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

Abstract

The design of synthetic gene networks requires an extensive genetic toolbox to precisely control the activities and levels of protein components to achieve desired cellular functions. Recently, a novel class of RNA-based control modules, which act through posttranscriptional processing of transcripts by directed RNase III (Rnt1p) cleavage, were shown to provide predictable control over gene expression and unique properties for manipulating biological networks. Here, we increase the regulatory range of the Rnt1p control elements, by modifying a critical region for enzyme binding to its hairpin substrates, the binding stability box (BSB). We used a high-throughput, cell-based selection strategy to screen a BSB library for sequences that exhibit low fluorescence and thus high Rnt1p processing efficiencies. Sixteen unique BSBs were identified that cover an intermediate range of protein expression levels (25%-75%), due to the ability of the sequences to affect the hairpin cleavage rate and to form active cleavable complexes with Rnt1p. We further demonstrated that the activity of synthetic Rnt1p hairpins can be rationally programmed by combining the synthetic BSBs with a set of sequences located within a different region of the hairpin that directly modulate cleavage rates, providing a modular assembly strategy for this class of RNA-based control elements.

3.1. Introduction

The field of synthetic biology encompasses the engineering of new cellular functions through the design of synthetic gene networks. The precise tuning of protein levels is critical for proper functioning of integrated genetic networks. For example, the optimization of metabolic networks often requires the precise tuning and regulation of enzyme levels and activities to avoid undesired consequences associated with metabolic burden due to gene overexpression[1-2], the accumulation of toxic intermediates[3-5], and the redirection of metabolic flux from pathways critical to cell growth and viability[6-8]. Altered levels of protein components can be achieved by controlling transcription[9-13], posttranscriptional stability and translation[14-17], and protein stability[18-19]. In addition, libraries of genetic control elements have been generated to increase the precision with which protein levels can be modulated[3, 10-11, 16-17]. However, the majority of gene regulatory tools developed to date function in bacterial hosts, such as *Escherichia coli*. Therefore, extending toolsets of genetic control elements to other cellular chassis is essential to supporting the design of more complex, integrated genetic networks in those organisms.

The budding yeast, *Saccharomyces cerevisiae*, is a relevant cellular chassis in industrial bioprocessing[20-25]. The current genetic toolbox for *S. cerevisiae* gene regulation relies primarily on transcriptional control mechanisms such as inducible and constitutive promoter systems. Many inducible promoters depend on accurately controlling the level of the exogenously-applied inducer molecule, where intermediate expression levels are determined through the partitioning of cells in the population between either being fully repressed or expressing the desired protein[26]. While engineered variants have been

constructed that offer more tunable responses to varying inducer concentrations[11-12], these systems can exhibit other undesirable properties, such as pleiotropic effects of the inducer molecules, undesired effects of altering the natural regulatory networks associated with the native promoter system, and the cost associated with the inducing molecule in scale-up processes. RNA-based control modules based on posttranscriptional mechanisms may offer an advantage since their activities are independent of the choice of promoter. Moreover, RNA-based controllers can be combined with transcriptional controllers to expand the design of integrated regulatory networks and thus provide more sophisticated control strategies.

We previously developed a novel class of RNA control modules that act through posttranscriptional cleavage by the *S. cerevisiae* Rnt1p enzyme (Chapter II). Rnt1p recognizes RNA hairpins that contain a consensus AGNN tetraloop, which forms a predetermined fold that is recognized by the dsRNA-binding domain (dsRBD) of Rnt1p[27-29]. RNA hairpins cleaved by Rnt1p have 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]. Rnt1p hairpins were inserted as genetic control elements within the 3' untranslated region (UTR) of a transcript in order to direct cleavage to that region, thereby inactivating the transcript and lowering target protein levels. We designed an initial library based on randomization of the Rnt1p substrate CEB and screened this library to identify a set of Rnt1p control modules that tune expression levels through differential Rnt1p processing rates (Chapter II). The utility of the Rnt1p control modules was

demonstrated for achieving predictable control over protein levels and manipulating biological networks.

Here, we examined the role of a different critical region of the Rnt1p substrate, the BSB, on Rnt1p processing efficiencies and thus gene regulatory activities. We generated a library of hairpins based on randomization of the BSB to identify sequences that modulated Rnt1p binding affinity. Rigid structural constraints imposed by the BSB resulted in a low percentage of sequence variants in the library that exhibit Rnt1p binding activity, and thus required the development of a selection strategy based upon fluorescence-activated cell sorting (FACS) by enriching for cells exhibiting low fluorescence. In total, 16 unique BSBs were identified that span an intermediate range of protein expression levels. *In vitro* characterization assays indicated that altered expression levels are due to the ability of BSBs to determine the hairpin cleavage rate and to form active cleavable complexes with Rnt1p. The integration of the synthetic BSB sequences with different synthetic CEB modules demonstrated that the BSB sequences function as modules that retain their relative activities under the context of different CEBs. Further characterization indicated that proportional deviation from the 'parent' BSB was inversely related to the strength of the coupled CEB. Our work establishes a set of BSB sequences and a previously developed set of CEB sequences as modular units that can be implemented combinatorially to build synthetic Rnt1p hairpins exhibiting precisely tuned processing properties and an extended range of gene regulatory activities.

3.2. Results

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

Rnt1p is an RNase III enzyme that cleaves hairpin structures in *S. cerevisiae*. An Rnt1p substrate can be divided into three critical regions: the IBPB, the BSB, and the CEB [29] (Figure 3.1A). The BSB has a reported structural requirement in which the three nucleotides immediately below the tetraloop must form Watson-Crick base-pairs and the nucleotides in the fourth position must also base-pair, in either a Watson-Crick or wobble conformation, for optimal activity[30]. Rnt1p initially binds to the tetraloop and then cleaves 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. Naturally-occurring Rnt1p hairpins have been identified in numerous noncoding RNAs, where Rnt1p plays a critical role in noncoding RNA processing and editing[31-33], and in transcripts, where Rnt1p was shown to play a role in controlling gene expression[34-36].

We previously developed a genetic system in which Rnt1p-mediated cleavage was used to regulate gene expression in yeast through the placement of Rnt1p hairpins in the 3' UTR of a target transcript (Figure 3.1B) (Chapter II). We developed a set of synthetic Rnt1p hairpins based on sequence modification within the CEB that exhibit a broad range of cleavage rates and thus gene regulatory activities. Since the BSB and CEB are required elements for Rnt1p binding and cleavage, respectively, we hypothesized that a similar library screening approach could be applied to generate synthetic BSBs exhibiting different Rnt1p binding affinities. The synthetic CEB and BSB elements are anticipated to act as modular units such that they could be implemented combinatorially

to build synthetic Rnt1p hairpins exhibiting more precise tuning and an extended range of gene regulatory activities.

