Published in final edited form as:

Metab Eng. 2013 September ; 19: 88–97. doi:10.1016/j.ymben.2013.07.001.

Use of High Capacity Terminators in *Saccharomyces cerevisiae* to Increase mRNA half-life and Improve Gene Expression Control for Metabolic Engineering Applications

Kathleen A. Curran¹, Ashty S. Karim¹, Akash Gupta¹, and Hal S. Alper^{1,2,*}

¹Department of Chemical Engineering, The University of Texas at Austin, 200 E Dean Keeton St. Stop C0400, Austin, TX 78712

²Institute for Cellular and Molecular Biology, The University of Texas at Austin, 2500 Speedway Avenue, Austin, TX 78712

Abstract

Control of gene and protein expression of both endogenous and heterologous genes is a key component of metabolic engineering. While a large amount of work has been published characterizing promoters for this purpose, less effort has been exerted to elucidate the role of terminators in yeast. In this study, we characterize over 30 terminators for use in metabolic engineering applications in *Saccharomyces cerevisiae* and determine mRNA half-life changes to be the major cause of the varied protein and transcript expression level. We demonstrate that the difference in transcript level can be over 6.5-fold even for high strength promoters. The influence of terminator selection is magnified when coupled with a low-expression promoter, with a maximum difference in protein expression of 11-fold between a high-capacity terminator and the parent plasmid terminator and over 35-fold difference when compared with a no-terminator baseline. This is the first time that terminators have been investigated in the context of multiple promoters spanning orders of magnitude in activity. Finally, we demonstrate the utility of terminator selection for metabolic engineering by using a mutant xylose isomerase gene as a proof-of-concept. Through pairing a high-capacity terminator with a low-expression promoter, we were able to achieve the same phenotypic result as with a promoter considerably higher in strength. Moreover, we can further boost the phenotype of the high-strength promoter by pairing it with a high-capacity terminator. This work highlights how terminator elements can be used to control metabolic pathways in the same way that promoters are traditionally used in yeast. Together, this work demonstrates that terminators will be an important part of heterologous gene expression and metabolic engineering for yeast in the future.

Keywords

terminator; yeast; *Saccharomyces cerevisiae*; metabolic engineering

© 2013 Elsevier Inc. All rights reserved.

*Corresponding author: Phone: (512) 471-4417, Fax: (512) 471-7060, halper@che.utexas.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

The yeast *Saccharomyces cerevisiae* is well established as a metabolic engineering platform host capable of producing an array of valuable chemicals, fuels, and pharmaceuticals in a safe and sustainable manner (Curran and Alper, 2012; Liu et al., 2013). However, achieving high yields and titers of these products is predicated on the capacity to precisely control both endogenous and heterologous gene and protein expression levels. Currently, this control is primarily achieved through changing the copy number of the gene and through altering the promoter driving expression. However, many factors (both genetic and context related) influence the final expression level of a gene, and ultimately, the protein. To this end, most of the components of a standard yeast “expression cassette” have been demonstrated to exert control on net protein output. Extensive work has illustrated the importance of promoter strength and regulation on net output from these expression cassettes (Blazeck and Alper, 2013; Da Silva and Srikrishnan, 2012). Additionally, studies have demonstrated the influence of additional factors including the origin of replication and selection marker on vector plasmids (Karim et al., 2013), the genomic integration site (Flagfeldt et al., 2009), and the 5' UTR region (Crook et al., 2011). However, with the exception of only a few recent studies, the 3' region after the gene, known as the terminator, has been largely overlooked in yeast. Here, we seek to identify and characterize key “high capacity” terminators that enable superior net protein output from an expression cassette. For this study, “high capacity” terminators are defined as ones that enable increased protein expression over conventionally-used elements such as the *CYC1* terminator.

The importance of terminator choice has not been as widely studied as promoter activity. Usually only a few default terminators, such as those from the *CYC1* or *ADHI* genes, are used in yeast. The importance of 3' UTR regions as RNA stability elements has been well-established for bacterial systems. Efforts in prokaryotic systems have recently demonstrated that both terminators and designed 3' UTR elements can fundamentally change heterologous expression level (Cambray et al., 2013; Pflieger et al., 2006). However, a similar level of fundamental understanding has yet to be applied to a fungal system. Different terminators have been selected in other studies either to pair with a corresponding promoter being used or to minimize the chances for undesired homologous recombination in large heterologous cassettes (Shao et al., 2009). However, in such cases, the terminators were not chosen on the basis of any particular characteristics nor were they functionally characterized. Recently, Yamanishi *et al.* demonstrated that fluorescent protein expression can be increased 50% by using the terminator from the *TPS1* gene in place of the commonly used *CYC1* terminator (Yamanishi et al., 2011). Further work by this group characterized the heterologous protein expression level using over 3000 terminators from the yeast genome and demonstrated that the choice of terminator can more than double protein expression compared to another commonly used terminator, *PGK1* (Yamanishi et al., 2013). However, none of these studies have established a concerted mechanism for the terminators behavior or demonstrated their applicability in a metabolic engineering application in yeast. Beyond bacterial and fungal systems, it has also been demonstrated in human cells that the lack of a terminator in a heterologous expression cassette results in dramatically decreased gene and protein expression (West and Proudfoot, 2009). Taken together, these studies provide strong evidence that terminator choice in a heterologous expression cassette should be strongly considered to achieve desirable and/or maximal final protein output in yeast.

