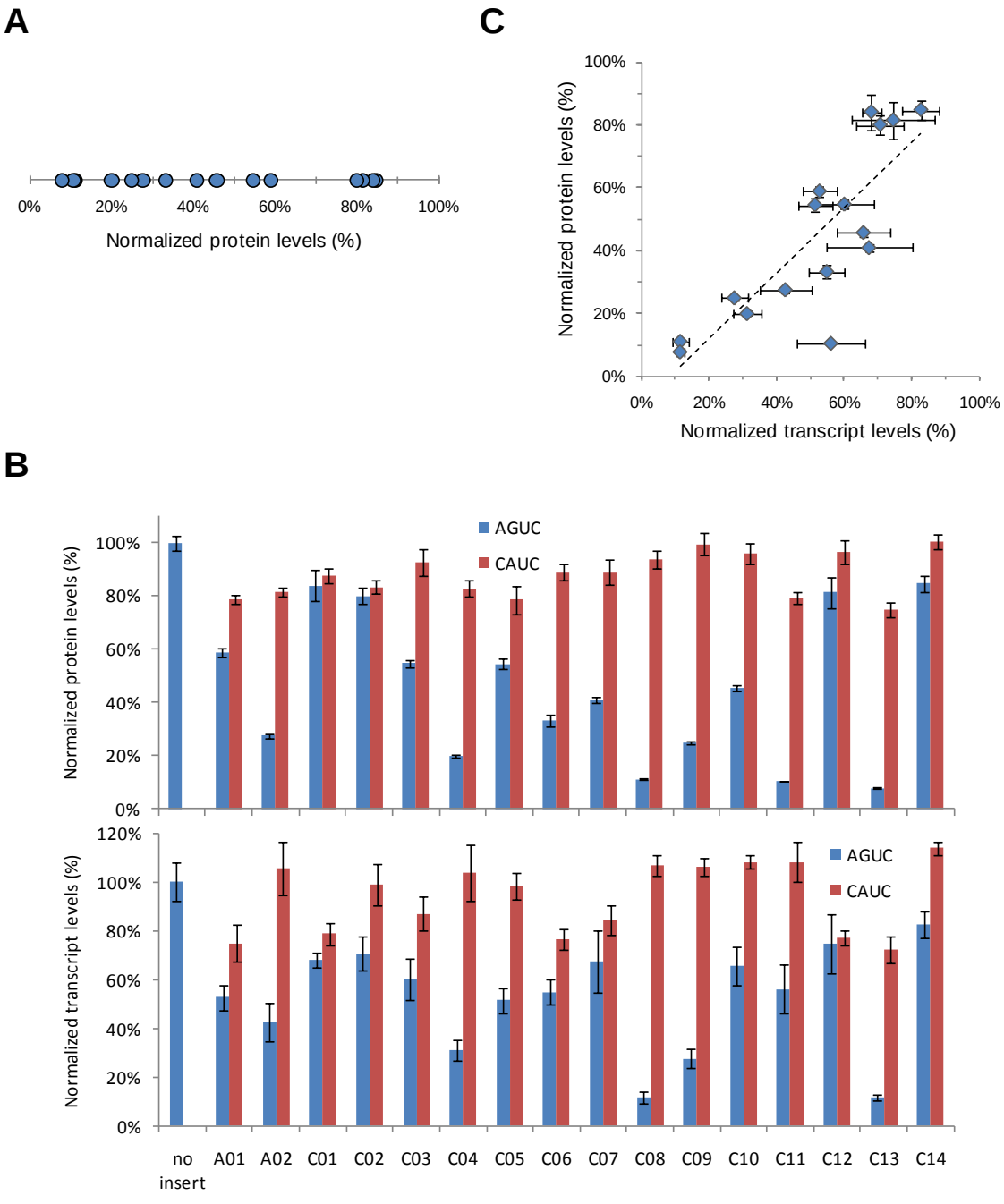


Figure 2.3. *In vivo* characterization of the selected Rnt1p cleavage library. (A) The gene regulatory range of the Rnt1p library spans a broad range of protein expression levels. (B) The transcript and protein levels associated with all Rnt1p library members and their corresponding mutated tetraloop (CAUC) controls supports that the observed gene regulatory activity is due to Rnt1p processing. (C) Correlation analysis of protein and transcript levels from the Rnt1p hairpin library members supports a strong correlation between the two measures of gene regulatory activity. All normalized protein and transcript levels and their error are determined as described in Figure 2.1D.

The activity of the synthetic Rnt1p hairpins was further confirmed by monitoring steady-state transcript levels in cells harboring the Rnt1p constructs (Table 2.1, Figure 2.3B). Rnt1p hairpins generally resulted in reduced transcript levels compared to a construct harboring no Rnt1p hairpins and to a construct harboring a mutated tetraloop. In addition, a plot of normalized *yEGFP3* expression levels versus normalized *yEGFP3* transcript levels indicates that there is a strong positive correlation ($r = 0.817$) between the two measures (Figure 2.3C). Specifically, with decreasing transcript levels a similar decrease in protein levels was generally observed, as further supported from a Spearman's rank correlation coefficient (ρ) value of 0.818. Unintended effects of the hairpins on translation and transcript stability caused by interference of the structures on the machinery controlling those processes may contribute to deviations from linearity. However, observed deviations are not entirely due to structural variability between the library members, as hairpins with similar secondary structure (i.e., A01, A02, C13) do not demonstrate an exact linear relationship between transcript and protein levels.

2.2.4. Rnt1p library hairpins maintain regulatory activity in a different genetic context

The utility of any genetic control element requires that the control module retain its activity under different genetic contexts. The modular function of the synthetic Rnt1p

hairpins, as measured by maintenance of gene regulatory activity, may be impacted by differences in 3' UTR, promoter, and gene sequences. We cloned the synthetic Rnt1p hairpins into a second construct harboring a different promoter (TEF1), terminator (CYC1), and gene (*ymCherry*), and measured the regulatory activities of each hairpin through flow cytometry and qRT-PCR assays (Supplementary Table 2.1). The data indicate a strong positive correlation ($r = 0.897$) and a strong preservation of rank order ($\rho = 0.882$) between the *ymCherry* protein and transcript levels, confirming that gene regulatory activity by the Rnt1p modules is due to the reduction of steady-state transcript levels (Figure 2.4A). Flow cytometry histograms confirm that regulatory activity associated with the Rnt1p hairpins in this second construct causes a population shift with reduced mean fluorescence levels (Supplementary Figure 2.3). The functional modularity of the hairpins was determined by performing a correlation analysis between expression data for the hairpins in the *yEGFP3* and *ymCherry* constructs (Figure 2.4B), which demonstrated a strong positive correlation ($r = 0.856$) between the two data sets. A lack of functional modularity was observed for one hairpin (C06), which did not maintain its gene regulatory activity in the second construct (33% yEGFP3 vs. 89% ymCherry) (Figure 2.4B; red point). The data suggests that the flanking sequences in the *ymCherry* construct may be disruptive to the structural integrity of C06, thereby affecting cleavage efficiency in the CEB, although RNA folding software does not predict alternative hairpin structures. The majority of data indicate that *ymCherry* expression tended to be slightly higher than that of *yEGFP3*, likely due to the differences in transcriptional strength between the TEF1 and GAL1 promoters, where TEF1 resulted in a greater absolute number of transcripts per cell (data not shown). The data indicate that there was

a strong preservation of rank order (ρ = 0.848) between the two different genetic contexts, supporting the functional modularity of the Rnt1p hairpin library.

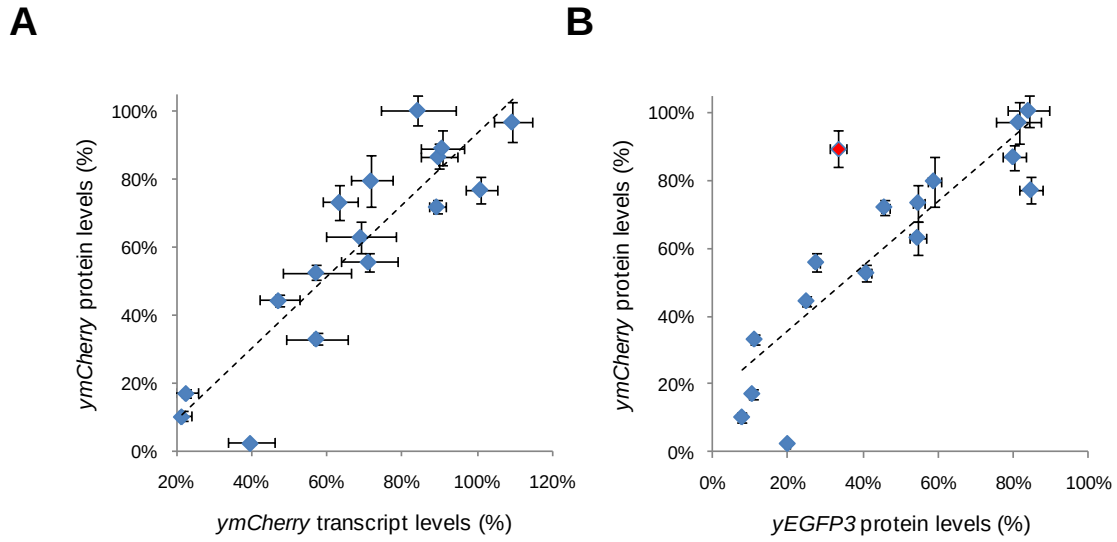


Figure 2.4. Demonstration of functional modularity of the hairpin library in the context of a different genetic construct. (A) Correlation analysis of ymCherry protein and transcript levels from the Rnt1p hairpin library members supports a strong correlation between the two measures of gene regulatory activity. Normalized protein and transcript levels and their error are determined as described in Figure 2.1D with the mean ymCherry fluorescence used for the protein level measurement. (B) Correlation analysis of ymCherry and yEGFP3 protein levels from the Rnt1p hairpin library members demonstrates a strong correlation between gene regulatory activities in different genetic contexts and preservation of library rank-order. Red data point, C06.

2.2.5. In vitro *characterization demonstrates that Rnt1p library members achieve differential activity through alterations in Rnt1p cleavage rates*

We hypothesized that the variation in transcript processing and subsequent protein expression levels exhibited by the Rnt1p hairpin library is due to alterations of Rnt1p cleavage rates through alterations of the CEB sequence and/or structure. We analyzed the reaction through a Michaelis-Menten model, with the substrate (S) being the hairpin

transcript, the enzyme (E) being Rnt1p, and the product (P) being the cleaved pieces of the transcript. Under these conditions, the following reaction occurs:

$$E + S \underset{k_{-1}}{\overset{k_1}{\leftrightharpoons}} ES \xrightarrow{k_2} E + P$$

The rate of product formation (V) is modeled as:

$$V = \frac{V_{max} * [S]}{K_M + [S]} = \frac{k_2 * [E]_0 * [S]}{K_M + [S]}$$

The maximum rate of product formation (V_{max}) is the product of the total enzyme concentration ($[E]_0$) and k_2. Alterations in the cleavage efficiency will have an effect on the value of k_2 and thus V_{max}.