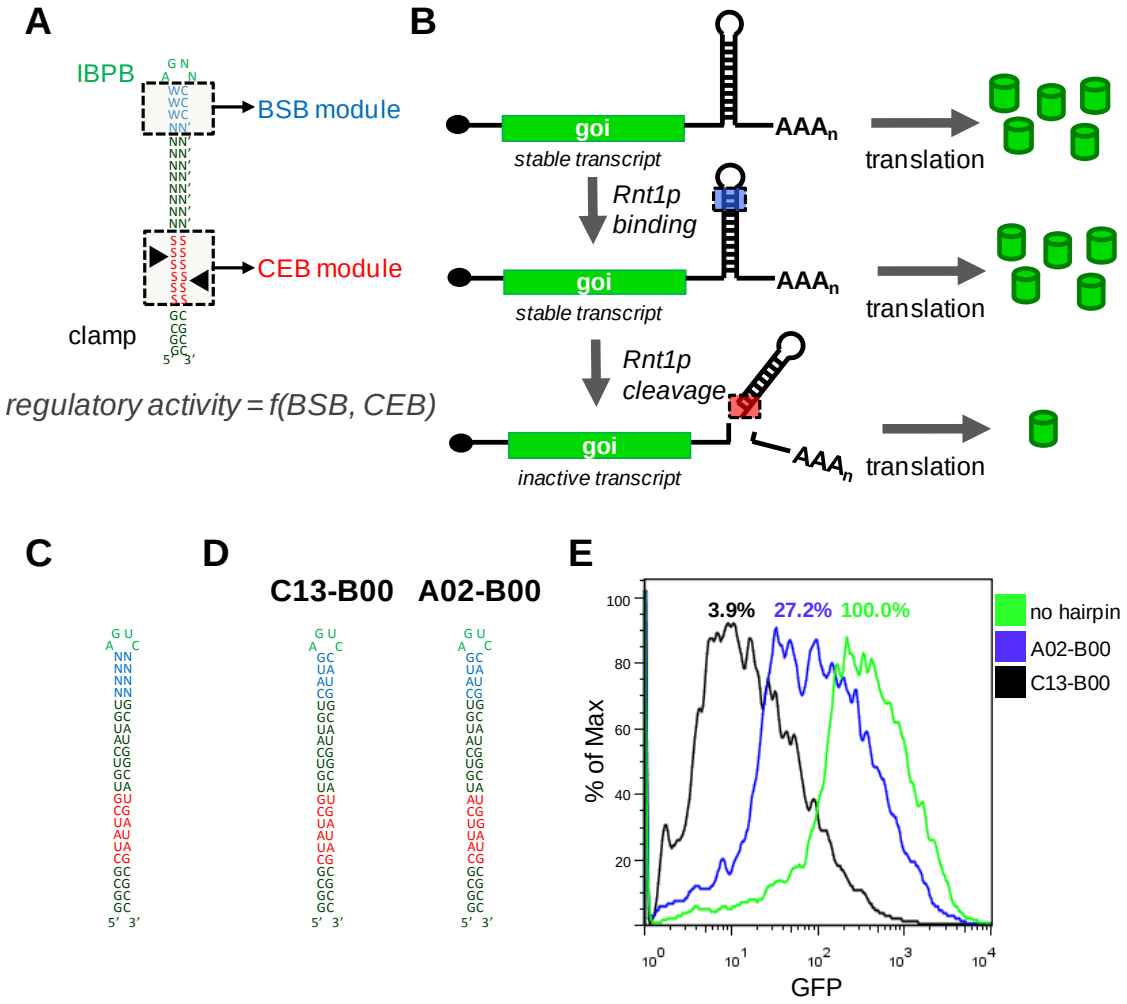


Figure 3.1. Implementation of Rnt1p hairpins as posttranscriptional genetic control elements and binding library design. (A) Consensus regions of an Rnt1p hairpin. Color scheme is as follows: cleavage efficiency box (CEB), red; binding stability box (BSB), blue; initial binding and positioning box (IBPB), green. Black triangles indicate locations of cleavage sites. The clamp region is a synthetic sequence that acts to insulate and maintain the structure of the control element. The regulatory activity of our synthetic hairpins is a function of the modular elements in the CEB and the BSB. (B) Rnt1p hairpins in the 3' UTR of a gene of interest (goi) reduce protein levels through transcript destabilization by endonucleolytic cleavage. Barrels represent protein molecules. Stable transcripts are exported out of the nucleus to the cytoplasm where translational processes occur. Unstable transcripts caused by Rnt1p cleavage will have reduced protein expression levels. (C) Sequence and structure of an Rnt1p binding library based on the randomization of 8 nts in the BSB and containing the C13 CEB. (D) Sequences and

structures of synthetic CEB sequences (Chapter II) used in the screening and characterization of the binding library. ‘B00’ refers to the ‘parent’ BSB used in the cleavage library. (E) Histograms of GFP fluorescence for cell populations containing the Rnt1p constructs based on the ‘parent’ BSB. Numbers above each peak represent the median GFP fluorescence for that sample normalized to a control sample with a construct containing no Rnt1p hairpin.

We designed an Rnt1p binding library based on randomizing the BSB (8 nt) to generate Rnt1p hairpins that exhibit different gene regulatory activities due to altered binding affinity between the hairpin and Rnt1p (Figure 3.1C). One of the variables in the design of the binding library was the CEB to place within the stem, as we had previously described a set of synthetic CEBs with modified gene regulatory activity (Chapter II). As it was unknown how the Rnt1p hairpins would respond to changes in the BSB, we selected a synthetic CEB contained in the hairpin that demonstrated the lowest level of gene expression (C13). All hairpins with active BSBs containing the CEB of C13 were expected to have comparable gene expression levels, improving their probability of being identified in the screen. C13 is also fully base-paired, such that its integration into an Rnt1p substrate stem results in a stable structure that is less susceptible to changes in flanking sequences and has a greater probability of maintaining the desired hairpin structure (Figure 3.1D). Flow cytometry analysis also indicates that C13 achieves the greatest population separation from a no hairpin control of all synthetic CEBs (Figure 3.1E). This separation is representative of that expected for the active and inactive binding populations in the binding library, thus, increasing the enrichment of the cell-based sort. We refer to the ‘parent’ BSB that was used in the cleavage library as B00, where synthetic Rnt1p hairpins are identified by their CEB and BSB as Cxx–Bxx (i.e., C13–B00) or Axx–Bxx (i.e., A02–B00).

Synthetic Rnt1p substrates with altered binding affinities were identified through a cell-based fluorescence screen. The designed library (N8) has a diversity of 65,536 different sequences. Due to the rigid structural requirements of the BSB, we predicted that at most 1,024 sequences (1.56%) would be actively bound and subsequently cleaved by Rnt1p. A manual plate-based screening strategy was previously used to efficiently identify active sequences from the cleavage library. However, due to the loose structural requirements for the CEB, the cleavage library had a substantially higher positive rate than that anticipated for the binding library, such that the plate-based screen was not feasible for the number of colonies that would have to be screened to identify a reasonable diversity of active BSBs. As a higher throughput method we employed FACS to efficiently identify cells with diminished fluorescence.

The binding library was transformed into yeast through a library-scale gap-repair strategy and clones exhibiting strong gene regulatory activity were selected through FACS by gating for cells with low fluorescence levels (Figure 3.2A). We performed two different single FACS screens based on a single-color (pCS1585; *yEGFP3*) system and a two-color (pCS1748; *ymCherry* and *yEGFP3*) system (Liang, J.C. et al., in preparation). Three fractions (A, B, and C) were collected for the pCS1585 system around the expression level of C13-B00. Following the initial sort, each fraction was regrown and subsequently characterized by flow cytometry (Figure 3.2B, Supplementary Figure 3.1). Only fraction A retained a low level of fluorescence, whereas the regrown populations in fractions B and C shifted substantially towards inactive (high fluorescence) levels. The results suggest that sorted fractions B and C contain a large percentage of false positives, or clones harboring plasmids containing inactive Rnt1p hairpins that exhibited low

fluorescence levels due to noise in gene expression profiles. Based on the fraction profiles, fraction A was selected for further testing. Most of the false positives due to genetic noise were removed with the two-color pCS1748 system. A gate representing diminished GFP fluorescence was determined based on a GFP positive control lacking an Rnt1p hairpin. All library clones with decreased GFP levels were collected into a single fraction (D). Fraction D was regrown, characterized by flow cytometry, and demonstrated to retain low fluorescence levels (Figure 3.2C, Supplementary Figure 3.2). Therefore, fraction D was also selected for further characterization.

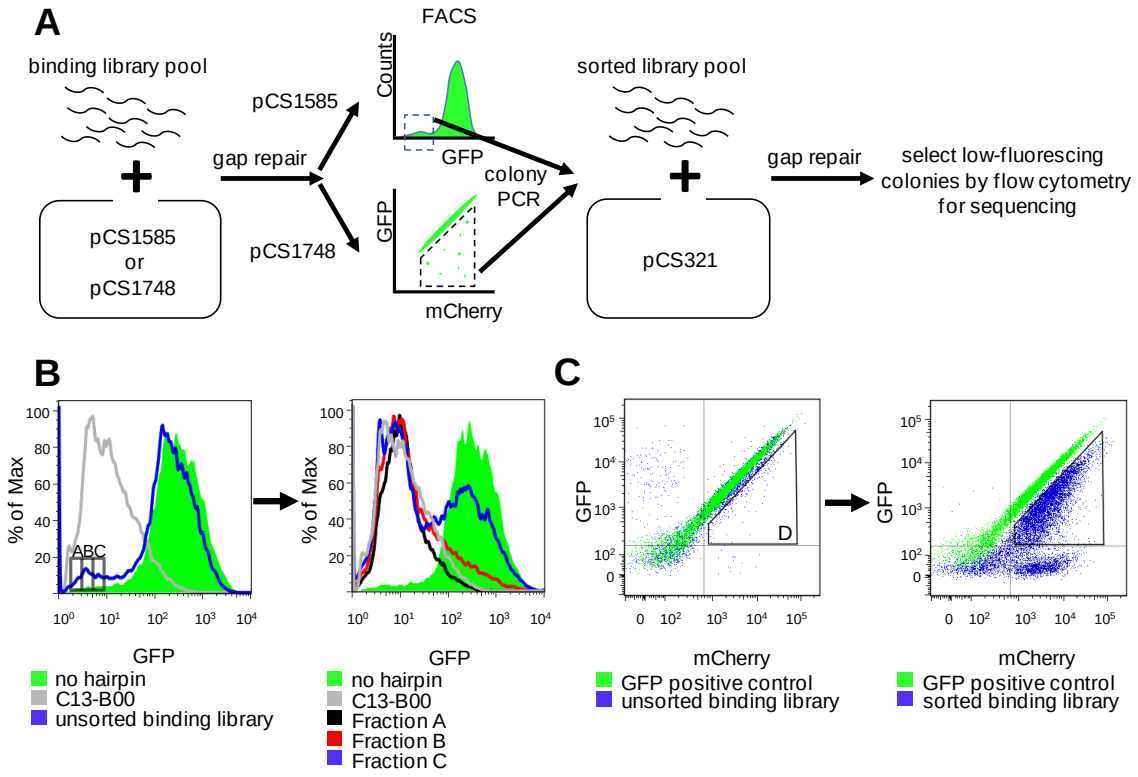


Figure 3.2. *In vivo* screening of an Rnt1p binding library. (A) A high-throughput, *in vivo*, fluorescence-based screen for Rnt1p hairpin activity. The library was cloned through gap-repair into yeast in two different plasmid systems. Clones exhibiting low GFP fluorescence were sorted from the population through FACS. A sorted library pool was generated through colony PCR from collected cellular fractions and gap-repaired into the characterization plasmid. Clones that maintained low GFP fluorescence levels were selected for sequencing and further characterization. (B) FACS procedure for the single-

color (pCS1585-based) system. Fractions A, B, and C were collected based on exhibiting GFP levels similar to the median fluorescence of C13-B00 (left panel). Only fraction A maintained a low level of GFP expression after the fractions were regrown (right panel). (C) FACS procedure for the two-color (pCS1748-based) system. Fraction D was collected based on a gate set to collect all cells exhibiting GFP fluorescence levels below a cells containing a positive GFP control construct lacking an Rnt1p hairpin module (left panel). Fraction D maintained a low level of GFP expression when regrown (right panel).