Despite a lack of characterization for metabolic engineering purposes, fundamental studies of yeast termination processes and 3' end-processing have been ongoing for several decades. Several well-documented sequence elements are shown to play a role in this process, including an upstream AU-rich efficiency element, an upstream A-rich positioning element, the cleavage site, and upstream and downstream U-rich elements (Mischo and Proudfoot,

2013). Using these polyadenylation site terminator elements, Guo et al. have demonstrated a synthetic, minimum sequence necessary to achieve termination (Guo and Sherman, 1996). Furthermore, it has recently been discovered that yeast may have additional termination processes beyond the canonical polyadenylation site terminators with the consensus sequences described above. Specifically, newly discovered NRD-dependent terminators, which contain binding sites for Nrd1p and Nab3p (Carroll et al., 2007) as well as terminators that utilize the Rnt1p termination pathway (Rondon et al., 2009) are found in yeast. Additionally, evidence has shown that terminators may also play an important role in transcription initiation and elongation through the formation of “gene loops” between the RNA Polymerase II complex and the promoter and terminator of a gene (O’Sullivan et al., 2004). Finally, since the terminator sequence determines the length and sequence of the 3’ un-translated region (UTR) as well as the level of polyadenylation, the terminator may influence mRNA half-life. These observations suggest that terminators are more germane to both transcription and translation processes than previously considered and that different terminators likely have different performance characteristics.

Here, we characterize over 30 different terminators and identify many “high-capacity” terminators for use in metabolic engineering applications in yeast. We define “high-capacity” terminators as those that allow for increased protein expression over conventional terminators such as the terminator from *CYC1*. Furthermore, we determine half-life to be the major factor influenced by terminators and demonstrate that the difference in relative output between terminators can vary as much as 11-fold over the commonly used *CYC1* terminator depending on the promoter selected. This is the first study to determine the impact of pairing varying strength promoters with terminators in yeast. Finally, we present the first demonstration for the utility of using terminators for metabolic engineering in yeast through tuning the over-expression of a mutant *xyIA* gene for increased growth on xylose. These results demonstrate that, like promoters, terminators can be used to tune gene expression and thus influence pathway performance.

2. Materials and Methods

2.1 Strains and media

Saccharomyces cerevisiae strains BY4741 (*MAT a; his3⁻1; leu2⁻0; met15⁻0; ura3⁻0*) and BY4741 *gre3::KanMX* from the Yeast Deletion Database (Giaever et al., 2002) were used in this study. Yeast strains were propagated at 30°C in yeast complete synthetic medium (CSM). CSM is composed of 6.7 g/L yeast nitrogen base, 20 g/L glucose, galactose or xylose, and either CSM-HIS, CSM-LEU, or CSM-HIS-URA-LEU supplement (MP Biomedicals, Solon, OH), depending on the required auxotrophic selection. *Escherichia coli* strain *DH10* was used for all cloning and plasmid propagation. *DH10* was grown at 37°C in Luria-Bertani (LB) media supplemented with 50 µg/mL of ampicillin. All strains were cultivated with 225 RPM orbital shaking. Yeast and bacterial strains were stored at –80°C in 15% glycerol.

2.2 Plasmid construction

All plasmids used in this study were based on the p4XX-vectors described in Mumberg et al. (1995). The *TEFmut3* promoter, the *GAL1* promoter, *xyIA**3, and *yECitrine* genes were cloned via PCR from plasmids described previously (Alper et al., 2005; Hegemann and Heick, 2011; Lee et al., 2012; Nevoigt et al., 2006; Sheff and Thorn, 2004). The *CPS1*, *HIS5*, and *VPS13* promoters and all terminators were cloned via PCR from extracted BY4741 gDNA obtained using the Wizard Genomic DNA Extraction Kit from Promega (Madison, WI). Promoter lengths for *CPS1*, *HIS5* and *VPS13* were chosen based on the location of known transcription factor binding sites as reported in the Yeast Promoter Atlas

(Chang et al., 2011). With the exception of the *CYC1**, *ADH1*, *PRM9*, and *TPS1* terminators, all terminator lengths were determined by using the 3' UTR length of the gene as previously reported (Tuller et al., 2009) and adding 50 base-pairs (Figure 1). In cases where the 3' UTR was unknown, such as for *PRM9*, a length of 250 base-pairs was used. For *TPS1* and *ADH1*, terminator length was chosen based on previous work (Chien et al., 1991; Yamanishi et al., 2011). The *CYC1** terminator is the terminator from the parent plasmid as reported in Mumberg et al. (1995), which was originally cloned from the D311-3A strain (Montgomery et al., 1978). Cloned terminators were inserted into the plasmid using the *SalI* and *EagI* restriction sites. The *SalI* site, in addition to the *SpeI* site at the beginning of the multicloning site, was also used to insert the genes, *yECitrine* or *xyIA*3*, so there was only a single restriction site between the end of the gene and the terminator. To create constructs with no terminator, the gene was cloned in using the *EagI* site instead of *SalI*. Promoters *GAL1*, *TEFmut3*, *CPS1*, *HIS5* and *VPS13* were cloned into the plasmid using the *SacI* and *XbaI* or *SpeI* restriction sites. The *CYC1* and *TDH3* promoters were available in the parent plasmid set (Mumberg et al., 1995). See Supplementary Table 1 for primers and sequences.

Standard cloning and bacterial transformations were performed according to Sambrook and Russell (Sambrook and Russell, 2001). PCR reactions used Phusion High-Fidelity DNA Polymerase from New England Biolabs (Ipswich, MA) and followed supplier instructions; primers were purchased from Integrated DNA Technologies (Coralville, Iowa). Antarctic phosphatase and all restriction enzymes were purchased from New England Biolabs. Fermentas T4 DNA ligase and all other enzymes and chemicals were purchased through Thermo Fisher Scientific (Waltham, MA). Vectors were isolated using the Zippy Plasmid Miniprep kit from Zymo Research Corp. (Irvine, CA) and DNA purification was performed with a Qiaquick PCR Cleanup kit (Qiagen, Valencia, CA). Plasmids were transformed using the EZ Yeast Transformation II Kit from Zymo Research Corp. according to manufacturer's instructions.