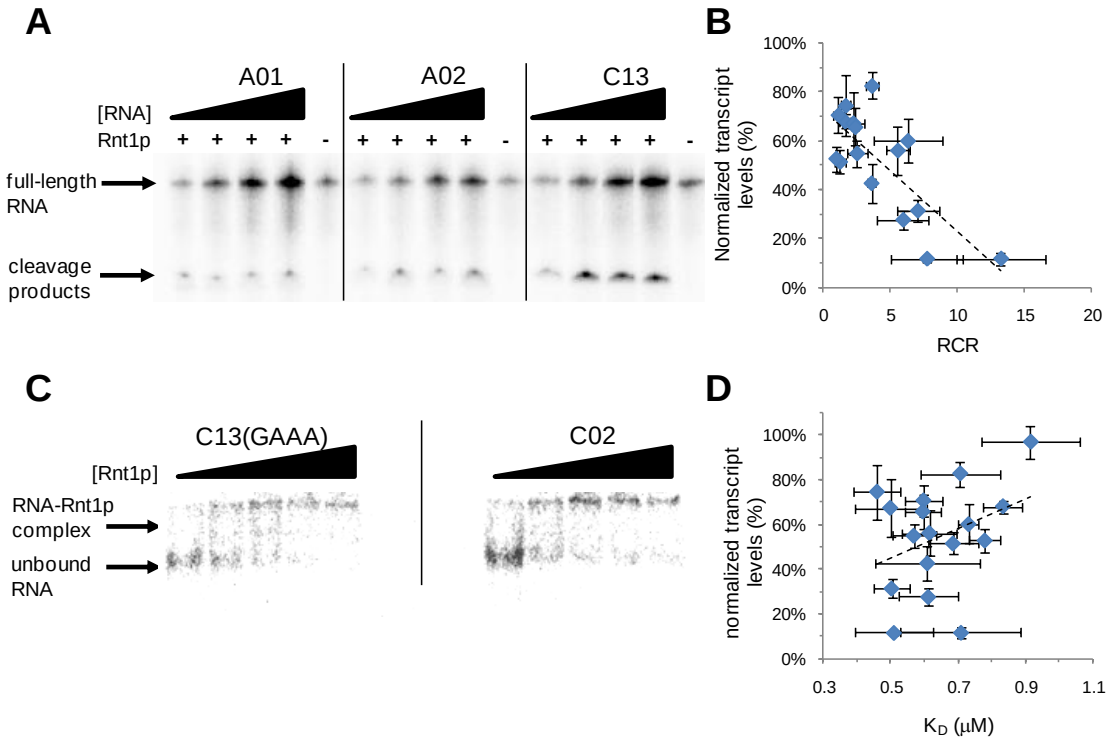


Figure 2.5. *In vitro* characterization of the Rnt1p library supports the tuning of gene regulatory activity through modulation of cleavage rates. (A) Representative cleavage reaction assays and analyses by denaturing polyacrylamide gel electrophoresis on hairpins A01, A02, and C13. The top band corresponds to full-length RNA; the bottom

band corresponds to the three cleavage products expected from Rnt1p processing. Due to added sequences flanking the Rnt1p hairpin for insulation, the three cleavage products differ in size by 1 nt and cannot be resolved into individual bands under the assay conditions. RNA is added to the following final concentrations in each reaction (left to right; in μM): 0.2, 0.35, 0.5, 0.6–0.8. Reactions lacking Rnt1p are with 0.2 μM RNA. (B) Correlation analysis of relative cleavage rate (RCR) and normalized *yEGFP3* transcript levels supports a strong correlation between cleavage rate and gene regulatory activity. Reported RCR values are determined from a Michaelis-Menten model parameter fit using Prism 5 (GraphPad) and standard error was calculated from the software. (C) Representative mobility shift assays and analyses by nondenaturing polyacrylamide gel electrophoresis on the mutated tetraloop (C13-GAAA) and C02. The top band corresponds to RNA-Rnt1p complexes; the bottom band corresponds to unbound RNA. Rnt1p is added to the following final concentrations in each reaction (left to right; in μM): 0, 0.42, 0.83, 1.25, 1.66. (D) Correlation analysis of binding affinity (K_D) and normalized *yEGFP3* transcript levels indicates a very weak correlation between binding affinity and gene regulatory activity. Reported K_D values are determined from a modified Scatchard model parameter fit using Prism 5 (GraphPad) and standard error was calculated from the software.

We performed *in vitro* RNA cleavage reactions with purified Rnt1p to determine relative values of k_2 for each synthetic Rnt1p hairpin. Reactions were run with varying concentrations of *in vitro* synthesized radiolabeled RNA encoding an Rnt1p hairpin flanked by A-rich sequences (see Materials and Methods) and a constant concentration of purified Rnt1p. Reaction products were separated by denaturing polyacrylamide gel electrophoresis and quantified through phosphorimaging analysis (Figure 2.5A). The resulting data were fit to the Michaelis-Menten model to calculate a relative cleavage rate (RCR), which is directly proportional to V_{max}. The RCR value for A01 is set to 1 and the rest of the reported values normalized to A01. The RCR values for each synthetic Rnt1p hairpin were determined through this analysis method (Table 2.2). There is a direct relationship ($r = –0.763$) between the measured RCR and gene regulatory activity for the synthetic Rnt1p hairpins (Figure 2.5B). Specifically, increasing Rnt1p's ability to cleave a substrate results in lowered transcript levels and thus lower protein expression levels. Notably, the transcript levels saturate at high RCRs, indicating that increasing the

cleavage rates above a certain threshold results in limiting decreases in transcript levels. The trend saturates at approximately 10% transcript levels, suggesting that increasing *in vitro* cleavage rates beyond an RCR value of approximately 8 will not result in an increase in the amount of transcript being processed *in vivo* and that we have approached the maximum amount of knockdown that can be achieved with a single substrate in this system. In initial control tests, we found that the mutant tetraloop (CAUC) was cleaved *in vitro* under excessive protein concentrations (i.e., nine times greater than that used in the cleavage assay). As such, another mutant tetraloop (GAAA) that exhibited no cleavage under excessive protein concentrations *in vitro* was used as a control for the cleavage assays.

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

Substrate	RCR	K_D (μM)
C01	1.66 ± 0.43	0.83 ± 0.06
C02	1.17 ± 0.41	0.60 ± 0.05
C03	6.33 ± 2.51	0.73 ± 0.03
C04	7.06 ± 1.56	0.50 ± 0.05
C05	1.28 ± 0.36	0.68 ± 0.07
C06	2.55 ± 0.73	0.57 ± 0.06
C07	2.27 ± 0.81	0.50 ± 0.11
C08	13.25 ± 3.29	0.71 ± 0.18
C09	5.98 ± 1.92	0.61 ± 0.09
C10	2.42 ± 0.25	0.60 ± 0.06
C11	5.58 ± 0.83	0.61 ± 0.08
C12	1.71 ± 0.37	0.46 ± 0.07
C13	7.75 ± 2.64	0.51 ± 0.12
C14	3.66 ± 0.44	0.71 ± 0.12
A01	1.00 ± 0.12	0.78 ± 0.05
A02	3.62 ± 0.32	0.61 ± 0.16
C13 (GAAA)	0*	0.91 ± 0.15

*Immeasurable due to lack of product formation

While changes in the CEB are anticipated to result in changes to the Rnt1p processing efficiencies, it is also possible that the introduced sequence alterations may result in changes to the binding affinities between the hairpins and Rnt1p. To examine

whether the synthetic Rnt1p hairpins exhibit any changes in binding affinity to Rnt1p, we performed *in vitro* binding assays with purified Rnt1p. Binding reactions were run with 20 nM of *in vitro* synthesized radiolabeled RNA encoding an Rnt1p hairpin and varying concentrations of purified Rnt1p in the absence of magnesium. As magnesium and other divalent metal ions are essential to Rnt1p function[41], these reaction conditions allow for Rnt1p to bind to the substrates without subsequent cleavage. Bound products were separated by nondenaturing polyacrylamide gel electrophoresis and quantified through phosphorimaging analysis (Figure 2.5C). We analyzed the reaction through a modified Scatchard equation in which the fraction of unbound RNA (R) to total RNA (R_0) is plotted against the enzyme (E) concentration. The equation is as follows:

$$Z = \frac{R}{R_o} = \frac{K_D}{K_D + [E]}$$

The dissociation constant, K_D, for each synthetic Rnt1p hairpin was determined through this analysis method (Table 2.2). The data indicate that there is no correlation between K_D and *in vivo* gene regulatory activity (Figure 2.5D). The values of K_D for the synthetic Rnt1p hairpins span a narrow range and are weakly correlated with transcript knockdown ($r = 0.351$). A lack of correlation was anticipated due to expectations of mutations in the CEB primarily affecting cleavage. For a series of four hairpins with approximately 70% transcript levels, the reported K_D cover the range of the entire library, suggesting that nucleotide modifications in the CEB can have an effect on protein binding. The mutant tetraloop control binds with a similar K_D as the library hairpins, albeit the binding is weaker than the library. While the mutant tetraloop does not impact binding greatly, it severely impacts the ability of Rnt1p to bind in a conformation that allows cleavage, which has been previously reported[39].

2.2.6. Control of endogenous ERG9 expression by 3' UTR replacement with Rnt1p library members

In many metabolic engineering applications, there is a balance that must be maintained between diverting cellular metabolites to the production of desired compounds and the conversion of those metabolites to molecules required for cell growth and viability. In these cases, completely knocking out endogenous genes to remove the drain caused by native cellular pathways is not an option and genetic tools that allow for precise titration of enzyme expression are desired such that flux through the endogenous pathway can be minimized to that required to maintain cell viability. Farnesyl pyrophosphate (FPP) is one such cellular metabolite that is an important precursor both to industrially-relevant molecules and to molecules required for cell viability in yeast[19, 42-43]. Squalene synthase, encoded by the *ERG9* gene, is responsible for catalyzing the conversion of two molecules of FPP to squalene, the first precursor in the ergosterol biosynthetic pathway in *S. cerevisiae*[44] (Figure 2.6A). In a series of 14 catalytic steps, squalene is converted to ergosterol, the analogue of cholesterol in mammalian cells. Ergosterol is an essential component of yeast cells due to its effect on structural stability of the cell membrane[22]. Therefore, controlled reduction of ergosterol levels will allow metabolic flux to be diverted from sterol synthesis to value-added products from FPP.

