RE: Assistant Professor, Systems Biology/Microbiology

October 3, 2005

From: Gábor Balázsi, Ph.D.
Applied BioDynamics Laboratory-Biomedical Engineering
Boston University, 44 Cummington Street, Boston, MA 02215

To: Prof. Yves Brun, Systems Biology/Microbiology Faculty Search
Department of Biology, Indiana University
Jordan Hall 142, 1001 E 3rd St
Bloomington IN 47405-7005

Dear Prof. Yves Brun,

I have excellent qualifications for the faculty position in Systems Biology/Microbiology in the Department of Biology and Biocomplexity Institute, Indiana University, Bloomington. My research experience is in modeling signal transduction and processing in biological systems, large-scale biological data analysis, and modeling complex biochemical and neuronal network dynamics. I have more than eight years of experience in modeling and mathematically/computationally analyzing biological systems.

Currently, I am a Postdoctoral Research Associate at the Applied BioDynamics Laboratory at Boston University. My current work includes analysis and normalization of microarray data, analysis of small- and large-scale network topology and the computational modeling of biochemical reaction network dynamics. I work in close collaboration with experimentalists, and have a good understanding of experimental procedures.

I have excellent lecturing and communication skills. I speak fluently three languages, and am able to communicate in a few more. I have worked in Europe and the USA, with people from everywhere around the world. As an undergraduate I was trained to be a Physics teacher.

I am a work-authorized permanent resident in the United States.

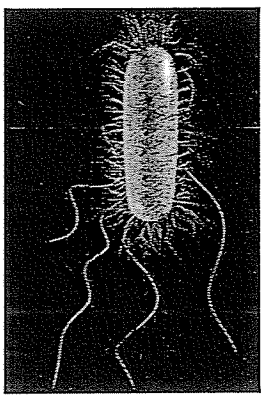
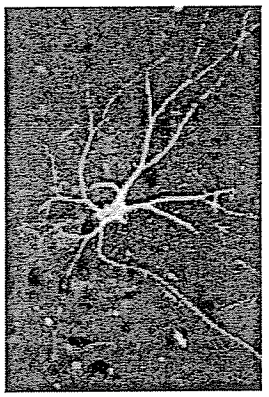
Thank you for considering my application, and I look forward to hearing from you.

Sincerely yours,

Gábor Balázsi

Tel: +1-617-353-5463
Fax: +1-617-353-5462
Email: balazsi@bu.edu

Signal Processing in Biological Networks

Research Proposal

Gábor Balázsi, Ph.D.

Applied BioDynamics Laboratory
Biomedical Engineering
Boston University
44 Cummington Street
Boston, MA 02215

Table of Contents

Signal Processing in Biological Networks

The quest to understand the world and use the acquired information for our benefit is coexistent with humanity. Human knowledge has developed from the phenomenological description of natural phenomena to a more quantitative understanding through mathematical modeling. Many of the spectacular advances in the fields of Physics and Chemistry were motivated by or due to the development of new quantitative methods.

Nowadays Biology is undergoing a similar revolution. New disciplines, such as Biophysics, Computational Biology, Bioengineering and Bioinformatics are emerging as a result of applying methods from Mathematics, Engineering and Computer Science in Biology. The quantitative approach to Biology is transforming today's science, leading to the development of new medications, biotechnological methods, etc. In addition, it provides a novel way to understand of what life is [1] – one of the most fundamental questions in science.

Research plan

Cells sense and respond to environmental perturbations and internal changes through the activity of various biochemical networks that operate over several orders of magnitude in time. The fastest of these comprise various signal transduction networks of the cell, ranging from two-component systems of prokaryotes to the highly complex signal transduction networks of mammalian cells, e.g., G-protein coupled signaling and Mitogen-Activated Protein Kinase (MAPK) cascades. Fast signaling is usually followed by a series of slower, 'transcriptional regulatory' events, during which regulatory gene products, Transcription Factors (TFs), alter the rate of transcription of other genes or operons. The totality of all potential transcriptional regulatory (TR) interactions forms a TR network [2] that, following stimulation, directs the reorganization of gene expression programs that alter metabolic activities, or governs distinct cellular programs such as the cell cycle, apoptosis, cell migration, or sporulation.

Recent evidence indicates that cellular networks (as mapped out by large-scale experiments [2] or literature searches [3]) represent only potential interactions among metabolites, proteins and genes. Various parts of these networks are dynamically utilized

[4,5,6], depending on the environmental conditions in which the cell exists. However, the principles governing this dynamic utilization are unknown.

