

Commentary

Open Access

## Models for synthetic biology

Yiannis N Kaznessis

Address: Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455, USA

Email: Yiannis N Kaznessis - yiannis@cems.umn.edu

Published: 6 November 2007

Received: 18 October 2007

*BMC Systems Biology* 2007, 1:47 doi:10.1186/1752-0509-1-47

Accepted: 6 November 2007

This article is available from: <http://www.biomedcentral.com/1752-0509/1/47>

© 2007 Kaznessis; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

Synthetic biological engineering is emerging from biology as a distinct discipline based on quantification. The technologies propelling synthetic biology are not new, nor is the concept of designing novel biological molecules. What is new is the emphasis on system behavior.

The objective is the design and construction of new biological devices and systems to deliver useful applications. Numerous synthetic gene circuits have been created in the past decade, including bistable switches, oscillators, and logic gates, and possible applications abound, including biofuels, detectors for biochemical and chemical weapons, disease diagnosis, and gene therapies.

More than fifty years after the discovery of the molecular structure of DNA, molecular biology is mature enough for real quantification that is useful for biological engineering applications, similar to the revolution in modeling in chemistry in the 1950s. With the excitement that synthetic biology is generating, the engineering and biological science communities appear remarkably willing to cross disciplinary boundaries toward a common goal.

Synthetic biological engineering is emerging from biology as a distinct discipline based on quantification [1-5]. The objective is the design and construction of new biological devices and systems to deliver useful applications. Numerous synthetic gene circuits have been created in the past decade, including bistable switches, oscillators, and logic gates [[1-5], and references therein], and possible applications abound, ranging from biofuels, to detectors for biochemical and chemical weapons, to disease diagnosis, to gene therapies.

Certainly, the technologies propelling synthetic biology are not new, nor is the concept of designing novel biological molecules [6,7]. What is perhaps new is the emphasis on system behavior, designing DNA sequences with synthetic phenotypes exhibiting prescribed dynamic responses.

Despite the initial successes of synthetic designs [1-5], the paradigm of biological sciences as descriptive disciplines may not rapidly assist in rationally engineering novel gene networks, despite the increasing volume of components that can be used in constructing synthetic networks. Genome projects identify the components of gene networks in biological organisms, gene after gene, and DNA microarray experiments discover the network connections. Yet, the static pictures of networks these experiments provide cannot adequately explain biomolecular phenomena or enable rational engineering of dynamic gene expression regulation. In other words, as an engineering discipline, synthetic biology cannot rely on endless trial and error methods driven by verbal description of biomolecular interaction networks.

The challenge facing the scientific and engineering communities is then to reduce the enormous volume and

complexity of biological data into concise theoretical formulations with predictive ability, ultimately associating synthetic DNA sequences to dynamic phenotypes. The paradigm is not new either: In the 1940s and 1950s chemistry was a well matured discipline for pioneers like Neil Amundson, Byron Bird and Rutherford Aris to develop mathematical models that captured the enormous complexity of chemical processes in a way useful for chemical engineering applications [8-10]. Quantitative models of chemical processes led to the establishment of the chemical engineering discipline and the emergence of a strong chemical/petroleum industry. Although arguments can be made about the detrimental role of this industry on the environment, there can be no doubt of the overall positive effects on human life.

But what types of models are appropriate for synthetic biology? Because of the large number of participating species and the complexity of their interactions, only detailed modeling can allow the investigation of dynamic gene expression in a way fit for analysis and design. Designs can be detailed at the molecular level with dynamic models of all the biomolecular interactions involved in transcription, translation, regulation, transport and induction. We contrast this to *a posteriori* modeling of synthetic networks. For example in their seminal 2000 paper [11], Gardner and co-workers developed a very elegant model that captures and explains the observed dynamic behavior of the bistable switch and provides additional insight in the biological mechanism. This formalism may abide well with Occam's razor, but cannot guide the choice of specific DNA sequences and their regulatory relations to achieve a bistable switch. More specifically, it will be challenging to use reduced models to choose, for example, between lactose, arabinose or tetracycline operators, or any one of dozens of their mutant variants, for building a new, different bistable switch.