2.3 Flow Cytometry

Fluorescence from strains expressing the *yECitrine* gene was measured using a FACS Fortessa (BD Biosciences) using a YFP fluorochrome in biological triplicate. Cells were grown to mid-log phase overnight from a starting OD₆₀₀=0.005 and 10,000 events were collected using the Fortessa for each strain. Day to day voltage variability was mitigated by measuring all comparable strains on the same day. FlowJo (Tree Star Inc., Ashland, OR) was used to analyze data and to compute mean fluorescence values.

2.4 Real-time PCR and mRNA half-life measurements

The relative abundance of heterologous mRNA was determined using quantitative RT-PCR. RNA was extracted from mid-log phase cells using the Zymo YeaStar RNA Kit. Contaminating DNA was removed using the Ambion DNA-free kit and cDNA was prepared using the Applied Biosystems High Capacity Reverse Transcription Kit (Life Technologies, Carlsbad, CA). Primers were obtained from Integrated DNA Technologies (see Supplementary Table 1 for primers). Quantitative PCR was performed on a ViiA7 Real Time PCR System (Life Technologies) using Fast Start SYBR Green Master Mix (Roche, Penzberg, Germany), following the manufacturer's instructions with an annealing temperature of 58°C. For single time-point experiments, *ALG9* was used as the reference housekeeping gene. For multi-time-point experiments to determine mRNA half-life, *SCR1* was used instead. mRNA half-life was determined by using plasmids expressing *yECitrine* driven by the *GAL1* inducible promoter. Strains were grown to mid-log phase in CSM-HIS + galactose media, then resuspended in an equal amount of CSM-HIS + glucose media. Once the media had been changed, samples were taken every 5 to 20 minutes and the RNA

extraction was performed. Half-lives were then determined as described previously (Coller, 2008).

2.5 Xylose growth experiments

BY4741 *gre3* strains co-expressing an evolved *xyIA**3 gene (Lee et al., 2012) and overexpressing *XKS1* and *TAL1* were grown in CSM-HIS-LEU-URA + glucose media to stationary phase, then re-inoculated into 5 mL aliquots of CSM-HIS-LEU-URA + xylose media at OD₆₀₀=0.25. Growth measurements were taken and exponential growth rates were calculated and averaged between biological triplicates.

3. Results

3.1 Initial characterization of terminators using a fluorescent protein output

Initially, potential “high-capacity terminators” capable of supporting increased net protein expression were identified through a literature search. Ultimately, the metabolic engineering toolkit will require a library of such terminators to pair with varied promoter strengths for total control of gene and protein expression. Since factors such as 3' UTR sequence, length, and polyadenylation are likely to influence mRNA half-life, it was hypothesized that genes with long reported mRNA half-lives could possess high-capacity terminators. However, since genome-wide half-life studies vary significantly in their results due to differences in growth conditions and measurement techniques, data from six separate studies was utilized. Five of the studies measured mRNA half-life across the transcriptome (Grigull et al., 2004; Holstege et al., 1998; Miller et al., 2011; Shalem et al., 2008; Wang et al., 2002) and one predicted terminator location and strength based on a bio-informatics approach (Graber et al., 2002). For the *in vivo* studies, terminators from the five genes with the highest reported mRNA half-lives were selected (Table 1). For the *in silico* study, the five strongest predicted terminator regions were chosen (Table 1). Interestingly, there was only one gene, *HSP26*, which was represented in more than one of these datasets. Terminator length was determined by adding 50 base pairs to the reported length of the 3' UTR (Tuller et al., 2009) (Figure 1).

In addition to these selected, potential high-capacity terminators, several commonly used terminators were cloned for comparison. These included the *ADHI* terminator, the *TPS1* terminator (Yamanishi et al., 2011), and the *CYCI* terminator. Two versions of the *CYCI* terminator were compared in this study, one cloned from the genome as described in Figure 1, and one available in the parent plasmid (Mumberg et al., 1995), referred to as the *CYCI** terminator. This terminator was originally cloned from *S. cerevisiae* strain D311-3A and varies in length and contains several point mutations compared to the one from the s288C genome (see Supplementary Table 1 for sequences). The length of the *ADHI* and *TPS1* terminators was the same as previously described (Chien et al., 1991; Yamanishi et al., 2011). Finally, a construct lacking a terminator was included as a comparison of effectiveness for terminators.

All 34 of these constructs were cloned into the plasmid p413-*TEFmut3-yECitrine*, which contains the yellow fluorescent protein gene *yECitrine* under control of the moderate strength *TEFmut3* promoter (Alper et al., 2005; Nevoigt et al., 2006). Terminator function was evaluated by measuring the relative fluorescence output from these expression cassettes in *S. cerevisiae* strain BY4741. The construct demonstrating the highest reported fluorescence was over 3-fold greater than the cassette using the commonly used *CYCI** terminator (Figure 2). Moreover, terminator capacity spanned a dynamic range of over 13-fold when comparing between the most effective terminator and the construct lacking a terminator. These results underscore the importance of terminator selection and provided characterization of several potential high-capacity terminators.

3.2 Terminators influence net mRNA abundance through modulating mRNA half-life

Next, to determine the principle through which these terminators yielded varied protein expression, a subset of the 34 terminator constructs including the four yielding highest fluorescence, two from strains with medium fluorescence and both variants of the *CYC1* terminators were chosen for further study. These terminators included the *PRM9*, *CPS1*, *HIS5*, *SPG5*, *SPO1*, *VPS13*, *CYC1** and *CYC1* terminators. In addition, the construct with no terminator was included. First, total mRNA abundance was determined for these strains, and was found to correlate to protein expression (Figure 3). The non-linear correlation at higher expression levels is likely due to translation limitations or signal saturation at high levels of yellow fluorescent protein. Nonetheless, this data indicates that the terminator is primarily changing protein expression through changes impacting the transcript level rather than the translation rate. Furthermore, the highest capacity terminator construct had almost 6.5 times more transcript than that with the *CYC1** terminator.