After the completion of the FACS screens, the sorted constructs were recloned to remove false positives due to mutations in the plasmid or the yeast background that would cause reduced GFP levels independent of Rnt1p activity. We retrieved the selected Rnt1p hairpin sequences from fractions A and D by colony PCR and gap-repaired the recovered hairpin constructs into pCS321. Individual clones were initially characterized for gene regulatory activity by measuring cellular fluorescence through a plate reader assay. Colonies positive for GFP knockdown were sequenced to determine the BSB sequence. In total 16 unique BSB sequences were identified including the 'parent' BSB (Supplementary Figure 3.3; Supplementary Table 3.1). The predicted secondary structure of the hairpins was determined by RNAstructure (http://rna.chem.rochester.edu/RNAstructure.html). The binding library structures deviate from structural requirements that were previously established through *in vitro* studies[30]. All of the BSB structures contain Watson-Crick base-pairing in the first three nucleotides below the tetraloop; however, in the fourth position certain sequences exhibit mismatching. The results from our cell-based BSB library screen suggest that the *in vivo* structural requirements for the BSB are not as stringent as previously described.

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

The initial screen for active BSB sequences was performed in the context of a CEB exhibiting high cleavage activity (C13). As a result, most of the recovered binding library hairpins exhibited low gene expression levels, such that differences in activity between the synthetic BSBs were difficult to resolve with the flow cytometry assay at low fluorescence levels (Supplementary Table 3.1). To gain better resolution on the differences in BSB activity and examine the activity of the synthetic BSBs in the context of a different CEB, we integrated the selected BSB sequences within the context of an Rnt1p hairpin containing a different synthetic CEB (A02) (Figure 3.1D). The range of regulatory activities spanned by the binding library in the context of the A02 CEB was measured at the protein and transcript levels. Flow cytometry analysis of the synthetic Rnt1p hairpins indicated that the selected set of hairpins spanned an intermediate gene regulatory range – from 25% (A02–B05) to 75% (A02–B01) (Table 3.1, Figure 3.3A). The regulatory activities of the selected hairpins are not evenly distributed across this range, with the majority exhibiting activities in the range of 25–45% relative protein levels. The results suggest that the binding library achieves a smaller regulatory range than that observed with the cleavage library and may be more appropriate for fine-tuning (Chapter II). We built negative controls for several binding library hairpins and for the ‘parent’ hairpins (A02–B00 and C13–B00) by mutating the tetraloop sequence (CAUC or GAAA) to impede Rnt1p activity while maintaining the secondary structure of the hairpins. The negative controls demonstrated that the majority of knockdown observed from each hairpin is due to Rnt1p processing (Figure 3.3B).

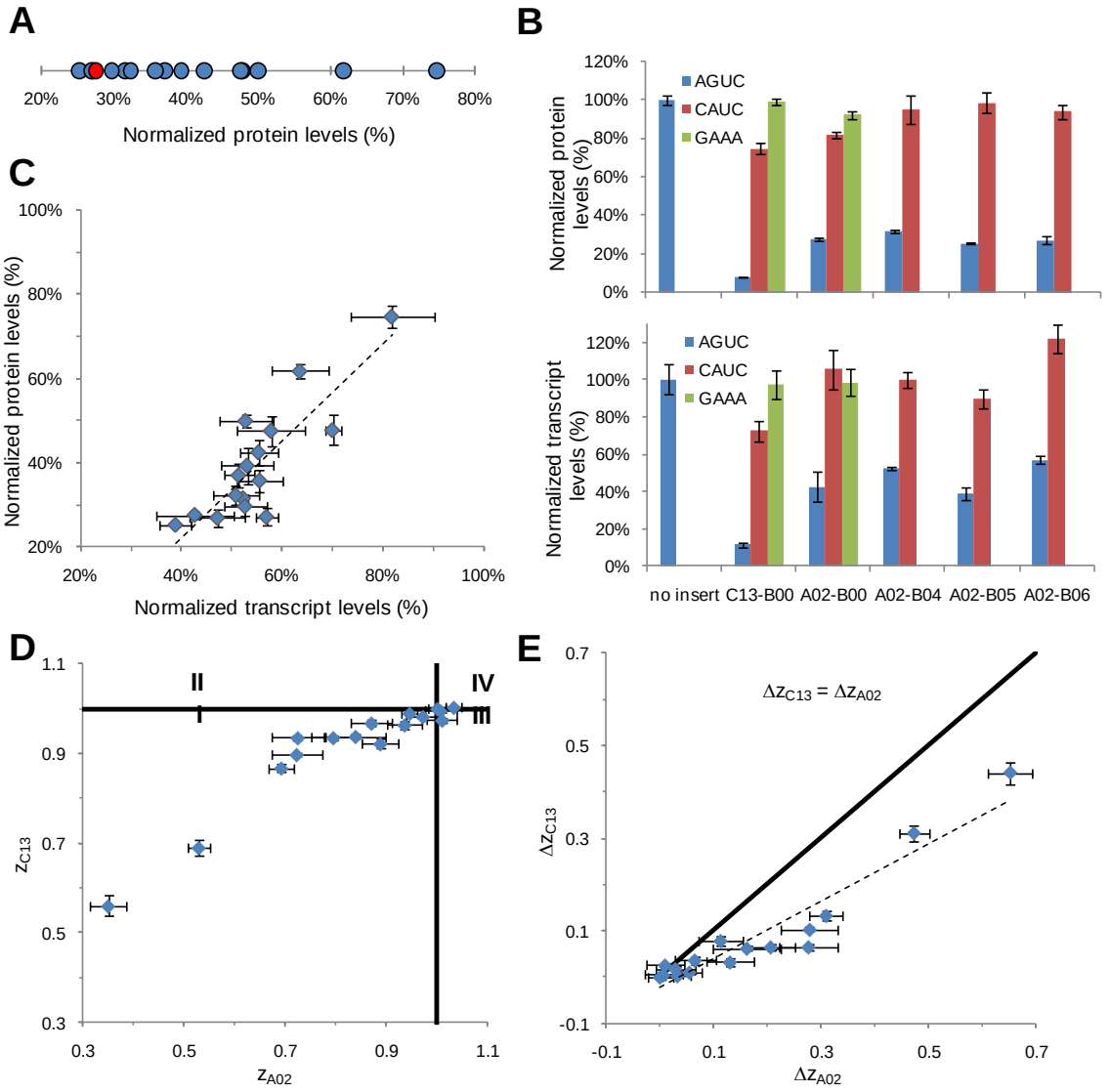


Figure 3.3. *In vivo* characterization of the selected Rnt1p binding library and demonstration of the modularity of the BSB sequences. (A) The gene regulatory activities of the binding library spans an intermediate range of protein expression levels. Normalized protein expression levels were determined by measuring the median GFP levels from a cell population containing the appropriate construct through flow cytometry analysis and values are reported relative to that from an identical construct lacking a hairpin module. The 'parent' BSB is indicated in red. (B) The transcript and protein levels associated with several binding library members and their corresponding mutated tetraloop (CAUC or GAAA) controls support that the observed gene regulatory activity is due to Rnt1p processing. Transcript levels were 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 for each construct are reported relative to that from an identical construct lacking a hairpin module ('no insert'). (C) Correlation analysis of protein and transcript levels from

the binding library members demonstrates a strong correlation between the two measures of gene regulatory activity. (D) The gene regulatory activity of synthetic BSBs is conserved in the context of different CEB modules. The ratio of the knockdown exhibited from a binding library hairpin to that exhibited from the 'parent' (B00 BSB), z_C, was determined in the context of two different CEBs and plotted against each other. Regions I and IV represent hairpins whose activities relative to 'parent' remain consistent when combined with different CEBs. Regions II and III represent hairpins whose activity varies relative to 'parent' when combined with different CEBs. (E) Synthetic BSBs modules generally exhibit higher relative activities in the context of weaker CEB modules. A variable representing the departure from 'parent' activity, Δz_C, was calculated in the context of two different CEBs and plotted against each other. The solid line indicates where the values of Δz_C are the same for both hairpins.

Table 3.1. *In vivo* characterization data for the binding library. All normalized protein and transcript levels were determined as described in Figure 3.3A and Figure 3.3B, respectively.

Substrate	Normalized protein levels (%)	Normalized transcript levels (%)
A02-B00	28% ± 1%	43% ± 8%
A02-B00 (CAUC)	81% ± 2%	106% ± 11%
A02-B01	75% ± 3%	82% ± 8%
A02-B02	62% ± 2%	64% ± 6%
A02-B03	50% ± 2%	53% ± 5%
A02-B04	32% ± 1%	52% ± 1%
A02-B05	25% ± 0%	39% ± 3%
A02-B06	27% ± 2%	57% ± 2%
A02-B07	37% ± 3%	51% ± 3%
A02-B08	30% ± 2%	53% ± 4%
A02-B09	36% ± 3%	56% ± 5%
A02-B10	42% ± 3%	55% ± 4%
A02-B11	32% ± 2%	51% ± 4%
A02-B12	27% ± 2%	47% ± 5%
A02-B13	39% ± 4%	53% ± 5%
A02-B14	48% ± 4%	70% ± 2%
A02-B15	48% ± 4%	58% ± 7%
no insert	100% ± 3%	100% ± 8%

The reduced protein expression levels observed from the Rnt1p binding library is expected to be due to a reduction in the steady-state transcript levels due to rapid degradation of the transcript following endonucleolytic cleavage by Rnt1p. We measured relative transcript levels for each Rnt1p hairpin by quantitative real-time PCR (qRT-PCR) (Table 3.1). A plot of normalized *yEGFP3* expression levels versus normalized *yEGFP3* transcript levels indicates that there is a strong positive correlation (r = 0.847) between the two measures of activity (Figure 3.3C). A preservation of rank order was also observed between protein and transcript levels as indicated by the Spearman rank correlation coefficient (ρ = 0.668). Specifically, with decreasing transcript levels a similar decrease in protein levels is generally observed, confirming that the fluorescence observed was due to changes in the steady-state transcript levels. The negative controls based on mutating the tetraloop confirmed that Rnt1p cleavage is the cause of the observed transcript knockdown (Figure 3.3B).