We examined the ability of our posttranscriptional genetic control modules to modulate flux through the ergosterol biosynthetic pathway in a predictable manner by incorporating several members of the Rnt1p library into the 3' UTR of the *ERG9* gene. *ERG9* is in close proximity to *CTF8* on the reverse strand of the chromosome (~50 nts

between stop codons), and there is a lack of information on the transcription terminators

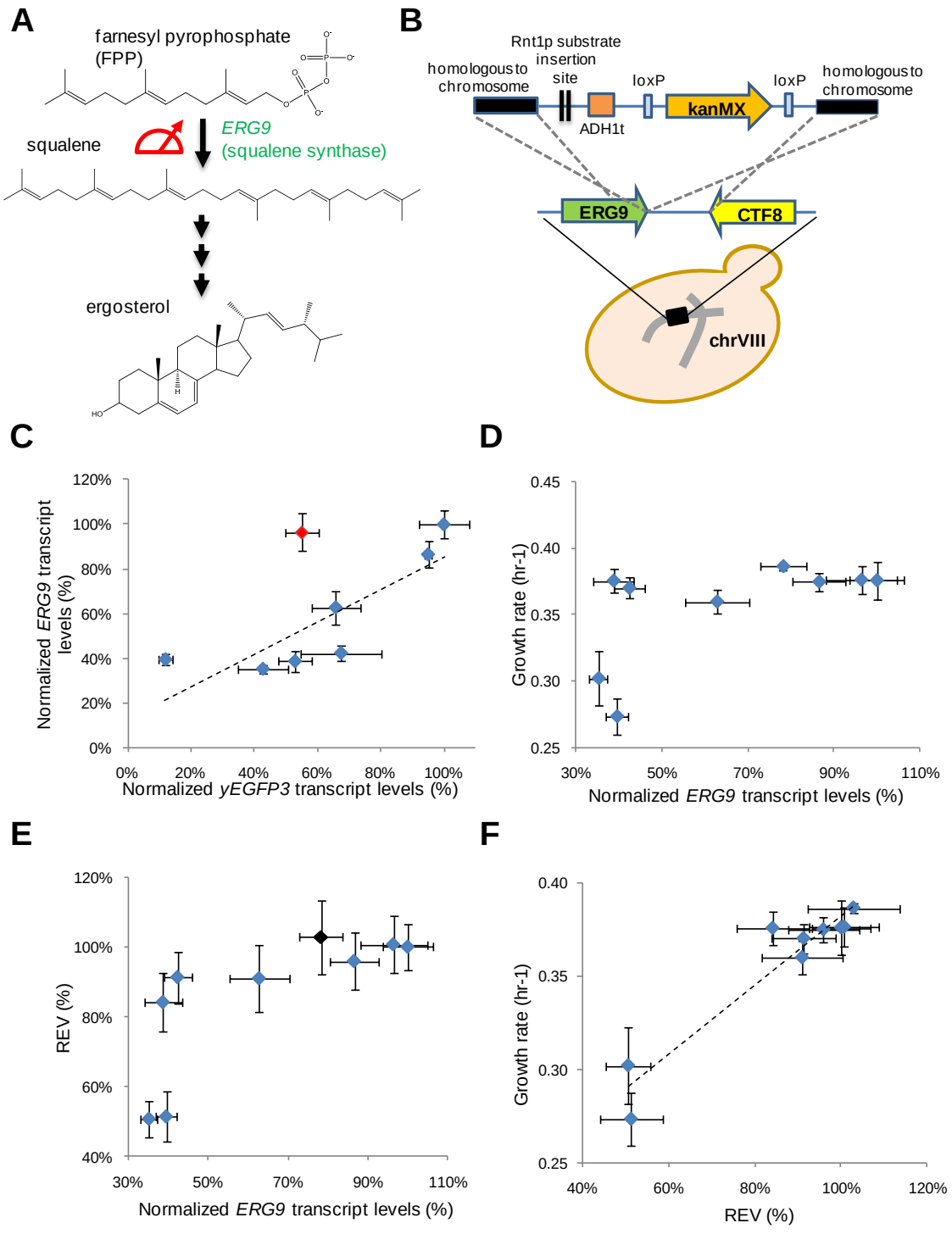


Figure 2.6. Synthetic Rnt1p hairpins enable posttranscriptional control over endogenous *ERG9* expression levels. (A) Simplified schematic of ergosterol biosynthesis from FPP showing key components for this work. Squalene is converted to ergosterol through 14

enzymatic steps. The dial highlights that *ERG9* levels are tuned with the synthetic Rnt1p control modules. (B) Schematic of the construct and strategy utilized for introducing the synthetic Rnt1p control modules into the 3' UTR of the endogenous *ERG9* gene. The construct is designed to replace the native *ERG9* 3' UTR with a synthetic 3' UTR harboring an Rnt1p hairpin through homologous recombination between the integration cassette and chrVIII. The illustrated strategy maintains the native feedback regulation acting through transcriptional mechanisms to control *ERG9* levels. (C) Correlation analysis of *yEGFP3* transcript levels and *ERG9* transcript levels indicates that the synthetic Rnt1p hairpins maintain their gene regulatory activity in a different genetic context. Normalized *ERG9* transcript levels and their error are determined as described in Figure 2.1D. Red data point, C06. (D) Correlation analysis of cellular growth rate and *ERG9* transcript levels indicates that the titration of *ERG9* levels results in two distinct phenotypic regimes – 'fast-growing' and 'slow-growing'. Growth rates are determined by measuring the OD_{600} during a time course and fitting the data to an exponential growth curve using Prism 5 and standard error was calculated from the software. Black data point, wild-type yeast strain. (E) Correlation analysis of relative ergosterol values (REVs) and *ERG9* transcript levels indicates that ergosterol levels remain relatively consistent across varying *ERG9* levels above a certain threshold value (approximately 40% normalized transcript levels). REVs are determined by extracting unsaponified sterols and measuring the absorbance of signature peaks associated with ergosterol in the UV spectrum. Reported REV values and their error are calculated from the mean and standard deviation from the three identical aliquots from sterol extractions, respectively (F) Correlation analysis of cellular growth rate and REV indicates that the two phenotypic measures of *ERG9* levels are strongly correlated.

of both genes. Thus, we designed a construct to integrate the entire 3' UTR and ADH1 terminator from the library plasmid (pCS321) immediately following the *ERG9* stop codon and before the intervening nucleotides between *ERG9* and *CTF8* (Figure 2.6B). We built an Rnt1p control module integration plasmid based on the library plasmid, where *yEGFP3* was replaced by *ERG9* and the marker *loxP-KanMX-loxP*[45] was inserted downstream of the ADH1 terminator to provide resistance against G418. PCR-amplification from the integration plasmid results in DNA cassettes that can be directly integrated into the desired location of the yeast genome. Six members of the synthetic Rnt1p library (A01, A02, C06, C07, C08, C10) along with a mutant tetraloop control (C13, GAAA tetraloop) were integrated into the 3' UTR of *ERG9* to cover the regulatory range of the library.

The regulatory activity of the seven examined Rnt1p hairpins for *ERG9* was initially assessed by measuring *ERG9* transcript levels (Table 2.3). The *ERG9* transcript levels were directly compared to the transcript levels measured from the *yEGFP3* construct. The Rnt1p hairpins resulted in reduced transcript levels compared to the integrant harboring no Rnt1p hairpins and to the construct harboring a mutated tetraloop. The transcript levels of the no hairpin integrant were higher than the transcript levels from the wild-type yeast strain, indicating that the 3' UTR replacement results in increased *ERG9* transcript levels, likely due to altered transcript stability. In addition, a plot of the normalized *yEGFP3* transcript levels versus normalized *ERG9* transcript levels for each synthetic Rnt1p hairpin reveals a strong positive correlation ($r = 0.844$) between the two measures when the hairpin C06 is excluded from the analysis, supporting the ability of the synthetic Rnt1p hairpins to act as predictable genetic control modules (Figure 2.6C). The modularity of the hairpin set (excluding C06) was further supported by a ρ-value of 0.771, indicating a preservation of rank order. C06 was previously determined to not maintain function in the context of the *ymCherry* construct (Figure 2.4B; red point). The difference in regulatory activity observed from the C06 hairpin in the context of the endogenous *ERG9* gene (Figure 2.6C; red point) further suggests that C06 may not be as well-insulated from different genetic contexts as the other tested hairpins, although RNA folding software does not predict alternative structures. In addition, the *ERG9* levels flatten and do not drop below ~40%, indicating that the natural feedback regulation associated with the *ERG9* promoter may act to maintain levels at this minimum value.

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

Substrate	Normalized *ERG9* transcript levels (%)	Growth rate (hr^{-1})	REV (%)
wild-type	78% ± 5%	0.386 ± 0.003	103% ± 11%
no insert	100% ± 6%	0.376 ± 0.014	100% ± 7%
C13 (GAAA)	87% ± 6%	0.375 ± 0.007	96% ± 8%
C06	96% ± 8%	0.376 ± 0.010	101% ± 8%
C07	42% ± 4%	0.370 ± 0.008	91% ± 7%
C08	40% ± 3%	0.273 ± 0.014	51% ± 7%
C10	63% ± 7%	0.360 ± 0.009	91% ± 9%
A01	39% ± 5%	0.375 ± 0.009	84% ± 8%
A02	35% ± 2%	0.302 ± 0.021	51% ± 5%

While sterol synthesis is vital to cell growth, knockouts of enzymes in the downstream ergosterol biosynthetic pathway are viable due to the retained ability to incorporate intermediate sterols into the cellular membrane[46]. However, knockouts of enzymes in the early part of the pathway, including *ERG9* and enzymes leading up to the production of FPP, are lethal. As such, we examined the effect of decreased *ERG9* expression on the cell growth rate. The OD_{600} of yeast strains harboring the different Rnt1p hairpins in the 3' UTR of *ERG9* was measured during the exponential growth phase. The growth rate, k, was calculated by fitting the OD_{600} data to an exponential growth curve (Table 2.3). A plot of the growth rate versus *ERG9* transcript levels for each strain reveals that above a certain threshold level of *ERG9* production, differences in growth rates in this 'fast-growing' regime are negligible (Figure 2.6D) This data suggest that decreasing the amount of *ERG9* in this regime does not significantly impact the flux through the ergosterol biosynthetic pathway due to overproduction of *ERG9* in the wild-type strain or feedback control of *ERG9*[47]. Below a certain threshold level a 'slow-growing' regime is observed, characterized by a substantial drop-off in the cell growth

rate. Interestingly, the two slow-growing strains harbor the Rnt1p hairpins exhibiting the strongest gene silencing activities, but demonstrate similar *ERG9* levels to two of the fast-growing strains. It is possible that feedback regulation is acting to increase *ERG9* expression to the desired setpoint in the slow-growing strains, but the perturbations introduced in these strains result in other impacts on the pathway that inhibit the endogenous control systems from restoring cellular growth to wild-type rates.

In order to better understand the cellular processes linking *ERG9* production and cell growth rate, we measured the amount of ergosterol, the end product of the *ERG9* biosynthetic pathway. Each culture was inoculated with the same amount of cells and allowed to grow for 8 hr before the cultures were saponified. The UV spectrum of unsaponified sterols was determined and used to calculate a relative ergosterol value (REV), which is normalized against the control containing no Rnt1p substrate (set to 100%) (Table 2.3). A plot of REV versus *ERG9* transcript levels reveals a similar relationship as observed between growth rate and *ERG9* levels, with the exception that the fast-growing strains exhibit a slight positive correlation between REV and *ERG9* transcript levels (Figure 2.6E). A plot of growth rate versus REV illustrates the two regimes (slow-growing and fast-growing) and highlights the strong positive correlation between these two phenotypic measures ($r = 0.953$) (Figure 2.6F). There is little difference in the amount of ergosterol per cell (determined as the ratio of REV and OD_{600}), indicating that alteration of flux through the *ERG9* pathway results in changes in the time required for a cell to produce sufficient ergosterol to duplicate and not decreased levels of ergosterol molecules in the cell membrane.