In the future, I would like to focus on the following lines of research:

i. In close collaboration with experimentalists, develop theoretical and computational models of biochemical networks, to understand their dynamics and the causes of their malfunction (*1.* and *2.*);

ii. Investigate the role of noise in the perception, transmission and processing of information in biochemical systems (*3.*);

iii. Assemble networks and databases relevant for cellular information processing (*4.*).

iv. Model multi-cell systems to understand the effect of inter-cell communication and its possible effect on oncogenesis.

1. Sensing the environment

Biological organisms interact continuously with their surroundings, in the quest for food and mating partners, or trying to avoid predators and harmful factors. However, it would be a waste to dedicate equal amounts of resource to detect every aspect of environmental changes. Instead, intricate systems have evolved to specifically monitor variations crucial for survival. For example, some insects and birds can see in the ultraviolet domain, bats can hear ultrasounds, while humans can not. On the other hand, little is known about the type of temporal signals to which bacteria (or other single cells) are most sensitive.

Cells detect environmental changes by special sensor molecules, which, after changing their conformation or activity, and inducing a series of reactions, in the end alter protein expression levels, or initiate cellular programs. In collaboration with experimentalists, by mathematical and computational modeling and by engineering sensor-induced promoters upstream of fluorescent reporters, I will investigate how external signals of various amplitude and shape are transduced by sensor membrane receptor proteins. Sensor protein-encoding gene mutants will be created and the signal transduction properties of mutant sensor proteins will be investigated. This will uncover how temporal environmental perturbations are reflected within the organism after passing the level of sensors.

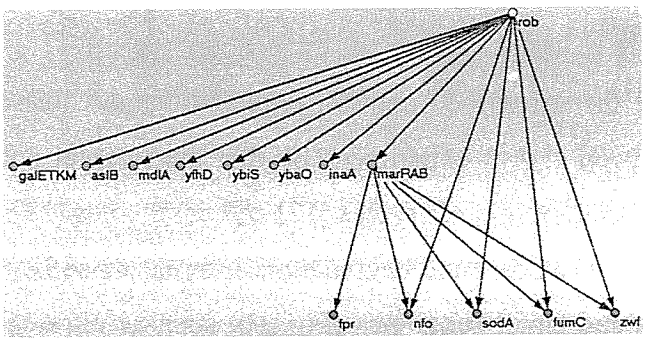
2. Signal processing in transcriptional networks

After passing the sensors, perturbations eventually reach the transcriptional regulatory network, which processes the incoming information and develops an appropriate response. It is unclear how dynamical changes of the environment are represented within the transcriptional-regulatory network.

In a recent study [7], we identified transcriptional modules with peculiar topology ('*origons*') within the *E. coli* transcriptional regulatory network that respond individually to different environmental stimuli (Fig. 1a). The emerging picture is that specific (elementary) perturbations propagate from layer to layer within one (or a few) hierarchically layered origons. Complex perturbations affect a larger number of origons, and the corresponding responses are implemented topologically by the convergence of origons near their output layers.

Perturbations pass through small subgraphs while propagating from the input towards the output within origons (Fig. 1b). Therefore, we must understand signal processing of small subgraphs. The potential role of certain subgraphs ('network motifs') is indicated by the fact that they are over-represented in transcriptional networks compared to randomized versions of the same networks [8,9].

a.



b.

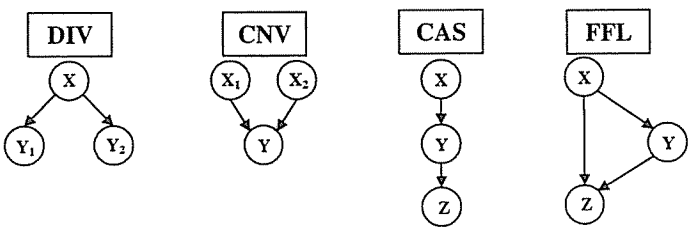


Figure 1. a. The rob origon. Circles represent transcriptional units (genes and operons); green and red arrows correspond to transcriptional activation and repression, respectively. **b.** Examples of three-node subgraphs that connect consecutive layers within the *E. coli* transcriptional network: divergence (DIV), convergence (CNV), cascade (CAS),and feed-forward loop (FFL).

Topological structures similar to origons might also underly the response of eukaryotic cells to external perturbations and inter-cellular cues (such as growth factors). I will use the yeast TR network [2,4,6] to identify the topological units of signal processing in this eukaryote. I will construct and study the topology of mammalian TR networks based on literature and databases.

To understand signal processing within TR networks, I plan to study the dynamics of transcriptional network motifs, subgraphs and origons in face of specific and combined environmental perturbations. I will collaborate with experimentalists to confirm the theoretical observations using fluorescent reporters and microarrays. Once the dynamical properties of small subgraphs have been elucidated, I will model the response of increasingly complex structures (a complete origon) to external perturbations, followed by modeling combinatorial response from two origons.

