Single-Cell Chemical Lysis Method for Analyses of Intracellular Molecules Using an Array of Picoliter-Scale Microwells

Yasuhiro Sasuga,[†,‡] Tomoyuki Iwasawa,[†] Kayoko Terada,[†,‡] Yoshihiro Oe,[†,§] Hiroyuki Sorimachi,[†] Osamu Ohara,*[,∥,⊥] and Yoshie Harada[†,‡,§,#]

The Tokyo Metropolitan Institute of Medical Science, 3-18-22 Honkomagome, Bunkyo-ku, Tokyo 113-8613, Japan, Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency (JST), Honcho, Kawaguchi, Saitama 332-0012, Japan, Department of Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Building FSB-401, 5-1-5, Kashiwanoha, Kashiwa, Chiba 277-8562, Japan, Department of Human Genome Research, Kazusa DNA Research Institute, 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan, Research Center for Allergy and Immunology, RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan, and Institute for Integrated Cell-Material Sciences (iCeMS), Kyoto University, Yoshida-konoecho 69, Sakyo-ku, Kyoto 6068501, Japan.

Analyzing the intracellular contents and enzymatic activities of single cells is important for studying the physiological and pathological activities at the cellular level. For this purpose, we developed a simple single-cell lysis method by using a dense array of microwells of 10−30-pL volume fabricated by poly(dimethylsiloxane) (PDMS) and a commercially available cell lysis reagent. To demonstrate the performance of this single-cell lysis method, we carried out two different assays at the single-cell level: detection of proteins by antibody conjugated microbeads and measurement of protease activity by fluorescent substrates. The results indicated that this method readily enabled us to monitor protein levels and enzymatic activities in a single cell. Because this method required only an array of PDMS microwells and a fluorescence microscope, the simplicity of this platform opens a way to explore the biochemical characteristics of single cells even by those who are not familiar with microfluidic technology.

A cell is the fundamental unit of life, and all functions in multicellular organisms are ultimately attributed to those of the cell. Although most conventional biochemical assays are performed using a considerable number of cells to determine their quantitative biomolecular profiles, such bulk assays only provide their averaged values in the analyzed ensemble and thus often overlook important information regarding their fluctuations among individual cells.[1,2] Because the differences among individual cells in the same ensemble are highly critical in some cases such as cell differentiation, quantitative measurement of biomolecular profiles at the single-cell level is a matter of concern in the biological community. Should it be possible to determine the biochemical parameters of each of multiple single cells in parallel, we would be able to describe the behavior of individual cells in an ensemble for obtaining deeper insights into cell−cell signaling, genetic heterogeneity, and heterotypic biological activities.[3] Furthermore, if single-cell biochemical analysis becomes feasible, we will be able to save much time, labor, and cost because we no longer have to purify a cell ensemble to homogeneity.

To develop a single-cell analysis in quantitative biology, various methods have been actively explored.[4−8] However, these newly emerging methods have not been fully applied in the biological sciences yet. One of the reasons for this is the fact that these methods are too sophisticated and integrated to be used without appropriate investment of time, money, and labor. Thus, there is a strong need to simplify and make these methods more biologist-friendly for realization of quantitative biology at the single-cell level. We believed that the difficulty lies in the fact that quantitative analysis of intracellular biological contents at the single-cell level must seamlessly integrate multiple steps: cell isolation/trapping, cell lysis, and quantitative assays of biochemical contents in the lysate. As for multiplexed single-cell analyses, most methods reported thus far take advantage of highly sophisticated and automated instruments for integration of these multiple steps on a single platform, e.g., single-cell capture followed by chemical lysis in a closed volume of 50 pL recently reported in a micro-

* To whom correspondence should be addressed. Phone: +81-438-52-3913. Fax: +81-438-52-3914. E-mail: ohara@kazusa.or.jp.
† The Tokyo Metropolitan Institute of Medical Science.
‡ CREST, JST.
§ The University of Tokyo.
∥ Kazusa DNA Research Institute.
⊥ RIKEN Yokohama Institute.
Kyoto University.

(1) Ferrells, J. E.; Machleder, E. M. *Science* **1998**, *280*, 895–898.
(2) Levsky, J. M.; Singer, R. H. *Trends Cell Biol.* **2003**, *13*, 4–6.
(3) Rubin, M. A. *Science* **2002**, *296*, 1329–1330.
(4) Levsky, J. M.; Shenoy, S. M.; Pezo, R. C.; Singer, R. H. *Science* **2002**, *297*, 836–840.
(5) Hong, J. W.; Studer, V.; Hang, G.; Anderson, W. F.; Quake, S. R. *Nat. Biotechnol.* **2004**, *22*, 435–439.
(6) Kurimoto, K.; Yabuta, Y.; Ohinata, Y.; Ono, Y.; Uno, K. D.; Yamada, R. G.; Ueda, H. R.; Saitou, M. *Nucleic Acids Res.* **2006**, *34*, e42.
(7) Huang, B.; Wu, H.; Bhaya, D.; Grossman, A.; Granier, S.; Kobilka, B. K.; Zarel, R. N. *Science* **2007**, *315*, 81–84.
(8) Newman, J. S.; Ghaemmaghami, S.; Ihmels, J.; Breslow, D. K.; Noble, M.; DeRisi, J. L.; Weissman, J. S. *Nature* **2006**, *441*, 840–846.

10.1021/ac8016423 CCC: $40.75 © XXXX American Chemical Society
Published on Web 11/05/2008

Monolayers of Silver Nanoparticles Decrease Photobleaching: Application to Muscle Myofibrils

P. Muthu,* N. Calander,* I. Gryczynski,† Z. Gryczynski,* J. M. Talent,* T. Shtoyko,‡ I. Akopova,* and J. Borejdo*
*Department of Molecular Biology and Immunology, and †Department of Cell Biology and Genetics, University of North Texas Health Science Center, Fort Worth, Texas; and ‡Department of Chemistry, University of Texas at Tyler, Tyler, Texas

ABSTRACT Studying single molecules in a cell has the essential advantage that kinetic information is not averaged out. However, since fluorescence is faint, such studies require that the sample be illuminated with the intense light beam. This causes photodamage of labeled proteins and rapid photobleaching of the fluorophores. Here, we show that a substantial reduction of these types of photodamage can be achieved by imaging samples on coverslips coated with monolayers of silver nanoparticles. The mechanism responsible for this effect is the interaction of localized surface plasmon polaritons excited in the metallic nanoparticles with the transition dipoles of fluorophores of a sample. This leads to a significant enhancement of fluorescence and a decrease of fluorescence lifetime of a fluorophore. Enhancement of fluorescence leads to the reduction of photodamage, because the sample can be illuminated with a dim light, and decrease of fluorescence lifetime leads to reduction of photobleaching because the fluorophore spends less time in the excited state, where it is susceptible to oxygen attack. Fluorescence enhancement and reduction of photobleaching on rough metallic surfaces are usually accompanied by a loss of optical resolution due to refraction of light by particles. In the case of monolayers of silver nanoparticles, however, the surface is smooth and glossy. The fluorescence enhancement and the reduction of photobleaching are achieved without sacrificing the optical resolution of a microscope. Skeletal muscle myofibrils were used as an example, because they contain submicron structures conveniently used to define optical resolution. Small nanoparticles (diameter ~60 nm) did not cause loss of optical resolution, and they enhanced fluorescence ~500-fold and caused the appearance of a major picosecond component of lifetime decay. As a result, the sample photobleached ~20-fold more slowly than the sample on glass coverslips.

INTRODUCTION

Recently, it has become possible to study single protein molecules in a cell (1,2). The advantage of single-molecule detection (SMD) is that it studies the behavior of proteins in their native (crowded) environment, and avoids problems associated with averaging responses of an assembly of molecules with different kinetics. However, the single-molecule approach is complicated by photobleaching, which arises because the sample must be illuminated with the intense laser beam to assure an adequate signal/noise ratio.

In an earlier work, we were able to decrease photobleaching by making measurements on coverglasses coated with nanoparticles known as surface island films (SIF) (3). SIFs are nanoparticles (4–10) that can support confined charge density oscillations (11) called localized surface plasmon (LSP) polariton modes. Excitation of LSPs causes a strong enhancement of local electric field and a substantial decrease of fluorescence lifetime caused by the distance-dependent changes in the radiative decay rates (12). Enhancement of the local field allows attenuation of illumination, which leads to reduction of photodamage. At the same time, enhancement causes an effective decrease of photobleaching: although it does not change the characteristics of a fluorophore, and therefore has no effect on the intrinsic resistance to photobleaching, it increases the time that a fluorophore remains unbleached, because the sample can be illuminated with a weaker light. Decrease of fluorescence lifetime, on the other hand, leads to a decrease of the intrinsic photobleaching rate. The decrease in bleaching is mainly caused by the fact that the molecule spends less time in the excited state, where it is subject to oxygen attack. For a fluorophore in a normal nonsaturation limit (i.e., the increase in laser power causes a corresponding increase in fluorescence intensity), the decrease of fluorescence lifetime also causes an increase in quantum yield, Q. However, since the Q of rhodamine is near 1 anyway, this effect is not very significant in our experiments. For a fluorophore in a saturation limit (i.e., the increase in laser power does not cause an increase in fluorescence intensity), there is an additional effect: the decrease in fluorescence lifetime causes an increase of effective cross section for absorption, because short-lived fluorophores can now be excited, whereas long-lived ones cannot be.

As a consequence of SIF-induced field enhancement and decrease of lifetime, the rate of photobleaching of rhodamine-phalloidin-labeled actin in a myofibril placed on glass coverslips coated with SIF decreased approximately twofold in comparison with that of myofibrils placed on uncoated coverslips (3). The price of the reduction in bleaching, however, was a loss of optical resolution and loss of nonuniformity of illumination, because SIF refracted the exciting and fluorescent light, and because nanoparticles were dis-

Submitted February 4, 2008, and accepted for publication April 7, 2008.

Address reprint requests to Julian Borejdo, Dept. of Biochemistry, Health Science Center, University of North Texas, 3500 Camp Bowie Blvd., Fort Worth, TX 76107-2699. Tel.: 817-735-2106; E-mail: jborejdo@hsc.unt.edu.

Editor: Elliot L. Elson.

0006-3495/08/10/3429/10 $2.00

doi: 10.1529/biophysj.108.130799

Cell

Leading Edge
Minireview

A Social Life for Discerning Microbes

Sam P. Brown[1,*] and Angus Buckling[1]
[1]Department of Zoology, University of Oxford, Oxford OX1 3PS, UK
*Correspondence: sam.brown@zoo.ox.ac.uk
DOI 10.1016/j.cell.2008.10.030

Microbes are not only extremely social but also extremely discerning about whom they socialize with. Recent research has uncovered some of the evolutionary explanations behind these feats of social sophistication in bacteria (Ackermann et al., 2008; Diggle et al., 2007) and, most recently, has provided insights into the molecular mechanisms of discrimination in yeast (Smukalla et al., 2008).

Cooperative acts ranging from minor help to marked self-sacrifice are seen in all corners of biology, from microbes to man. Given the presence of "freeloading cheaters," why is it that cooperative behaviors persist? Following the pioneering work of Hamilton (1964), a consensus has emerged that cooperative behaviors evolve and are maintained through a mix of self-interest and nepotism, which maximizes an individual's inclusive fitness, that is, the reproductive success of an individual and its close relatives. Cooperation may be self-interested if it directly benefits the actor as well as the recipients (for instance, increasing the success of one's own group). More extreme forms of cooperation—for example, altruistic cooperation where individuals experience a direct fitness cost in helping others—may be favored because the behavior helps recipients who are likely to share the altruistic gene (Hamilton, 1964). In other words, natural selection will cause a gene to spread if it confers an advantage on the individual in which the gene is present, as well as if it confers an advantage on other individuals with the same gene. But for this latter process to operate, altruistic acts must be preferentially directed toward other altruists. The most common scenario is that altruism is expressed blindly to neighbors, who will tend to be relatives (with an overrepresentation of similar genes) due simply to population viscosity (there is not complete mixing of individuals within the population) (Hamilton, 1964). A more complex scenario involves active processes of kin recognition and discrimination. However, this is not a perfect system as kin may still be "cheaters" that lack the vital gene for altruism.

Hamilton (1964) realized that one foolproof way to avoid wasting help on cheaters is for an individual to clearly display an altruistic or social gene and to recognize directly the same gene in others, rather than relying on proxy identifiers such as location or kinship. An individual social gene displays itself and enables recognition of the same gene in others, thus ensuring that help is only conferred on individuals expressing that gene. Richard Dawkins (1976) popularized this notion as the "green beard" gene: a gene that can be recognized externally because it confers a green beard on its carrier, ensuring that only green-bearded individuals are helped.