2.3. Discussion

We have developed a novel class of genetic control modules in *S. cerevisiae* based on Rnt1p cleavage. Although an Rnt1p substrate has been shown to play a role in regulating the expression of the endogenous *MIG2* gene[36], our work describes the first synthetic gene regulatory system based on engineered Rnt1p hairpins. A library of synthetic Rnt1p hairpins that span a wide range of gene regulatory activities was generated to act as posttranscriptional control modules by placing these elements in the 3' UTR of a target gene. To ensure the modularity of the synthetic Rnt1p substrates, two design strategies were implemented. First, a 'clamp' region was added to the base of each hairpin. Second, only sequences that formed single predicted hairpin structures at the lowest free energies were included within the Rnt1p control module set. These properties minimize any potential of flanking sequences to disrupt desired folding of the control modules, thereby reducing the likelihood of varying function within different genetic contexts. We observed a significant improvement in the correlation between transcript and protein levels for hairpins that exhibited both properties (Supplementary Figure 2.4), suggesting that undesired interactions between hairpin and flanking sequences can affect translation. The functional modularity of the resulting hairpin library, as measured by maintenance of regulatory activity and rank order, was demonstrated under three different genetic contexts (Figure 2.4B, Figure 2.6C), where in each case one hairpin did not retain its expected activity, likely due to improper folding. Our studies indicate that for any given genetic system, one of the library members (less than 10%) may not exhibit regulatory activity, where the coverage of the reported library will allow a researcher to select hairpins that span the desired regulatory range.

Previous *in vitro* studies identified three regions of the hairpin substrates critical to Rnt1p cleavage activity: the CEB, BSB, and IBPB[29]. The BSB and IBPB are regions of the hairpin that affect the overall binding of the protein, such that modifications of the nucleotides in these regions can inhibit Rnt1p binding and subsequent cleavage. The CEB is a region of the hairpin that affects the processing of the stem by Rnt1p, such that nucleotide modifications in this region are expected to specifically modulate the cleavage rate. We developed an Rnt1p library based on randomization of the CEB region and screened this library *in vivo* for substrates with altered processing efficiencies. While earlier *in vitro* studies retained base-pairing within the CEB of modified Rnt1p substrates[29], our library screen demonstrated that the CEB has substantial structural flexibility in maintaining function *in vivo*, as the majority of library members retained processability. Cleavage assays with purified Rnt1p support that variations in the CEB within the set of synthetic Rnt1p substrates alter the processing efficiency in a manner that is directly correlated with the observed gene regulatory activity, whereas binding assays indicate no relationship between binding affinity and processing efficiency (or gene regulatory activity). The data support that modifications in the CEB directly influence the ability of Rnt1p to cleave the hairpin, whereas interactions between the nucleotides in the CEB and RBDs of Rnt1p likely lead to the small variation in observed K_D values[29].

The set of synthetic Rnt1p substrates developed in this work represents the first engineered library of transcript stability control modules in the yeast *S. cerevisiae*. While other posttranscriptional regulatory elements, such as internal ribosome entry sites (IRESes) and AU-rich elements (AREs), have been applied to regulate heterologous gene

expression in yeast, such genetic elements have exhibited substantial variability in activity and have not been engineered as synthetic libraries of control modules exhibiting a wide range of activities[48-50]. In addition, a library of short synthetic internal ribosome entry sites (IRESes) that act through translation initiation was previously developed for yeast[51]. However, these short IRESes result in substantially reduced expression levels compared with cap-dependent translation mechanisms, such that the resulting library spans a much narrower range of regulatory activities than exhibited by the synthetic Rnt1p library. The most similar control module library is one that was developed based on mutating a constitutive promoter (TEF1) in yeast, which spans a similar range of gene regulatory activities as the described Rnt1p hairpin library (TEF: 8–120%; Rnt1p: 2–100%)[26]. However, the two libraries exhibit different coverage of these ranges, where the Rnt1p library provides greater coverage of expression levels between 2–60% and the TEF promoter library provides greater coverage between 60–100%.

A unique advantage of control modules based on posttranscriptional processes is that such elements can be readily used in combination with one another and with other genetic control modules, such as promoter elements and other transcriptional regulators, to achieve more finely-tuned and expanded regulatory schemes. As one example, inducible promoters are commonly used to turn on and set the expression levels of genes by controlling the concentration of the inducing molecule exogenously added to the system. However, such transcriptional control modules on their own are limited to applying identical regulatory activities to multiple gene targets within a given system. The combination of an inducible promoter and our engineered Rnt1p substrates will allow for the relative gene regulatory activities at a given inducer concentration to be

modulated (based on the Rnt1p substrate), thus enabling ligand-mediated control over multiple genes with different expression levels.

In addition to expanding the utility of inducible promoters in the context of multi-gene circuits, the ability to combine posttranscriptional control modules with any promoter element has key advantages in the context of endogenous networks. Endogenous networks often have critical control strategies in place, such as feedback regulation, that commonly operate at the level of transcriptional processes. The combination of the synthetic Rnt1p substrates with endogenous genetic targets allows specific engineered control strategies to be added to a system, while retaining native regulatory schemes that may play an important role in the overall system operation. Therefore, these posttranscriptional control modules provide a useful toolset for predictably modulating specific components in complex biological systems and can be further used to probe and study native regulatory networks. One consideration in the implementation of these genetic modules is that their gene regulatory activities may be affected by variation in the ratio of cellular levels of Rnt1p to transcript levels. While absolute activities of the synthetic Rnt1p hairpins are expected to vary with substantial changes in Rnt1p levels, the rank order of the hairpin activities is expected to be maintained.

To demonstrate the utility of these posttranscriptional control modules, we implemented a synthetic control strategy directed to modulating a key enzyme component of the endogenous ergosterol synthesis network by combining the Rnt1p control modules with the *ERG9* genetic target. Previous work had replaced the endogenous *ERG9* promoter with a MET3 repressible promoter, and demonstrated a

sharp decrease in ergosterol levels with full transcriptional repression[12]. In contrast, our engineered control strategy was anticipated to allow the system to retain previously identified transcriptional feedback control around the *ERG9* gene[47], while allowing for titration of *ERG9* transcript levels. Generally, transcript levels between *ERG9* and *yEGFP3* with a given hairpin correlated strongly. However, relative *ERG9* levels did not fall below ~40% regardless of the Rnt1p hairpin strength, indicating an endogenous feedback mechanism that maintains *ERG9* expression levels at that threshold value. Interestingly, the data indicate a 'buffer' region in the endogenous control strategy around *ERG9* levels, where wild-type levels are set substantially higher than that threshold value. The synthetic Rnt1p hairpin set allowed for systematic titration of *ERG9* expression levels and the identification of two regimes for the system. Strains expressing over ~40% *ERG9* transcript levels exhibited high ergosterol levels and growth rates. Strains harboring two synthetic Rnt1p hairpins resulting in the lowest expression levels exhibited a significant reduction in the amount of ergosterol produced and growth rate. Interestingly, these 'slow-growing' strains have similar levels of *ERG9* to two strains in the 'fast-growing' regime. One possible explanation for these observations is that the diminished cellular sterol levels result in positive regulation of the *ERG9* promoter below a certain threshold value, maintaining expression at a minimum level. However, dialing down *ERG9* levels below a critical value can affect the cell through unknown mechanisms that do not permit restoration of ergosterol levels or cell growth rate through the endogenous control system[47]. Therefore, this work supports the unique ability of the synthetic Rnt1p hairpin library to systematically titrate pathway enzyme levels while maintaining native cellular control strategies acting through transcriptional mechanisms.

In summary, we have developed a library of genetic control modules for yeast that can be implemented with different genetic targets and promoters to predictably tune gene expression levels. The Rnt1p library provides a key tool for synthetic biology applications in yeast, which can function to rationally dial in expression levels similar to well-developed control modules in bacteria such as ribosome binding sites (RBS). However, unlike RBS elements the structural and functional insulation of the synthetic Rnt1p controllers provide for more successful maintenance of regulatory activities across different genetic contexts[52]. We show here that the synthetic controllers can be applied to predictably modulate flux through metabolic pathways and probe regulation schemes in endogenous networks by introducing precise perturbations around major control points. With growing interests in eukaryotic hosts and complex networks in synthetic biology, and more specifically yeast in bioprocessing and biosynthesis applications, the synthetic controllers developed here will provide an important foundational toolset for the rapidly growing field.

2.4. Materials and Methods

2.4.1. Plasmid construction

Standard molecular biology techniques were utilized to construct all plasmids[53]. DNA synthesis was performed by Integrated DNA Technologies (Coralville, IA) or the Protein and Nucleic Acid Facility (Stanford, CA). All enzymes, including restriction enzymes and ligases, were obtained through New England Biolabs (Ipswich, MA) unless otherwise noted. Pfu polymerases were obtained through Stratagene. Ligation products were electroporated into *Escherichia coli* DH10B (Invitrogen, Carlsbad, CA), where cells

harboring cloned plasmids were maintained in Luria-Bertani media containing 50 mg/ml ampicillin (EMD Chemicals). Clones were initially verified through colony PCR and restriction mapping. All cloned constructs and chromosomal integrations were sequence verified by Laragen (Los Angeles, CA) or the Protein and Nucleic Acid Facility (Stanford, CA). Plasmid maps are available in Supplementary Figure 2.5.

A yeast-enhanced GFP gene, *yEGFP3*, was PCR-amplified from pSVA13[38] using forward and reverse primers GFP.mono.di.fwd (5' GCAAGCTTGGAGATCTAAAAGA AATAATGTCT) and GFP.mono.rev (5' CGCTCGAGGCCTAGGCTTTATTTGTACA ATT), respectively. The plasmid pCS182 was constructed by inserting the *yEGFP3* PCR product into a modified version of pRS316[54] harboring the GAL1-10 promoter via the unique restriction sites HindIII and XhoI located in the multiple cloning site (MCS) downstream of the GAL1-10 promoter. The ADH1 terminator was PCR-amplified from pSVA13[38] using the forward and reverse primers ADH1t_fwd (5' GCACCTCGAGAGG GCGCGCCACTTC) and ADH1t_rev (5' GCACGGTACCTATATTACCCTGTTATCC CTAGCGG), respectively. The base Rnt1p substrate characterization plasmid (pCS321) was constructed by inserting the ADH1 terminator PCR product into pCS182 via the unique restriction sites XhoI and KpnI located in the MCS. A *ymCherry* characterization plasmid, pCS1749, was constructed from pCS321 by replacing the GAL1-10 promoter with the endogenous TEF1 promoter and by replacing the *yEGFP3* open reading frame (ORF) the ADH1 terminator with the ORF of *ymCherry* and the CYC1 terminator (J. Liang et al., in preparation).

The endogenous *S. cerevisiae* gene *RNT1* was PCR-amplified directly from the yeast genome by colony PCR using forward and reverse primers Rnt1p_prmr_fwd (5'

GACCATGGGGATGGGCTCAAAAGTAGCAGGTAAAAAGAAAACC) and Rnt1p_ prmr_rev (5’ TGCTCGAGTCAGCTTGTATCTGAGAATTTTCTTTTCTTATTCTTTT GTGAG), respectively. The Rnt1p expression plasmid (pRNT1) was constructed by inserting the *RNT1* PCR product into the pProEx HT plasmid (Stratagene) via the unique restriction sites NcoI and XhoI located in the MCS downstream of the 5’ $(His)_6$ tag, the spacer region, and rTEV protease cleavage site.