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

We next examined the gene expression data for the Rnt1p hairpins harboring the synthetic BSBs in the context of two CEBs (C13, A02) for trends in regulatory activity across the binding library. We defined a new variable z_C as the ratio of the knockdown from a binding library member (Bxx) to that of the 'parent' BSB (B00) for a specific CEB (Cxx or Axx):

$$z_C(Bxx) = \frac{knockdown_{C-Bxx}}{knockdown_{C-B00}}$$

A *z* value greater than unity indicates increased knockdown due to the synthetic BSB, whereas a *z* value less than unity indicates decreased knockdown due to the BSB. We

calculated *z* values for each BSB in the context of the CEBs C13 and A02 (z_{C13} and z_{A02}, respectively) and plotted the two variables against each other (Figure 3.3D). For ease in interpretation, we divided the graph into four regions with the point (1,1) at the intersection of the quadrants. Regions I and IV indicate BSBs for which activities relative to parent are conserved between the different CEBs, whereas regions II and III indicate BSBs that exhibit varying activities in the context of different CEBs. Nearly all BSBs are located in regions I and IV, with the majority falling in region I. The data indicate that if a BSB causes increased knockdown in the context of one CEB, it will likely exhibit the same activity in the context of another CEB.

To further examine the gene regulatory activities of the synthetic BSBs, we determined a new variable Δz_C, which is the difference between the z_C value of the 'parent' (which by definition is 1) and the *z* value of the BSB:

$$\Delta z_C(Bxx) = z_C(B00) - z_C(Bxx) = 1 - z_C(Bxx)$$

We calculated Δz for each BSB in the context of the CEBs C13 and A02 and plotted the variables against each other (Figure 3.3E). Data points that fall on the $\Delta z_{C13} = \Delta z_{A02}$ line would indicate BSBs that have the same proportional effect on knockdown for both CEBs. The data fall beneath the $\Delta z_{C13} = \Delta z_{A02}$ line in the region where Δz_{A02} is greater than Δz_{C13}, indicating that Rnt1p hairpins with weaker CEBs are affected more by changes in binding affinity through modification to the BSBs. The data exhibit a strong positive correlation ($r = 0.946$) and can be fit with a trendline by linear regression that passes close to the origin (0,0), suggesting that we see a consistent ratio between Δz_{C13} and Δz_{A02} values, where this ratio is dependent on the CEBs. The data also exhibit a strong preservation of rank order ($\rho = 0.929$), demonstrating that the relative activity

between BSB modules is maintained regardless of the CEB module present in the hairpin. Taken together, the results show the maintenance of BSB activity in connection with different CEB stems and also that the proportional deviation from the 'parent' BSB is determined by the strength of the CEB.

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

We hypothesized that the variation in transcript processing and subsequent protein expression levels exhibited by the binding library is due to variations in binding affinity resulting from alterations in the BSB sequence and/or structure. To examine whether the synthetic Rnt1p binding library members exhibit differences in binding affinity to Rnt1p, we performed *in vitro* binding assays with purified Rnt1p. Binding reactions were ran with 20 nM of *in vitro* synthesized radiolabeled RNA encoding an Rnt1p hairpin and varying concentrations of purified Rnt1p. The reactions were ran in the absence of magnesium and other divalent metal ions that are essential for cleavage to allow Rnt1p to bind to the substrates without subsequent cleavage[37]. Bound products were separated by nondenaturing polyacrylamide gel electrophoresis and quantified through phosphorimaging analysis (Figure 3.4A). 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 3.2). The data indicate that there is a moderate positive

correlation (r = 0.486) between K_D and *in vivo* gene regulatory activity (Figure 3.4B). While we observe several data points demonstrating similar transcript levels for different K_D values, hairpins that bind less tightly to Rnt1p (i.e., higher K_D) generally tend to have higher transcript levels as anticipated. The binding library has an expanded range of K_D values compared to the cleavage library, due to several library members having K_D values greater than those previously reported with the cleavage library (Babiskin, A.B. and Smolke, C.D., in submission). *In vivo*, we observe that most binding library members have increased gene expression levels greater than 'parent'. *In vitro*, this same phenomenon is experienced as most binding library members have decreased affinity for Rnt1p. Binding library members that have decreased gene expression levels than 'parent' *in vivo* also have K_D values comparable or less than 'parent' *in vitro*. However, we also observed that the mutant tetraloop control bound Rnt1p with a similar K_D as the library hairpins, although cleavage was not evident (Figure 3.3B). It has been previously reported that Rnt1p is able to bind its substrates in inactive and active conformations *in vitro*[38]. Therefore, it is plausible that the binding observed in these *in vitro* assays is due to both types of complexes with Rnt1p. Under this situation, the reported K_D may not be solely related to complexes that can be processed. It has also been shown that changes in the BSB affect both binding affinity and hairpin processing by Rnt1p *in vitro*[29]. Therefore, it is important to examine the effects of the binding library on Rnt1p processing rates.

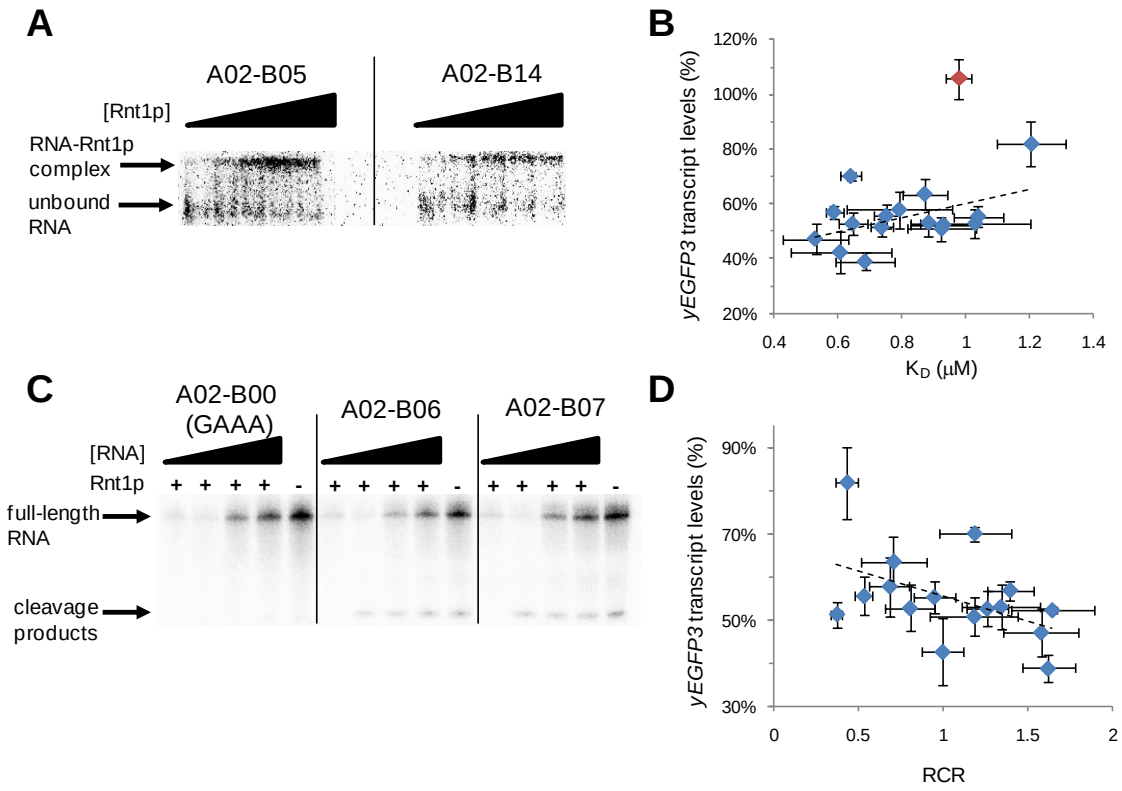


Figure 3.4. *In vitro* characterization of the binding library demonstrates that the observed tuning of gene regulatory activity is achieved through modulation of cleavage rates and binding affinities. (A) Representative mobility shift assays and analyses by nondenaturing polyacrylamide gel electrophoresis of two binding library members: A02-B05 and A02-B14. The top band corresponds to RNA-Rnt1p complexes; the bottom band corresponds to unbound RNA. Rnt1p was added to the following final concentrations in each reaction (left to right; in μM): 0, 0.42, 0.83, 1.25, 1.66. (B) Correlation analysis of binding affinity (K_D) and *yEGFP3* transcript levels indicates a moderate positive correlation between binding affinity and gene regulatory activity. The data point for the A02-B00 (GAAA) negative control is indicated in red. (C) Representative cleavage reaction assays and analyses by denaturing polyacrylamide gel electrophoresis on hairpins A02-B00 (GAAA), A02-B06, and A02-B07. The top band corresponds to unreacted full-length RNA; the bottom band corresponds to the three cleavage products expected from Rnt1p processing. The three cleavage products differ in size by 1 nt and cannot be resolved into individual bands under the conditions used for this assay. RNA was added to the following final concentrations in each reaction (left to right; in μM): 0.2, 0.5, 1.0, 2.0. Reactions lacking Rnt1p were performed with 0.2 μM of RNA. (D) Correlation analysis of relative cleavage rate (RCR) and *yEGFP3* transcript levels demonstrates a moderate positive correlation between cleavage rate and gene regulatory activity.