Hunting for Green Beard Genes

Green beard genes remained a plausible thought experiment until a remarkable empirical example was reported by Keller and Ross (1998): They discovered a gene cluster linked to the display and discrimination of identity in red fire ants. Next, Queller et al. (2003) identified an even tighter association between display, discrimination, and social traits in the first clear documentation of a single green beard gene in the slime mold *Dictyostelium discoideum*. When food is plentiful, *Dictyostelium* amoebae lead a single-celled, individualistic life of consumption and (asexual) reproduction. However, once starvation arrives they aggregate to form multicellular assemblies, replete with complex signaling mechanisms, division of labor, and individual sacrifice. The multicellular slugs differentiate into a sacrificial stalk, on top of which sit the lucky spores, which can then hitch a ride on invertebrates to potentially more favorable destinations. Queller et al. (2003) demonstrated that a single gene is responsible for holding nonstalk-forming "cheaters" at bay, by controlling entry to the spores. The *csaA* gene encodes a cell adhesion protein anchored in the cell membrane (the discernable green beard), which binds to homologous adhesion proteins (discrimination), building a multicellular aggregate preferentially of *csaA* carriers (targeted cooperation). This simple mechanism highlights the ability of microbes to harness molecular tricks to build sophisticated social behaviors (Foster et al., 2007), raising the possibility that these feats may be more common in microbes than previously suspected. Indeed, the bacterium *Proteus mirabilis* will only form motile swarms with those of the same strain. Although the molecular mechanism behind this recognition has not yet been fully deduced, a set of genes required for the recognition of "like" cells has been identified in *P. mirabilis* (Gibbs et al., 2008).

A Discriminating Yeast

The latest addition to the green beard gene stable, as Smukalla et al. (2008) report in this issue of *Cell*, appears in an unexpected microbe, the budding yeast *Saccharomyces cerevisiae*. Most strains of *S. cerevisiae* display an aggregating response to environmental stress called flocculation. Smukalla et al. (2008) now reveal that the expression of a single gene, *FLO1*, restores flocculation to a laboratory strain of *S. cerevisiae*, S288C, generating, at an individual cost, a social protection for S288C yeast within the aggregate (floc) against diverse environmental stresses such as ethanol and fungicides. But what about the potential for nonexpressing *flo1* cells to act as cheaters, exploiting the protection of established aggregates without paying the cost associated with *FLO1* expression? Smukalla et

Molecular Cell
Article

Molecular Basis for Regulation of the Heat Shock Transcription Factor σ^{32} by the DnaK and DnaJ Chaperones

Fernanda Rodriguez,[1,3] Florence Arsène-Ploetze,[1,4] Wolfgang Rist,[1,5] Stefan Rüdiger,[1,6] Jens Schneider-Mergener,[2] Matthias P. Mayer,[1,*] and Bernd Bukau[1,*]
[1]Zentrum für Molekulare Biologie der Universität Heidelberg, DKFZ-ZMBH Alliance, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany
[2]Institut für Medizinische Immunologie, Universitätsklinikum Charité, Schumannstrasse 20-21, 10098 Berlin, Germany, and the Jerini AG, Invalidenstrasse 130, 10115 Berlin, Germany
[3]Present address: Growth and Development, Biozentrum, University of Basel, Klingelbergstrasse 50/70, CH-4056 Basel, Switzerland
[4]Present address: UMR7156, Université Louis Pasteur/CNRS, Laboratoire de Génétique Moléculaire, Génomique et Microbiologie, Département Micro-organismes, Génomes, Environnement, 28, rue Goethe, 67083 Strasbourg Cedex, France
[5]Present address: Boehringer Ingelheim Pharma GmbH, Department of Respiratory Research, Proteomics Laboratory, Birkendorfer Strasse 65, 88397 Biberach, Germany
[6]Present address: Cellular Protein Chemistry, Department of Chemistry, Utrecht University, Kruytgebouw, Room O-701, Padualaan 8, 3584CH, Utrecht, the Netherlands
*Correspondence: m.mayer@zmbh.uni-heidelberg.de (M.P.M.), bukau@zmbh.uni-heidelberg.de (B.B.)
DOI 10.1016/j.molcel.2008.09.016

SUMMARY

Central to the transcriptional control of the *Escherichia coli* heat shock regulon is the stress-dependent inhibition of the σ^{32} subunit of RNA polymerase by reversible association with the DnaK chaperone, mediated by the DnaJ cochaperone. Here we identified two distinct sites in σ^{32} as binding sites for DnaK and DnaJ. DnaJ binding destabilizes a distant region of σ^{32} in close spatial vicinity of the DnaK-binding site, and DnaK destabilizes a region in the N-terminal domain, the primary target for the FtsH protease, which degrades σ^{32} in vivo. Our findings suggest a molecular mechanism for the DnaK- and DnaJ-mediated inactivation of σ^{32} as part of the heat shock response. They furthermore demonstrate that DnaK and DnaJ binding can induce conformational changes in a native protein substrate even at distant sites, a feature that we propose to be of general relevance for the action of Hsp70 chaperone systems.

INTRODUCTION

The heat shock response is an evolutionary conserved protective mechanism of cells against stress-induced damage of proteins. In *E. coli*, this response is mediated by the *rpoH* gene product, the heat shock transcription factor σ^{32}, that binds as an alternative σ subunit to the RNA polymerase (RNAP) core enzyme and targets it to the promoters of heat shock genes (Bukau, 1993; Gross, 1996; Yura and Nakahigashi, 1999). Stress-dependent changes in heat shock gene expression are mediated by changes in the activity and stability of σ^{32}. In cells growing at 30°C, σ^{32} is rapidly degraded ($t_{1/2} < 1$ min) primarily by the membrane-bound AAA-protease FtsH, which accounts for its very low intracellular levels (10–30 molecules per cell). Upon temperature upshift from 30°C to 42°C, the levels and half-life of σ^{32} increase transiently (induction phase). The induction phase is followed by a shut-off phase during which the synthesis of heat shock proteins is reduced to a new steady-state level. This shut-off results from rapid inactivation and destabilization of σ^{32} (Herman et al., 1995; Kanemori et al., 1999; Straus et al., 1987; Tilly et al., 1989; Tomoyasu et al., 1995).

Genetic and biochemical evidences indicate that the major Hsp70 chaperone in *E. coli*, DnaK, and its cochaperones DnaJ and GrpE play a central role for the adjustment of the steady-state levels of heat shock proteins and the rapid shut-off of the heat shock response (Grossman et al., 1987; Straus et al., 1990; Tilly et al., 1983). In vitro, σ^{32} in its native state is recognized as a substrate by DnaK and its cochaperone DnaJ (Gamer et al., 1992, 1996; Liberek et al., 1992, 1995). The interaction of DnaK with σ^{32} is transient and controlled by DnaK's nucleotide status. In the ADP state, the affinity of DnaK to σ^{32} is high and association and dissociation rates are low, while in the ATP state the affinity is low and association and dissociation rates are high (Gamer et al., 1996; Liberek et al., 1995; Mayer et al., 2000). DnaJ binds σ^{32} with high affinity, and simultaneous interaction of DnaJ and σ^{32} with DnaK·ATP stimulates DnaK's ATPase activity several thousand-fold, leading to the tight binding of DnaK·ADP to σ^{32} (Gamer et al., 1992, 1996; Laufen et al., 1999; Liberek et al., 1995).

The precise mechanism by which DnaK and DnaJ interact with σ^{32} and regulate its activity and stability is still unclear. By screening libraries of peptides derived from many different chaperone substrates, earlier work established the consensus motif that DnaK and DnaJ recognize in substrates. For DnaK it is composed of a core of five amino acids enriched in hydrophobic residues flanked by sequences enriched in positively charged amino acids (Rüdiger et al., 1997). By screening a σ^{32}-derived peptide library for DnaK-binding sites, seven sites were identified within the σ^{32}

Theory

Cell

Cytokine-Induced Signaling Networks Prioritize Dynamic Range over Signal Strength

Kevin A. Janes,[1,2,3,4] H. Christian Reinhardt,[1,3] and Michael B. Yaffe[1,*]
[1]Koch Institute for Integrative Cancer Research, Center for Cell Decision Processes, Department of Biology, and Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[2]Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA
[3]These authors contributed equally to this work
[4]Present address: Department of Biomedical Engineering, University of Virginia, Charlottesville, VA 22908, USA
*Correspondence: myaffe@mit.edu
DOI 10.1016/j.cell.2008.08.034

SUMMARY

Signaling networks respond to diverse stimuli, but how the state of the signaling network is relayed to downstream cellular responses is unclear. We modeled how incremental activation of signaling molecules is transmitted to control apoptosis as a function of signal strength and dynamic range. A linear relationship between signal input and response output, with the dynamic range of signaling molecules uniformly distributed across activation states, most accurately predicted cellular responses. When nonlinearized signals with compressed dynamic range relay network activation to apoptosis, we observe catastrophic, stimulus-specific prediction failures. We develop a general computational technique, "model-breakpoint analysis," to analyze the mechanism of these failures, identifying new time- and stimulus-specific roles for Akt, ERK, and MK2 kinase activity in apoptosis, which were experimentally verified. Dynamic range is rarely measured in signal-transduction studies, but our experiments using model-breakpoint analysis suggest it may be a greater determinant of cell fate than measured signal strength.

INTRODUCTION

Changes in cell behavior are determined by an interconnected set of proteins that actively transmit signaling information as a network (Irish et al., 2004; Jordan et al., 2000; Pawson, 2004). Modifications of the posttranslational state, enzymatic activity, or total level of key proteins can act as "molecular signals" that are relayed and interpreted to control cell function. The challenge of identifying which observed molecular signals determine a cell response is complicated because many signaling proteins appear to send mixed or opposing messages. For example, the transcription factor nuclear factor-κB (NF-κB) is widely regarded as a prosurvival protein because nuclear relocalization and DNA binding upregulate expression of apoptosis inhibitors such as c-IAP2, Bcl-x_L, and c-FLIP (Karin and Lin, 2002). In response to DNA-damaging agents, however, nuclear NF-κB can promote cell death by recruiting histone deacetylases that silence antiapoptotic genes (Campbell et al., 2004). Molecular signals can not only change their phenotypic meaning but also the relative importance of their message. Tumor cells, for instance, become addicted to chronically activated mitogenic pathways that are used only transiently in normal cells (Weinstein, 2002). Tools that could predict or explain such context-specific roles of molecular signals would be valuable for designing better-targeted therapies against disease (Blume-Jensen and Hunter, 2001; Miller-Jensen et al., 2007).

Many data-driven approaches exist for grouping, separating, or predicting outcomes on the basis of complex quantitative patterns of signaling or gene expression (D'Haeseleer, 2005; Janes and Yaffe, 2006; Noble, 2006). The problem with all of them is that they cannot distinguish molecules that are mechanistically linked to a phenotype from biomarkers that are correlative but not causative (Sawyers, 2008). This difficulty can be avoided by creating models from data sets that consist of molecular signals with recognized but complicated roles in the outcome that is to be predicted (Janes et al., 2005; Miller-Jensen et al., 2007). The drawback is that one's interpretation of such a model is biased toward the recognized roles of the molecular signals and away from more-surprising correlations with phenotype that could indicate new mechanisms. Data-driven models often identify hundreds of correlations in large data sets, making it impractical to perturb each one experimentally. Thus, an additional means for filtering correlation-based hypotheses is greatly needed.

Here, we develop a general approach, called "model-breakpoint analysis," which involves globally perturbing the measurements used to build a data-driven model and then quantifying the loss of model accuracy. We altered signaling-network measurements by manipulating each molecular signal's "dynamic range," defined as the responsiveness of cell outcomes to incremental changes in signal activation. Dynamic range has been understudied, because signaling networks are typically measured in either their basal (minimum) or hyperstimulated (maximum) states

Cell

Leading Edge
Review

Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences

Arjun Raj[1] and Alexander van Oudenaarden[1,*]
[1]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
*Correspondence: avano@mit.edu
DOI 10.1016/j.cell.2008.09.050

Gene expression is a fundamentally stochastic process, with randomness in transcription and translation leading to cell-to-cell variations in mRNA and protein levels. This variation appears in organisms ranging from microbes to metazoans, and its characteristics depend both on the biophysical parameters governing gene expression and on gene network structure. Stochastic gene expression has important consequences for cellular function, being beneficial in some contexts and harmful in others. These situations include the stress response, metabolism, development, the cell cycle, circadian rhythms, and aging.

Introduction

Life is a study in contrasts between randomness and determinism: from the chaos of biomolecular interactions to the precise coordination of development, living organisms are able to resolve these two seemingly contradictory aspects of their internal workings. Scientists often reconcile the stochastic and the deterministic by appealing to the statistics of large numbers, thus diminishing the importance of any one molecule in particular. However, cellular function often involves small numbers of molecules, of which perhaps the most important example is DNA. It is this molecule, usually present in just one or few copies per cell, that gives organisms their unique genetic identity. But what about genetically identical organisms grown in homogenous environments? To what degree are they unique? Increasingly, researchers have found that even genetically identical individuals can be very different and that some of the most striking sources of this variability are random fluctuations in the expression of individual genes. Fundamentally, this is because the expression of a gene involves the discrete and inherently random biochemical reactions involved in the production of mRNAs and proteins. The fact that DNA (and hence the genes encoded therein) is present in very low numbers means that these fluctuations do not just average away but can instead lead to easily detectable differences between otherwise identical cells; in other words, gene expression must be thought of as a stochastic process.