The endogenous *S. cerevisiae* gene *ERG9* was PCR-amplified directly from the yeast genome by colony PCR using forward and reverse primers ERG9_321_prmr_fwd (5’ GCGAAGCTTGGAGATCTAAAAGAAATAATGGGAAAGCTATTACAATTGGC ATTGCATCC and ERG9_321_prmr_rev (5’ GGCTCGAGGCCTAGGCTTCACGCTCT GTGTAAAGTGTATATATAATAAAACCCAAGAAGA), respectively. The plasmid pCS321-ERG9 was constructed by replacing the *yEGFP3* sequence in pCS321 with *ERG9* by cloning the *ERG9* PCR product into the unique restriction sites HindIII and XhoI. The plasmid pUG6[45] was modified by removing the unique XhoI restriction site by site-directed mutagenesis via the oligonucleotides c1546g (5’ GTGTCGAAAACGAGCT CTGGAGAACCCTTAATATAAC) and c1546g_antisense (5’ GTTATATTAAGGGTT CTCCAGAGCTCGTTTTCGACAC) and the PfuUltra II polymerase (Stratagene). A PCR product harboring the full *ERG9* coding sequence through the 3’ end of the ADH1 terminator was amplified from pCS321-ERG9 using forward and reverse primers ERG9hpADH1t-SalI_fwd_prmr (5’ CAACGTCGACATGGGAAAGCTATTACAATTG GCA) and ERG9hpADH1t-SalI_rev_prmr (5’ AAGTGTCGACTATATTACCCTGTTA TCCCTAGCGG), respectively. The ERG9-RNT1 integration plasmid (pCS1813) was

constructed by inserting this PCR product into the modified pUG6 plasmid via the unique restriction site SalI located directly upstream of the first loxP site.

Insertion of engineered Rnt1p substrates and appropriate controls into the 3' UTR of pCS321, pCS1749, and pCS1813 was facilitated through either digestion with the appropriate restriction endonucleases and ligation-mediated cloning or homologous recombination-mediated gap-repair during transformation into *S. cerevisiae* strain W303 (*MATa, his3-11,15 trp1-1 leu2-3 ura3-1 ade2-1*) through standard lithium acetate procedures[55]. The Rnt1p substrates were amplified for insertion into pCS321 and pCS1813 with both techniques using the forward and reverse primers RntGap321_fwd (5' ACCCATGGTATGGATGAATTGTACAAATAAAGCCTAGGTCTAGAGGCG) and RntGap321_rev2 (5' TAAGAAATTCGCTTATTTAGAAGTGGCGCGCCCTCTCG AGGGCG), respectively. The Rnt1p substrates were amplified for insertion into pCS1749 by gap-repair using the forward and reverse primers mCherry_gap_fwd_prmr (5' GGTGGCATGGATGAACTATACAAATAATAAAGCCTAGGTCTAGAGGCG) and mCherry_gap_rev_prmr (5' TGACATAACTAATTACATGATGCGGCCCTCCCC TCTCGAGGGCG). In the case of digestion and ligation, the PCR products were digested with the unique restriction sites AvrII and XhoI, which are located 3 nts downstream of the *yEGFP3* or *ERG9* stop codon and upstream of the ADH1 terminator. Following construction and sequence verification of the desired vectors, 100–500 ng of each plasmid was transformed into W303. In the case of gap-repair (for pCS321 and pCS1749), 250–500 ng of the PCR product and 100 ng of pCS321 digested with AvrII and XhoI were transformed into the yeast strain. All yeast strains harboring cloned

plasmids were maintained on synthetic complete media with an uracil dropout solution containing 2% dextrose at 30°C.

2.4.2. 3' UTR replacement cassette and integration

The ERG9-RNT1 replacement cassettes were synthesized through PCR amplification from the appropriate pCS1813-based plasmids using forward and reverse primers ERG9-1150_fwd_prmr (5' AATTACCTCCTAACGTGAAGCCAAATGAAAC TCCAATTTTCTTGAAAGTT) and Rnt1p_cassette_rev_prmr2 (5' GGCCTCTACCTAT TATGTAAGTACTTAGTTATTGTTCGGAGTTGTTTGTTAATACGACTCACTATA GGGAGACCGGCAGA), respectively. These PCR products extend from 1150 nts into the *ERG9* gene to the end of the second loxP site with an overhang extension comprising 50 nts of homology to the native *ERG9* 3' UTR. Each integration cassette (~1–5 μg) was transformed into yeast as previously described. The integrants were selected and maintained on YPD plates with 200 mg/ml G418.

2.4.3. Rnt1p substrate characterization assays

S. cerevisiae cells harboring pCS321-based and pCS1749-based plasmids were grown on synthetic complete media with an uracil dropout solution and the appropriate sugars (2% raffinose and 1% sucrose for pCS321; 2% dextrose for pCS1749) overnight at 30ºC. The cells were back-diluted the following morning into fresh media (4.5 ml total volume in test tubes and 450 μl in deep-well plates) to an optical density at 600 nm (OD_{600}) of 0.1 and grown again at 30ºC. For pCS321-based plasmids, after 1 hr, 0.5 ml (test tubes) or 50 μl (plates) of 20% galactose (2% final concentration) or water (non-

induced control) was added to the cell cultures. The cells were grown for another 4.5 hr before measuring the fluorescence levels or collecting cells for RNA extraction.

S. cerevisiae integrated with Rnt1p hairpins or its controls were grown on YPD overnight at 30ºC. The cells were back-diluted the following morning into fresh media (5 ml total volume in test tubes) and grown again for 3 hr at 30ºC. After 3 hr, the cells were back-diluted to an OD_{600} of 0.1 (for RNA extraction and growth rate determination, 5 ml total volume) or 0.05 (for ergosterol quantification, 7 ml total volume) and grown for an appropriate length of time dependent on the application.

2.4.4. Fluorescence quantification

On a SAFIRE plate reader (TECAN, Männedorf, Switzerland), GFP fluorescence was read from 200 µl of cells with an excitation wavelength of 485 nm, an emission wavelength of 515 nm, and a gain of 100. The population-averaged fluorescence readings were normalized to the amount of cells by dividing the relative fluorescence units (RFU) by the OD_{600} of the sample. On the Quanta flow cytometer (Beckman Coulter, Fullerton, CA), the distribution of GFP fluorescence was measured with the following settings: 488-nm laser line, 525-nm bandpass filter, and photomultiplier tube setting of 5.83. Data were collected under low flow rates until 10,000 viable cell counts were collected. A non-induced cell population was used to set a gate to represent GFP-negative and GFP-positive populations. The median fluorescence of the positive population was measured from three identically grown samples. The LSRII flow cytometer (Becton Dickinson Immunocytometry Systems) was used to measure ymCherry fluorescence from p1749-based plasmids. ymCherry was excited at 532 nm and measured with a splitter of 600 nm

LP and a bandpass filter of 610/20 nm. A DAPI stain (excited at 405 nm and measured with a bandpass filter of 450/50 nm) was utilized to gate for cell viability. The mean fluorescence was measured from three identically grown samples and baseline-subtracted with an empty vector control. Reported values and their error are calculated from the mean and standard deviation from the triplicate data, respectively.

2.4.5. Quantification of cellular transcript levels

Total RNA from *S. cerevisiae* was collected by a standard hot acid phenol extraction method[56] and followed by DNase I (New England Biolabs) treatment to remove residual plasmid DNA according to manufacturer's instructions. cDNA was synthesized from 5 μg of total RNA with gene-specific primers for *yEGFP3*, *ymCherry*, *ERG9*, and *ACT1*[57] (rnt1p_rtpcr_rev2 and ACT1_rtpcr_rev, respectively) and SuperScript III Reverse Transcriptase (Invitrogen) according to manufacturer's instructions. The forward and reverse primers for *yEGFP3* quantification are rnt1p_rtpcr_fwd2 (5' CGGT GAAGGTGAAGGTGATGCTACT) and rnt1p_rtpcr_rev2 (5' GCTCTGGTCTTGTAGT TACCGTCATCTTTG), respectively; for *ymCherry* quantification are mCherry_qrtpcr_ fwd (5' AAGGGTTTAAGTGGGAGCGTGTGA) and mCherry_qrtpcr_rev (5' AAGGC ACCATCTTCAGGGTACATTCG), respectively; for *ERG9* quantification are erg9_ rtpcr_fwd (5' AACTGTTGAACTTGACCTCCAGATCGTTTG) and erg9_rtpcr_rev (5' GGCTCTGTCCTTCACATCGGGGGCATTTCC), respectively; for *ACT1* quantification are ACT1_rtpcr_fwd (5' GGCATCATACCTTCTACAACGAAT) and ACT1_rtpcr_rev (5' GGAATCCAAAACAATACCAGTAGTTCTA), respectively. Relative transcript levels and their error were quantified in triplicate from three identical reactions from the

cDNA samples by using an appropriate primer set and iQ SYBR Green Supermix (Bio-Rad, Hercules, CA) on an iCycler iQ qRT-PCR machine (Bio-Rad) according to the manufacturer's instructions. For each run, a standard curve was generated for either *yEGFP3*, *ymCherry*, or *ERG9* and a house-keeping gene, *ACT1*, using a dilution series for a control representing no insertion of an Rnt1p substrate. Relative transcript levels were first individually determined for each sample and then the values for *yEGFP3*, *ymCherry*, and *ERG9* were normalized by their corresponding *ACT1* values.

2.4.6. Cell growth rate determination

At multiple time points during a course of 7 hr, 200 μl were taken from a yeast culture and the OD_{600} measured on a SAFIRE plate reader. The growth rate, k, and its standard error were analyzed using Prism 5 by fitting the data to an exponential growth curve.

2.4.7. Cellular ergosterol quantification

The method for quantification of cellular ergosterol levels was adapted from previously developed protocols[58-59]. Briefly, yeast cells were harvested after 8 hr with the OD_{600} recorded and collected by centrifugation at 3,000 rpm for 5 min. The cells were washed with water and centrifuged again. 10 ml of 25% alcoholic KOH [25% KOH, 60% (v/v) ethanol] was added to the cell pellet and vortexed. The suspension was transferred to a 50 ml Falcon tube and saponified by incubating at 90°C for 3 hr. After cooling to room temperature, the nonsaponified sterols were extracted by adding 5 ml of heptane, vortexing, and collecting the heptane layer once it had clarified. The heptane layer was

directly applied to a 96-well UV plate (Greiner Bio-One) and its absorbance was read in the UV spectrum on a SAFIRE plate reader. A relative ergosterol value (REV) was calculated from the following equation:

$$REV = \frac{OD_{281.5}}{290} - \frac{OD_{230}}{518}$$

The reported value and error were determined from the mean and standard deviation, respectively, from three individual heptanes aliquots.

2.4.8. In vitro *transcription of Rnt1p substrates*

All Rnt1p substrates were PCR-amplified to include an upstream T7 promoter site and A-rich sequences flanking the hairpin using forward and reverse primers Rnt1p-T7-PCR_fwd_prmr (5' TTCTAATACGACTCACTATAGGGACCTAGGAAACAAACAA AGTTGGGC) and Rnt1p-T7-PCR_rev_prmr (5' CTCGAGTTTTTATTTTTCTTTTTGC CGGGCG), respectively. 1–2 μg of PCR product was transcribed with T7 Polymerase (New England Biolabs) in the presence and absence of α-P^{32}-GTP. The 25-μl reaction consisted of the following components: 1x RNA Pol Reaction Buffer (New England Biolabs), 3 mM rATP, 3 mM rCTP, 3 mM rUTP, 0.3 mM rGFP, 1 μl RNaseOUT (Invitrogen), 10 mM $MgCl_2$, 2 mM DTT, 1 μl T7 Polymerase, and 0.5 μCi α-P^{32}-GTP. Unincorporated nucleotides were removed from the reactions by running the samples through NucAway Spin Columns (Ambion, Austin, TX) according to the manufacturer's instructions.