Next, to determine whether this transcript level was caused by increased transcription rate or decreased degradation, the mRNA half-life for each of these transcripts was determined. To do so, the *TEFmut3* promoter was replaced with the galactose inducible and glucose repressed *GAL1* promoter. Since the *GAL1* promoter shows no detectable fluorescence output in the presence of glucose, transcription can be turned off by switching the media from galactose to glucose, thus allowing for the measurement of mRNA degradation over time. Similar to the transcript level, the mRNA half-life correlated non-linearly with fluorescence (Figure 4a). Again, the non-linear correlation at higher half-life values is likely due to translation limitations or signal saturation at high levels of yellow fluorescent protein. However, the mRNA half-life and the mRNA abundance showed a direct linear correlation implicating the selection of terminator as a sole determinant of mRNA half-life and total transcript level (Figure 4b). Furthermore, the highest capacity terminator construct had over 2.5 times higher mRNA half-life compared to transcripts using the *CYC1** terminator.

3.3 The influence of the terminator is most pronounced for low-expression promoters

Very little is known about the interaction between terminators and promoters in heterologous expression cassettes. In a previous study on terminators, Yamanishi et al. (2013) primarily used the *TDH3* (also known as *GPD*) promoter and tested a subset of terminators using the *ACT1* promoter (another relatively strong promoter), which showed a similar trend to that of *TDH3*. However, these promoters are quite high in expression strength and thus do not represent the dynamic range seen across all possible yeast promoters. Here, the set of eight terminators and a no-terminator control described above were combined with a collection of seven different promoters, spanning two orders of magnitude in expression level—more representative of the typical range used in metabolic engineering applications (Figure 5a). Four of the promoters in this set corresponded to the natural promoter paired with a terminator from the set. Terminators were paired with both their native and non-native promoters to test potential transcriptional coupling as well as study the influence of promoter-terminator pairs. These tests once again used the *yECitrine* fluorescent protein as an output for terminator capacity (Figure 5b). These results illustrate that protein expression varied most significantly with terminator choice for the lower strength promoters. Specifically, the most dramatic differences were seen with the *CPS1* promoter, which had a difference in expression of 11-fold when paired with the *VPS13* terminator instead of the *CYC1** terminator and over 35-fold difference when compared with a no-terminator baseline. While the fold difference decreased as a function of promoter strength, significant differences in protein output can be seen even for strong promoters such as *GAL1*. Similar non-linear behavior has been observed for other expression cassette elements including promoters and plasmid copy numbers (Hajimorad et al., 2011; Zucca et al., 2012). Specifically, the system is more responsive to changes from the terminator at low

promoter expression levels. Furthermore, it is likely that the influence of the terminator is muted for higher strength promoters due to translational limitations at such high transcript levels and saturation of the fluorescent signal in the cell. It is possible that some of the stronger promoters have reached a regime that is translationally limited and in this case, there is simply more mRNA than can be translated. Despite this overall trend, there were some exceptions that indicate that specific promoter-terminator interactions may indeed be occurring, although further work is needed to understand the specific mechanisms involved. In general, there did not appear to be any substantial interaction observed between native promoter-terminator pairs.

3.4 Specific terminator sequence can influence performance

In this work, terminator length was determined as based on previously reported 3' UTR length with an additional 50 base pairs added to the 3' side (see Figure 1). The influence in exact terminator sequence and length can be seen by the various versions of the *CYC1* terminators used in this study. The *CYC1** terminator present in the parent plasmid (Mumberg et al., 1995) begins 32 bp after the stop codon of the *CYC1* gene, stops 100 bp after the end of the 3' UTR, and contains several point mutations in comparison to the s288C genomic terminator region (see Supplementary Table 1 for sequences). In addition, due to the cloning scheme used in this work, the *CYC1* terminator construct contained only one restriction enzyme site between the gene and terminator whereas the *CYC1** terminator contained multiple sites since it was left in the parent plasmid as originally constructed. To determine whether the length, mutations, or restriction sites were responsible for the difference in expression level, two additional *CYC1* terminator variants were made. The *CYC1a* terminator has the same sequence as the *CYC1** terminator, but was re-cloned to remove the excess restriction site. The *CYC1b* terminator used the same cloning strategy as the *CYC1a* terminator with the mutations observed in the *CYC1** terminator reverted to the s288C genomic sequence (Figure 6a). Unexpectedly, both of these constructs demonstrated an additional increase in fluorescence when compared to the s288C *CYC1* terminator (Figure 6b). This result demonstrates that the difference in length (and therefore sequence) between the *CYC1* terminator and the other three variants influenced terminator performance (p-value < 0.05 in each case). Therefore the specific region that is defined and cloned as the “terminator” is critical to determining the final output from a heterologous expression cassette.

3.5 Terminator selection can be used to tune metabolic pathway flux

As a proof-of-concept for the utility of high-capacity terminators in metabolic engineering applications, terminator selection was used to influence pathway flux and performance in a heterologous expression cassette. Often, promoters have been used to alter the expression of heterologous enzymes; here terminator and promoter pairings are used to accomplish the same goal. Previously, we have reported on a mutant xylose isomerase gene that provided significant improvements to the xylose catabolic capacity of *S. cerevisiae*. Specifically, the identified mutant, *xyIA*3*, was shown to increase aerobic growth rate by 61 fold and ethanol production and xylose uptake by 8 fold when expressed in a *gre3* background with *XKS1* and *TAL1* overexpressed (Lee et al., 2012). This feat was accomplished by strong overexpression of the xylose isomerase mutant through a high-strength promoter, *TDH3*. Here, the *TEFmut3* promoter, which is considerably weaker than *TDH3*, was used instead. When paired with the *TEFmut3* promoter, the cassettes expressing the high-capacity terminators of *CPS1* and *PRM9* showed a significant increase in growth rate compared to the *CYC1** terminator (p-value=0.0014 and 0.0037, respectively) (Figure 7a). Furthermore, the highest growth rates achieved using the *TEFmut3* promoter were equal to that from the *TDH3* promoter when it was paired to the *CYC1** terminator. When using the strong *TDH3* promoter, the *CPS1* high-capacity terminator was able to further increase pathway flux and

enabled the highest growth rate we have achieved with this xylose isomerase mutant of 0.094 hr^{-1} . Quantitative PCR confirmed that the different terminator constructs resulted in changed transcript levels and were therefore responsible for the change in growth rate (Figure 7b). This result demonstrates the power of using terminators to modulate protein expression and thus pathway output. In this case, a much lower strength promoter, when paired with a high-capacity terminator, can achieve the same phenotypic results as a high-strength promoter with standard terminators.