The experimental observation that the levels of gene expression vary from cell to cell is certainly not new. In 1957, Novick and Weiner showed that the production of beta-galactosidase in individual cells was highly variable and random, with induction increasing the proportion of cells expressing the enzyme rather than increasing every cell's expression level equally (Novick and Weiner, 1957). Such early studies were hindered, however, by the lack of reliable single-cell assays of gene expression. One of the first studies to use an expression reporter in single cells to examine the stochastic underpinnings of expression variability was the pioneering work of Ko et al. (1990). They examined the effect of different doses of glucocorticoid on the expression of a glucocorticoid-responsive transgene encoding beta-galactosidase and found that the cell-to-cell variability in the expression of the transgene was surprising large. Moreover, increasing the dose led to an increased frequency of cells displaying a high level of expression rather than a uniform increase in expression in every cell; that is, dose dependence was a consequence of changing the probability that an individual cell would express the gene at a high level.

Yet, despite the potential biological consequences of random cellular variability (Spudich and Koshland, 1976), several years would pass before theoretical work ignited much of the present interest in stochastic gene expression (McAdams and Arkin, 1997; Arkin et al., 1998). They modeled gene expression using a stochastic formulation of chemical kinetics derived by Gillespie (1977), predicting that in some biologically realistic parameter ranges, protein numbers could fluctuate markedly within individual cells. They then extended their analysis to model the circuit underlying the decision between lysis and lysogeny of the phage lambda, showing that stochastic effects in the expression of key regulators could explain why some cells activated the lytic pathway whereas others followed the lysogenic pathway. The notion that stochastic effects in gene expression could have important biological implications has motivated much research in the field and has only recently been explored experimentally.

Since this early research, the study of stochastic gene expression has blossomed into a rich field, with researchers from a diverse set of backgrounds working on a wide range of problems. The field is also notable for its strong interplay between theory and experiment, with many scientists making significant contributions to both. In this review, we will describe these researchers' efforts to characterize the underlying phenomenon through a host of organisms using a variety of experimental and theoretical methods. We will then highlight some recent endeavors trying to tie stochastic gene expression to biological phenomena.

Cell

The MatP/*matS* Site-Specific System Organizes the Terminus Region of the *E. coli* Chromosome into a Macrodomain

Romain Mercier,[1] Marie-Agnès Petit,[2] Sophie Schbath,[3] Stéphane Robin,[4] Meriem El Karoui,[2] Frédéric Boccard,[1,*] and Olivier Espéli[1,*]
[1]Centre de Génétique Moléculaire du CNRS, 91198 Gif-sur-Yvette, France
[2]INRA, UR888, Unité Bactéries Lactiques et Pathogènes Opportunistes, F-78350 Jouy-en-Josas, France
[3]INRA, UR1077 Mathématique, Informatique, et Génome, F-78350 Jouy-en-Josas, France
[4]AgroParisTech/INRA, UMR518 Unité Mathématiques et Informatique Appliquées, F-75005 Paris, France
*Correspondence: boccard@cgm.cnrs-gif.fr (F.B.), espeli@cgm.cnrs-gif.fr (O.E.)
DOI 10.1016/j.cell.2008.08.031

SUMMARY

The organization of the *Escherichia coli* chromosome into insulated macrodomains influences the segregation of sister chromatids and the mobility of chromosomal DNA. Here, we report that organization of the Terminus region (Ter) into a macrodomain relies on the presence of a 13 bp motif called *matS* repeated 23 times in the 800-kb-long domain. *matS* sites are the main targets in the *E. coli* chromosome of a newly identified protein designated MatP. MatP accumulates in the cell as a discrete focus that colocalizes with the Ter macrodomain. The effects of MatP inactivation reveal its role as main organizer of the Ter macrodomain: in the absence of MatP, DNA is less compacted, the mobility of markers is increased, and segregation of Ter macrodomain occurs early in the cell cycle. Our results indicate that a specific organizational system is required in the Terminus region for bacterial chromosome management during the cell cycle.

INTRODUCTION

The large size of genomes compared to cellular dimensions imposes extensive compaction of chromosomes compatible with DNA metabolism during replication, transcription, and segregation. Compaction of the chromosome results in the formation of a structure called the nucleoid. Work performed in different bacteria has revealed a number of processes involved in bacterial DNA condensation and organization; they include unrestrained DNA supercoiling, formation of a chromatin-like structure through the interaction of nucleoid-associated proteins (NAPs) with DNA, condensation by structural maintenance of chromosomes (SMC)-like proteins, and macromolecular crowding (Thanbichler and Shapiro, 2006). Cytological analyses have revealed that bacterial circular chromosomes are organized with a specific disposition within growing cells that preserves the linear order of loci in the DNA (Viollier et al., 2004; Nielsen et al., 2006b; Wang et al., 2006). In *E. coli*, cytological (Niki et al., 2000) and genetic (Valens et al., 2004) analyses based on long distance DNA interactions revealed a structuring process that spatially insulates large regions of the chromosome called macrodomains (MDs). Collisions between DNA sites belonging to different MDs occur at low frequency, and two particular regions, called nonstructured (NS) regions, can interact with both flanking MDs (Valens et al., 2004). The Ori MD containing *oriC* is centered on *migS*, a centromere-like site involved in bipolar positioning of *oriC* (Yamaichi and Niki, 2004). Opposite the Ori MD, the Ter MD containing the replication terminus is centered on the chromosome dimer resolution site *dif*. The Ter MD is flanked by the Left and Right MDs, whereas the Ori MD is flanked by the two NS regions. MD organization was directly visualized by analysis of the positioning, the segregation pattern, and the dynamics of markers belonging to various MDs. Markers in MDs showed much lower mobility than markers in NS regions (Espéli et al., 2008).

Understanding of chromosome segregation has improved considerably in recent years (Thanbichler and Shapiro, 2006). Replication initiates from a single origin, *oriC*, and progresses bidirectionally to the terminus of replication located opposite the origin. In *E. coli*, replication initiation occurs in specific replication factories where both replisomes are colocalized. As replication progresses, foci representing the two replisomes follow separate paths (Bates and Kleckner, 2005; Reyes-Lamothe et al., 2008). A colocalization step between sister chromatids of the newly duplicated foci has been reported and appears to vary with growth conditions (Sunako et al., 2001; Bates and Kleckner, 2005; Nielsen et al., 2006a; Adachi et al., 2008; Espéli et al., 2008). This process seems to be more persistent for loci located near the terminus of replication (Li et al., 2003; Bates and Kleckner, 2005; Espéli et al., 2008). The loss of colocalization occurs simultaneously at various positions on the chromosome when much of the chromosome is replicated (Bates and Kleckner, 2005; Espéli et al., 2008).

The mechanisms responsible for structuring the chromosome in MDs are largely unknown. A likely model postulates that internal organization of the different MDs involves recognition of a domain-specific repeated motif by a protein that would isolate

Cell

FLO1 Is a Variable Green Beard Gene that Drives Biofilm-like Cooperation in Budding Yeast

Scott Smukalla,[1,8] Marina Caldara,[1,8] Nathalie Pochet,[1,2,8] Anne Beauvais,[3] Stephanie Guadagnini,[4] Chen Yan,[1] Marcelo D. Vinces,[1] An Jansen,[5,6] Marie Christine Prevost,[4] Jean-Paul Latgé,[3] Gerald R. Fink,[5] Kevin R. Foster,[1] and Kevin J. Verstrepen[1,6,7,*]

[1]FAS Center for Systems Biology, Harvard University, Northwest Labs, 52 Oxford Street, Cambridge, MA 02138, USA

[2]Bioinformatics and Evolutionary Genomics & Quantitative Genomics Groups, Department of Plant Systems Biology, Flanders Institute for Biotechnology (VIB), and Department of Molecular Genetics, Ghent University, Technologiepark 927 B-9052 Ghent, Belgium

[3]Institut Pasteur, Aspergillus Unit

[4]Institut Pasteur, Electron Microscopy Platform

25 rue du Docteur Roux, 75724 Paris cedex 15, France

[5]Whitehead Institute for Biomedical Research/M.I.T., Nine Cambridge Center, Cambridge, MA 02142, USA

[6]Centre of Microbial and Plant Genetics, Department of Molecular and Microbial Systems, K.U.Leuven, Kasteelpark Arenberg 22, B-3001 Leuven (Heverlee), Belgium

[7]Laboratory for Systems Biology, Flanders Institute for Biotechnology (VIB), K.U.Leuven, Heverlee, Belgium

[8]These authors contributed equally to this work

*Correspondence: kevin.verstrepen@biw.kuleuven.be

DOI 10.1016/j.cell.2008.09.037

SUMMARY

The budding yeast, *Saccharomyces cerevisiae*, has emerged as an archetype of eukaryotic cell biology. Here we show that *S. cerevisiae* is also a model for the evolution of cooperative behavior by revisiting flocculation, a self-adherence phenotype lacking in most laboratory strains. Expression of the gene *FLO1* in the laboratory strain S288C restores flocculation, an altered physiological state, reminiscent of bacterial biofilms. Flocculation protects the *FLO1* expressing cells from multiple stresses, including antimicrobials and ethanol. Furthermore, *FLO1*$^+$ cells avoid exploitation by nonexpressing *flo1* cells by self/non-self recognition: *FLO1*$^+$ cells preferentially stick to one another, regardless of genetic relatedness across the rest of the genome. Flocculation, therefore, is driven by one of a few known "green beard genes," which direct cooperation toward other carriers of the same gene. Moreover, *FLO1* is highly variable among strains both in expression and in sequence, suggesting that flocculation in *S. cerevisiae* is a dynamic, rapidly evolving social trait.

INTRODUCTION

Since Darwin, evolutionary biologists have been troubled by cooperative behavior. Darwin systematically identified the phenomena that were the greatest challenge to his ideas. Cooperation was, and remains (Pennisi, 2005), one of these: "*If it could be proved that any part of the structure of any one species had been formed for the exclusive good of another species, it would annihilate my theory, for such could not have been produced through natural selection*" (Darwin, 1859). Cooperation is a problem for evolution by natural selection because individuals are predicted to act in a way that maximizes their personal reproduction. Costly behaviors that invest in a common good, therefore, are expected to be disrupted by so-called "cheaters" that save on the cost of cooperation but reap in the benefits of the investment of others. Such cheaters will be fitter than cooperators and take over the population, ultimately resulting in the loss of the cooperative behavior.

Why then do organisms frequently evolve behaviors that help others? For example, honeybee workers labor their whole life without reproducing, birds make alarm calls, and humans often help one another. This fundamental question has received considerable attention over the last 50 years with the development of the field of sociobiology. Following the work of Hamilton (1964), it is now widely accepted that cooperative behaviors evolve because they directly help the actor alongside any recipients, or they help individuals who share more alleles with the actor than predicted by chance (genetic relatedness), or both (Dawkins, 1976; Hamilton, 1964; Queller, 1984; West et al., 2006). In extreme cases, therefore, cooperators can successfully transmit their genes by helping another individual that carries these alleles, as occurs when near-sterile honeybee workers help their mother to reproduce. Typically, it is assumed that the correlation in genotype among individuals is generated by family, as is the case for sister workers in the social insects. However, Hamilton also engaged in a thought experiment, in which he proposed that cooperation is also possible if a single gene that drives the tendency to cooperate can also preferentially direct cooperation to other carriers of the gene. Such a (hypothetical) gene was later named a "green beard gene" by Dawkins, the green beard

Cell

Mistranslation of Membrane Proteins and Two-Component System Activation Trigger Antibiotic-Mediated Cell Death

Michael A. Kohanski,[2,3] Daniel J. Dwyer,[2,4] Jamey Wierzbowski,[2] Guillaume Cottarel,[2] and James J. Collins[1,2,3,*]
[1]Howard Hughes Medical Institute
[2]Department of Biomedical Engineering, Center for BioDynamics, and Center for Advanced Biotechnology
Boston University, 44 Cummington Street, Boston, MA 02215, USA
[3]Boston University School of Medicine, 715 Albany Street, Boston, MA 02118, USA
[4]Department of Math and Statistics, 111 Cummington Street, Boston University, Boston, MA 02215, USA
*Correspondence: jcollins@bu.edu
DOI 10.1016/j.cell.2008.09.038

SUMMARY

Aminoglycoside antibiotics, such as gentamicin and kanamycin, directly target the ribosome, yet the mechanisms by which these bactericidal drugs induce cell death are not fully understood. Recently, oxidative stress has been implicated as one of the mechanisms whereby bactericidal antibiotics kill bacteria. Here, we use systems-level approaches and phenotypic analyses to provide insight into the pathway whereby aminoglycosides ultimately trigger hydroxyl radical formation. We show, by disabling systems that facilitate membrane protein traffic, that mistranslation and misfolding of membrane proteins are central to aminoglycoside-induced oxidative stress and cell death. Signaling through the envelope stress-response two-component system is found to be a key player in this process, and the redox-responsive two-component system is shown to have an associated role. Additionally, we show that these two-component systems play a general role in bactericidal antibiotic-mediated oxidative stress and cell death, expanding our understanding of the common mechanism of killing induced by bactericidal antibiotics.