2.4.9. Rnt1p expression and purification

The pRNT1 plasmid was transformed into *E. coli* strain BL21 using the Z-competent *E. coli* Transformation Kit and Buffer Set (Zymo Research, Orange, CA) according to manufacturer's instructions. Rnt1p was collected as a protein extract as previously described[60]. Briefly, an overnight culture of BL21 cells harboring pRNT1 was back-diluted to an OD_{600} of 0.5. Once the culture reached an OD_{600} of 1.1–1.4, it was induced with 1 mM IPTG and grown for an additional 3 hr. The cells were centrifuged at 2,500g for 12 min at 4°C and the resulting cell pellet was frozen in a -80°C freezer. After weighing the frozen cell pellet, the cells were resuspended in 4 ml Ni_{2+} buffer [25% (v/v) glycerol, 1 M NaCl, 30 mM Tris pH 8.0] per gram of harvested cells. The resuspension was sonicated (Heat Systems-Ultrasonics, Inc.) twice with the following settings: 2 x 30 sec, output control 5, and 50% duty cycle. Cellular debris was removed by centrifugation at 20,000g for 30 min at 4°C and the supernatant was filtered through a 0.2-μm pore size Acrodisc 25 mm syringe filter (Pall Life Sciences, Ann Arbor, MI).

Rnt1p was purified from the resulting supernatant with one 1-ml HisTrap HP column (GE Healthcare) on an AKTA FPLC machine (GE Healthcare). Elution of the protein was performed with an imidazole concentration of 150 mM in Ni_{2+} buffer and the protein was collected in 6 1-ml fractions. Protein purification was confirmed by analyzing an aliquot of each fraction on a SDS-PAGE gel (NuPAGE 4–12% Bis-Tris Gel, Invitrogen) and protein function was confirmed by incubating an aliquot of each fraction with a control Rnt1p substrate and analyzing the resulting cleavage products on an 8% denaturing polyacrylamide gel. Positive fractions were pooled and concentrated to less than a 3-ml volume using a Centricon Centrifugal Filter Device (10,000 MWCO;

Millipore) according to the manufacturer's instructions. The concentrated protein was then injected into a Slide-A-Lyzer Dialysis Cassette (10,000 MWCO; Pierce Biotechnology) and buffer-exchanged twice with Rnt1p Storage Buffer [50% (v/v) glycerol, 0.5 M KCl, 30 mM Tris pH 8.0, 0.1 M DTT, 0.1 M EDTA] at 4ºC. The first buffer exchange took place for 4 hr and the second buffer exchange occurred overnight. The purified Rnt1p was stored in aliquots at –20°C.

2.4.10. In vitro *Rnt1p substrate cleavage assay*

Cleavage assays were performed on Rnt1p substrates as previously described[39, 60]. Briefly, a 10-µl mixture of RNA and Rnt1p was incubated at 30ºC for 15 min in Rnt1p reaction buffer [30 mM Tris (pH 7.5), 150 mM KCl, 5 mM spermidine, 20 mM $MgCl_2$, 0.1 mM DTT, and 0.1 mM EDTA (pH 7.5)]. RNA concentrations were varied from 0.1 to 1.0 µM and the Rnt1p concentration was 2.3 µM. The cleavage reaction products were separated on an 8% denaturing polyacrylamide gel run at 35 W for 30 min. Gels were transferred to filter paper and analyzed for relative substrate and product levels through phosphorimaging analysis on a FX Molecular Imager (Bio-Rad). The levels of cleaved RNA product were determined and fit to a Michaelis-Menten model using Prism 5 (GraphPad), where a relative V_{max} was calculated and reported with the standard error determined by the fit of the model.

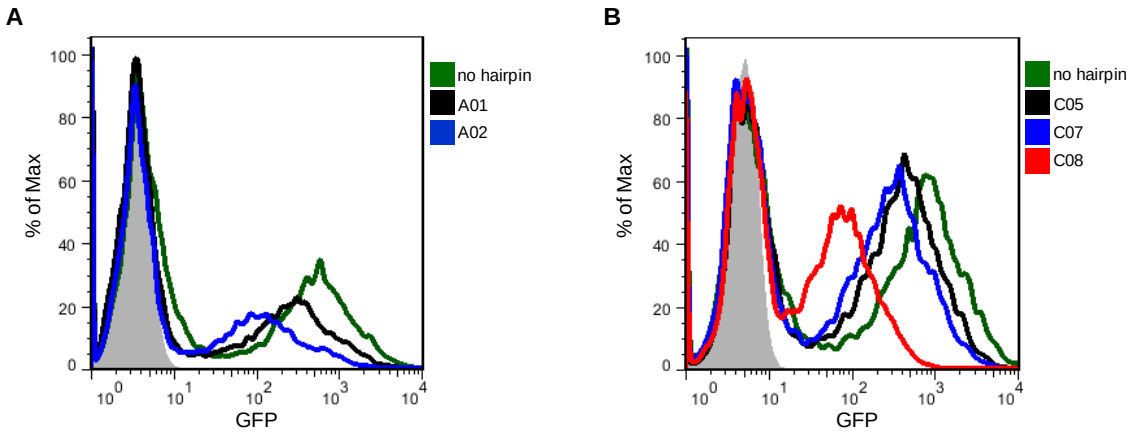
2.4.11. In vitro *Rnt1p substrate mobility shift assay*

Mobility shift assays were performed as previously described[39, 60]. Briefly, a 10-µl mixture of RNA and Rnt1p were incubated on ice for 10 min in Rnt1p binding buffer

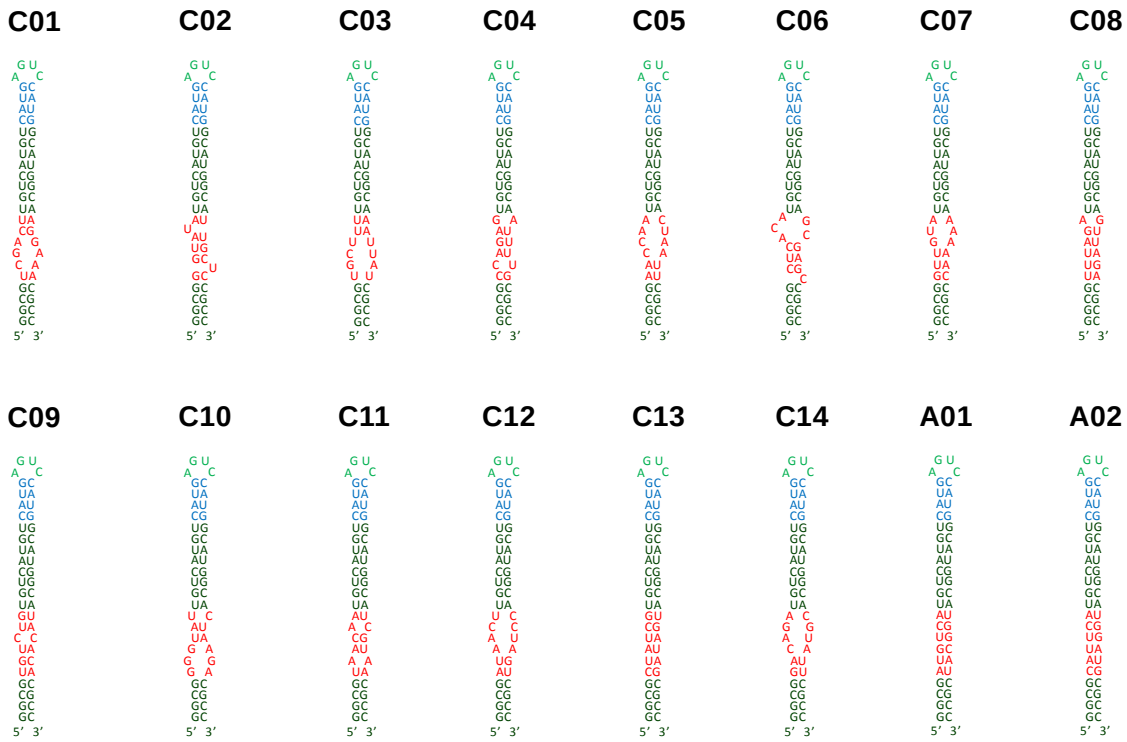
[20% (v/v) glycerol, 30 mM Tris (pH 7.5), 150 mM KCl, 5 mM spermidine, 0.1 mM DTT, and 0.1 mM EDTA (pH 7.5)]. The RNA concentration in all samples was 200 nM and the Rnt1p concentration ranged from 0 to 1.7 μM. The binding reaction products were separated on a 6% native polyacrylamide gel run at 350 V until the samples entered the gel and then at 150 V for 2 hr. Gels were transferred to filter paper and analyzed for free RNA and RNA-Rnt1p complex levels through phosphorimaging analysis on a FX Molecular Imager. The fraction of unbound RNA to total RNA was determined and fit to a modified Scatchard model using Prism 5, where a K_D value was calculated and reported with the standard error determined by the fit of the model.

2.5. Supplementary Information

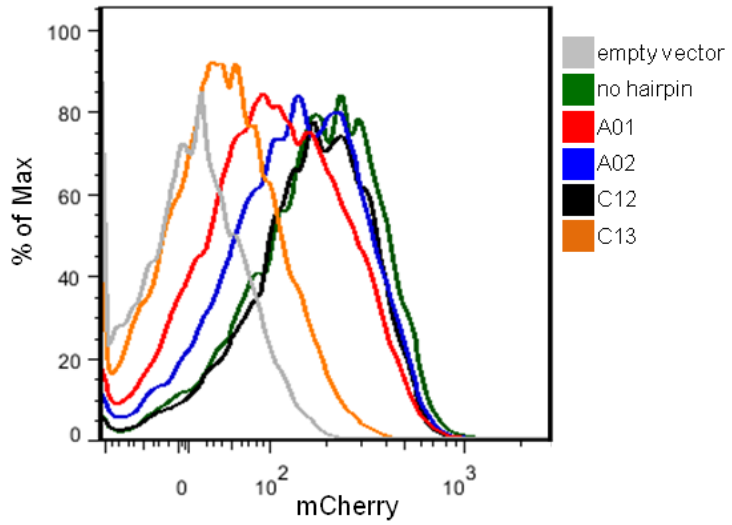
Supplementary Figures and Tables



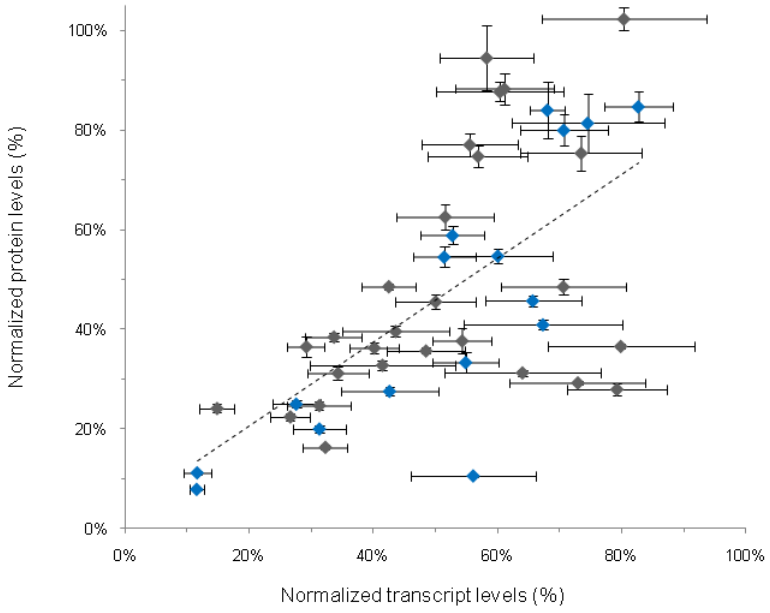
Supplementary Figure 2.1. Flow cytometry histograms of pCS321-based constructs bearing control and library Rnt1p hairpins. The shaded population (grey) indicates the noninduced cell population generated by analyzing cells harboring the 'no insert' control in the absence of galactose. Histograms are representative of three independent experiments. (A) Histograms of constructs bearing the A01 and A02 control hairpins. (B) Histograms of constructs bearing the C05, C07, and C08 library hairpins.