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

Substrate	RCR	K_D (μM)
A02-B00	1.00 ± 0.12	0.61 ± 0.16
A02-B00 (GAAA)	0*	0.98 ± 0.04
A02-B01	0.44 ± 0.07	1.21 ± 0.11
A02-B02	0.71 ± 0.20	0.87 ± 0.07
A02-B03	0.81 ± 0.15	1.03 ± 0.17
A02-B04	1.65 ± 0.25	0.93 ± 0.10
A02-B05	1.62 ± 0.15	0.69 ± 0.09
A02-B06	1.40 ± 0.14	0.59 ± 0.03
A02-B07	0.37 ± 0.03	0.74 ± 0.04
A02-B08	1.26 ± 0.12	0.65 ± 0.05
A02-B09	0.53 ± 0.05	0.75 ± 0.04
A02-B10	0.95 ± 0.12	1.04 ± 0.08
A02-B11	1.19 ± 0.26	0.93 ± 0.11
A02-B12	1.58 ± 0.22	0.53 ± 0.10
A02-B13	1.34 ± 0.23	0.88 ± 0.06
A02-B14	1.19 ± 0.21	0.64 ± 0.03
A02-B15	0.69 ± 0.12	0.79 ± 0.16

*Immeasurable due to lack of product formation

We analyzed the cleavage reaction between Rnt1p and the binding library 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}. We performed *in vitro* RNA cleavage reactions with a constant concentration of purified Rnt1p against a range of *in vitro* synthesized radiolabeled Rnt1p hairpins to determine the relative values of k_2 for each synthetic Rnt1p hairpin. Reaction products were separated by denaturing polyacrylamide gel electrophoresis and quantified through phosphorimaging analysis (Figure 3.4C). 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 A02–B00 is set to 1 and the rest of the reported values are normalized to A02–B00.

The RCR values for each synthetic Rnt1p hairpin were determined through this analysis method (Table 3.2). The results confirm that the mutant tetraloop control is not processed by Rnt1p *in vitro* (supporting *in vivo* observations). There is a moderate positive correlation ($r = 0.480$) between the measured RCR and gene regulatory activity for the synthetic Rnt1p hairpins (Figure 3.4D). Generally, increases in Rnt1p's ability to cleave a substrate result in greater transcript knockdown. Compared to the cleavage library (Babiskin, A.B. and Smolke, C.D., in submission), we observe a smaller range of RCR values with the binding library, due to the observed decreased range in transcript knockdown. In fact, we observe that cleavage library members exhibiting gene regulatory activities within the range exhibited by the binding library members have similar RCR values. However, a correlation analysis between the K_D and RCR values indicates that there is no correlation between the binding affinity and relative cleavage rates for the binding library members (data not shown). Changing the BSB does result in changes in

binding affinity, but it can also result in variation in the cleavage rate. For any given hairpin, the two properties contribute to the observed transcript levels. In particular, the BSB may affect the hairpin's ability to form active or inactive complexes or the BSB may affect the hairpin's processing rate.

3.3. Discussion

We utilized a cell-based library screening approach to develop a set of synthetic BSB sequences to modulate the gene regulatory activity of engineered Rnt1p hairpins. Previous *in vitro* studies showed that the BSB contains nucleotides critical to Rnt1p binding as mutations in the region resulted in reduced affinity[29]. These studies established a consensus BSB structure, where three Watson-Crick base-pairs were required immediately below the tetraloop followed by another base-pair that could include the wobble guanine-uracil pair. Based on these reported structural requirements, we estimated that a small percentage (~1.6%) of the randomized BSB library would contain hairpins cleavable by Rnt1p. Therefore, a high-throughput FACS-based screen was employed to enrich the library for hairpins resulting in reduced GFP fluorescence. The BSB library was designed in the context of a synthetic CEB from the cleavage library that produced the greatest amount of knockdown (C13) (Chapter II). This design biased the library such that any positive hits would exhibit the lowest possible expression levels to enhance the separation, and thus the selection, of the population of cells containing active hairpins versus the larger population containing inactive hairpins (~98.4% of the population). In total, 16 unique BSBs were identified. However, in contrast to previous *in*

vitro work[29], several of the synthetic BSBs did not contain a base-pair in the fourth position from the tetraloop, suggesting that this structural requirement is relaxed *in vivo*.

The selected BSB sequences were further characterized in the context of a CEB that exhibited weaker gene regulatory activity (A02) (Chapter II) to better resolve differences in the BSB activities. The regulatory activities of the BSB sequences are distributed across an intermediate range of 25% to 75% relative protein levels with the majority exhibiting activities in the range of 25-45%. The binding library exhibited a decreased range of activity relative to the cleavage library (Chapter II), suggesting that the binding library is more appropriate for the tuning of gene expression. For example, the cleavage library can be employed initially to first identify regulatory ranges of interest. More focused regulatory activities can then be explored through implementation of the synthetic BSBs with the appropriate synthetic CEBs. As such, the combinatorial application of the synthetic BSB and CEB elements can be used to extend and tune the regulatory range accessible through the engineered Rnt1p hairpins. To demonstrate the ability to predictably combine BSB and CEB sequences, we examined the synthetic BSB sequences in the context of two synthetic CEBs generated in the cleavage library. Our experimental results show that the synthetic CEB and BSB elements act as modular units, that the BSBs maintain their activity under the context of different CEBs, and that the proportional deviation from the 'parent' BSB is determined by the strength of the CEB.

In vitro characterization studies determined the relationship between binding affinity, cleavage rate, and gene-regulatory activity for the binding library. Changes in the BSB sequence are expected to result in changes in Rnt1p cleavage rate as well as to affect Rnt1p binding[29]. We observed a moderate correlation between the binding affinity

and transcript levels and between the cleavage rate and transcript levels for the binding library. In contrast, although slight changes in affinity were detected with the cleavage library, these changes were not correlated with gene regulatory activity (Chapter II). However, there was a stronger correlation observed between cleavage rate and transcript levels for the cleavage library. The data indicate that nucleotide modifications in the BSB cause changes in affinity and cleavage rate; however, the exact contribution of each of these variables to the observed gene regulatory activity is unclear. In addition, Rnt1p is known to bind in active and inactive complexes where the inactive complex is magnesium-independent[38-39]. The inactive complex may be more prevalent in the binding assay due to the absence of magnesium in the reaction buffer, which is also critical to the proper folding of RNA molecules[40]. The changes in K_D are reflective of how much total RNA is bound regardless of conformation. Thus, it is possible that the changes to the BSB are affecting the partitioning between inactive and active states.

This work extends the regulatory capacity of the first set of posttranscriptional control elements in yeast by developing a set of BSB modules that can be integrated with a previously described set of CEB modules to rationally design synthetic Rnt1p substrates. With 16 CEB and 16 BSB modules developed, 256 different Rnt1p hairpins can be generated. The two classes of modules can be combined to predictably set and precisely tune levels of gene expression in *S. cerevisiae*. The engineered Rnt1p control modules can be used in combination with engineered or native promoter systems, as well as other posttranscriptional elements, to provide a powerful tool for programming genetic regulatory networks in yeast, and thus advancing the application of this cellular chassis in biomanufacturing and biosynthesis processes.

3.4. Materials and Methods

3.4.1. Plasmid construction

Standard molecular biology techniques were utilized to construct all plasmids[41]. 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 with a GenePulser XCell (Bio-Rad, Hercules, CA) 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 were sequence verified by Elim Biopharmaceuticals (Hayward, CA) or the Protein and Nucleic Acid Facility (Stanford, CA). Plasmid maps are available in Supplementary Figure 3.4.

The construction of the Rnt1p characterization plasmid, pCS321, and the Rnt1p expression plasmid, pRNT1, have been previously described (A Babiskin and C Smolke, in submission). A screening plasmid (pCS1585) was constructed from pCS321 by replacing the GAL1-10 promoter with the endogenous TEF1 promoter (JC Liang, unpublished data, 2008). A second screening plasmid (pCS1748) was constructed from pCS1585 by inserting an additional open reading frame (ORF) containing the yeast enhanced mCherry gene, *ymCherry*, flanked by a TEF1 promoter and a CYC1 terminator (JC Liang et al., in preparation).

Insertion of engineered Rnt1p substrates and appropriate controls into the 3’ UTR of *yEGFP3* in pCS321 and pCS1585 was performed through either digestion with 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[42]. The Rnt1p substrates were amplified for insertion with both techniques using the forward and reverse primers RntGap321_fwd (5’ ACCCATGGTATGGATGA ATTGTACAAATAAAGCCTAGGTCTAGAGGCG) and RntGap321_rev2 (5’ TAAGA AATTCGCTTATTTAGAAGTGGCGCGCCCTCTCGAGGGCG), respectively. 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* 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 strain W303. In the case of gap-repair, 250–500 ng of the PCR product and 100 ng of plasmid 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 and 2% dextrose at 30°C.

3.4.2. Library-scale yeast transformation

Yeast transformations with the binding library were performed on Rnt1p substrates as previously described[43]. The C13-based binding library was amplified with template BndLib_C13 (5’ AGCCTAGGTCTAGAGGCGCTATCGTGTCATGTNNNNA GTCNNNNGCATGGCATGATAGCGCCCTCGAGAGGG) and forward and reverse

primers C13BindLibgap_fwd_prmr (5’ GTATTACCCATGGTATGGATGAATTGTAC AAATAAAGCCTAGGTCTAGAGGCGCTATC) and C13BindLibgap_rev_prmr (5’ AATCATAAGAAATTCGCTTATTTAGAAGTGGCGCGCCCTCTCGAGGGCGCTA TCA), respectively. The reaction was scaled-up to 800 μl to obtain roughly 40–50 μg of PCR product. 8 μg of plasmid (either pCS1585 or pCS1748) was digested overnight with AvrII and XhoI in 400 μl total reaction volume. Two tubes of DNA were made with 375 μl of PCR product (~20 μg) and 150 μl of digested plasmid (~3 μg). A third tube acting as a negative control contained 450 μl of water and 75 μl of digested plasmid (~1.5 μg). Each tube was extracted with phenol-chloroform (1:1) and ethanol-precipitated into fresh tubes.