3. Signal processing in the presence of noise

From the organism's point of view, some environmental variations are useful signals, while the others are noise. In addition to environmental (external) noise, biological systems also generate their own (internal) noise [10] which necessarily affects signal transduction- and processing. How organisms survive in a highly noisy environment is still an open question.

A recent review discusses the "control, exploitation and tolerance of intracellular noise" [11]. It has been shown experimentally that the internal noise generated during transcription and translation in prokaryotes [12] and eukaryotes [13] depends on the rates of protein and mRNA synthesis and degradation. Internal noise is generated and shaped by individual nodes and small subgraphs (Fig. 1b) such as ultrasensitive cascades [14]. However, the effect of external noise on internal fluctuations has not been considered in these studies. I will investigate how internal noise depends on external fluctuations by computational modeling [15,16], and verify the results by experimentation (fluorescent cell cytometry and observation through the microscope). I will study how noise in prokaryotes [12] and eukaryotes [13] differs at increasing levels of complexity (from the single gene to multi-gene networks), and correlate the results with the type of small subgraphs present in both networks.

4. Integration of biochemical networks relevant for cellular signal processing

Nowadays, a large number of databases are available listing genes, proteins and their interactions. After proper definition of "nodes" and "links", networks are assembled from these databases to reflect transcriptional regulation [2,6,8,9], metabolic reactions [17], protein-protein interactions [18,19], etc. In spite of data availability [20], integrated networks including transcriptional regulation and protein modifications (phosphorylation, methylation, etc.) are still lacking.

I will assemble an integrated transcriptional regulatory - protein modification network for *E. coli*, study its small- and large-scale topological properties, and investigate the dynamic utilization of the network's components and their signal processing capabilities by modeling and experimentation.

An additional set of regulatory interactions awaiting quantitative modeling and integration is regulation by non-coding RNAs. The number of regulatory non-coding RNAs has increased steadily over the past years, together with the development of large-scale tools for their identification in prokaryotes [21] and eukaryotes [22,23]. I will integrate ncRNA-based regulation with other regulatory interactions into a common network. I will study the topology of this integrated network, and develop dynamical models of the subgraphs and motifs resulting from the topological analysis to understand their biological role.

5. Multi-cell systems and cancer

The transcriptional response of an individual bacterium is an important clue of how external perturbations are perceived and processed by the cell. However, even prokaryotic cells have inter-cellular signalling, which makes the behavior of a single cell to external changes insufficient to predict the fate of the culture. For example, quorum sensing might affect the biofilm formation capability of bacteria [24]. Inter-cellular communication is even more pronounced between mammalian cells, which form tissues, allowing them to communicate a large variety of signals through small volumes between adjacent cells. Inter-cellular communication can diminish or amplify small signals, the response of the tissue being drastically different from that of the individual cell exposed to the same stimulus. How inter-cell communication modifies single-cell signalling and

noise is largely unknown. I will study the effect of cell-cell communication on these phenomena, by varying the degree of inter-cell communication through mutating and deleting membrane-spanning channels. Based on the example of well-known multi-cell (neuronal) networks, a wide range of interesting spatio-temporal and noise-affected phenomena are expected to arise. One of the possible outcomes of intercell communication could be oncogenesis. Unbalanced expression of a growth factor or sensor due to mutation might not lead to unconstrained division and outgrowth if cells are not in a tissue, amplifying each other's signals. Understanding the role of such collective phenomena is crucial when modeling the development of cancer.

I will develop mathematical/computatonal models of mammalian regulatory networks relevant in inter-cell communication and cell division, and develop multi-cell models to investigate how different strengths of cell-cell coupling lead to altered processing (e.g., nonlinear amplification or attenuation) of inter-cell signals. In parallel, I will develop more phenomenological models, to understand spatio-temporal dynamics of cells arranged on a grid.

References

[1] Erwin Schrödinger, *What Is Life?: 'The Physical Aspect of the Living Cell'*, ed. by Roger Penrose (Cambridge, 1992).

[2] Lee, T.I., et al., *Science* **298**: 799-804 (2002).

[3] Salgado, H., et al., *Nucleic Acids Res.* **32** (Database issue): D303-6 (2004).

[4] Luscombe, N. M., et al., *Nature* **431**(7006), 308-12 (2004).

[5] Han, J. D., et al., *Nature* **430**(6995), 88-93 (2004).

[6] Harbison, C. T., et al, *Nature* **431**(7004), 99-104 (2004).

[7] Balázsi, G., Barabási, A.-L, and Oltvai, Z. N., *Proc. Nat. Acad. Sci., USA*, **102**(22), 7841-7846 (2005).