4. Discussion

In this study, over 30 potential “high-capacity” terminators from *S. cerevisiae* were characterized for use in a metabolic engineering and synthetic biology context. It was demonstrated that high-capacity terminators act by increasing mRNA half-life and thus increase net transcript levels in cells. This effect can be used in a metabolic engineering context to boost expression from given promoters. In doing so, similar phenotypes can be achieved with much lower promoter strengths, and hence lower transcriptional loads to the cell. This study is the first to demonstrate the importance of terminator selection for metabolic engineering applications in yeast.

In this study, candidate high-capacity terminators were identified using aggregated data from reported half-life experiments. From these six studies (Table 1), the one best able to identify high-capacity terminators was the study by Miller et al. (2011). In fact, three of the top five high-capacity terminators came from this study. The other two terminators selected from this set were still effective for tuning protein expression above that of the *CYC1** terminator. As a result, additional candidates from this complete dataset may prove successful in identifying additional high-capacity terminators for yeast. Given the results that mRNA half-life can be directly influenced by terminator selection, it is evident that certain terminators are intrinsically high-capacity. However, there is no correlation between the reported mRNA half-life of the native gene and the output from the synthetic heterologous cassette containing the terminator. As a result, a crude estimation of terminator strength can be gained from genome-wide mRNA half-life studies, but actual strength and capacity as a high-capacity terminator will likely depend on the gene and surrounding DNA context.

The data and results in this study vary from those reported for similar terminators in another recent study (Yamanishi et al., 2013). In that study, a green fluorescent protein was expressed by the *TDH3* promoter in coordination with different terminators. The results presented here are compared to the results in the Yamanishi et al. (2013) study for the same terminators in Figure 8. There are several possible reasons for the difference in expression level, with the primary difference likely due to differences in the terminator length and sequence. The comparison of different *CYC1* terminator variants described above demonstrates that even subtle differences in terminator sequence and length can yield surprising differences in performance and protein output (Figure 6). In the work by Yamanishi et al. (2013), all terminators were cloned to be approximately 500 bp in length whereas in this study terminator length was defined as 50 base pairs beyond the 3' UTR cut site. Once again, this demonstrates that the exact sequence of the terminator can influence function and this finding must be considered when designing optimal expression cassettes. It should be noted that the length itself may not be the only important factor in determining the output and net protein expression. These two studies utilized different fluorescent protein genes and different promoters. It is certainly possible to expect some interplay between the gene and terminator as specific *cis* elements embedded in the gene sequence could enhance or minimize a particular terminator's effectiveness. Similarly, there is some limited evidence of interactions between specific promoters and terminators (Figure 5b) and it has been

previously reported that promoter-terminator gene loops can form (O'Sullivan et al., 2004). However, in general terminator capacity was relatively independent of promoters (Figure 5b) and genes (using the *yECitrine* gene and xylose isomerase gene, Figures 2 and 7). As such, while it is possible that these variables contribute to some of the variation in terminator capacity, they are not likely the dominant causes. Furthermore, there was no apparent correlation between the relative protein output of the terminator constructs and the length, GC content, or calculated free energy of the mRNA species. However, the length of the poly-A tail for each mRNA species in this study was unknown. Further study is needed to determine if the mechanism of increased mRNA half-life is the result of the level of polyadenylation.

One important premise elucidated by this work is that DNA context can be extremely important when characterizing genetic “parts” for metabolic engineering and synthetic biology purposes (Lanza et al., 2012; Young and Alper, 2010). Previous work by our group has demonstrated the importance of restriction sites and 5' UTR sequence on the performance of a heterologous expression cassette (Crook et al., 2011). As a result, it is not unexpected to see a similar effect by the 3' UTR. Furthermore, based on this work, potential interactions between the promoter and terminator must be taken into account as well—demonstrating that components of an expression cassette are not independently acting. The consequences of these studies as a whole underscores the fact the formal characterization of genetic “parts” must take into account the specific sequence and genetic context. The results here demonstrate that promoter characterization can be strongly influenced by terminator choice. Specifically, the results in Figure 5b highlight that relative promoter strengths and rankings would change depending on the terminator selected. Thus, these collective results explain why many studies characterizing promoters in yeast do not agree on the relative strength of the same set of promoters (Da Silva and Srikrishnan, 2012)—namely, the surrounding cassette, including restriction sites, reporter genes, and terminators, vary. Indeed, a parts-to-function synthetic biology design process is difficult to achieve due to these interactions.