500 μl of Tris-DTT [2.5 M DTT, 1 M Tris (pH 8.0)] was added to a 50-ml culture of yeast strain W303 that was grown in YPD to an OD_{600} of 1.3-1.5 at 30°C. Following 10–15 minutes of additional incubation, the cells were collected and washed in 25 ml of ice-cold Buffer E [10 mM Tris (pH 7.5), 2 mM $MgCl_2$, 270 mM sucrose] and washed again in 1 ml of Buffer E before being resuspended to a final volume of 300 μl in Buffer E. 60 μl of this cell mixture was added to the negative control tube and 120 μl was added to the two tubes containing digested plasmid and the library. After allowing the precipitated DNA to resuspend, 50 μl of the negative control or the library suspension was transferred to a chilled 2-mm gap cuvette and electroporated (540 V, 25 μF, infinite resistance, 2 mm gap). Each 120-μl tube of library suspension contained enough material for two electroporations. Following electroporation, the cells were resuspended in 1 ml of prewarmed YPD and added to a fresh 15-ml Falcon tube. The cuvette was washed a second time with a fresh 1-ml aliquot of YPD, which was added to the same Falcon tube.

Library electroporations were collected in the same Falcon tube (8 ml total). The Falcon tubes were incubated with shaking for 1 hour at 30°C. After incubation, the cells were collected and resuspended in 1 ml of synthetic complete media with an uracil dropout solution and 2% dextrose. The resuspension was added to 6 ml of fresh media to prepare for FACS.

3.4.3. FACS and sorted library retransformation

The transformed binding library was grown for 2–3 days in liquid culture. Following this growth period, the library and appropriate control cultures were collected and suspended in 1x PBS with 1% BSA and either 7-amino-actinomycin D (7-AAD; Invitrogen) or 4’,6-diamidino-2-phenylindole (DAPI; Invitrogen) was added as a viability stain. The cell suspension was passed through a 40-μm Cell Strainer (BD Falcon) prior to analysis on a FACSAria or FACSAria II flow cytometry cell sorter (Becton Dickinson Immunocytometry Systems, San Jose, CA). On the FACSAria, GFP was excited at 488 nm and measured with a bandpass filter of 530/30 nm. 7-AAD was excited at 488 nm and measured with a bandpass filter of 695/40 nm. On the FACS Aria II, GFP was excited at 488 nm and measured with a splitter of 505 nm and bandpass filter of 525/50 nm. mCherry was excited at 532 nm and measured with a splitter of 600 nm and a bandpass filter of 610/20 nm. DAPI was excited at 355 nm and measured with a bandpass filter of 450/50 nm. Detailed sorting procedures are presented in Supplementary Figures 3.1 and 3.2. The collected fractions were diluted to 100 ml in synthetic complete media with an uracil dropout solution and 2% dextrose and grown until an OD_{600} of approximately 1.5. The culture was continually back-diluted and grown in successively decreasing culture

volume for two days at which time freezer stocks were made of the fractions. 100 μl of the culture was collected and the library hairpins amplified by colony PCR with forward and reverse primers RntGap321_fwd and RntGap321_rev2, respectively. The PCR products representing the sorted binding library were recloned through a gap-repair method by transforming the DNA with the pCS321 plasmid in yeast strain W303.

3.4.4. Rnt1p substrate characterization assays

S. cerevisiae cells harboring pCS321-based plasmids were grown on synthetic complete media with an uracil dropout solution and the appropriate sugars (2% raffinose, 1% sucrose) 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. 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. Cells harboring pCS1585-based and pCS1748-based plasmids followed the same procedure as pCS321-based plasmids, except 2% dextrose was the only sugar in the media and no induction was required.

3.4.5. Fluorescence quantification

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 (pCS321-based) of 4.50

(pCS1585-based). Data were collected under low flow rates until 10,000 viable cell counts were collected. For pCS321-based plasmids, a non-induced cell population was used to set a gate to represent GFP-negative and GFP-positive populations. For pCS1585-based plasmids, a plasmid harboring the same backbone as pCS1585 but with no fluorescence gene was used to set the GFP-negative and GFP-positive gates.

The LSRII flow cytometer (Becton Dickinson Immunocytometry Systems) was used to measure mCherry and GFP fluorescence from pmCh-Y-based plasmids. GFP was excited at 488 nm and measured with a splitter of 505 nm and a bandpass filter of 525/50. mCherry was excited at 532 nm and measured with a splitter of 600 nm LP and a bandpass filter of 610/20 nm. DAPI was excited at 405 nm and measured with a bandpass filter of 450/50 nm.

3.4.6. Quantification of cellular transcript levels

Total RNA from *S. cerevisiae* was collected by a standard hot acid phenol extraction method[44] 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* and *ACT1*[45] (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' CGGTGAAGGTGA AGGTGATGCTACT) and rnt1p_rtpcr_rev2 (5' GCTCTGGTCTTGTAGTTACCGTCA TCTTTG), respectively. The forward and reverse primers for *ACT1* quantification are ACT1_rtpcr_fwd (5' GGCATCATACCTTCTACAACGAAT) and ACT1_rtpcr_rev (5'

GGAATCCAAAACAATACCAGTAGTTCTA), respectively. Relative transcript levels 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) on an iCycler iQ qRT-PCR machine (Bio-Rad) according to the manufacturer's instructions. For each run, a standard curve was generated for *yEGFP3* and a house-keeping gene, *ACT1*, using a dilution series for a control representing no insertion of an Rnt1p substrate. Relative *yEGFP3* and *ACT1* levels were first individually determined for each sample and then the *yEGFP3* values were normalized by their corresponding *ACT1* values.

3.4.7. In vitro *transcription of Rnt1p substrates*

All Rnt1p substrates were PCR-amplified to include an upstream T7 promoter site using forward and reverse primers Rnt1p-T7-PCR_fwd_prmr (5' TTCTAATACGACTC ACTATAGGGACCTAGGAAACAAACAAAGTTGGGC) and Rnt1p-T7-PCR_rev_ prmr (5' CTCGAGTTTTTATTTTTCTTTTTGCCGGGCG), respectively. 1–2 μg of PCR product was transcribed with T7 RNA 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.

3.4.8. 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[46]. 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 20,000 g 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.

3.4.9. In vitro *Rnt1p substrate cleavage assay*

Cleavage assays were performed on Rnt1p substrates as previously described[38, 46]. Briefly, a 10-μl mixture of RNA and Rnt1p were 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 0.2 to 2.0 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.

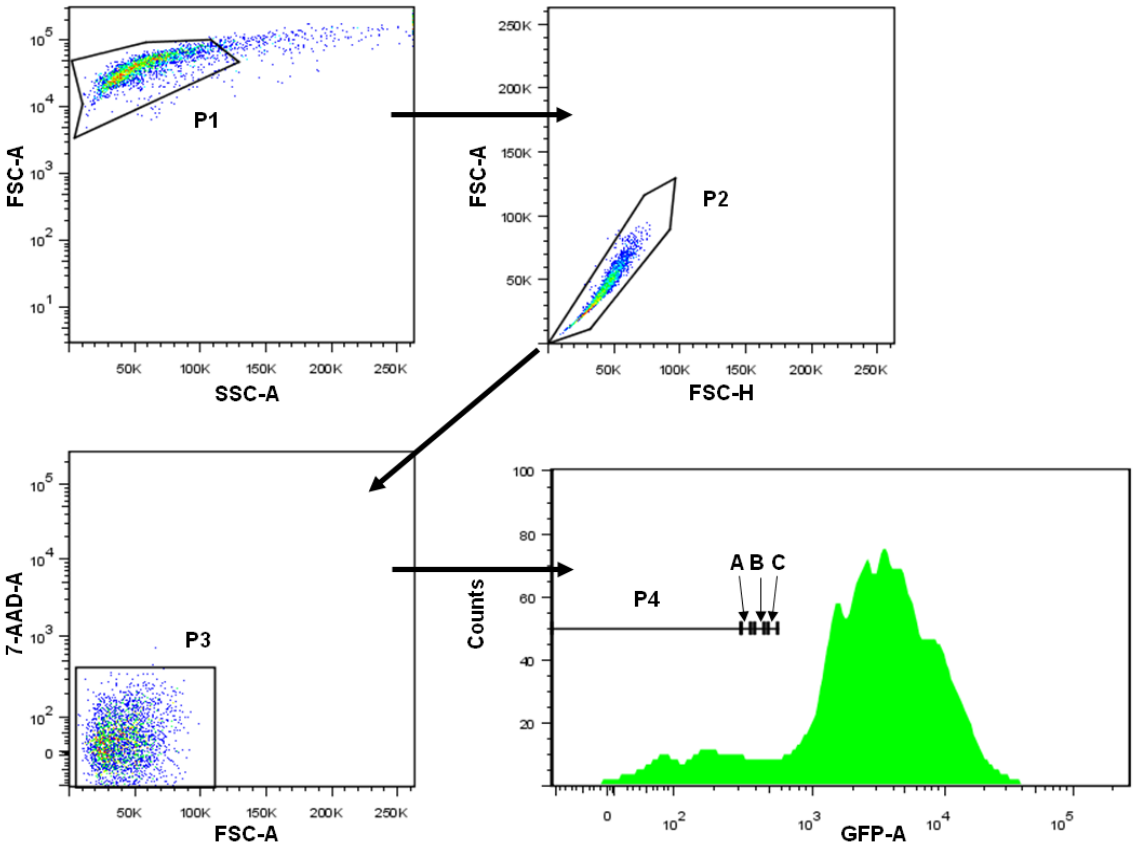
3.4.10. In vitro Rnt1p substrate mobility shift assay