INTRODUCTION

Aminoglycosides, such as kanamycin and gentamicin, are a powerful class of bactericidal antibiotics that target the 30S subunit of the ribosome (Davis, 1987). The aminoglycoside family of antibiotics falls into the larger aminocyclitol group of 30S ribosome inhibitors, which also includes the bacteriostatic antibiotic spectinomycin. Protein mistranslation through tRNA mismatching is one of the hallmark phenotypes separating the bactericidal aminoglycosides from the other classes of ribosome inhibitors, including spectinomycin, which are bacteriostatic against *Escherichia coli* (*E. coli*) (Davis, 1987; Weisblum and Davies, 1968). Other phenotypes associated with aminoglycoside uptake and lethality include changes in membrane potential and permeability (Bryan and Kwan, 1983; Taber et al., 1987).

The lethal mode of action of aminoglycosides is thought to be due to either insertion of misread proteins into the inner membrane of *E. coli* (Bryan and Kwan, 1983; Davis et al., 1986) or irreversible uptake of aminoglycosides leading to complete inhibition of ribosome function (Davis, 1987); however, it is still unclear how this latter proposed mechanism contributes to aminoglycoside-mediated killing (Vakulenko and Mobashery, 2003). It has also been suggested that the mechanism of aminoglycoside-induced lethality is a function of more than ribosome inhibition and may be due to inhibition of multiple cellular targets (Hancock, 1981). In addition, there is a correlation between abnormal protein synthesis, such as mistranslation, and protein carbonylation, a form of oxidative stress (Dukan et al., 2000).

We have proposed that bactericidal antibiotics, including aminoglycosides, induce reactive oxygen species formation, which contributes to drug-mediated cell death (Kohanski et al., 2007). In this proposed model, bactericidal antibiotics perturb metabolism and respiration, leading to increased superoxide production and release or exposure of ferrous iron, which interacts with endogenous hydrogen peroxide to form lethal hydroxyl radicals (Dwyer et al., 2007; Kohanski et al., 2007). The sequence of events following the initial drug-target interaction that generate an intracellular environment promoting hydroxyl radical formation remains unknown for the different classes of bactericidal antibiotics.

In this study, we have elucidated the biological events following aminoglycoside interaction with the ribosome that lead to reactive oxygen species formation and contribute to cell death. Through systems-level approaches together with phenotypic and biochemical studies, we show that the protein translocation machinery and the Cpx envelope stress-response two-component system play important roles in oxidative stress-related cell death induced by aminoglycosides. We present evidence that the redox-responsive Arc two-component system is involved in this event, possibly via crosstalk between the Cpx and Arc two-component systems. Our results indicate that aminoglycoside-induced free radical formation is triggered by two-component stress-response system sensing of misfolded proteins in

JOURNAL OF BACTERIOLOGY, Nov. 2008, p. 6983–6995 Vol. 190, No. 21
0021-9193/08/$08.00+0 doi:10.1128/JB.00722-08

The Origins of 168, W23, and Other *Bacillus subtilis* Legacy Strains[▽]†

Daniel R. Zeigler,[1]* Zoltán Prágai,[2] Sabrina Rodriguez,[2] Bastien Chevreux,[2] Andrea Muffler,[2] Thomas Albert,[3] Renyuan Bai,[2]‡ Markus Wyss,[2] and John B. Perkins[4]

Bacillus Genetic Stock Center, The Ohio State University, Columbus, Ohio[1]; DSM Nutritional Products, Kaiseraugst, Switzerland[2]; NimbleGen Systems Inc., Madison, Wisconsin[3]; and DSM Anti-Infectives, DAI Innovation, Delft, The Netherlands[4]

Received 21 May 2008/Accepted 13 August 2008

***Bacillus subtilis* is both a model organism for basic research and an industrial workhorse, yet there are major gaps in our understanding of the genomic heritage and provenance of many widely used strains. We analyzed 17 legacy strains dating to the early years of *B. subtilis* genetics. For three—NCIB 3610[T], PY79, and SMY—we performed comparative genome sequencing. For the remainder, we used conventional sequencing to sample genomic regions expected to show sequence heterogeneity. Sequence comparisons showed that 168, its siblings (122, 160, and 166), and the type strains NCIB 3610 and ATCC 6051 are highly similar and are likely descendants of the original Marburg strain, although the 168 lineage shows genetic evidence of early domestication. Strains 23, W23, and W23SR are identical in sequence to each other but only 94.6% identical to the Marburg group in the sequenced regions. Strain 23, the probable W23 parent, likely arose from a contaminant in the mutagenesis experiments that produced 168. The remaining strains are all genomic hybrids, showing one or more "W23 islands" in a 168 genomic backbone. Each traces its origin to transformations of 168 derivatives with DNA from 23 or W23. The common prototrophic lab strain PY79 possesses substantial W23 islands at its *trp* and *sac* loci, along with large deletions that have reduced its genome 4.3%. SMY, reputed to be the parent of 168, is actually a 168-W23 hybrid that likely shares a recent ancestor with PY79. These data provide greater insight into the genomic history of these *B. subtilis* legacy strains.**

Bacillus subtilis, a model organism for gram-positive bacteria, is the focus of diverse research interests in both academic and industrial settings (54, 63, 64). One primary emphasis is sporulation, an archetypical form of cell development (25, 58, 66). Of equal interest, however, is the production of commercially attractive levels of small metabolites and enzymes (46, 57, 61). Investigations with *B. subtilis* benefit from the ease of its genetic manipulation (32), the wealth of available physiological and biochemical data (64), and the accessibility of a well-annotated genome sequence (40). Associated technologies are leading to a growing understanding of the proteome (33, 39, 74), the transcriptome (43), the metabolome (51), and the metabolic flux patterns (26, 41, 59) of this organism.

B. subtilis strains used in virtually all academic research and many industrial processes derive from a single tryptophan-requiring auxotroph, strain 168. Despite its central importance to research, our knowledge of this strain's genomic heritage is incomplete. Strain 168 was isolated after *B. subtilis* Marburg was mutagenized with X-rays by two Yale University botanists, Paul Burkholder and Norman Giles (15). Their experiments showed that for both vegetative cells and spores of *B. subtilis*, sublethal doses of UV or X-rays caused high frequencies of auxotrophy among survivors. For many mutants, the auxotrophic requirement could be met by a single nutrient, allowing the "one gene-one enzyme" model previously developed for *Neurospora* (8) and *Escherichia coli* (42) to be extended to this gram-positive spore former as well. Soon afterwards, the Yale group abandoned its *B. subtilis* studies to pursue other research interests (24). Sadly, most of its *B. subtilis* collection, including the wild-type parent, was subsequently lost. At least five mutants, however—auxotrophs requiring threonine (strain 23), nicotinic acid (strain 122), or tryptophan (strains 160, 166, and 168)—were preserved and transferred to the possession of Charles Yanofsky. Nearly a decade later, Yanofsky provided the mutants to John Spizizen (65), who in a landmark publication demonstrated that three of them—122, 166, and 168—could be transformed to prototrophy when exposed to DNA from strain 23 (65). The highly transformable strain 168 became the subject of follow-up studies detailing this phenomenon (6, 79). Strain 168 was subsequently disseminated around the world. Researchers soon developed classical genetic methods—and later, recombinant and genomic technologies—to elucidate the physiology and spore development of this strain. By the mid-1970s, so many mutants had been developed from 168 that a centralized repository, the Bacillus Genetic Stock Center (BGSC), was established to maintain them (81).

Strain 168 and its siblings are not the only legacy strains surviving from the earliest years of *B. subtilis* genetics, however. By the early 1960s, strain W23 began appearing regularly in the literature (20, 69, 77, 78). Its origin has remained a mystery (34). A few researchers have explicitly stated, without citing evidence, that W23 is derived from Burkholder and Giles strain 23 (14). The connection is not intuitively obvious, however: strain 23 requires threonine (49, 65), while W23 is prototrophic and streptomycin resistant (69). Although W23 was initially considered a wild-type equivalent to 168, it was soon discovered that these strains differed significantly in bacterio-

* Corresponding author. Mailing address: Bacillus Genetic Stock Center, The Ohio State University, 484 W. 12th Ave., Columbus, OH 43210. Phone: (614) 292-5550. Fax: (614) 292-6773. E-mail: zeigler.1@osu.edu.

† Supplemental material for this article may be found at http://jb.asm.org/.

‡ Present address: The Johns Hopkins University, School of Medicine, Department of Neurosurgery, Baltimore, MD.

▽ Published ahead of print on 22 August 2008.

JOURNAL OF BACTERIOLOGY, Dec. 2008, p. 7786–7796 Vol. 190, No. 23
0021-9193/08/$08.00+0 doi:10.1128/JB.00715-08

Processing of a Membrane Protein Required for Cell-to-Cell Signaling during Endospore Formation in *Bacillus subtilis*[▽]

Mónica Serrano,[1] Filipe Vieira,[1] Charles P. Moran, Jr.,[2] and Adriano O. Henriques[1]*

Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Avenida da República, Apartado 127, 2781-901 Oeiras Codex, Portugal,[1] and Emory University School of Medicine, Department of Microbiology and Immunology, Atlanta, Georgia 30322[2]

Received 21 May 2008/Accepted 12 September 2008

Activation of the late prespore-specific RNA polymerase sigma factor σ^G during *Bacillus subtilis* sporulation coincides with completion of the engulfment process, when the prespore becomes a protoplast fully surrounded by the mother cell cytoplasm and separated from it by a double membrane system. Activation of σ^G also requires expression of *spoIIIJ*, coding for a membrane protein translocase of the YidC/Oxa1p/Alb3 family, and of the mother cell-specific *spoIIIA* operon. Here we present genetic and biochemical evidence indicating that SpoIIIAE, the product of one of the *spoIIIA* cistrons, and SpoIIIJ interact in the membrane, thereby linking the function of the *spoIIIJ* and *spoIIIA* loci in the activation of σ^G. We also show that SpoIIIAE has a functional Sec-type signal peptide, which is cleaved during sporulation. Furthermore, mutations that reduce or eliminate processing of the SpoIIIAE signal peptide arrest sporulation following engulfment completion and prevent activation of σ^G. SpoIIIJ-type proteins can function in cooperation with or independently of the Sec system. In one model, SpoIIIJ interacts with SpoIIIAE in the context of the Sec translocon to promote its correct localization and/or topology in the membrane, so that it can signal the activation of σ^G following engulfment completion.

Two important issues in developmental biology are how adjacent cells communicate so that coordinated responses emerge and how these signal transduction complexes are assembled at appropriate, discrete sites within the cell. The process of spore differentiation by the gram-positive bacterium *Bacillus subtilis* provides several examples of cell-cell communication pathways that initiate in one cell and influence the behavior of an adjacent sister cell (11, 15, 39). Sporulation involves the cooperation between two cells formed at the onset of the process by asymmetric division of a sporangial cell. This polar division produces a polar prespore, the future spore, and an adjacent larger mother cell, a terminal cell line, which contributes to spore morphogenesis but that lyses at its completion (11, 15, 39). The programs of gene expression in the prespore and the mother cell are controlled by the successive activation of cell-type-specific RNA polymerase sigma factors. The activation of three of these development-specific sigma factors requires signals conveyed by the adjacent cell. In this way, close coordination between the programs of gene expression in the two cells is achieved. The activity of σ^F in the prespore is required for activation of σ^E in the mother cell. Because activation of σ^F is coupled to the completion of septation, activation of σ^E is also coupled to this morphological landmark (11, 15, 39).

Following activation of σ^E, the mother cell initiates a phagocytic-like process whereby it engulfs the prespore, eventually encapsulating it in a double membrane system formed by the inner and outer prespore membranes. Upon engulfment completion, σ^F is replaced by σ^G, itself produced under the control of σ^F, which then controls late stages in development in the prespore (11, 15, 39). σ^G in turn triggers a signaling pathway that leads to the activation of the late mother cell-specific regulator σ^K, which administers the final stages of development, including lysis of the mother cell to discharge the mature spore into the environment (11, 15, 39). Activation of σ^G does not follows its synthesis under σ^F control but depends upon and coincides with completion of the engulfment process (reviewed in references 11, 15, and 39). However, the mechanism by which σ^G is kept inactive prior to engulfment completion remains elusive (5, 6, 7, 13, 33, 41). Two anti-sigma factors are known that can bind to and antagonize σ^G, SpoIIAB (6, 13, 22, 23, 41) and the product of the σ^F-controlled *csfB* gene (also called *gin*) (5, 8, 9, 21). However, making σ^G refractory to the action of SpoIIAB, or deletion of *csfB/gin*, even in a strain where σ^G is insensitive to SpoIIAB, does not cause its massive premature activity prior to engulfment completion (5, 8, 13, 41). The process by which σ^G is activated following conclusion of the engulfment process is also unknown. It is clear, however, that activation of σ^G in the engulfed prespore requires a signaling pathway emanating from the mother cell and involving the activity of σ^E (5, 6, 7, 8, 21, 22, 33, 41, 42, 45).

