Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Advances and current limitations in transcript-level control of gene expression

John M Leavitt[1] and Hal S Alper[1,2]

Gene expression control is critical to increase production of recombinant proteins, fine-tune metabolic pathways and reliably express synthetic pathways. The importance of transcriptional control seems to be most important in eukaryotic systems. In this review, we highlight recent developments in the field of transcriptional engineering with an emphasis on the opportunities and challenges. We discuss the engineering of 'parts' that influence transcriptional throughput including promoters, terminators, and transcription factors as well as the genetic context of the expression cassette. While great strides have been made in the area, the robustness of these parts has been largely untested. This review highlights the importance of considering robustness in biological systems and the limitations that current synthetic parts possess.

Addresses

[1] Institute for Cellular and Molecular Biology, The University of Texas at Austin, 2500 Speedway Avenue, Austin, TX 78712, United States

[2] McKetta Department of Chemical Engineering, The University of Texas at Austin, 200 E Dean Keeton St. Stop C0400, Austin, TX 78712, United States

Corresponding author: Alper, Hal S (halper@che.utexas.edu)

Current Opinion in Biotechnology 2015, **34**:98–104

This review comes from a themed issue on **Systems biology**

Edited by **Sarah Maria Fendt** and **Costas D Maranas**

http://dx.doi.org/10.1016/j.copbio.2014.12.015

Introduction

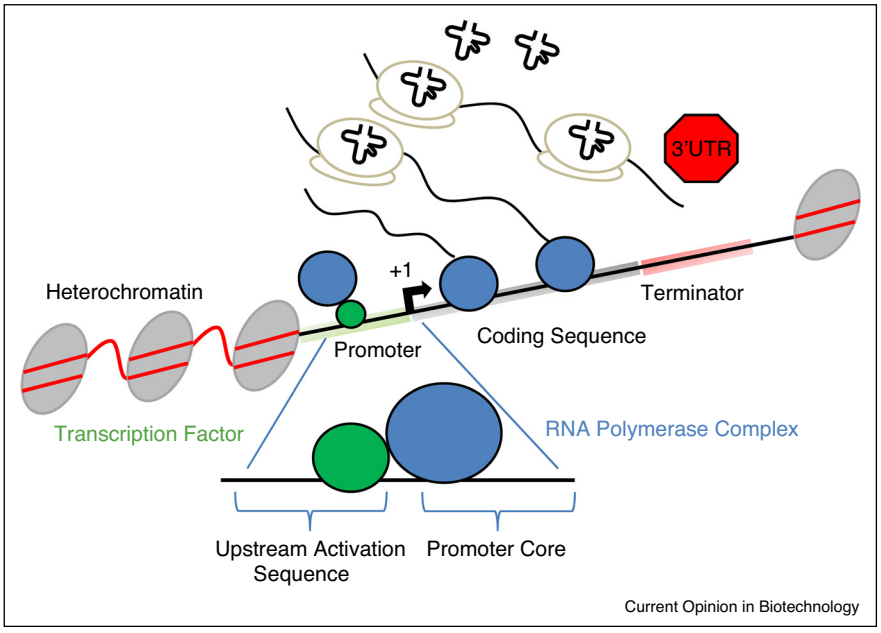
Controlling gene expression is a paramount, and often foremost, goal of most biological endeavors — from therapeutic antibody production [1] to the production of industrial enzymes [2] to the expression of heterologous metabolic pathways [3,4]. While most of these efforts initially focus on the need for high expression, further work (especially in optimizing these processes) requires a more sophisticated, tighter control of gene expression. The need for control at many levels obviates the necessity of libraries of synthetic parts capable of controlling transcript levels. However, not all parts are created equal and not all have been tested adequately enough to ensure function in a new system. Specifically, the current synthetic biology 'parts on a shelf' model seemingly necessitates interoperability and robustness of parts, yet relies on community sourced databases to assemble experimental tools [5,6]. This reality provides both opportunities for rapid advancement as well as a limitation in the field. In this review, we highlight the advances in synthetic parts for controlling transcript levels and address inherent challenges and considerations in more accurately defining their robustness.