Mobility shift assays were performed as previously described[38, 46]. Briefly, a 10-μl mixture of RNA and Rnt1p were incubated on ice for 10 min in Rnt1p binding buffer

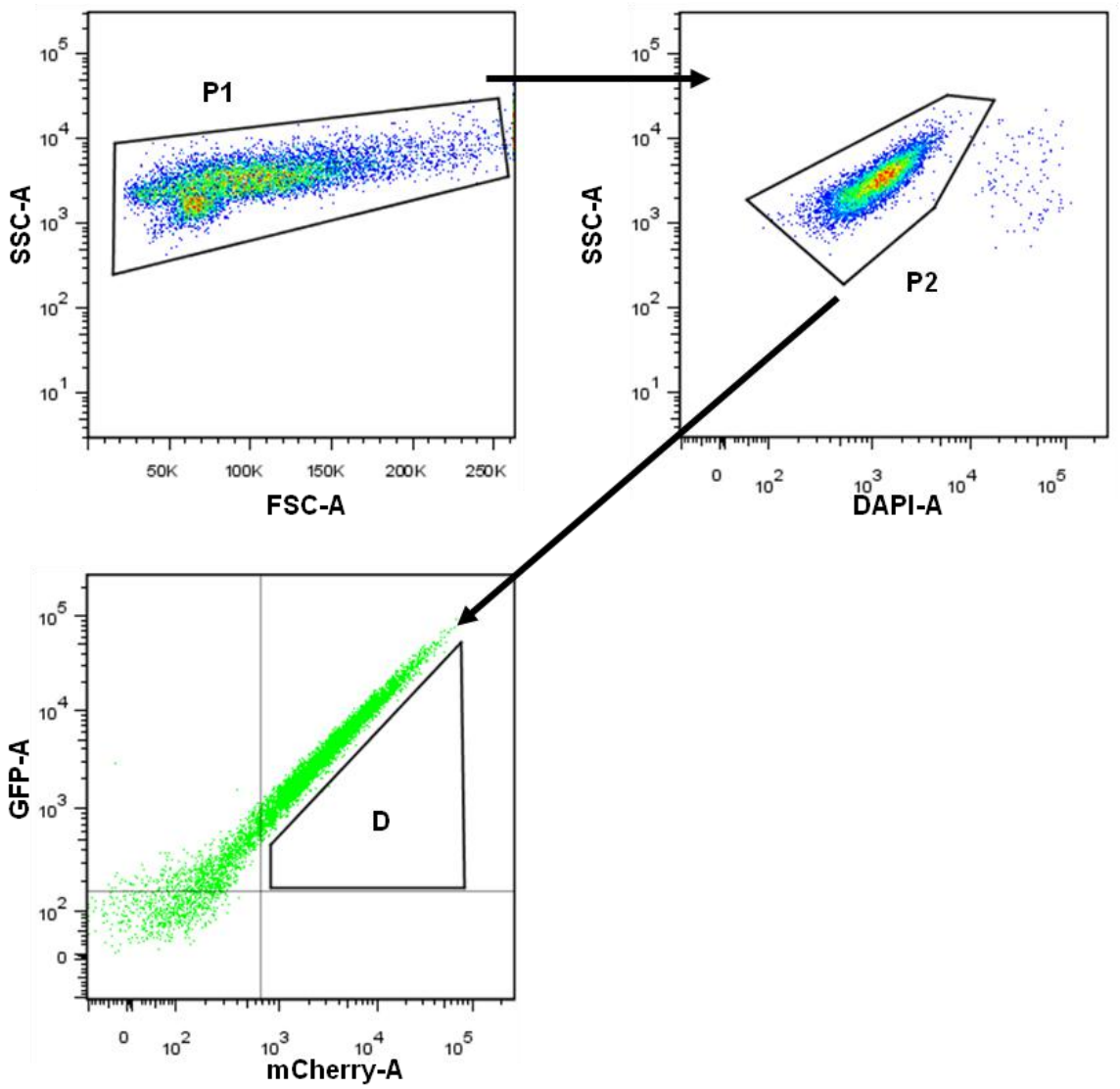
[20% (v/v) glycerol, 30 mM Tris (pH 7.5), 150 mM KCl, 5 mM spermidine, 0.1 mM DTT, and 0.1 mM EDTA (pH 7.5)]. The RNA concentration in all samples was 200 nM and the Rnt1p concentration ranged from 0 to 1.7 μM. The binding reaction products were separated on a 6% native polyacrylamide gel run at 350 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.

3.5. Supplementary Information

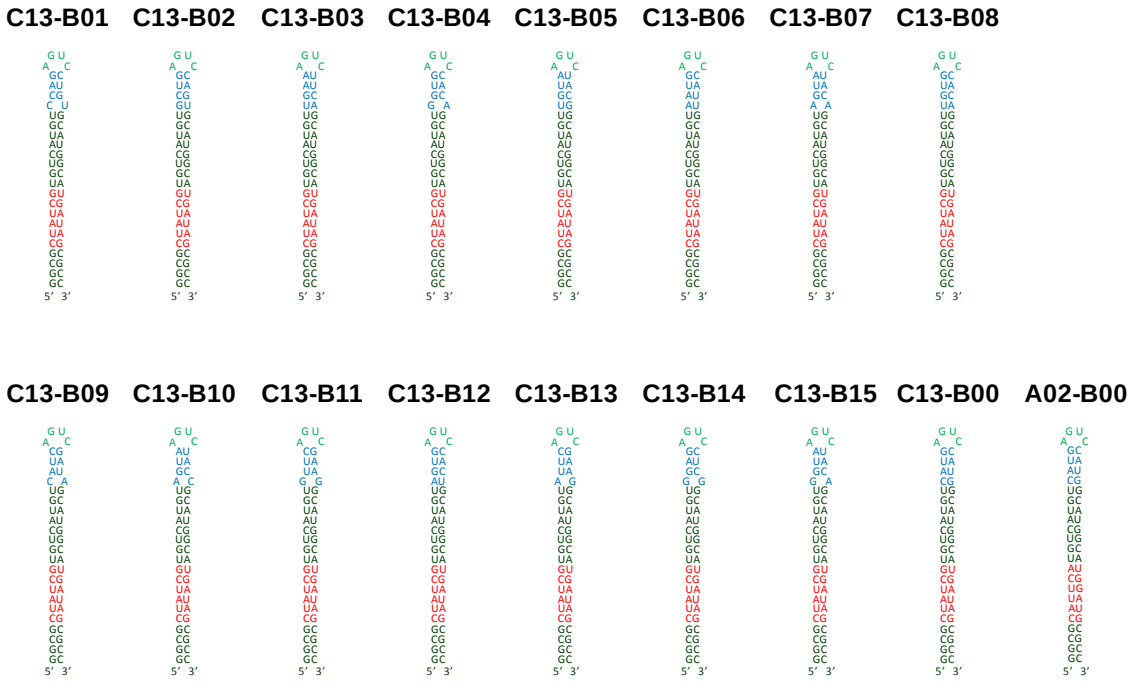
Supplementary Figures and Tables



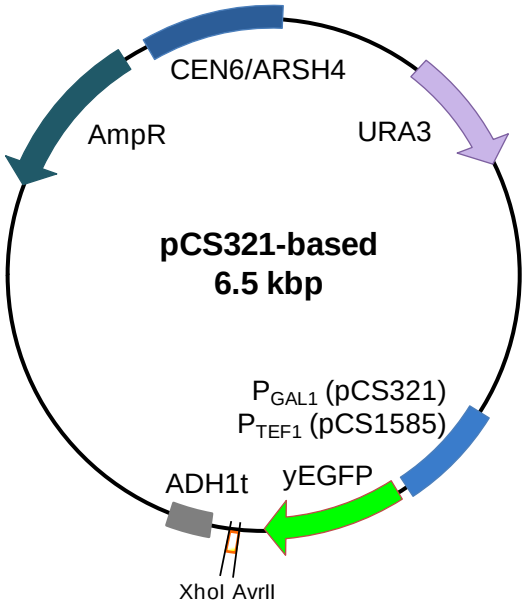
Supplementary Figure 3.1. FACS analysis and gating procedure for pCS1585 system on FACSAria. As an example, data for the construct baring no Rnt1p hairpin is presented. Dot plots show initial gating of stable cells (P1), followed by gating for cell uniformity (P2), and finally gating for live cells with the 7-AAD stain (P3). GFP-negative cells (P4) were gated initially with a construct lacking a fluorescent gene (empty vector). Cells outside of P4 represent GFP-positive cells. Fractions A, B, and C are represented on this graph, but collections were only performed with the binding library sample. The fractions cover the range of expression seen with the C13-B00 hairpin. With the binding library sample, ~120,000 cells were analyzed with 719, 841, and 943 cells collected in fractions A, B, and C respectively.



Supplementary Figure 3.2. FACS analysis and gating procedure for pCS1748 system on FACSAria II. As an example, data for the GFP positive construct (no Rnt1p hairpin) is presented. Dot plots show initial gating of stable cells (P1) and subsequent gating of uniform, live cells (P2) with DAPI used for the viability stain. A construct lacking fluorescent genes (empty vector) was used to set the gates for mCherry- and GFP-positive cells. A single gate (D) was set to collect all GFP-positive cells that exhibited lower GFP fluorescence than cells containing the positive construct. With the binding library sample, 1,000,000 cells were analyzed with 18,416 cells collected in fraction D.