σ^E directs the expression of the *spoIIIA* locus, which encodes eight proteins predicted to be membrane associated and all of which are required for the activation of σ^G (5, 16, 22, 45). The SpoIIIA-encoded proteins (SpoIIIAA to SpoIIIAH) may assemble into a complex that links engulfment completion to the activation of σ^G (2, 3, 5, 10, 19, 22, 45). Consistent with this idea, SpoIIIAH localizes to the outer prespore membrane through a direct interaction of its C-terminal extracellular domain with the extracellular domain of the inner prespore membrane protein SpoIIQ (2, 3, 5, 10, 19, 28), which has been implicated in the synthesis and activation of σ^G (5, 19, 26, 37, 47). The two proteins colocalize and form discrete foci along

* Corresponding author. Mailing address: Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Avenida da República, Apartado 127, 2781-901 Oeiras Codex, Portugal. Phone: 351-21-4469521. Fax: 351-21-4411277. E-mail: aoh@itqb.unl.pt.

▽ Published ahead of print on 26 September 2008.

JOURNAL OF BACTERIOLOGY, Nov. 2008, p. 7479–7490 Vol. 190, No. 22
0021-9193/08/$08.00+0 doi:10.1128/JB.00823-08

Escherichia coli Harboring a Natural IncF Conjugative F Plasmid Develops Complex Mature Biofilms by Stimulating Synthesis of Colanic Acid and Curli[▽]†

Thithiwat May and Satoshi Okabe*

Graduate School of Engineering, Hokkaido University, Kita-13, Nishi-8, Kita-Ku, Sapporo 060-8628, Japan

Received 12 June 2008/Accepted 29 August 2008

It has been shown that *Escherichia coli* harboring the derepressed IncFI and IncFII conjugative F plasmids form complex mature biofilms by using their F-pilus connections, whereas a plasmid-free strain forms only patchy biofilms. Therefore, in this study we investigated the contribution of a natural IncF conjugative F plasmid to the formation of *E. coli* biofilms. Unlike the presence of a derepressed F plasmid, the presence of a natural IncF F plasmid promoted biofilm formation by generating the cell-to-cell mating F pili between pairs of F^+ cells (approximately two to four pili per cell) and by stimulating the formation of colanic acid and curli meshwork. Formation of colanic acid and curli was required after the initial deposition of F-pilus connections to generate a three-dimensional mushroom-type biofilm. In addition, we demonstrated that the conjugative factor of F plasmid, rather than a pilus synthesis function, was involved in curli production during biofilm formation, which promoted cell-surface interactions. Curli played an important role in the maturation process. Microarray experiments were performed to identify the genes involved in curli biosynthesis and regulation. The results suggested that a natural F plasmid was more likely an external activator that indirectly promoted curli production via bacterial regulatory systems (the EnvZ/OmpR two-component regulators and the RpoS and HN-S global regulators). These data provided new insights into the role of a natural F plasmid during the development of *E. coli* biofilms.

Surface-attached microbial communities (biofilms) have distinct morphological and biochemical properties that distinguish them from free-living planktonic cells (30). The formation of a three-dimensional (3D) mushroom-type mature biofilm by *Pseudomonas* sp. has been described as a stepwise process with at least four developmental stages: (i) initial reversible attachment of planktonic bacteria to the surface, (ii) a transition to irreversible attachment involving specific adhesive factors and the formation of microcolonies, (iii) maturation of the microcolonies into a thin film by production of extracellular polymeric substances in an adhesive matrix, and (iv) detachment of cells from the biofilm and return of cells to the planktonic state (41).

In addition to morphological descriptions of biofilm formation, there is also increasing interest in the horizontal gene transfer that controls biofilm formation (39). Mobile genetic elements mediate horizontal gene transfer between bacteria in natural habitats. These elements can be conjugative plasmids, transposons, or bacteriophages (24). Although a laboratory strain of *Escherichia coli* does not form extensive biofilms spontaneously, conjugative plasmid expression promotes the development of thick mature biofilms (18, 35). In previous studies of *E. coli* harboring a conjugative plasmid, the workers identified IncFI and IncFII conjugative plasmids that potentially promoted formation of a 3D mushroom-type mature biofilm similar to a *Pseudomonas* sp. biofilm (6, 34, 46). Notably, the plasmids previously tested (pOX38Km, an IncFI plasmid [34]; F′traD36, an IncFI plasmid [6]; and R1drd19, an IncFII plasmid [46]) are classified as derepressed conjugative plasmids which constitutively express F pili at all times (15). However, constitutively expressed F pili are required for development of highly organized mature *E. coli* biofilms (18) and support biofilm maturation even in the absence of flagella, type 1 fimbriae, curli, the Ag-43 autotransport protein, or AI-2-mediated cell-to-cell communication (34). These findings raise the question of how horizontal gene transfer can promote the formation of dense biofilms, especially via a natural IncF F plasmid, in which F-pilus formation and the conjugative mechanism are repressed. The mechanistic basis of the putative connection between natural F-pilus formation and biofilm maturation is poorly understood. Furthermore, it is also not known how natural IncF F-plasmid-containing biofilms are different in terms of ultrastructure and gene expression.

A natural IncF F plasmid that was originally obtained from a host *E. coli* K-12 strain is a pioneer conjugative plasmid belonging to the IncFI incompatibility group (15; H. Shimizu, Y. Saitoh, Y. Suda, K. Uehara, G. Sampei, and K. Mizobuchi, unpublished data). This plasmid is referred to as the fertility factor (also known as F factor or sex factor) that allows bacteria to produce F pili necessary for bacterial conjugation. All the sequences required for conjugative transmission (20 *tra* genes) of a natural IncF F plasmid are present in the 33.3-kb transfer region (the conjugative factor) of this 99.2-kb plasmid, while the other 65.9 kb (the nonconjugative factor) contains a number of other genetic sequences responsible for incompatibility, replication, and other functions (27).

* Corresponding author. Mailing address: Graduate School of Engineering, Hokkaido University, Kita-13, Nishi-8, Kita-Ku, Sapporo 060-8628, Japan. Phone and fax: 81-(0)11-706-6266. E-mail: sokabe@eng.hokudai.ac.jp.

† Supplemental material for this article may be found at http://jb.asm.org/.

▽ Published ahead of print on 12 September 2008.

J|A|C|S
ARTICLES

Published on Web 10/29/2008

Jammed Acid–Base Reactions at Interfaces

Julianne M. Gibbs-Davis, Jennifer J. Kruk, Christopher T. Konek, Karl A. Scheidt, and Franz M. Geiger*

Department of Chemistry, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208

Received June 6, 2008; E-mail: geigerf@chem.northwestern.edu

Abstract: Using nonlinear optics, we show that acid–base chemistry at aqueous/solid interfaces tracks bulk pH changes at low salt concentrations. In the presence of 10 to 100 mM salt concentrations, however, the interfacial acid–base chemistry remains jammed for hours, until it finally occurs within minutes at a rate that follows the kinetic salt effect. For various alkali halide salts, the delay times increase with increasing anion polarizability and extent of cation hydration and lead to massive hysteresis in interfacial acid–base titrations. The resulting implications for pH cycling in these systems are that interfacial systems can spatially and temporally lag bulk acid–base chemistry when the Debye length approaches 1 nm.

Introduction

Interfacial acid–base processes are ubiquitous in nature.[1–6] Physical chemistry tells us that acid–base properties at interfaces should track acid–base properties in the bulk under equilibrium conditions. However, charge-balancing at the interface[7–9] as well as the molecular environments of the bulk and the interfacial species can shift interfacial acid–base equilibria by multiple pK_a units from their corresponding bulk solution values.[7,10–16] While a molecular-level understanding of these dramatic effects has now emerged for equilibrium, or steady-state conditions, their influence on the time dependence of interfacial acid–base reactions is just beginning to be understood. For instance, dynamic measurements using a surface force balance show that, following progressively longer waiting times after separation from contact, mica loses protons to approach its equilibrium charge state in water via an activated process that can be affected by the salt concentration in the surrounding aqueous solution.[17] While pH-sensitive chromophores diluted into surfactants at air/water interfaces have been studied as a function of electrolyte concentration,[7] time delays or hysteresis was not reported for those highly mobile two-dimensional systems. Hysteresis, however, was observed in electroosmotic measurements of silica capillaries[18] and calorimetric bulk equilibrium measurements on colloidal silica,[19] the latter of which can be rationalized using statistical rate theory.[20]

Here, we apply nonlinear optics to show that interfacial acid–base chemistry tracks the bulk pH at low salt concentrations. In the presence of 10 to 100 mM salt concentrations, however, the interfacial acid–base chemistry remains jammed for hours, until it finally occurs within minutes at a rate that follows the kinetic salt effect. For various alkali halide salts, the delay times increase with increasing anion polarizability and extent of cation hydration and lead to massive hysteresis in interfacial acid–base titrations.

Experimental Section

We use a flow system (Figure 1A) described previously[21,22] to monitor the interfacial protonation state in real time and in situ using an optical probe that is based on the "$\chi^{(3)}$ method".[12] The square root of the measured second harmonic generation (SHG)[23,24] intensity yields the SHG E-field, which depends on the interfacial potential, produced by the interfacial charges, whose density is given by the number density of the interfacial SiOH, SiO^-, and $SiOH_2^+$ groups (Figure 1B). For charged interfaces, the second harmonic generation (SHG) E-field can be modeled with the familiar second-order response to which a third-order term is added, according to $E_{SHG} = \chi^{(2)}E_\omega E_\omega - \chi^{(3)}E_\omega E_\omega \Phi_o$.[12] The third-order response is given by the product of the third-order nonlinear susceptibility, $\chi^{(3)}$, the

(1) Hynes, J. T. *Nature* **1999**, *397*, 565.
(2) Eigen, M. *Angew. Chem., Int. Ed. Engl.* **1964**, *3*, 1.
(3) Pearson, R. G. *J. Am. Chem. Soc.* **1963**, *85*, 3533.
(4) Bordwell, F. G. *Acc. Chem. Res.* **1988**, *21*, 456.
(5) Stumm, W.; Morgan, J. J. *Aquatic Chemistry, Chemical Equilibria and Rates in Natural Waters*, 3rd ed.; John Wiley & Sons: New York, 1996.
(6) Raviv, U.; Laurat, P.; Klein, J. *Nature* **2001**, *413*, 51.
(7) Xiao, X. D.; Vogel, V.; Shen, Y. R. *Chem. Phys. Lett.* **1989**, *163*, 555.
(8) Uibel, R. H.; Harris, J. M. *Anal. Chem.* **2002**, *74*, 5112.
(9) Huang, X.; Kovaleski, J. M.; Wirth, M. J. *Anal. Chem.* **1996**, *68*, 4119.
(10) Bhattacharyya, K.; Sitzmann, E. V.; Eisenthal, K. B. *J. Chem. Phys.* **1987**, *87*, 1442.
(11) Bain, C. D.; Whitesides, G. M. *Langmuir* **1989**, *5*, 1370.
(12) Zhao, X.; Subrahmanyan, S.; Eisenthal, K. B. *Chem. Phys. Lett.* **1990**, *171*, 558.
(13) Zhao, X.; Ong, S.; Wang, H.; Eisenthal, K. B. *Chem. Phys. Lett.* **1993**, *214*, 203.
(14) Hu, K.; Bard, A. J. *Langmuir* **1997**, *13*, 5114.
(15) Gershevitz, O.; Sukenik, C. N. *J. Am. Chem. Soc.* **2004**, *126*, 482.
(16) Konek, C. T.; Musorrafiti, M. J.; Al-Abadleh, H. A.; Bertin, P. A.; Nguyen, S. T.; Geiger, F. M. *J. Am. Chem. Soc.* **2004**, *126*, 11754.
(17) Raviv, U.; Laurat, P.; Klein, J. *J. Chem. Phys.* **2002**, *116*, 5167.
(18) Lambert, W. J.; Middleton, D. L. *Anal. Chem.* **1990**, *62*, 1585.
(19) Machesky, M. L.; Anderson, M. A. *Langmuir* **1986**, *2*, 582.
(20) Piasecki, W. *Langmuir* **2003**, *19*, 9526.
(21) Mifflin, A. L.; Gerth, K. A.; Geiger, F. M. *J. Phys. Chem. A* **2003**, *107*, 9620.
(22) Al-Abadleh, H. A.; Mifflin, A. L.; Musorrafiti, M. J.; Geiger, F. M. *J. Phys. Chem. B* **2005**, *109*, 16852.
(23) Shen, Y. R. *The Principles of Nonlinear Optics*; John Wiley & Sons: New York, 1984.
(24) Eisenthal, K. B. *Chem. Rev.* **1996**, *96*, 1343.