In this case study, the importance of terminators in heterologous expression cassettes in yeast is clearly demonstrated. All of the terminators selected as potential high-capacity terminators had fluorescence expression levels at least 4-fold higher than the cassette with no terminator and over 2-fold higher than the native *CYC1* terminator and are therefore capable of supporting higher capacity expression than the majority of terminators included in current state-of-the-art expression cassettes. Furthermore, the potential to modulate gene expression was demonstrated to a greater extent when comparing transcript level between constructs, which varied almost 6.5-fold between the construct with the *PRM9* terminator and that with the *CYC1** terminator when coupled with the *TEFmut3* promoter, as compared to the 3.5 fold difference in fluorescence level. Even more compelling was the difference in fluorescence level between the strains using the *CPS1* promoter, which showed an 11-fold difference when paired with the *VPS13* terminator compared to the *CYC1** terminator and over 35-fold difference when compared with a no-terminator baseline. Furthermore, this study demonstrated that terminator selection can increase the flux of a heterologous metabolic pathway using xylose catabolism as a case study. A combination of a lower strength promoter with a high-capacity terminator was able to achieve significant increase in growth rate. Moreover, the achieved growth rate was equivalent to using a strong promoter with the parental terminator. In this strain, improved flux and protein expression is being achieved using a lower transcriptional rate and instead increasing mRNA half-life. The result is a decreased transcriptional energy expended per protein. When considering large, complex heterologous pathways, it is expected that the impact will be even more profound. Thus, this work has immediate implications for nearly all metabolic engineering projects where more efficient transcription would result in higher yields. Moreover, while *S.*

cerevisiae is an important host for metabolic engineers, it is likely that this work and approach would be applicable to other eukaryotic hosts as well.

5. Conclusion

This study demonstrated the importance of carefully characterizing and selecting terminators for heterologous expression cassettes in *S. cerevisiae*. We identified several “high-capacity” terminators that are quite potent for metabolic engineering applications. The major cause of the difference in output seen between terminators was due to changes in the mRNA half-life, likely due to the 3' UTR and polyadenylation resulting from each terminator. Furthermore, we demonstrated that the relative difference in output between terminators is magnified when coupled with a low expression promoter, with a maximum difference of 11-fold between a high-capacity terminator and the parent plasmid terminator and upwards of 35-fold compared to a construct lacking a terminator. Finally, we demonstrated a metabolic engineering application whereby terminator selection can increase flux in a heterologous metabolic pathway. In conclusion, terminators are an important synthetic part and should be a key consideration for metabolic engineering projects in the future.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by a NSF Graduate Research Fellowship to K. Curran and by the National Institutes of Health (grant number R01GM090221). The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

References

- Alper H, Fischer C, Nevoigt E, Stephanopoulos G. Tuning genetic control through promoter engineering. *Proc Natl Acad Sci U S A*. 2005; 102:12678–12683. [PubMed: 16123130]
- Blazeck J, Alper HS. Promoter engineering: Recent advances in controlling transcription at the most fundamental level. *Biotechnol J*. 2013; 8:46–58. [PubMed: 22890821]
- Cambray G, Guimaraes JC, Mutalik VK, Lam C, Mai QA, Thimmaiah T, Carothers JM, Arkin AP, Endy D. Measurement and modeling of intrinsic transcription terminators. *Nucleic Acids Res*. 2013; 41:5139–5148. [PubMed: 23511967]
- Carroll KL, Ghirlando R, Ames JM, Corden JL. Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. *RNA*. 2007; 13:361–373. [PubMed: 17237360]
- Chang DT-H, Huang C-Y, Wu C-Y, Wu W-S. YPA: an integrated repository of promoter features in *Saccharomyces cerevisiae*. *Nucleic Acids Res*. 2011; 39:D647–D652. [PubMed: 21045055]
- Chien CT, Bartel PL, Sternglanz R, Fields S. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc Natl Acad Sci U S A*. 1991; 88:9578–9582. [PubMed: 1946372]
- Coller J. Methods to determine mRNA half-life in *Saccharomyces cerevisiae*. In: Maquat, LE.; Kiledjian, M., editors. *RNA Turnover in Eukaryotes: Nucleases, Pathways and Analysis of mRNA Decay*. Vol. 448. Elsevier Academic Press Inc; San Diego: 2008. p. 267-284.
- Crook NC, Freeman ES, Alper HS. Re-engineering multicloning sites for function and convenience. *Nucleic Acids Res*. 2011; 39:92.
- Curran KA, Alper HS. Expanding the chemical palate of cells by combining systems biology and metabolic engineering. *Metab Eng*. 2012; 14:289–297. [PubMed: 22595280]
- Da Silva NA, Srikrishnan S. Introduction and expression of genes for metabolic engineering applications in *Saccharomyces cerevisiae*. *FEMS Yeast Res*. 2012; 12:197–214. [PubMed: 22129153]