Supplementary Figure 3.3. Sequences and structures of the selected Rnt1p binding library and control hairpins containing the 'parent' BSB. The binding library was initially sequenced when in the context of the C13 CEB.



Supplementary Figure 3.4. Plasmid maps of pCS321-based vectors. pCS321 is the characterization plasmid and GFP expression is driven by the GAL1 promoter. pCS1585 is a screening plasmid used with FACS and GFP expression is driven by the TEF1 promoter.

Supplementary Table 3.1. Sequence and *in vivo* characterization data of all tested Rnt1p hairpins. The nucleotides of the BSB are indicated in blue. The CEB sequences in the 'parent' hairpins (xx-B00) are indicated in red. All normalized protein and transcript levels were determined as described in Figure 3.3A and Figure 3.3B, respectively.

Substrate	Sequence	Normalized protein levels (%)	Normalized transcript levels (%)
A02-B00	GGCGCAUUCAUGUCAUGUCAUGAG UCCAUGGCAUGGCAUGGAUGCGCC	28% ± 1%	43% ± 8%
A02-B01	GGCGCAUUCAUGUCAUGUCCAGAG UCCUGUGCAUGGCAUGGAUGCGCC	75% ± 3%	78% ± 11%
A02-B02	GGCGCAUUCAUGUCAUGUGCUGAG UCCAGUGCAUGGCAUGGAUGCGCC	62% ± 2%	64% ± 6%
A02-B03	GGCGCAUUCAUGUCAUGUUGAAAG UCUUCAGCAUGGCAUGGAUGCGCC	50% ± 2%	53% ± 5%
A02-B04	GGCGCAUUCAUGUCAUGUGGUGAG UCCACAGCAUGGCAUGGAUGCGCC	32% ± 1%	57% ± 7%
A02-B05	GGCGCAUUCAUGUCAUGUUGUAAG UCUACGGCAUGGCAUGGAUGCGCC	25% ± 0%	36% ± 5%
A02-B06	GGCGCAUUCAUGUCAUGUAAUGAG UCCAUUGCAUGGCAUGGAUGCGCC	27% ± 2%	60% ± 6%
A02-B07	GGCGCAUUCAUGUCAUGUAGUAAG UCUACAGCAUGGCAUGGAUGCGCC	37% ± 3%	51% ± 3%
A02-B08	GGCGCAUUCAUGUCAUGUUGUGAG UCCACAGCAUGGCAUGGAUGCGCC	30% ± 2%	53% ± 4%
A02-B09	GGCGCAUUCAUGUCAUGUCAUCAG UCGAUAGCAUGGCAUGGAUGCGCC	36% ± 3%	60% ± 9%
A02-B10	GGCGCAUUCAUGUCAUGUAGUAAG UCUACCGCAUGGCAUGGAUGCGCC	42% ± 3%	55% ± 4%
A02-B11	GGCGCAUUCAUGUCAUGUGUUCAG UCGAAGGCAUGGCAUGGAUGCGCC	32% ± 2%	51% ± 4%
A02-B12	GGCGCAUUCAUGUCAUGUAGUGAG UCCACUGCAUGGCAUGGAUGCGCC	27% ± 2%	47% ± 5%
A02-B13	GGCGCAUUCAUGUCAUGUAUUCAG UCGAAGGCAUGGCAUGGAUGCGCC	39% ± 4%	53% ± 5%
A02-B14	GGCGCAUUCAUGUCAUGUGGAGAG UCCUCGGCAUGGCAUGGAUGCGCC	48% ± 4%	74% ± 8%
A02-B15	GGCGCAUUCAUGUCAUGUGGUAAG UCUACAGCAUGGCAUGGAUGCGCC	48% ± 4%	58% ± 7%
C13-B00	GGCGCUAUCGUGUCAUGUCAUGAG UCCAUGGCAUGGCAUGAUAGCGCC	8% ± 0%	12% ± 1%
C13-B01	GGCGCUAUCGUGUCAUGUCCAGAG UCCUGUGCAUGGCAUGAUAGCGCC	48% ± 2%	
C13-B02	GGCGCUAUCGUGUCAUGUGCUGAG UCCAGUGCAUGGCAUGAUAGCGCC	37% ± 2%	
C13-B03	GGCGCUAUCGUGUCAUGUUGAAAG UCUUCAGCAUGGCAUGAUAGCGCC	20% ± 1%	
C13-B04	GGCGCUAUCGUGUCAUGUGGUGAG UCCACAGCAUGGCAUGAUAGCGCC	9% ± 0%	
C13-B05	GGCGCUAUCGUGUCAUGUUGUAAG UCUACGGCAUGGCAUGAUAGCGCC	8% ± 0%	
C13-B06	GGCGCUAUCGUGUCAUGUAAUGAG UCCAUUGCAUGGCAUGAUAGCGCC	8% ± 0%	
C13-B07	GGCGCUAUCGUGUCAUGUAGUAAG UCUACAGCAUGGCAUGAUAGCGCC	11% ± 1%	
C13-B08	GGCGCUAUCGUGUCAUGUUGUGAG UCCACAGCAUGGCAUGAUAGCGCC	10% ± 0%	

C13-B09	GGCGCUAUCGUGUCAUGUCAUCAG UCGAUAGCAUGGCAUGAUAGCGCC	15% ± 1%
C13-B10	GGCGCUAUCGUGUCAUGUAGUAAG UCUACCGCAUGGCAUGAUAGCGCC	14% ± 1%
C13-B11	GGCGCUAUCGUGUCAUGUGUUCAG UCGAAGGCAUGGCAUGAUAGCGCC	11% ± 1%
C13-B12	GGCGCUAUCGUGUCAUGUAGUGAG UCCACUGCAUGGCAUGAUAGCGCC	10% ± 0%
C13-B13	GGCGCUAUCGUGUCAUGUAUUCAG UCGAAGGCAUGGCAUGAUAGCGCC	14% ± 0%
C13-B14	GGCGCUAUCGUGUCAUGUGGAGAG UCCUCGGCAUGGCAUGAUAGCGCC	17% ± 0%
C13-B15	GGCGCUAUCGUGUCAUGUGGUAAG UCUACAGCAUGGCAUGAUAGCGCC	14% ± 1%

Acknowledgements

We thank J. Liang and A. Chang for assistance with FACS and for providing the pCS1585 and pCS1748 plasmids; K. Hoff for assistance in the expression and purification of Rnt1p; and S. Bastian and F.H. Arnold for assistance in sonication and FPLC. This work was supported by the National Science Foundation (CAREER award CBET-0917705 to C.D.S.); and the Alfred P. Sloan Foundation, fellowship (to C.D.S.).

References

1. 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).
2. 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).

3. 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).

4. Zhu, M.M., Lawman, P.D. & Cameron, D.C. Improving 1,3-propanediol production from glycerol in a metabolically engineered Escherichia coli by reducing accumulation of sn-glycerol-3-phosphate. *Biotechnol Prog* **18**, 694-699 (2002).

5. 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).

6. 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).

7. Alper, H., Miyaoku, K. & Stephanopoulos, G. Construction of lycopene-overproducing E. coli strains by combining systematic and combinatorial gene knockout targets. *Nat Biotechnol* **23**, 612-616 (2005).

8. 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).

9. Lutz, R. & Bujard, H. Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* **25**, 1203-1210 (1997).

10. 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).

11. 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).

12. 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).

13. 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).

14. 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).

15. Beisel, C.L., Bayer, T.S., Hoff, K.G. & Smolke, C.D. Model-guided design of ligand-regulated RNAi for programmable control of gene expression. *Mol Syst Biol* **4**, 224 (2008).

16. 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).

17. Anderson, J.C., Clarke, E.J., Arkin, A.P. & Voigt, C.A. Environmentally controlled invasion of cancer cells by engineered bacteria. *J Mol Biol* **355**, 619-627 (2006).

18. Keiler, K.C., Waller, P.R. & Sauer, R.T. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**, 990-993 (1996).

19. 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).

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

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

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

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

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

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

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

27. 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).

28. Wu, H. et al. A novel family of RNA tetraloop structure forms the recognition site for Saccharomyces cerevisiae RNase III. *EMBO J* **20**, 7240-7249 (2001).

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

30. Lamontagne, B., Hannoush, R.N., Damha, M.J. & Abou Elela, S. Molecular requirements for duplex recognition and cleavage by eukaryotic RNase III: discovery of an RNA-dependent DNA cleavage activity of yeast Rnt1p. *J Mol Biol* **338**, 401-418 (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. Ge, D., Lamontagne, B. & Elela, S.A. RNase III-mediated silencing of a glucose-dependent repressor in yeast. *Curr Biol* **15**, 140-145 (2005).

35. Danin-Kreiselman, M., Lee, C.Y. & Chanfreau, G. RNAse III-mediated degradation of unspliced pre-mRNAs and lariat introns. *Mol Cell* **11**, 1279-1289 (2003).

36. Lee, A., Henras, A.K. & Chanfreau, G. Multiple RNA surveillance pathways limit aberrant expression of iron uptake mRNAs and prevent iron toxicity in S. cerevisiae. *Mol Cell* **19**, 39-51 (2005).

37. 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).

38. 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).

39. Lavoie, M. & Abou Elela, S. Yeast ribonuclease III uses a network of multiple hydrogen bonds for RNA binding and cleavage. *Biochemistry* **47**, 8514-8526 (2008).

40. Misra, V.K. & Draper, D.E. The linkage between magnesium binding and RNA folding. *J Mol Biol* **317**, 507-521 (2002).

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

42. 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).

43. Chao, G. et al. Isolating and engineering human antibodies using yeast surface display. *Nat Protoc* **1**, 755-768 (2006).

44. 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).

45. 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).

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