[8] S. Shen-Orr, R. Milo, S. Mangan, and U. Alon, *Nature Gen.* **31**, 64-68 (2002).

[9] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, *Science* **298**, 824-827 (2002).

[10] Swain, P. S., Elowitz, M. B., Siggia, E. D. *Proc Natl Acad Sci U S A.* **99**(20):12795-800 (2002).

[11] C. V. Rao, D. M. Wolf, and A. P. Arkin, *Nature* **420**, 231-237 (2002).

[12] E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden, *Nat Genet.* **31**(1), 69-73 (2002).

[13] W. J. Blake, M. Kærn, C. R. Cantor, and J. J. Collins, *Nature* **422** (6932), 633-637 (2003).

[14] M. Thattai, and A. van Oudenaarden, *Biophys J.* **82**(6), 2943-2950 (2002).

[15] D. T. Gillespie, *J.Comp. Phys.* **22**, 403-434 (1976).

[16] D. T. Gillespie, *J. Phys. Chem.* **81**, 2340-2361 (1977).

[17] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.-L. Barabási, *Nature* **407**, 651-654 (2000).

[18] Gavin A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M. et al. *Nature* **415**, 141–147 (2002).

[19] von Mering C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S. and Bork, P., *Nature* **417**, 399–403 (2002).

[20] http://vigen.biochem.vt.edu/xpd/xpd.htm

[21] Gottesman, S., *Annu Rev Microbiol.* **58**, 303-328 (2004). Review.

[22] Gustafson, A. M., et al., *Nucleic Acids Res.* **33**, D637-640 (2005).

[23] Pang, K. C., et al. *Nucleic Acids Res.* **33**, D125-130 (2005).

[24] Parsek, M. R., and Greenberg, E. P., *Trends Microbiol.* **13**(1), 27-33 (2005).

Teaching Philosophy

Gábor Balázsi

Teaching is probably the most crucial factor in the maintenance of human culture and civilization. It is the process by which our foregoers relay their experience tous. It is also an efficient means mediating the exchange of knowledge in today's society. Teachers are responsible for the continuity of human culture and knowledge.

As a curious child, I realized early how important teachers are in conveying knowledge. Subjects that were not interesting to me at once became fascinating and exciting when a new teacher started teaching them. I remember myself taking note of each teacher's actions, trying to imagine the 'ideal teacher'.

Since then, I have had many opportunities to teach. Before starting graduate school in the USA, I was studying to be a physics teacher. I took classes in Methods of Physics Teaching, Pedagogy, Psychology and Logic. I also taught high school students from freshmen to senior level. I taught Physics laboratory and discussion for freshmen and sophomore students at the Babes-Bolyai University of Cluj, Romania for two years. In the USA, I worked as a Teaching Assistant for three years. I taught various laboratory and discussion sessions (Basic Physics, Mechanics, Thermodynamics, Elctricity and Magnetism), and I tutored students, when they needed extra help. These were excellent opportunities to examine how far I was from my 'ideal teacher' image, and to improve myself whenever possible in its direction.

Specific to the scientific understanding of the world is objectivity: there is a unique reflection of the world in the mind, conceived the same way by every scientifically thinking person. It is like an enormous puzzle that the scientific community tries to piece together. The goal of teaching science to students is to reveal this picture to them. I believe that a teacher has to keep students interested, by continuously involving them into discussions, encouraging them to ask questions, and assigning interesting research projects.

As a teacher, I am careful to keep theory and application in balance. I give practical examples to every theoretical law or principle that I teach. Due to my background in Biophysics, Neuroscience and Molecular Biology, I easily find applications for students

with a major other than Physics or Mathematics. I tell anecdotes about famous scientists, stories on how physical laws were discovered, etc. I know all my students by name, and treat every one of them as an individual with his/her special background, learning style and set of interests. I believe that the principle of resonance applies in teaching: one can teach most effectively by transferring knowledge at the students' own wavelength.

Teaching is a two-way street. Leading discussions helps me to better understanding of the subject. Students' questions, can help me understand a phenomenon at a different level, or see it from a new perspective. I hope I will have the opportunity to practice teaching and experience again its benefits for the students and as well for myself.

If my aplication is successful, I would like to develop a course focusing on the topology and dynamical modeling of cellular (biochemical) networks. This is an extremely important part of a student's curriculum if he/she becomes involved in time-series measurements (which could be at single-cell level, as in confocal microscopy, or at multi-cell level, such as in microarray experiments). The course would cover currently used methods of modeling (Mass Action Kinetics, delay differential equations, stochastic simulation), including the necessary background in Mathematics and Computer Science.