Two major processes contribute to protein expression level: transcriptional rates and translational rates. Translation-level control (especially through tools such as ribosomal binding site calculators [7–9] and codon optimization) allow users to forward engineer the ribosomal efficiency for their gene of interest. This approach has been successfully demonstrated in prokaryotic systems where strong, orthogonal viral promoters and simpler translational mechanisms exist. In this context, translation-level control can span a 10^5-fold range [7] by editing a relatively small sequence space (such as the 5′UTR containing an RBS). Recent work on translational control in eukaryotes has focused on codon optimization to allow for improved protein expression, but the level of control of translation is not nearly as high. As an example, by optimizing the codon usage of the heterologous catechol 1,2-dioxygenase gene to be better expressed at stationary phase in *Saccharomyces cerevisiae*, a 2.9-fold increase in titer was achieved [10]. In contrast, for yeast and higher eukaryotes, tuning transcription rates through promoters imparts a higher level of control and can achieve between a 10^2-fold dynamic range [11•] and 10^4 range for orthogonal transcription factors [12••]. By comparison, the native range in transcript levels for yeast spans a roughly 10^3–10^4 dynamic range [13]. A similar range is also achievable in prokaryotic systems. The mRNA level of a transcript is controlled by many factors including the promoter, terminator, plasmid/expression cassette copy number, and the surrounding DNA context of the plasmid or genomic locus of integration (Figure 1). Given the success of transcription-level control (esp. in eukaryotic systems), we focus the rest of this review on this area by first considering the synthetic parts that lead to control and then addressing the issues of robustness.

Promoters

Promoters have one of the largest impacts on gene expression and were among the first synthetic parts to be studied and diversified via random mutagenesis [14]. These initial efforts were marked by a more robust definition of promoter strength taking into account

Figure 1



Factors influencing transcriptional control of genes. The above expression cassette is color coded to represent promoter (green), coding sequence (gray) and terminator regions (red). Heterochromatic regions are used as an example of potential complications provided by the genetic context in which the cassette is located. Ribosomes moving along the mRNA and movement are limited by factors such as tRNA availability. In the expanded region of the figure, a simplified promoter diagram is provided that depicts upstream activation sequence containing a bound transcription factor (green) along with RNA polymerase complex (blue) bound to the promoter core.

dilutions by growth, the promoter's ability to impact multiple proteins, measurement of mRNA levels, and utility in heterologous pathway expression. More recent efforts aim at creating more novel promoters (independent of a native scaffold) to increase the range of transcriptional capacity (Table 1). Developing synthetic promoters for eukaryotic systems will increase the number and diversity of promoter parts available in these systems. Comparable efforts to increase part diversity in *Escherichia coli* have been undertaken using genome mining [15] and screening of promoter and translation initiation libraries [16].

The galactose inducible promoter (GAL) is the strongest yeast inducible promoter; however it suffers from complete repression by glucose. Liang and coworkers developed a novel gene switch that coupled the inductive strength of the GAL promoter with the tight binding affinity of estradiol for the estrogen receptor protein. This ultimately led to a series of parts capable of inducing a multistep pathway using 10 nM estradiol in the presence of glucose and resulting in a 50-fold improvement in zeaxanthin production over previous efforts using constitutive promoters [17•].

The strongest yeast promoters have been constructed through a hybrid approach by coupling upstream activating sequences (UAS) with a core promoter. Adjusting the composition of the UAS elements enables upwards of 50–300-fold dynamic range in expression strength, reaching the highest reported strength of a promoter in *S. cerevisiae* [11•,18]. Among these, the strongest constitutive hybrid promoter exhibits a 2.5-fold improvement over the TDH3 promoter with respect to mRNA levels making this promoter as strong as the GAL promoter [11•]. However, it should be noted that there is no real statistically significant improvement in fluorescent protein production, illustrating the limitations of relying on reporter proteins alone without more robust, comprehensive measurements like mRNA levels. Improved core promoters could lead to even greater transcriptional control in these systems. Core promoters were investigated in the yeast *Pichia pastoris* and synthetic core promoters were designed using common sequence motifs and transcription factor binding sites. These synthetic core promoters were combined with the methanol inducible promoter pAOX1 to generate diverse activity between 10% and 117% of the wild-type promoter, however only fluorescent protein expression was reported [19].

The field is also quickly moving to de novo synthetic promoters that lack homology to anything else in the genome. In the case of Chinese Hamster Ovary (CHO) cell lines, Brown and coworkers used an enrichment