- Flagfeldt DB, Siewers V, Huang L, Nielsen J. Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*. *Yeast*. 2009; 26:545–51. [PubMed: 19681174]
- Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, Arkin AP, Astromoff A, El-Bakkoury M, Bangham R, Benito R, Brachat S, Campanaro S, Curtiss M, Davis K, Deutschbauer A, Entian KD, Flaherty P, Foury F, Garfinkel DJ, Gerstein M, Gotte D, Guldener U, Hegemann JH, Hempel S, Herman Z, Jaramillo DF, Kelly DE, Kelly SL, Kotter P, LaBonte D, Lamb DC, Lan N, Liang H, Liao H, Liu L, Luo C, Lussier M, Mao R, Menard P, Ooi SL, Revuelta JL, Roberts CJ, Rose M, Ross-Macdonald P, Scherens B, Schimmack G, Shafer B, Shoemaker DD, Sookhai-Mahadeo S, Storms RK, Strathern JN, Valle G, Voet M, Volckaert G, Wang CY, Ward TR, Wilhelmy J, Winzeler EA, Yang Y, Yen G, Youngman E, Yu K, Bussey H, Boeke JD, Snyder M, Philippsen P, Davis RW, Johnston M. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature*. 2002; 418:387–91. [PubMed: 12140549]
- Graber JH, McAllister GD, Smith TF. Probabilistic prediction of *Saccharomyces cerevisiae* mRNA 3'-processing sites. *Nucleic Acids Res*. 2002; 30:1851–1858. [PubMed: 11937640]
- Grigull J, Mnaimneh S, Pootoolal J, Robinson MD, Hughes TR. Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors. *Mol Cell Biol*. 2004; 24:5534–47. [PubMed: 15169913]
- Guo ZJ, Sherman F. Signals sufficient for 3'-end formation of yeast mRNA. *Mol Cell Biol*. 1996; 16:2772–2776. [PubMed: 8649385]
- Hajimorad M, Gray PR, Keasling JD. A framework and model system to investigate linear system behavior in *Escherichia coli*. *J Bio Eng*. 2011; 5:3. [PubMed: 21510907]
- Hegemann JH, Heick SB. Delete and Repeat: A Comprehensive Toolkit for Sequential Gene Knockout in the Budding Yeast *Saccharomyces cerevisiae*. *Strain Engineering: Methods and Protocols*. 2011; 765:189–206.
- Holstege FCP, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell*. 1998; 95:717–728. [PubMed: 9845373]
- Karim AS, Curran KA, Alper HS. Characterization of plasmid burden and copy number in *Saccharomyces cerevisiae* for optimization of metabolic engineering applications. *FEMS Yeast Res*. 2013; 13:107–116. [PubMed: 23107142]
- Lanza AM, Crook NC, Alper HS. Innovation at the intersection of synthetic and systems biology. *Curr Opin Biotechnol*. 2012; 23:712–717. [PubMed: 22265125]
- Lee SM, Jellison T, Alper HS. Directed evolution of xylose isomerase for improved xylose catabolism and fermentation in the yeast *Saccharomyces cerevisiae*. *Appl Environ Microbiol*. 2012; 78:5708–16. [PubMed: 22685138]
- Liu L, Redden H, Alper H. Frontiers of yeast metabolic engineering: diversifying beyond ethanol and *Saccharomyces*. *Curr Opin Biotechnol*. 2013; 24:1016/j.copbio.2013.03.005
- Miller C, Schwalb B, Maier K, Schulz D, Dumcke S, Zacher B, Mayer A, Sydow J, Marcinowski L, Dolken L, Martin DE, Tresch A, Cramer P. Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. *Mol Syst Biol*. 2011; 7:458. [PubMed: 21206491]
- Mischo HE, Proudfoot NJ. Disengaging polymerase: Terminating RNA polymerase II transcription in budding yeast. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 2013; 1829:174–185.
- Montgomery DL, Hall BD, Gillam S, Smith M. Identification and isolation of the yeast cytochrome c gene. *Cell*. 1978; 14:673–680. [PubMed: 210956]
- Mumberg D, Muller R, Funk M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene*. 1995; 156:119–122. [PubMed: 7737504]
- Nevoigt E, Kohnke J, Fischer CR, Alper H, Stahl U, Stephanopoulos G. Engineering of promoter replacement cassettes for fine-tuning of gene expression in *Saccharomyces cerevisiae*. *Appl Environ Microbiol*. 2006; 72:5266–5273. [PubMed: 16885275]
- O'Sullivan JM, Tan-Wong SM, Morillon A, Lee B, Coles J, Mellor J, Proudfoot NJ. Gene loops juxtapose promoters and terminators in yeast. *Nat Genet*. 2004; 36:1014–8. [PubMed: 15314641]

- Pflegler BF, Pitera DJ, Smolke CD, Keasling JD. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nat Biotechnol.* 2006; 24:1027–32. [PubMed: 16845378]
- Rondon AG, Mischo HE, Kawauchi J, Proudfoot NJ. Fail-safe transcriptional termination for protein-coding genes in *S. cerevisiae*. *Mol Cell.* 2009; 36:88–98. [PubMed: 19818712]
- Sambrook, J. *Molecular Cloning: A Laboratory Manual*. Russell, DW., editor. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY: 2000.
- Shalem O, Dahan O, Levo M, Martinez MR, Furman I, Segal E, Pilpel Y. Transient transcriptional responses to stress are generated by opposing effects of mRNA production and degradation. *Mol Syst Biol.* 2008; 4:223. [PubMed: 18854817]
- Shao ZY, Zhao H, Zhao HM. DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res.* 2009; 37:10.
- Sheff MA, Thorn KS. Optimized cassettes for fluorescent protein tagging in *Saccharomyces cerevisiae*. *Yeast.* 2004; 21:661–670. [PubMed: 15197731]
- Tuller T, Ruppian E, Kupiec M. Properties of untranslated regions of the *S. cerevisiae* genome. *BMC Genomics.* 2009; 10:11. [PubMed: 19133128]
- Wang Y, Liu CL, Storey JD, Tibshirani RJ, Herschlag D, Brown PO. Precision and functional specificity in mRNA decay. *Proc Natl Acad Sci U S A.* 2002; 99:5860–5865. [PubMed: 11972065]
- West S, Proudfoot NJ. Transcriptional termination enhances protein expression in human cells. *Mol Cell.* 2009; 33:354–64. [PubMed: 19217409]
- Yamanishi M, Ito Y, Kintaka R, Imamura C, Katahira S, Ikeuchi A, Moriya H, Matsuyama T. A Genome-Wide Activity Assessment of Terminator Regions in *Saccharomyces cerevisiae* Provides a “Terminatome” Toolbox. *ACS Synth Biol.* 2013; 10.1021/sb300116y
- Yamanishi M, Katahira S, Matsuyama T. TPS1 Terminator Increases mRNA and Protein Yield in a *Saccharomyces cerevisiae* Expression System. *Biosci Biotechnol Biochem.* 2011; 75:2234–2236. [PubMed: 22056446]
- Young E, Alper H. *Synthetic Biology: Tools to Design, Build, and Optimize Cellular Processes*. J Biomed Biotechnol. 2010; 2010:130781. [PubMed: 20150964]
- Zucca S, Pasotti L, Mazzini G, De Angelis MG, Magni P. Characterization of an inducible promoter in different DNA copy number conditions. *BMC Bioinformatics.* 2012; 13(Suppl 4):S11. [PubMed: 22536957]

Highlights

- Characterized over 30 high-capacity terminators for improved protein expression.
- High-capacity terminators modulated protein levels up to 35-fold.
- Expression modulation by terminators is due to changes in mRNA half-life.
- Demonstrated first use of high-capacity terminators in a metabolic pathway.