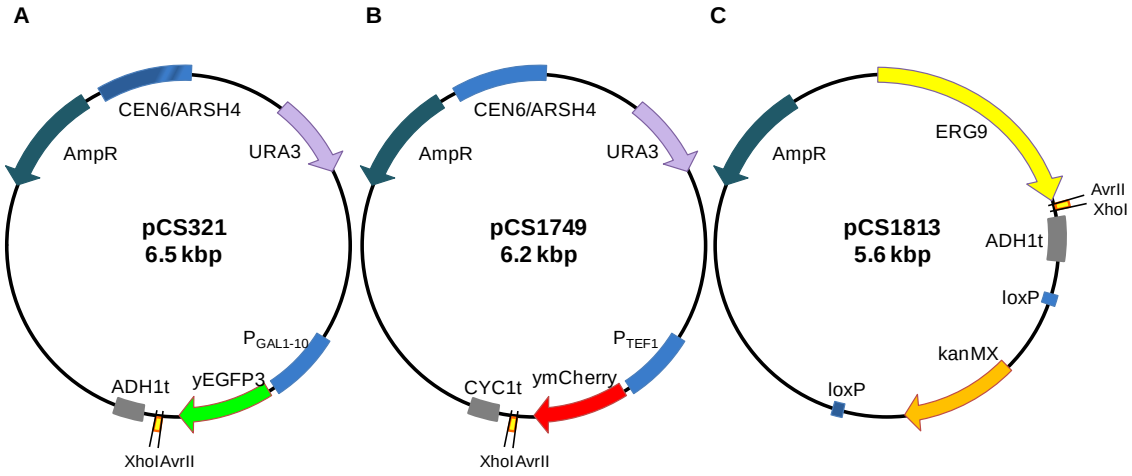
Supplementary Figure 2.2. Sequences and structures of the final Rnt1p cleavage library and the positive control Rnt1p hairpins.



Supplementary Figure 2.3. Flow cytometry histograms of pCS1749-based constructs bearing control and library Rnt1p hairpins. The empty vector histogram (grey) lacks a fluorescent gene and is indicative of the cell and media autofluorescence. Histograms are representative of three independent experiments.



Supplementary Figure 2.4. Correlation analysis of protein and transcript levels for all hairpins identified from the fluorescence-based *in vivo* screening assay. Color scheme is as follows: final library hairpins selected for structural stability to act as modular control elements, blue; library hairpins not selected for inclusion in the final Rnt1p hairpin library, black. The regression line is determined from the entire data set and indicates an r-value of 0.629. The r-value for an identical analysis performed on the data set in blue indicates a stronger correlation (r value = 0.817, Figure 2.3C). All normalized protein and transcript levels are determined as described in Figure 2.1D.



Supplementary Figure 2.5. Plasmid maps for key constructs used in this work. (A) Plasmid map of pCS321, the *yGFP3* characterization plasmid. (B) Plasmid map of pCS1749, the *ymCherry* characterization plasmid. (C) Plasmid map of pCS1813, the Rnt1p hairpin integration plasmid.

Supplementary Table 2.1. Sequence and i*n vivo* characterization data for all screened Rnt1p hairpins. The nucleotides of the cleavage efficiency box are indicated in red. Constructs labeled with 'C' or 'A' represent the final selected cleavage library and positive controls. Constructs labeled with 'c' represent screened hairpins not selected for inclusion in the final Rnt1p hairpin library. All normalized protein and transcript levels are determined as described in Figure 2.1D.

Substrate	Sequence	Normalized protein levels (%)	Normalized transcript levels (%)
C01	GGCGUCGACUUGUCAU GUCAUGAGUCCAUGGC AUGGCAAGGAAACGCC	84% ± 6%	68% ± 3%
C02	GGCGGGUAUAUGUCAU GUCAUGAGUCCAUGGC AUGGCAUUGCUCCGCC	80% ± 3%	71% ± 7%
C03	GGCGUGCUUUUGUCAU GUCAUGAGUCCAUGGC AUGGCAAAUUAUCGCC	55% ± 1%	60% ± 9%
C04	GGCGCCAGAGUGUCAU GUCAUGAGUCCAUGGC AUGGCAAUUUUGCGCC	20% ± 1%	31% ± 4%
C05	GGCGAACCAAUGUCAU GUCAUGAGUCCAUGGC AUGGCACUAAUUCGCC	55% ± 2%	51% ± 5%
C06	GGCGCUCACAUGUCAU GUCAUGAGUCCAUGGC AUGGCAGCGAGCCGCC	33% ± 2%	55% ± 5%
C07	GGCGGUUGUAUGUCAU GUCAUGAGUCCAUGGC AUGGCAAAAAACCGCC	41% ± 1%	67% ± 13%
C08	GGCGUUUAGAUGUCAU GUCAUGAGUCCAUGGC AUGGCAGUUAGACGCC	11% ± 0%	12% ± 2%
C09	GGCGUGUCUGUGUCAU GUCAUGAGUCCAUGGC AUGGCAUACACACGCC	25% ± 1%	28% ± 4%
C10	GGCGGGGUAUUGUCAU GUCAUGAGUCCAUGGC AUGGCACUAAGACGCC	46% ± 1%	66% ± 8%
C11	GGCGUAACAAUGUCAU GUCAUGAGUCCAUGGC AUGGCAUCGUAACGCC	11% ± 0%	56% ± 10%
C12	GGCGAUAACUUGUCAU GUCAUGAGUCCAUGGC AUGGCACCUAGUCGCC	81% ± 6%	75% ± 12%
C13	GGCGCUAUCGUGUCAU GUCAUGAGUCCAUGGC AUGGCAUGAUAGCGCC	8% ± 0%	12% ± 1%
C14	GGCGGACAGAUGUCAU GUCAUGAGUCCAUGGC AUGGCACGUAUUCGCC	85% ± 3%	83% ± 6%
A01	GGCGAUGUCAUGUCAU GUCAUGAGUCCAUGGC AUGGCAUGGCAUCGCC	59% ± 2%	53% ± 5%
A02	GGCGCAUUCAUGUCAU GUCAUGAGUCCAUGGC AUGGCAUGGAUGCGCC	28% ± 1%	43% ± 8%
c01	GGCGACUUACUGUCAU GUCAUGAGUCCAUGGC AUGGCAUAUGCCCGCC	102% ± 2%	80% ± 13%

c02	GGCGAUUCGCUGUCAU GUCAUGAGUCCAUGGC AUGGCAAAACGCCGCC	95% ± 6%	58% ± 8%
c03	GGCGAACUUAUGUCAU GUCAUGAGUCCAUGGC AUGGCAUAAUAACGCC	31% ± 1%	64% ± 13%
c04	GGCGGGACAGUGUCAU GUCAUGAGUCCAUGGC AUGGCAUAGUUGCGCC	22% ± 1%	27% ± 3%
c05	GGCGUUAUGAUGUCAU GUCAUGAGUCCAUGGC AUGGCAAUGUUGCGCC	36% ± 1%	40% ± 4%
c06	GGCGGUCGCAUGUCAU GUCAUGAGUCCAUGGC AUGGCAGUCUAACGCC	37% ± 2%	29% ± 3%
c07	GGCGUUUGGCUGUCAU GUCAUGAGUCCAUGGC AUGGCAAUCAUGCGCC	16% ± 0%	32% ± 4%
c08	GGCGAUGAAAUGUCAU GUCAUGAGUCCAUGGC AUGGCAUUUACGCGCC	38% ± 2%	54% ± 5%
c09	GGCGGUAAAGUGUCAU GUCAUGAGUCCAUGGC AUGGCAUUAUAGCGCC	46% ± 1%	50% ± 6%
c10	GGCGUUUUAGUGUCAU GUCAUGAGUCCAUGGC AUGGCAAAAAUGCGCC	75% ± 3%	73% ± 10%
c11	GGCGGGUAGUUGUCAU GUCAUGAGUCCAUGGC AUGGCAAGUGGACGCC	24% ± 1%	15% ± 3%
c12	GGCGAUUCAGUGUCAU GUCAUGAGUCCAUGGC AUGGCAGGGUAUCGCC	25% ± 1%	31% ± 5%
c13	GGCGAAGCCGUGUCAU GUCAUGAGUCCAUGGC AUGGCAUCUACACGCC	88% ± 3%	61% ± 8%
c14	GGCGCCGGAAUGUCAU GUCAUGAGUCCAUGGC AUGGCAGUUGAGCGCC	49% ± 2%	71% ± 10%
c15	GGCGACAUUGUGUCAU GUCAUGAGUCCAUGGC AUGGCAACGUUGCGCC	33% ± 1%	41% ± 12%
c16	GGCGCCUGCAUGUCAU GUCAUGAGUCCAUGGC AUGGCAGGCCAUCGCC	40% ± 1%	44% ± 9%
c17	GGCGGAUCCAUGUCAU GUCAUGAGUCCAUGGC AUGGCAAUGAAGCGCC	36% ± 1%	48% ± 6%
c18	GGCGGUAGGGUGUCAU GUCAUGAGUCCAUGGC AUGGCAACGUAGCGCC	37% ± 1%	80% ± 12%
c19	GGCGAGUAGGUGUCAU GUCAUGAGUCCAUGGC AUGGCACGCACACGCC	28% ± 1%	79% ± 8%
c20	GGCGGUUUGAUGUCAU GUCAUGAGUCCAUGGC AUGGCAGCAGUUCGCC	29% ± 0%	73% ± 11%
c21	GGCGUUUUAAUGUCAU GUCAUGAGUCCAUGGC AUGGCAAUUGUUCGCC	38% ± 1%	34% ± 5%
c22	GGCGAUUAUGUGUCAU GUCAUGAGUCCAUGGC AUGGCAGUUAUGCGCC	49% ± 1%	42% ± 4%

c23	GGCGAUGUGUUGUCAU GUCAUGAGUCCAUGGC AUGGCACAGACACGCC	88% ± 2%	60% ± 10%
c24	GGCGAAUUUUUGUCAU GUCAUGAGUCCAUGGC AUGGCAGUAGGUCGCC	63% ± 2%	52% ± 8%
c25	GGCGCUAUCAUGUCAU GUCAUGAGUCCAUGGC AUGGCACGUAAUCGCC	31% ± 1%	34% ± 5%
c26	GGCGUACACAUGUCAU GUCAUGAGUCCAUGGC AUGGCAUUCUUGCGCC	77% ± 2%	55% ± 8%
c27	GGCGACUUAUUGUCAU GUCAUGAGUCCAUGGC AUGGCAUCAAUACGCC	75% ± 2%	57% ± 8%

Supplementary Table 2.2. *In vivo* characterization data for the Rnt1p cleavage library in the context of *ymCherry* (pCS1749). All normalized protein and transcript levels are determined as described in Figure 2.1D.