10.1021/ja804302s CCC: $40.75

J|A|C|S
COMMUNICATIONS

Published on Web 10/22/2008

Photoswitchable Nanoparticles Enable High-Resolution Cell Imaging: PULSAR Microscopy

Dehong Hu,[†] Zhiyuan Tian,[‡] Wuwei Wu,[‡] Wei Wan,[‡] and Alexander D. Q. Li*,[‡]

Pacific Northwest National Laboratory, Richland, Washington 99352, and Department of Chemistry, Washington State University, Pullman, Washington 99164

Received July 29, 2008; E-mail: dequan@wsu.edu

Fluorescence imaging has transformed biological sciences and opened a window to reveal biological mechanisms in real time despite the diffraction limit restricting the optical microscope resolution to ~300 nm. Recently, two ultrahigh-resolution fluorescence microscopic techniques have circumvented the diffraction limit by stochastically photoswitching fluorophores on and off.[1−4] The current selection of photoswitchable fluorophores for such high-resolution imaging techniques are extremely limited, and the photoswitching mechanisms have not been well studied. Two methods were reported: photoactivatable fluorescence protein (PFP) and a pair of Cy3 and Cy5 dyes in close proximity. However, the PFP is much larger than dye molecules, and Cy3−Cy5 pairs suffer from stringent requirements that the environments must be oxygen free. These will limit their cellular imaging applications. Thus, new photoswitchable fluorophore development becomes critically important to apply such ultrahigh-resolution microscopy broadly.[5]

The ideal photoswitchable probe should integrate high brightness, low molecular weight, and flexible chemical modifications into a single molecule so that it will easily adapt to the cellular environment. In this paper, spiropyran derivatives having the desired characteristics are presented as the excellent photoswitchable fluorophores. Using such a fluorophore, we have achieved photoactuated unimolecular logical switching attained reconstruction (PULSAR) microscopy, with resolution down to 10−40 nm, far beyond the diffraction limit. Herein, we observe nanostructures on glass surfaces and cellular organelles in fixed cells, which cannot be resolved by conventional fluorescence microscopy. Spiropyran requires no special treatment to cells and can be readily modified using chemical functional groups for specific targeting.[6,7]

Unlike the photoactivatable green fluorescent protein and the Cy3−Cy5 pair in very close proximity, the spiropyran−merocyanine photoswitching mechanism and photochemically induced structures are well understood (Figure 1a). [8−12] The spiropyran form has negligible visible absorption and consequently no fluorescence under visible excitation (Figure 1b). Its ring-opened form merocyanine, however, absorbs strongly at 570 nm and emits vivid red fluorescence (665 nm) in hydrophobic nanoparticle cores (Figure 1b).[6,7] UV illumination induces the spiropyran-to-merocyanine conversion and switches on fluorescence; visible photoexcitation of the merocyanine accelerates the back conversion and returns the photoswitching dye into the off state. Such desired fluorescence photoswitching properties will enable nanometer-resolution PULSAR microscopy provided that spiropyran/merocyanine dyes have a high emission rate, long photobleaching time, qualified emitted photon count, and efficient photochemical switching.

Accordingly, the single-molecule photophysics of merocyanine was studied first. Spin-coating dilute spiropyran and poly (methyl methacrylate) in dichloromethane resulted in thin films containing isolated photoswitchable dyes. Next, UV illumination converted the spiropyran dye to the merocyanine fluorophore. The detected merocyanine photons produce an emission rate of several kHz under 0.6 kW/cm^2 excitation, comparable to other high-quantum yield dyes such as rhodamine and strong enough for single-molecule studies.[13]

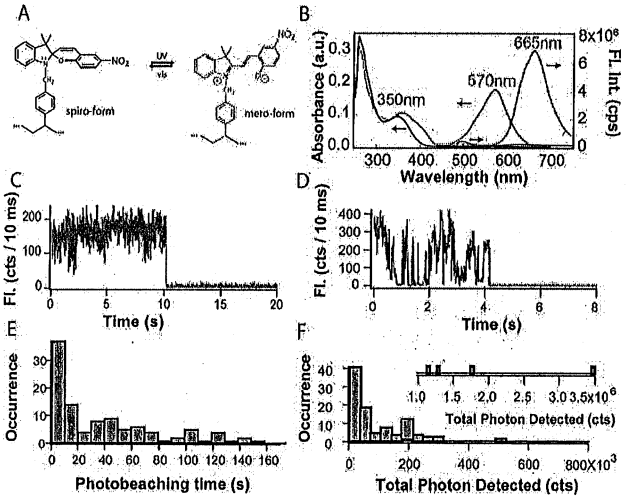


Figure 1. (A) Spiropyran (spiro-form) and merocyanine (mero-form) were polymerized into the hydrophobic cores of polymer nanoparticles, and thus their photochemical conversion occurs within isolated protected nanoparticles. (B) The optical absorption (left and bottom axes) of spiro- (black) and mero-form (red) and fluorescence spectra (bottom and right axes) of the spiro- (black) and mero-form (red) are plotted against the wavelength. (C and D) Two typical single-molecule intensity trajectories illustrate that the photoswitching dye in its emissive state; merocyanine behaves very much like normal high-quantum yield dyes such as rhodamine. (E) The photobleaching time histogram reveals that the photobleaching reaction follows first-order kinetics. (F) The histogram of total detected photon displays the number of photons detected from each individual molecule and such emission occurrence.

Figure 1c and d display two representative fluorescence intensity trajectories from 100 photoswitchable individual merocyanine molecules, and the trajectories exhibit typical single molecule signatures such as intensity blinking and one-step photobleaching. Permanent merocyanine photobleaching is evident. In Figure 1c and d, the molecules were bleached at 10.2 s and at 4.2 s, respectively. Moreover, Figure 1e reveals the measured photobleaching times in a histogram, which yields an average photobleaching time of 47 s. The histogram exponential decay suggests that photobleaching obeys first-order kinetics with a rate constant of ~0.02 s^{-1}.

The total number of photons that one molecule can emit before photobleaching is an important photophysical property because it determines the ultimate spatial resolution of the PULSAR imaging

[†] Pacific Northwest National Laboratory.
[‡] Washington State University.

10.1021/ja805948u CCC: $40.75

Coarse-grained kinetic Monte Carlo models: Complex lattices, multicomponent systems, and homogenization at the stochastic level

Stuart D. Collins, Abhijit Chatterjee,[a] and Dionisios G. Vlachos[b]
Department of Chemical Engineering, University of Delaware, Newark, Delaware 19716, USA

(Received 2 July 2008; accepted 30 September 2008; published online 10 November 2008)

On-lattice kinetic Monte Carlo (KMC) simulations have extensively been applied to numerous systems. However, their applicability is severely limited to relatively short time and length scales. Recently, the coarse-grained MC (CGMC) method was introduced to greatly expand the reach of the lattice KMC technique. Herein, we extend the previous spatial CGMC methods to multicomponent species and/or site types. The underlying theory is derived and numerical examples are presented to demonstrate the method. Furthermore, we introduce the concept of homogenization at the stochastic level over all site types of a spatially coarse-grained cell. Homogenization provides a novel coarsening of the number of processes, an important aspect for complex problems plagued by the existence of numerous microscopic processes (combinatorial complexity). As expected, the homogenized CGMC method outperforms the traditional KMC method on computational cost while retaining good accuracy. © *2008 American Institute of Physics*. [DOI: 10.1063/1.3005225]

INTRODUCTION

Traditional kinetic Monte Carlo (KMC) simulations have enjoyed impressive success in the engineering and computational scientific community due to their ability to capture, among others, noise, out-of-equilibrium processes, and complex particle interactions.[1–8] The KMC method allows the simulation of spatial heterogeneous systems with nanoscopic variation. However, many systems are too computationally demanding for KMC simulation due to multiple reasons. KMC simulations are capable of capturing roughly 10^4-10^6 lattice points (from 100×100 to 1000×1000 nm^2, assuming a lattice constant of 1 nm), putting many systems of interest with correlation lengths at or above the micrometer range out of reach. Long time scale events, such as pattern formation and aggregation, are also often intractable by KMC simulation.[9–11] Diffusion-controlled systems pose a particular difficulty for KMC simulation due to the hydrodynamic slowdown from the overwhelming number of small diffusion jumps that must be simulated.[12] Calculating long-distance interactions consumes a large fraction of CPU time. Finally, systems with large reaction networks involve too many individual processes for KMC to track, store, and search through. This problem, termed as combinatorial complexity in this paper, arises in many applications. Examples include biology, due to the huge number of conformations proteins can take,[13–15] and epitaxy of metals, due to numerous diffusion barriers arising from different local atomic environments.[16,17]

To extend the capabilities of the KMC method to longer time and length scales, the coarse-grained MC (CGMC) method has recently been developed.[8,12,18–20] In our approach, neighboring microscopic sites are grouped together into "coarse-grained" (CG) cells and a closure is applied at the stochastic level to resident atoms or molecules (here after termed adparticles) to describe their distribution in the cell.[12,19,20] In the simplest closure, the local mean-field (LMF) approximation, adparticles within cells are assumed to be well mixed.[12,19,20] Other closures are explored in Ref. 21 and strong interactions in Ref. 22 The adparticles of each cell are then allowed to interact with, react with, and diffuse to nearby cells.

The CGMC method efficiently addresses many of the stated weaknesses of the traditional KMC method. First, the grouping of microscopic sites simply reduces the number of lattice nodes to be individually tracked and the number of processes to be simulated. Second, as the size of the CG cells increases, the interaction potential length (relative to the CG lattice constant) shrinks, leading to a much faster calculation of the coarse interaction potential. Third, simulated diffusion jumps are much larger, overcoming (in part) the hydrodynamic slowdown in diffusion-controlled systems.

Previous CGMC simulations have focused on uniform surfaces comprised of a single type of microscopic site with a single type of adparticle. In this paper, we extend our previous CGMC method[12,19,20] to an arbitrary number of site types and/or adparticle species. Lattice-based simulations of multiple site type and adparticle species systems have been performed previously using a MF estimate.[23] This extension allows CGMC to be applied to a much wider range of systems of interest, such as catalytic reaction systems[24] (where various site types may represent different elements or lattice positions), diffusion on surfaces and in nanoporous materials,[4,5,25] and biological signaling[26–28] (where site types may represent distinct areas of a cell membrane). In order to overcome the problem of combinatorial complexity, the concept of homogenization at the stochastic level is introduced.

The organization of this paper is as follows. First, the

[a]Present address: Theoretical Division, T-12, MS B268, Los Alamos National Laboratory, Los Alamos, NM. 87545.
[b]Author to whom correspondence should be addressed. Electronic mail: vlachos@udel.edu. Tel.: 302-831-2830.

Lab on a Chip

RSC Publishing

Journals Lab on a Chip Advance Articles DOI: 10.1039/b808579j

Lab Chip, 2009 | DOI: 10.1039/b808579j | Paper

Simple, robust storage of drops and fluids in a microfluidic device†

Hakim Boukellal [a], Šeila Selimović [a], Yanwei Jia [a], Galder Cristobal [b] and Seth Fraden *[a]

[a] *Complex Fluids Group, Martin Fisher School of Physics, Brandeis University, Waltham, MA 02454, USA. E-mail: fraden@brandeis.edu*

[b] *Rhodia Asia Pacific Pte Ltd, Singapore Science Park II, 117586, Singapore*

Received 22nd May 2008 , Accepted 22nd September 2008

First published on the web 28th October 2008

We describe a single microfluidic device and two methods for the passive storage of aqueous drops in a continuous stream of oil without any external control but hydrodynamic flow. Advantages of this device are that it is simple to manufacture, robust under operation, and drops never come into contact with each other, making it unnecessary to stabilize drops against coalescence. In one method the device can be used to store drops that are created upstream from the storage zone. In the second method the same device can be used to simultaneously create and store drops from a single large continuous fluid stream without resorting to the usual flow focusing or T-junction drop generation processes. Additionally, this device stores all the fluid introduced, including the first amount, with zero waste. Transport of drops in this device depends, however, on whether or not the aqueous drops wet the device walls. Analysis of drop transport in these two cases is presented. Finally, a method for extraction of the drops from the device is also presented, which works best when drops do not wet the walls of the chip.

Introduction

In microfluidics, drops are considered as "micro-reactors", as each drop can be the site of an independent experiment.[1–3] Distinct processing steps include formulation, drop creation and mixing of the contents of the drop. This paper will focus on related subsequent steps: on-chip drop storage and extraction, which are useful for the class of devices where drop processing occurs on-chip.

Drop formulation and drop creation, which are conceptually distinct steps, are often intimately related. In this paper drops will always be an aqueous phase while an immiscible oil will be the continuous phase. Popular techniques for continuous drop generation are "flow focusing"[4] and the "T-junction"[5] where the flow of water solution is cut by a flow of oil. These two methods allow drop production at a frequency as high as tens of kiloHertz, but have several shortcomings. One is that formulation is limited to gradual composition variations in successive drops.[3] A second shortcoming is that there is a start-up time before the flows are steady enough to produce drops of the desired size and frequency and thus it is inevitable that some solution is wasted. Mixing the drop's contents is achieved by chaotic advection when the drop passes around a corner.[3] An alternate approach to

Vol 455|30 October 2008|doi:10.1038/nature07282

LETTERS

Non-random segregation of sister chromosomes in *Escherichia coli*

Martin A. White[1], John K. Eykelenboom[1], Manuel A. Lopez-Vernaza[1], Emily Wilson[1] & David R. F. Leach[1]

It has long been known that the 5′ to 3′ polarity of DNA synthesis results in both a leading and lagging strand at all replication forks[1]. Until now, however, there has been no evidence that leading or lagging strands are spatially organized in any way within a cell. Here we show that chromosome segregation in *Escherichia coli* is not random but is driven in a manner that results in the leading and lagging strands being addressed to particular cellular destinations. These destinations are consistent with the known patterns of chromosome segregation[2,3]. Our work demonstrates a new level of organization relating to the replication and segregation of the *E. coli* chromosome.