In engineering, descriptive models that are succinct and lucid are appreciated, but the ones used will be at the level of design degrees of freedom. For example, Bernoulli's equation can explain the aerodynamic lift of an airplane, but modern aircraft design is based on simulations that include all the components of flight in detail. Turning to synthetic biology, model-driven rational engineering of synthetic gene networks is possible at two levels:

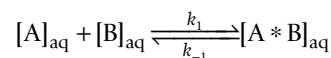
First, the level of network topologies, where biomolecules control the concentration of other biomolecules, e.g. DNA binding proteins regulate the expression of specific genes by either activation or repression. By combining simple regulatory interactions, such as negative and positive feedback and feed forward loops, one may create more intricate networks that precisely control the production of protein molecules, such as bistable switches, oscil-

lators, and filters. In the laboratory, these networks can be created using existing libraries of regulatory proteins and their corresponding operator sites. The now classical example is the aforementioned bistable switch Gardner and co-workers built [11]: they connected two regulatory proteins repressing one another and this resulted in a bistable switch they could control. Another is the repressilator of Elowitz and Leibler [12]: three regulatory proteins repressing one another in a sequential loop resulted in oscillating concentration profiles.

Secondly, the level of molecular components, which describes the kinetics and strengths of biomolecular interactions within the system. Indeed, the dynamical behavior of the system is a complex function of the kinetic interactions of the components. By altering the characteristics of the components, such as DNA-binding proteins and their corresponding DNA sites, one can modify the system's dynamical behavior without modifying the network topology. In the laboratory, the DNA sequences that yield the desired characteristics of each component can be engineered to achieve the desired protein-protein, protein-RNA, or protein-DNA binding constants and enzymatic activities. For example, Alon and co-workers [13] showed how simple mutations on the DNA sequence of the lactose operon can result in widely different phenotypic behaviors.

Ultimately, the large number of variants (interaction topologies and strengths) for these two types of design degrees of freedom requires sophisticated computational modeling, since the cost of experimentally changing these components and the kinetics of their interactions can quickly become prohibitive. Computer simulations enable exhaustive searches of different network connectivities and molecular thermodynamic/kinetic parameters, greatly advancing the development of design principles that seek to simplify the complicated behavior of the network into a brief, usable framework.

All gene expression molecular level events can be represented with reactions. For any two molecular species A and B (proteins, DNA, RNA, signaling molecules, etc.) interacting in solution to form a complex A\*B (e.g. a repressor protein and the corresponding DNA operator site) we can write

