Engineering transcription-based logic

Reshma P. Shetty RSHETTY@MIT.EDU
Thomas F. Knight, Jr. TK@CSAIL.MIT.EDU
MIT Computer Science and Artificial Intelligence Laboratory, 32 Vassar Street, Cambridge MA, 02139 USA

1. Introduction

Synthetic biology (http://www.syntheticbiology.org) is an emerging engineering discipline concerned with the design, fabrication and analysis of systems built from biological parts. Similar to the way electrical engineering takes advantage of the science of physics and chemical engineering takes advantage of chemistry to develop useful engineered systems, synthetic biology seeks to make use of biology. The potential application space for synthetic biology is enormous spanning areas such as chemical energy, materials and information. However, for the construction of synthetic biological systems to be routine, biology must be developed as a technology. Current systems are severely limited by the lack of many, well-characterized biological parts and devices.

The focus of this work is on engineering devices capable of implementing digital logic in cells. In these devices, information is encoded as transcription rates (or the rate at which DNA is transcribed to RNA). Thus, these transcription-based logic devices are composed of proteins that bind to DNA (termed a transcription factor or repressor) thereby regulating the transcription rate of that DNA (see figure 1). In this work, I describe a model that defines device behavior in terms of biochemical parameters like binding affinities and synthesis/degradation rates. Analysis of the model permits identification of which biochemical parameters have the greatest influence on device behavior. I also find target values for the key biochemical parameters. Determination of optimal parameter values informs the design of novel transcription-based logic devices.

2. Previous work

Current transcription-based logic devices are usually composed of bacterial repressor proteins (Elowitz & Leibler, 2000; Gardner et al., 2000). Since there are a limited number of these naturally occurring bacterial repressors, the scale and complexity of the systems that can be assembled from these devices is limited. To address this limitation, I will engineer synthetic transcription factors from zinc finger DNA binding domains and leucine zipper dimerization domains. Such an implementation change has several advantages. First, it should enable the eventual construction of a library of transcription-logic devices since there are large numbers of both kinds of domains available. Second, it will add modularity to the design of these devices since the two functions of the transcription factor, DNA binding and dimerization, are separated. This separation should enable independent tuning of the two domains. Third, since leucine zippers domains are capable of both homo- and heterodimerizing, a wider range of functions can be implemented in transcription-based logic.

The construction of logic devices from zinc fingers and leucine zippers has been proposed previously (Batten et al., 2004). This work differs from that presented here in several key ways. First, the devices proposed by Batten *et al.* encode signals as protein concentrations rather than transcription rates. The advantage of using transcription rates as the signal carrier is that devices are composable: any device may be connected to any other device. Devices whose output is protein concentration can only be connected to devices which take the same protein as input. Second, Batten *et al.* ask the question, given typical biological parameter values, what kind of device performance is expected? In this work, I instead ask, given that I as the device engineer have some measure of control over device design, how should I design the device in order to achieve the best possible device performance? By approaching the model from a purely design perspective, I obtain somewhat different results. Based on the results of my model analysis and previous work on dimeric zinc finger proteins, I have designed a novel implementation of transcription-based logic devices from zinc fingers and leucine zippers (Pomerantz et al., 1998).

3. Models inform device design

I develop a model that describes the device output in terms of the device input and relevant biochemical parameters. To enable the comparison of different device designs and to evaluate the affect of varying parameter values on device performance, I quantify device performance using existing metrics developed for digital logic devices like noise

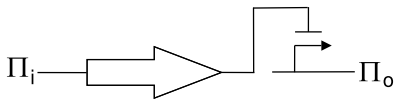

Figure 1. Schematic diagram of a transcription-based logic device: an inverter. The input signal Π_i causes the transcription of a gene encoding a repressor protein. The repressor protein dimerizes (two proteins bind to each other) and binds to its cognate promoter regulating the output transcription signal Π_o.

margins: the amount of noise a device can tolerate on its input signal without giving an erroneous output signal (Hill, 1968).

Several key observations which inform device design arise from the model analysis.

1. The parameter α_i, defined as the product of the ratio of mRNA and protein synthesis to their decay rates and copy number, determines the device input protein swing (the range of input protein concentration over which the device operates) as well as the device fan out (the maximum number of outputs the device can drive).

2. The protein-protein and protein-DNA binding affinities are the primary determinants of the shape of the device transfer curve. I obtain approximations for the values of these binding affinities that lead to good device behavior. Interestingly, it is not the absolute value of these parameters that determines device behavior (as most previous work suggests) but rather their value relative to α_i.

3. Inclusion of nonspecific DNA binding in the model leads to larger noise margins in the transfer curve.

4. An alternate device design in which several nonfunctional, high affinity protein binding sites are present yields a substantially improved transfer characteristic as measured by the noise margin.

4. Design of transcription-based logic devices

Using the model results as a guide, I selected previously characterized zinc fingers and leucine zippers to construct an initial set of synthetic transcription factors (Wolfe et al., 1999; Newman & Keating, 2003). I specify the DNA sequences encoding the DNA binding and dimerization domains so that each domain is a separate BioBricks part (see http://parts.mit.edu for more information). Designing the synthetic transcription factors in this way allows easy mixing and matching of DNA binding and dimerization domains via BioBricks standard assembly techniques. Previously, transcription factors were specified as a single BioBricks part and thus not very modular. Another contribution of this work is that many of the designed devices use heterodimerizing leucine zipper domains rather than the typical homodimerizing domains. Such devices will be capable of carrying out the logical NAND operation demonstrating that transcription-based logic is capable of implementing arbitrary logic operations.

5. Future Work

Having completed the DNA sequence specification of my transcription-based logic devices, I am in the process of fabricating the synthetic transcription factors (and cognate promoters) using standard molecular biology techniques. Analysis of device behavior should either yield working transcription-based logic devices or shed light on how to better engineer these devices in the future. Additionally, I am also developing methods for quantitatively characterizing device behavior. These methods will use both quantitative RNA measurements to characterize the device transfer curves and flow cytometry to assess variations in device performance between different cells. The measured transfer curves can be directly compared to the model results in order to validate the model. Moreover, insights from the model should aid in debugging issues in device function.

References

Batten, C., Krashinksky, R., & Knight Jr., T. (2004). A scalable cellular logic technology using zinc-finger proteins. *3rd Workshop on Non-Silicon Computing*.

Elowitz, M. B., & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature*, *403*, 335–8.

Gardner, T. S., Cantor, C. R., & Collins, J. J. (2000). Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, *403*, 339–42.

Hill, C. F. (1968). Noise margin and noise immunity in logic circuits. *Microelectronics*, *1*, 16–22.

Newman, J. R. S., & Keating, A. E. (2003). Comprehensive identification of human bZIP interactions with coiled-coil arrays. *Science*, *300*, 2097–101.

Pomerantz, J. L., Wolfe, S. A., & Pabo, C. O. (1998). Structure-based design of a dimeric zinc finger protein. *Biochemistry*, *37*, 965–70.

Wolfe, S. A., Greisman, H. A., Ramm, E. I., & Pabo, C. O. (1999). Analysis of zinc fingers optimized *via* phage display: evaluating the utility of a recognition code. *J Mol Biol*, *285*, 1917–34.