Prokaryotic cells were long considered to be featureless until recent advances in imaging revealed an array of internal structures and subcellular organizations[4]. One such example is the nuclear architecture of *E. coli*, where the domains of the left and right chromosome arms occupy distinct cellular locations[2]. During replication, these domains are progressively segregated by an unknown mechanism that results in translational symmetry of the chromosome arms (known as replichores)[2,3]. Counter-intuitively, to obtain this translational symmetry there must be mirror symmetry in the segregation of the leading and lagging strands of the two replication forks (Fig. 1A). This leads to three possible situations. Either the two lagging strands of replication are positioned at mid-cell while the leading strands migrate to the cell poles; the two lagging strands migrate to the cell poles while the leading strands are positioned at mid-cell; or a random combination of these two segregation patterns occurs within the population. The specific hypothesis that leading strands segregate to cell poles has been proposed[5,6] to explain the preponderance of highly expressed genes on the leading strands of the left (L) and right (R) replichores[7]. However, no evidence for or against this has yet been presented.

To distinguish these three possibilities, we designed a construct (Supplementary Fig. 1) that would allow us specifically to visualize a locus on the right replichore (*lacZ*) that was replicated on the leading strand of the replication fork (R_{lead}). This is accomplished by inducing SbcCD-mediated palindrome cleavage[8] in a *recA*$^-$ mutant that has the palindrome flanked upstream by an array of *tetO* sites and downstream by an array of *lacO* sites. These arrays are visualized by fluorescence microscopy upon the binding of TetR–yellow fluorescent protein (YFP) and LacI–cyan fluorescent protein (CFP), respectively, and allow us to follow the cellular position of the chromosomal region containing the palindrome as it is replicated and segregated. This is visible as a twin YFP–CFP spot, demonstrating the presence of DNA on both sides of the palindrome. Induction of SbcCD results in the specific cleavage of the palindrome that was replicated on the lagging strand[8] (R_{lag}) by formation of a DNA hairpin[9]. In a *recA*$^-$ mutant the broken chromosome is degraded[10] leaving behind the intact copy of the construct located on R_{lead} (Fig. 1B). As long as the labelling method does not disrupt the known pattern of chromosome segregation, then the location of only one of the chromosome arms needs to be known for the rest to be inferred (Fig. 1A).

The cellular position of this construct was followed during growth, and the number of segregated copies related to cell length. In the absence of induced double-strand breaks, 68.8% of exponentially growing cells longer than 1.50 μm had two visibly segregated copies of the construct. In these cells, the loci were segregated such that one was located at mid-cell and the other at a cell pole (Fig. 2a). This

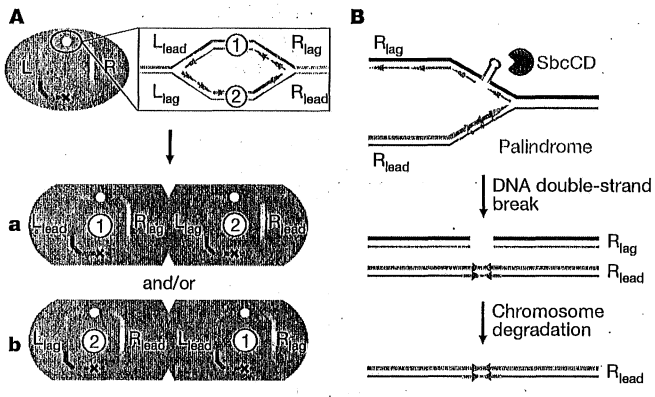


Figure 1 | Distinguishing leading and lagging strands. A, Cartoon demonstrating how mirror symmetry of leading and lagging strands accounts for translational symmetry of sister chromosomes. **Aa**, Segregation of leading strands towards the cell poles and of lagging strands towards mid-cell. **Ab**, Segregation of lagging strands towards the cell poles and of leading strands towards mid-cell. **B**, SbcCD cleaves a DNA hairpin formed by the palindrome on the lagging strand of replication, and the broken chromosome is degraded in a *recA*$^-$ mutant.

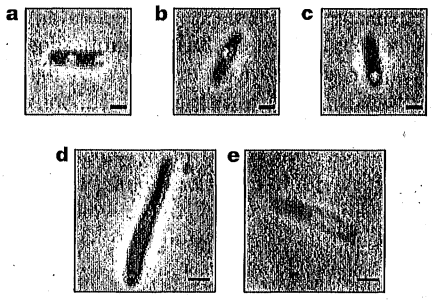


Figure 2 | Visualization of construct. *E. coli* cell containing: **a**, two segregated copies of the construct; **b**, one located at mid-cell; and **c**, one located at a cell pole. The CFP signal is pseudocoloured green, YFP magenta. Overlapping CFP and YFP signals appear white. Visualization of the nucleoid in cells subjected to SbcCD-mediated palindrome cleavage for 120 min by DAPI staining (DNA, blue) (**d**) and mounting on gelatin (DNA, white) (**e**). Scale bars, 1 μm (**a**–**c**) and 2 μm (**d**, **e**).

[1]Institute of Cell Biology, School of Biological Sciences, University of Edinburgh, King's Buildings, Edinburgh EH9 3JR, UK.

Activity motifs reveal principles of timing in transcriptional control of the yeast metabolic network

Gal Chechik[1,5], Eugene Oh[2], Oliver Rando[3], Jonathan Weissman[2], Aviv Regev[4] & Daphne Koller[1]

Significant insight about biological networks arises from the study of network motifs—overly abundant network subgraphs[1,2]—but such wiring patterns do not specify when and how potential routes within a cellular network are used. To address this limitation, we introduce activity motifs, which capture patterns in the dynamic use of a network. Using this framework to analyze transcription in *Saccharomyces cerevisiae* metabolism, we find that cells use different timing activity motifs to optimize transcription timing in response to changing conditions: forward activation to produce metabolic compounds efficiently, backward shutoff to rapidly stop production of a detrimental product and synchronized activation for co-production of metabolites required for the same reaction. Measuring protein abundance over a time course reveals that mRNA timing motifs also occur at the protein level. Timing motifs significantly overlap with binding activity motifs, where genes in a linear chain have ordered binding affinity to a transcription factor, suggesting a mechanism for ordered transcription. Finely timed transcriptional regulation is therefore abundant in yeast metabolism, optimizing the organism's adaptation to new environmental conditions.

Cellular processes are mediated through intricate networks of interacting molecules, whose local[1–3] and global[4,5] topology has been intensively studied. Analysis of network wiring patterns has revealed network motifs—local sets of interaction patterns that occur significantly more often than expected by chance and potentially reflect the functionality of the complex network. However, whereas such network motifs correspond to the static wiring of the network, networks are used dynamically and adapt to external conditions and internal states in functionally distinct ways.

Such dynamic activity is particularly important in metabolic processes, which are tightly controlled based on the cell's environment. Upon a change in environmental conditions, a cell may have to rapidly reconfigure its metabolism to produce or degrade compounds to ensure its survival in the new environment. Fluxes through metabolic reactions are controlled by enzymes, the activities and abundances of which are further controlled by post-transcriptional mechanisms. Furthermore, protein abundance is partly determined by transcriptional control[6], which modifies the mRNA level of the gene through binding of transcription factors. This complex hierarchy of regulatory mechanisms raises questions about the individual roles and interplay between the different regulation layers. For instance, is transcription regulation tuned to fit the usage patterns of enzymes in the metabolic network? Recent work argues for the predominance of hierarchical control in the metabolic network[7,8]. Yet the metabolic network exhibits significant changes in transcript levels in response to environmental perturbations, suggesting the use of transcriptional control. Moreover, recent work on transcriptional control of metabolism identified cases of finer-grained patterns of co-regulation in *S. cerevisiae*[9,10]. In one specific example in *Escherichia coli*, the genes in a linear pathway for amino acid biosynthesis were reported to show sequential transcriptional activation ("just-in-time" transcription)[11].

We developed an analysis framework based on the notion of an activity motif (**Fig. 1a**) to systematically study the dynamical behavior of such networks. Given a particular network structure (e.g., a linear cascade of enzymes; **Fig. 1b**), an activity motif describes a specific pattern of functional data, such as ordered timing of activation of the corresponding genes (**Fig. 1c**(i)). Unlike network motifs, which reflect the static wiring of the network (analogous to routes in a road map)[1,12], activity motifs reflect dynamic and functional patterns associated with their use (analogous to traffic patterns that emerge before, during and after rush hour). Activity motifs can be identified by assessing the enrichment of activity patterns given the network wiring structure.

We applied the framework of activity motifs to study the dynamics of regulation of gene expression in the metabolic network in *S. cerevisiae*. Using our systematic analysis, we identified abundant activity motifs involving timed gene expression regulation, recurring in different forms across many conditions. We demonstrate that the same timing behavior can be conserved in the dynamics of protein abundance, suggesting that timing patterns in mRNA expression can have a direct effect on the timing of metabolic processes. Finally, by studying activity motifs in transcription factor binding affinity, we show that evolution of quantitative transcription factor binding affinities provides a mechanism that can underlie some of this fine-grained control of transcription timing. Overall, the activity motif framework allows us to systematically investigate three levels of

[1]Department of Computer Science, Stanford University, Stanford, California 94305, USA. [2]Howard Hughes Medical Foundation and Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, California 94143, USA. [3]Departments of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01655,USA. [4]Department of Biology, Massachusetts Institute of Technology and the Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, Massachusetts 02142, USA. [5]Present address: Google Research, 1600 Amphitheater Parkway, Mountain View, California 94043, USA. Correspondence should be addressed to A.R. (aregev@broad.mit.edu) or D.K. (koller@cs.stanford.edu).

Published online 26 October 2008; doi:10.1038/nbt.1499

nature biotechnology

Protein microarrays with carbon nanotubes as multicolor Raman labels

Zhuo Chen[1,4], Scott M Tabakman[1,4], Andrew P Goodwin[1], Michael G Kattah[2], Dan Daranciang[1], Xinran Wang[1], Guangyu Zhang[1], Xiaolin Li[1], Zhuang Liu[1], Paul J Utz[2], Kaili Jiang[3], Shoushan Fan[3] & Hongjie Dai[1]

The current sensitivity of standard fluorescence-based protein detection limits the use of protein arrays in research and clinical diagnosis. Here, we use functionalized, macromolecular single-walled carbon nanotubes (SWNTs) as multicolor Raman labels for highly sensitive, multiplexed protein detection in an arrayed format. Unlike fluorescence methods, Raman detection benefits from the sharp scattering peaks of SWNTs with minimal background interference, affording a high signal-to-noise ratio needed for ultra-sensitive detection. When combined with surface-enhanced Raman scattering substrates, the strong Raman intensity of SWNT tags affords protein detection sensitivity in sandwich assays down to 1 fM—a three-order-of-magnitude improvement over most reports of fluorescence-based detection. We use SWNT Raman tags to detect human autoantibodies against proteinase 3, a biomarker for the autoimmune disease Wegener's granulomatosis, diluted up to 10^7-fold in 1% human serum. SWNT Raman tags are not subject to photobleaching or quenching. By conjugating different antibodies to pure ^{12}C and ^{13}C SWNT isotopes, we demonstrate multiplexed two-color SWNT Raman-based protein detection.

Several methods can detect proteins for proteomic and clinical diagnostic applications. These include enzyme-linked immunosorbent assays (ELISAs), fluorescence-based protein microarrays[1], electrochemistry[2], label-free optical methods[3,4], surface-enhanced Raman scattering (SERS)[5–7], microcantilevers[8], quantum dots[9,10] and nanotube-[11,12] or nanowire-based[13] field-effect transistors. Among these, protein microarrays[14,15] are the most common methods providing high-throughput, multiplexed protein detection for a range of applications[14,16,17]. Typically, undesirable background interference or autofluorescence resulting from assay reagents and materials limits the sensitivity of protein arrays based on fluorophore tags to ~1 pM[18]. Increasing the sensitivity of protein detection in arrayed format could enhance the capability of this technology for proteomics research. Moreover, identification of soluble biomarkers for many diseases has increased clinical demands for high sensitivity and selectivity in protein detection to facilitate minimally invasive risk assessment, early-stage disease diagnosis and monitoring of responses to therapeutic interventions[19]. In addition to improving detection sensitivity, protein sensor platforms for diagnosis and research applications would benefit greatly from expanded dynamic ranges, which allow more samples to be compared simultaneously with the same standard set, thereby increasing throughput and reducing the quantity of reagents required.

At least two methodologies under development show particular promise for highly sensitive protein detection in research and clinical applications. Although label-free, nanowire-based transistors[13] demonstrate femtomolar sensitivity, this sensitivity is limited to samples in pure water or low-salt solutions, and cannot be achieved in serum or other physiological fluids. The second strategy, amplified detection based upon multifunctional nanoparticles, has even greater sensitivity[10] but requires multiple reagents and is very time consuming. Our methodology, based upon Raman scattering, has simpler requirements, may be easily multiplexed and exhibits high sensitivity in clinically relevant samples over the nM to fM range.