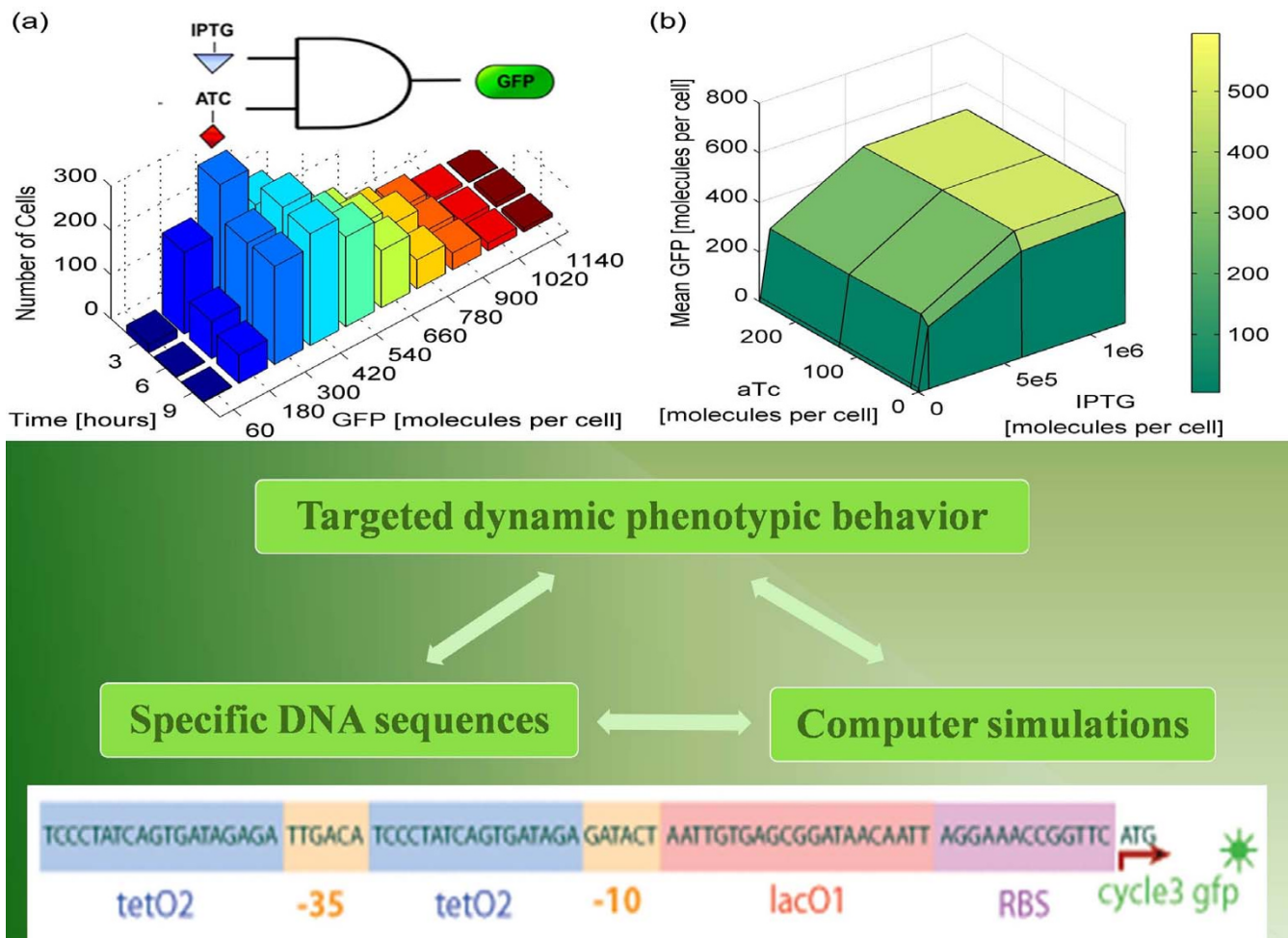


with  $k_1$  and  $k_{-1}$  the association and dissociation kinetic constants, respectively. If we considered the cell as a well-stirred reactor we could calculate the behavior of the network using a set of ordinary differential equations, which determine concentration changes as prescribed by kinetic

laws. However, the underlying assumption of such continuous-deterministic models, that the number of molecules approaches the thermodynamic limit (i.e. that the volume of the system is infinite), can be invalid for biological systems, since for some components (DNA for example) there are only a few copies available.

In the 1950s Oppenheim and McQuarrie, among others, explored stochasticity in kinetic models, developing the chemical Master equation formalism to capture discrete interaction events that occur with certain probability in time [14,15]. A numerical stochastic simulation algorithm (SSA) to calculate these probabilistic trajectories was described by Gillespie [16]. Gillespie's algorithm uses the system dynamics to simulate the occurrence of each individual reaction event. In general, given the current

state of the system, the SSA seeks the time until the next reaction occurs. It then executes that reaction, updates the state of the system, and increments the simulation time to the new value. Although accurate in capturing the dynamic of biomolecular interaction systems, SSA becomes computationally intractable, if the time scales of involved interaction events are disparate, because it simulates every single biomolecular interaction event, spending inordinate amounts on fast reactions for very few simulated occurrences of slow reactions. The modeling community was up to the challenge and in the last decade there have been numerous attempts to improve the efficiency of the SSA [17-23]. As a result, recently algorithms have appeared that successfully tackle biomolecular interaction phenomena with disparate time scales [24-29] (see Figure 1). Although work is still underway, there are now



**Figure 1**  
**A major challenge in synthetic biology is to rationally select DNA sequences that result in targeted dynamic phenotypes.** For example, with simulations using Hy3S [29] we are experimenting with multiple alternative promoter sequences to identify the optimal AND gate synthetic gene network, with tetracycline (atc) and IPTG as inputs and green fluorescence protein (GFP) as output.

exciting developments that the synthetic biology community can benefit from.

More than fifty years after the discovery of the molecular structure of DNA, molecular biology is mature enough for quantification useful for biological engineering applications, similar to chemistry in the 1950s. With the excitement synthetic biology is generating, the engineering and biological science communities appear remarkably willing to cross disciplinary boundaries toward this common goal.