Substrate	Normalized protein levels (%)	Normalized transcript levels (%)
C01	100% ± 5%	84% ± 10%
C02	87% ± 4%	90% ± 5%
C03	73% ± 5%	63% ± 5%
C04	2% ± 0%	40% ± 6%
C05	63% ± 5%	69% ± 9%
C06	89% ± 5%	91% ± 6%
C07	53% ± 2%	57% ± 9%
C08	33% ± 2%	57% ± 8%
C09	44% ± 2%	47% ± 5%
C10	72% ± 2%	89% ± 2%
C11	17% ± 1%	23% ± 3%
C12	97% ± 6%	109% ± 5%
C13	10% ± 1%	21% ± 2%
C14	77% ± 4%	101% ± 4%
A01	80% ± 7%	72% ± 5%
A02	56% ± 3%	71% ± 7%

Acknowledgements

We thank K Hoff, S Bastian, and FH Arnold for assistance in the purification of Rnt1p; and J Liang for assistance with *in vitro* assays. This work was supported by the National Science Foundation (CAREER award to CDS; CBET-0917705) and the Alfred P. Sloan Foundation (fellowship to CDS).

References

1. Elowitz, M.B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335-338 (2000).
2. Gardner, T.S., Cantor, C.R. & Collins, J.J. Construction of a genetic toggle switch in Escherichia coli. *Nature* **403**, 339-342 (2000).
3. Basu, S., Mehreja, R., Thiberge, S., Chen, M.T. & Weiss, R. Spatiotemporal control of gene expression with pulse-generating networks. *Proc Natl Acad Sci U S A* **101**, 6355-6360 (2004).
4. Pfleger, B.F., Pitera, D.J., Smolke, C.D. & Keasling, J.D. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat Biotechnol* **24**, 1027-1032 (2006).
5. Anderson, J.C., Voigt, C.A. & Arkin, A.P. Environmental signal integration by a modular AND gate. *Mol Syst Biol* **3**, 133 (2007).
6. Pitera, D.J., Paddon, C.J., Newman, J.D. & Keasling, J.D. Balancing a heterologous mevalonate pathway for improved isoprenoid production in Escherichia coli. *Metab Eng* **9**, 193-207 (2007).
7. Peralta-Yahya, P.P. & Keasling, J.D. Advanced biofuel production in microbes. *Biotechnol J* **5**, 147-162 (2010).
8. Jin, Y.S., Ni, H., Laplaza, J.M. & Jeffries, T.W. Optimal growth and ethanol production from xylose by recombinant Saccharomyces cerevisiae require moderate D-xylulokinase activity. *Appl Environ Microbiol* **69**, 495-503 (2003).

9. Jones, K.L., Kim, S.W. & Keasling, J.D. Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. *Metab Eng* **2**, 328-338 (2000).

10. Zhu, M.M., Skraly, F.A. & Cameron, D.C. Accumulation of methylglyoxal in anaerobically grown Escherichia coli and its detoxification by expression of the Pseudomonas putida glyoxalase I gene. *Metab Eng* **3**, 218-225 (2001).

11. Alper, H., Jin, Y.S., Moxley, J.F. & Stephanopoulos, G. Identifying gene targets for the metabolic engineering of lycopene biosynthesis in Escherichia coli. *Metab Eng* **7**, 155-164 (2005).

12. Paradise, E.M., Kirby, J., Chan, R. & Keasling, J.D. Redirection of flux through the FPP branch-point in Saccharomyces cerevisiae by down-regulating squalene synthase. *Biotechnol Bioeng* **100**, 371-378 (2008).

13. Alper, H., Fischer, C., Nevoigt, E. & Stephanopoulos, G. Tuning genetic control through promoter engineering. *Proc Natl Acad Sci U S A* **102**, 12678-12683 (2005).

14. Carrier, T.A. & Keasling, J.D. Library of synthetic 5' secondary structures to manipulate mRNA stability in Escherichia coli. *Biotechnol Prog* **15**, 58-64 (1999).

15. Jensen, P.R. & Hammer, K. The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl Environ Microbiol* **64**, 82-87 (1998).

16. Smolke, C., Martin, V. & Keasling, J. in Protein expression technologies: current status and future trends. (ed. F. Baneyx) (Horizon Bioscience, Wymondham, Norfolk; 2004).

17. Szczebara, F.M. et al. Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nat Biotechnol* **21**, 143-149 (2003).

18. Hawkins, K.M. & Smolke, C.D. Production of benzylisoquinoline alkaloids in Saccharomyces cerevisiae. *Nat Chem Biol* **4**, 564-573 (2008).

19. Ro, D.K. et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940-943 (2006).

20. Nguyen, H.T. et al. Engineering of Saccharomyces cerevisiae for the production of L-glycerol 3-phosphate. *Metab Eng* **6**, 155-163 (2004).

21. Ostergaard, S., Olsson, L. & Nielsen, J. Metabolic engineering of Saccharomyces cerevisiae. *Microbiol Mol Biol Rev* **64**, 34-50 (2000).

22. Veen, M. & Lang, C. Production of lipid compounds in the yeast Saccharomyces cerevisiae. *Appl Microbiol Biotechnol* **63**, 635-646 (2004).

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

24. Hawkins, K.M. & Smolke, C.D. The regulatory roles of the galactose permease and kinase in the induction response of the GAL network in Saccharomyces cerevisiae. *J Biol Chem* **281**, 13485-13492 (2006).

25. Nevoigt, E. et al. Engineering promoter regulation. *Biotechnol Bioeng* **96**, 550-558 (2007).

26. 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).

27. Caponigro, G. & Parker, R. Mechanisms and control of mRNA turnover in Saccharomyces cerevisiae. *Microbiol Rev* **60**, 233-249 (1996).

28. Filippov, V., Solovyev, V., Filippova, M. & Gill, S.S. A novel type of RNase III family proteins in eukaryotes. *Gene* **245**, 213-221 (2000).

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

30. Wu, H., Henras, A., Chanfreau, G. & Feigon, J. Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. *Proc Natl Acad Sci U S A* **101**, 8307-8312 (2004).

31. 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).

32. 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).

33. 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).

34. 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).

35. Lamontagne, B. & Abou Elela, S. Short RNA guides cleavage by eukaryotic RNase III. *PLoS One* **2**, e472 (2007).

36. Ge, D., Lamontagne, B. & Elela, S.A. RNase III-mediated silencing of a glucose-dependent repressor in yeast. *Curr Biol* **15**, 140-145 (2005).

37. Pelletier, J. & Sonenberg, N. Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency. *Cell* **40**, 515-526 (1985).

38. Mateus, C. & Avery, S.V. Destabilized green fluorescent protein for monitoring dynamic changes in yeast gene expression with flow cytometry. *Yeast* **16**, 1313-1323 (2000).

39. Lamontagne, B. & Elela, S.A. Evaluation of the RNA determinants for bacterial and yeast RNase III binding and cleavage. *J Biol Chem* **279**, 2231-2241 (2004).

40. Jeppsson, M., Johansson, B., Jensen, P.R., Hahn-Hagerdal, B. & Gorwa-Grauslund, M.F. The level of glucose-6-phosphate dehydrogenase activity strongly influences xylose fermentation and inhibitor sensitivity in recombinant Saccharomyces cerevisiae strains. *Yeast* **20**, 1263-1272 (2003).

41. Lamontagne, B., Tremblay, A. & Abou Elela, S. The N-terminal domain that distinguishes yeast from bacterial RNase III contains a dimerization signal required for efficient double-stranded RNA cleavage. *Mol Cell Biol* **20**, 1104-1115 (2000).

42. Khosla, C. & Keasling, J.D. Metabolic engineering for drug discovery and development. *Nat Rev Drug Discov* **2**, 1019-1025 (2003).

43. Bach, T.J. Some New Aspects of Isoprenoid Biosynthesis in Plants - a Review. *Lipids* **30**, 191-202 (1995).

44. Poulter, C.D. & Rilling, H.C. in 413-441 (Wiley, New York; 1981).

45. Guldener, U., Heck, S., Fielder, T., Beinhauer, J. & Hegemann, J.H. A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res* **24**, 2519-2524 (1996).

46. Daum, G., Lees, N.D., Bard, M. & Dickson, R. Biochemistry, cell biology and molecular biology of lipids of Saccharomyces cerevisiae. *Yeast* **14**, 1471-1510 (1998).

47. Kennedy, M.A., Barbuch, R. & Bard, M. Transcriptional regulation of the squalene synthase gene (ERG9) in the yeast Saccharomyces cerevisiae. *Biochim Biophys Acta* **1445**, 110-122 (1999).

48. Lautz, T., Stahl, U. & Lang, C. The human c-fos and TNFalpha AU-rich elements show different effects on mRNA abundance and protein expression depending on the reporter in the yeast Pichia pastoris. *Yeast* **27**, 1-9 (2010).

49. Vasudevan, S. & Peltz, S.W. Regulated ARE-mediated mRNA decay in Saccharomyces cerevisiae. *Mol Cell* **7**, 1191-1200 (2001).

50. Zhou, W., Edelman, G.M. & Mauro, V.P. Transcript leader regions of two Saccharomyces cerevisiae mRNAs contain internal ribosome entry sites that function in living cells. *Proc Natl Acad Sci U S A* **98**, 1531-1536 (2001).

51. Zhou, W., Edelman, G.M. & Mauro, V.P. Isolation and identification of short nucleotide sequences that affect translation initiation in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **100**, 4457-4462 (2003).

52. Salis, H.M., Mirsky, E.A. & Voigt, C.A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* **27**, 946-950 (2009).

53. Sambrook, J. & Russell, D.W. Molecular Cloning: A Laboratory Manual, Edn. 3rd. (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY; 2001).

54. Sikorski, R.S. & Hieter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics* **122**, 19-27 (1989).

55. Gietz, R. & Woods, R. in Guide to Yeast Genetics and Molecular and Cell Biology, Part B, Vol. 350. (eds. C. Guthrie & G. Fink) 87-96 (Academic, San Diego, CA; 2002).

56. Caponigro, G., Muhlrad, D. & Parker, R. A small segment of the MAT alpha 1 transcript promotes mRNA decay in Saccharomyces cerevisiae: a stimulatory role for rare codons. *Mol Cell Biol* **13**, 5141-5148 (1993).

57. Ng, R. & Abelson, J. Isolation and sequence of the gene for actin in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **77**, 3912-3916 (1980).

58. Asadollahi, M.A. et al. Production of plant sesquiterpenes in Saccharomyces cerevisiae: effect of ERG9 repression on sesquiterpene biosynthesis. *Biotechnol Bioeng* **99**, 666-677 (2008).

59. Arthington-Skaggs, B.A., Jradi, H., Desai, T. & Morrison, C.J. Quantitation of ergosterol content: novel method for determination of fluconazole susceptibility of Candida albicans. *J Clin Microbiol* **37**, 3332-3337 (1999).

60. Lamontagne, B. & Elela, S.A. Purification and characterization of Saccharomyces cerevisiae Rnt1p nuclease. *Methods Enzymol* **342**, 159-167 (2001).