Figure 1. Diagram of terminator cloning scheme

Each terminator was designed and cloned from genomic DNA from the end of the open reading frame (ORF) to 50 basepairs past the end of the 3' UTR.

Figure 2. Fluorescence output of yeast strains expressing yECitrine under the control of the TEFmut3 promoter with varied terminators

All strains are relative to the *CYC1** terminator, the standard terminator available in the plasmids described in Mumberg et al. (1995). Terminator selection is seen to influence the output of the expression cassette. Error bars represent standard deviation from three biological replicates.

Figure 3. mRNA level measurements of select constructs

Transcript level was measured for nine constructs representing the range of fluorescence values seen in Figure 2 and plotted against those values. All values are relative to the *CYC1** terminator. There is a good correspondence between mRNA and protein level indicating a transcriptionally-controlled process. Error bars for transcript level represent standard deviation for three technical replicates.

Figure 4a

Figure 4b

Figure 4. Half-life analysis of terminator constructs

mRNA half-life analysis was conducted for the subset of terminators used in Figure 3. **A.** mRNA half-life versus fluorescence values are plotted. Half-life values were generated using the *GAL1* promoter. Fluorescence values are the same as reported in Figure 2 using the *TEFmut3* promoter. There is a non-linear relationship between these two variables. **B.** mRNA half-life versus transcript abundance are plotted. Transcript abundance values are the same as reported in Figure 3 using the *TEFmut3* promoter. A direct linear correlation exists between the mRNA abundance and half-life. Error bars for fluorescence values represent standard deviation from three biological replicates. For transcript level, error bars represent

the standard deviation for three technical replicates. For mRNA half-life, error bars represent standard error calculated during linear regression to determine half-life value with the exception of the two lowest points, which represent standard deviation of two separate measurements due to the difficulty in measuring extremely short half-lives.

Figure 5a

Figure 5b

Figure 5. Influence of promoter and terminator combinations on expression cassettes
 Fluorescence from *yECitrine* expression with varied terminator and promoter cassettes is plotted with error bars that represent standard deviation from three biological replicates. Strains expressing the *GAL1* promoter were grown in minimal media containing 20 g/L galactose. All other strains were grown in comparable media with 20 g/L glucose. **A.** Relative promoter expression levels using the *CYC1** terminator. The selected promoters vary in expression by two orders of magnitude. **B.** Relative terminator expression levels with varied promoters. Each promoter data set is relative to its own *CYC1** terminator value. The influence of terminator selection is most profound for low expression promoters.

Figure 6a

Figure 6b

Figure 6. Comparison of *CYC1* terminator variants

A. Schematic of the different *CYC1* terminator variants. The *CYC1* terminator was made consistent to the scheme in Figure 1. The *CYC1** terminator is the terminator present in the parent plasmid. The short 3' UTR is missing the first 32 bp after the end of the *CYC1* gene. Blue bars represent mutations relative to the s288C genomic sequence. The *CYC1a*

terminator has the same sequence as the *CYC1** terminator, but was re-cloned into the plasmid in order to remove the excess restriction site between the gene and the terminator. The *CYC1b* terminator is the same as the *CYC1a* terminator with the addition of having the mutations observed in the *CYC1** terminator reverted to the s288C genomic sequence. See Supplementary Table 1 for sequences. **B.** Relative fluorescence of *yECitrine* using the *TEFmut3* promoter and *CYC1* terminator variants. Error bars represent standard deviation from biological triplicates. The *CYC1* terminator is significantly different from the other three variants (p-value < 0.05). Terminator length and sequence are seen to influence expression output.

Figure 7a

Figure 7b

Figure 7. Tuning a pathway flux through terminator selection

A. Growth rate in xylose media of strains expressing *xylA*3* using the *TEFmut3* or *TDH3* promoter and varied terminators. For *TEFmut3* values, error bars represent standard deviation from two separate experiments including three biological replicates each. For *TDH3* values, error bars represent standard deviation from three biological replicates. No growth was observed for the strain with the *TEFmut3* promoter and no terminator. **B.** Transcript level of the *xylA*3* gene during exponential phase growth relative to the level from the strain with the *TEFmut3* promoter and the *CYC1** terminator. Error bars represent standard deviation for three technical replicates.

Figure 8. Comparison of terminator activity across two studies

Expression output from terminators studied in this work is contrasted to those in Yamanishi et al. (2013). Results from this work are the values from Figure 2, using the *yECitrine* gene and the *TEFmut3* promoter. For the data from Yamanishi et al. (2013), fluorescence values were determined using a green fluorescent protein gene and the *TDH3* promoter. While there is a trend, these results highlight the importance of genetic context, sequence, and terminator length on function.

Table 1
List of genes and source from which terminators were taken

For each source the five terminators with the highest reported mRNA half-life for the corresponding gene were used. The exception is for Graber et al. (2002), for which the five highest reported predicted terminators were used.

Holstege et al. (1998)	Wang et al. (2002)	Shalem et al. (2008)	Miller et al. (2011)	Grigull et al. (2004)	Graber et al. (2002)
<i>GYP7</i>	<i>PRM5</i>	<i>GAT2</i>	<i>HSP26</i>	<i>FUM1</i>	<i>MRP4</i>
<i>TIP1</i>	<i>AIP1</i>	<i>GSY2</i>	<i>CPS1</i>	<i>GND1</i>	<i>YJR085C</i>
<i>YGR127W</i>	<i>ECM10</i>	<i>CYC7</i>	<i>PRM9</i>	<i>SED1</i>	<i>HIS5</i>
<i>SPG5</i>	<i>SPO1</i>	<i>UBX6</i>	<i>HUG1</i>	<i>VPS13</i>	<i>PSY4</i>
<i>GRE3</i>	<i>PDC6</i>	<i>HSP26</i>	<i>IDP1</i>	<i>YOL036W</i>	<i>LSC2</i>