## References

- Andrianantoandro E, Basu S, Karig DK, Weiss R: **Synthetic biology: new engineering rules for an emerging discipline.** *Mol Syst Biol* 2006, **2**:2006.0028.
- Church GM: **From systems biology to synthetic biology.** *Mol Syst Biol* 2005, **1**:2005.0032.
- Keasling J: **The promise of synthetic biology.** *Bridge Natl Acad Eng* 2005, **35**:18-21.
- Alon U: **Biological networks: the tinkerer as an engineer.** *Science* 2003, **301**:1866-1867.
- Kobayashi H, Kaern M, Araki M, Chung K, Gardner TS, Cantor CR, Collins JJ: **Programmable cells: interfacing natural and engineered gene networks.** *Proc Natl Acad Sci USA* 2004, **101**:8414-8419.
- Cohen SN, Chang ACY, Boyer H, Helling RB: **Construction of biologically functional bacterial plasmids in vitro.** *Proc Natl Acad Sci USA* 1973, **70**:3240-3244.
- Eisen H, Brachet P, Pereira da Silva L, Jacob F: **Regulation of repressor expression in  $\lambda$ .** *Proc Natl Acad Sci USA* 1970, **66**:855-862.
- Acrivos A, Amundson NR: **Applications of Matrix Mathematics to Chemical Engineering Problems.** *Ind Eng Chem* 1955, **47**(8):1533-1541.
- Vectors, Tensors and the Basic Equations of Fluid Mechanics.** Rutherford Aris, Prentice-Hall; 1962.
- Fredrickson A, Bird RB: **Non-Newtonian Flow in Annuli.** *Ind Eng Chem* 1958, **50**(3):347-352.
- Gardner TS, Cantor CR, Collins JJ: **Construction of a genetic toggle switch in *Escherichia coli*.** *Nature* 2000, **403**:339-342.
- Elowitz MB, Leibler S: **A synthetic oscillatory network of transcriptional regulators.** *Nature* 2000, **403**:335-338.
- Mayo AE, Setty Y, Shavit S, Zaslaver A, Alon U: **Plasticity of the cis-regulatory input function of a gene.** *PLoS Biol* 2006, **4**:e45.
- McQuarrie DA: **Stochastic Approach to Chemical Kinetics.** *Journal of Applied Probability* 1967, **4**(3):413-478.
- Grabert H, Hanggi P, Oppenheim I: **Fluctuations in Reversible Chemical Reactions.** *Physica* 1983, **117A**:300-316.
- Gillespie DT: **A general method for numerically simulating the stochastic time evolution of coupled chemical reactions.** *Journal of Computational Physics* 1976, **22**:403-434.
- Gibson MA, Bruck J: **Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels.** *J Phys Chem* 2000, **104**:1876-1889.
- Puchalka J, Kierzek AM: **Bridging the gap between stochastic and deterministic regimes in the kinetic simulations of the biochemical reaction networks.** *Biophys J* 2004:1357-1372.
- Rathinam M, Petzold LR, Cao Y, et al.: **Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method.** *J Chem Phys* 2005, **119**(24):12784-12794.
- Chatterjee A, Mayawala K, Edwards JS, Vlachos DG: **Time accelerated Monte Carlo simulations of biological networks using the binomial t-leap method.** *Bioinformatics* 2005, **21**(9):2136-2137.
- Haseltine EL, Rawlings JB: **Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics.** *J Chem Phys* 2002, **117**(15):6959-6969.
- Salis H, Kaznessis Y: **Accurate Hybrid Stochastic Simulation of a System of Coupled Chemical or Biochemical Reactions.** *J Chem Phys* 2005, **122**:054103.
- Salis H, Kaznessis YN: **An equation-free probabilistic steady-state approximation: dynamic application to the stochastic simulation of biochemical reaction networks.** *J Chem Phys* 2005, **123**(21):214106.
- Kaznessis Y: **Multi-Scale Models for Gene Network Engineering.** *Chemical Engineering Science* 2006, **61**(3):940-953.
- Salis H, Kaznessis Y: **Computer-aided design of modular protein devices: Boolean AND gene activation.** *Phys Biol* 2006, **3**(4):295-310.
- Sotiropoulos V, Kaznessis Y: **Synthetic tetracycline-inducible regulatory networks: computer-aided design of dynamic phenotypes.** *BMC Systems Biology* 2007, **1**:7.
- Kiehl TR, Mattheyses RM, Simmons MK: **Hybrid simulation of cellular behavior.** *Bioinformatics* 2004:316-322.
- De Jong H: **Modeling and simulation of genetic regulatory systems: A literature review.** *J Comput Biol* 2002, **9**(1):67-103.
- Salis H, Sotiropoulos V, Kaznessis YN: **Multiscale Hy3S: hybrid stochastic simulation for supercomputers.** *BMC Bioinformatics* 2006, **24**(7):93.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