The SERS effect provides the potential for rapid, high-throughput, sensitive protein detection. Although SERS has been applied for immobilized protein detection by coupling small Raman-active dyes to gold nanoparticles functionalized by ligands[5,6], the utility of these sensors is limited by the weak intensities of typical Raman labels, owing to their small Raman scattering cross-sections[20]. As a result, sensitivity is not quantitative[5], or is limited to the nM range[5], which does not compare favorably with fluorescence methods. For high-sensitivity Raman sensing with dye molecules[7], long acquisition times are needed and molecules are subject to degradation, resulting from laser radiation.

In contrast, SWNTs are ideal labels for SERS-based protein detection. SWNTs have a unique one-dimensional structure and exhibit distinct electrical and spectroscopic properties, including strong and simple resonance Raman signatures. SWNTs possess enormous Raman scattering cross-sections ($\sim 10^{-21}$ cm^2 sr^{-1} molecule^{-1}), have simple and tunable spectra and are more stable than other organic Raman labels[21,22]. Various schemes have been explored to develop

[1]Department of Chemistry and Laboratory for Advanced Materials, Stanford University, 333 Campus Drive, Mudd Building, Room 121, Stanford, California 94305, USA. [2]School of Medicine, Stanford University, 269 Campus Drive, Stanford, California 94305, USA. [3]Department of Physics, Tsinghua-Foxconn Nanotechnology Research Center, Tsinghua University, Beijing 100084, China. [4]These authors contributed equally to the work. Correspondence should be addressed to H.D. (hdai@stanford.edu).

Received 4 June; accepted 18 September; published online 26 October 2008; doi:10.1038/nbt.1501

Non-Markovian noise mediated through anomalous diffusion within ion channels

Sisir Roy,[1,*] Indranil Mitra,[2,†] and Rodolfo Llinas[3,‡]

[1]*Physics and Applied Mathematics Unit, Indian Statistical Institute, 203 Barrackpore Trunk Road, Kolkata 700108, India*

[2]*Brain & Behavior Program, Department of Physics, Georgia State University, 29 Peachtree Center Avenue, Atlanta, Georgia 30303, USA*

[3]*New York University School of Medicine, 530 First Avenue, New York, New York 10016, USA*

(Received 23 May 2008; published 28 October 2008)

It is evident from a wide range of experimental findings that ion channel gating is inherently stochastic. The issue of "memory effects" (diffusional retardation due to local changes in water viscosity) in ionic flow has been recently addressed using Brownian dynamics simulations. The results presented indicate such memory effects are negligible, unless the diffusional barrier is much higher than that of free solute. In this paper using differential stochastic methods we conclude that the Markovian property of exponential dwell times gives rise to a high barrier, resulting in diffusional memory effects that cannot be ignored in determining ionic flow through channels. We have addressed this question using a generalized Langevin equation that contains a combination of Markovian and non-Markovian processes with different time scales. This approach afforded the development of an algorithm that describes an oscillatory ionic diffusional sequence. The resulting oscillatory function behavior, with exponential decay, was obtained at the weak non-Markovian limit with two distinct time scales corresponding to the processes of ionic diffusion and drift. This will be analyzed further in future studies using molecular dynamics simulations. We propose that the rise of time scales and memory effects is related to differences of shear viscosity in the cytoplasm and extracellular matrix.

DOI: 10.1103/PhysRevE.78.041920 PACS number(s): 87.10.−e, 87.16.Xa

I. INTRODUCTION

Ion channels are transmembrane proteins that include a pore-forming subunit that allows ions to flow between the extracellular and intracellular and interior of a cell. The open or closed condition of such channels may be gated by either the transmembrane electric field, via a dipole moment sensitive moiety, or via ligand interactions with "chemical sensing" moieties in the channel proteins. Ion channel pores present a narrow cross section (100 Å) and define a path of low dielectric constant across the membrane. When open, the channel pore presents a rather specific ion selectivity filter where the lines of the electric field tend to be confined to the high dielectric interior of the pore. This paper addresses the problem of pore selectivity.

The continuity requirement for the orthogonal component of the electric displacement field between the interior of a channel and membrane is given by $\epsilon_w \mathbf{E}_w^n = \epsilon_p \mathbf{E}_m^n$. Since the lipid membrane has a dielectric constant $\epsilon_p \approx 2$, while the dielectric constant of water is $\epsilon \approx 80$, it becomes evident that the orthogonal component of the electric field at the membrane pore boundary must be very close to zero. Indeed, there is only a very slight penetration of the electric field into the interior of the phospholipid membrane. The situation, therefore, is very similar to the expulsion of the magnetic field by a superconductor. As an example, in a channel with a 3 Å radius and a channel of length $L=25$ Å, the barrier is about $6k_BT$. Although it is quite large, it should allow ionic conductivity. This is not too different from such conditions where water filled nanopores are introduced into silicon oxide films, polymer membranes, etc. [1,2].

Methods ranging from molecular dynamics (MD) and Brownian dynamics (BD) (which treat water implicitly as a uniform dielectric continuum) to the mean-field Poisson-Nernst Planck equation (PNP) (which treats both the ions and water implicitly as a uniform dielectric continuum) have been used to simulate the characteristics of ion channel phenomena. While MD simulations are clearly the most accurate, they are computationally very intensive [3]. Brownian dynamics are significantly faster than MD, but because of the dielectric discontinuities across the various interfaces, a new solution of the Poisson equation is required for each configuration of the ionic-pore profiles during permeation. The simplest approach to study the ionic conduction is based on the PNP theory [4,5]. This combines the continuity equation with the Poisson equation and Ohm's and Fick's laws. PNP is intrinsically mean field and is, therefore, bound to fail when ionic correlations become important.

For narrow channels, the cylindrical geometry, combined with the field confinement, results in a pseudo one-dimensional potential of very long range [6]. Under these conditions the correlational effects dominate, and the mean-field approximation fails [7]. Indeed, a recent comparison between the BD and the PNP showed that the PNP breaks down when the pore radius is smaller than about two debye lengths [8,9]. At the moment, therefore, for narrow pores it appears that a semicontinuum (implicit solvent) BD simulation is best compromised between computational load and accuracy [10–12]. If the interaction potential between the ions inside the channel were known, the simulation could proceed orders of magnitude faster.

Note that though BD simulations of ion channels have yielded suitable results, these studies have been confined to the use of the Langevin equation with Markovian random

*sisir@isical.ac.in
†imitra@gsu.edu
‡llinar01@med.nyu.edu

Extensive genomic copy number variation in embryonic stem cells

Qi Liang, Nathalie Conte, William C. Skarnes, and Allan Bradley[1]

Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom

Edited by Kathryn V. Anderson, Sloan-Kettering Institute, New York, NY, and approved September 12, 2008 (received for review June 10, 2008)

Recent analysis of the human and mouse genomes has revealed that highly identical duplicated elements account for >5% of the sequence content. These elements vary in copy number between individuals. Copy number variations (CNVs) contribute significantly to genetic differences among individuals and are increasingly recognized as a causal factor in human diseases with different etiologies. In inbred mouse strains, CNVs have been fixed by inbreeding, but they are highly variable among strains. Within strains, de novo germ-line CNVs can occur, leading to interindividual variation. By analyzing the genome of clonal isolates of mouse ES cells derived from common parental lines, we have uncovered extensive and recurrent CNVs. This variation arises during mitosis and can be cotransmitted into the mouse germ line along with engineered alleles, contributing to genetic variability. The frequency and extent of these genomic changes in ES cells suggests that all somatic tissues in individuals will be mosaics composed of variants of the zygotic genome. Human ES (hES) cells and derived somatic lineages may be similarly affected, challenging the concept of a stable somatic genome.

inbred mouse strains | comparative genomic hybridization | BAC arrays

Copy number variation (CNV) of DNA segments in the human genome can involve large segments of DNA (1–5) that occur in phenotypically normal individuals and these can be disease-associated (6). Within an inbred mouse strain, CNVs also occur among individuals (7, 8) and are presumed to arise de novo by meiotic homologous recombination between nearly identical duplicated sequences and through nonhomologous end-joining (9–12). Although the frequency of homologous recombination is several orders of magnitude greater than the single-nucleotide mutation rate, many duplications contain active genes; thus, a CNV arising de novo is expected to contribute more phenotypic variation (on average) than single-nucleotide alterations.

ES cell lines established from several different strains have been the major route through which thousands of new mutations have been established in the mouse germ line over the last 2 decades. To limit the impact of interstrain variation, the programs generating genome-wide resources of knockout alleles (13) use ES cells from a single genetic background, C57BL6/N. The underlying genetic stability of ES cell lines used for these resources is critical, because the modified ES cell clones used to establish new germ-line alleles should be genetically identical to the parental cell lines.

Compared with other cultured cell lines, ES cells are relatively stable. Karyotypic variants that arise, such as trisomies or loss of the Y chromosome (14, 15), preclude germ-line transmission; thus, they do not impact genetic studies. However, very little is known about other types of structural variation, such as CNV, which is likely to be compatible with germ-line transmission.

In humans, CNVs have been shown to have a meiotic origin (16). Recently, the comparison of monozygotic twins has identified CNVs, suggesting these genome alterations also occur during somatic development (17). Because recombination between duplicated sequences occurs at measurable frequencies in mouse ES cells (18), it would be expected that CNVs would be generated de novo in ES cells in culture. Given that ES cells undergo 30–40 mitotic divisions in vitro before resuming their normal developmental route into the germ line, we reasoned that ES cells may accumulate CNVs, which might be transmitted into the mouse germ line along with targeted alleles, contributing phenotypic variability to the analysis of mutant phenotypes.

Results

CNVs in ES Cell Clones. We examined the genomes of 50 different ES cell clones for evidence of CNV by comparative genomic hybridization (CGH) against BAC arrays. Given that some types of variation may be incompatible with germ-line transmission, and we wished to formulate an unbiased view of the extent of this variation, we analyzed clones with confirmed and compromised germ-line potential. To maximize the chance of discovering independent events, the clones examined in this study were derived from 3 different parental lines: 2 widely used lines, AB2.2, E14, and the JM8 line that is the basis for the EUCOMM and KOMP mutation resource.

Of 26 clones that could not contribute to the mouse germ line, trisomies were detected in 7 which involved chromosomes 1, 6, 8, and 11, [supporting information (SI) Fig. S1 and Tables S1 and S2]. In 5 cases, loss of the Y chromosome was detected. These types of aneuploidy have been observed previously, and they explain the germ-line transmission failure of nearly half of these ES clones. In addition to gains and losses of whole chromosomes, 14 germ-line-compromised clones exhibited subchromosomal changes of 3- to 5-Mb intervals (deletions or duplications) (Fig. S2 and Tables S1 and S2). These smaller genomic alterations may directly explain the germ-line transmission failure of these clones, or they may identify breakpoints of more substantial structural rearrangements, such as inversions or translocations that cannot be directly detected by CGH. In total, CGH analysis identified genetic changes in 19 of these 26 clones.

Next, we analyzed the 24 germ-line-competent clones and, as expected, none of these clones exhibited gains or losses of entire chromosomes. However, 7 of these had 1- to 2-Mb deletions and/or duplications (Table S2). Subclones derived from all 3 parental cell lines exhibited this type of variation, including 2 clonal isolates of the recently derived JM8 cell line. A total of 9 different CNVs were detected in 7 of the 24 germ-line-transmittable ES cell clones. Five of these variants were also detected in germ-line-compromised clones, indicating these specific changes could not explain the germ-line transmission failure of these clones. However, the number and size of the variants were greater in the germ-line-compromised clones.

Author contributions: A.B. designed research; Q.L. and A.B. performed research; N.C. and W.C.S. contributed new reagents/analytic tools; Q.L., N.C., W.C.S., and A.B. analyzed data; and Q.L. and A.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: abradley@sanger.ac.uk.

This article contains supporting information online at www.pnas.org/cgi/content/full/0805638105/DCSupplemental.

The chemistrode: A droplet-based microfluidic device for stimulation and recording with high temporal, spatial, and chemical resolution

Delai Chen[a,1], Wenbin Du[a,1], Ying Liu[a,1], Weishan Liu[a,1], Andrey Kuznetsov[b], Felipe E. Mendez[b], Louis H. Philipson[b], and Rustem F. Ismagilov[a,2]

[a]Department of Chemistry and Institute for Biophysical Dynamics, University of Chicago, 929 East 57th Street, Chicago, IL 60637; and [b]Department of Medicine and the Kovler Diabetes Center, University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637

Edited by George M. Whitesides, Harvard University, Cambridge, MA, and approved September 15, 2008 (received for review August 19, 2008)

Microelectrodes enable localized electrical stimulation and recording, and they have revolutionized our understanding of the spatiotemporal dynamics of systems that generate or respond to electrical signals. However, such comprehensive understanding of systems that rely on molecular signals—e.g., chemical communication in multicellular neural, developmental, or immune systems—remains elusive because of the inability to deliver, capture, and interpret complex chemical information. To overcome this challenge, we developed the "chemistrode," a plug-based microfluidic device that enables stimulation, recording, and analysis of molecular signals with high spatial and temporal resolution. Stimulation with and recording of pulses as short as 50 ms was demonstrated. A pair of chemistrodes fabricated by multilayer soft lithography recorded independent signals from 2 locations separated by 15 μm. Like an electrode, the chemistrode does not need to be built into an experimental system—it is simply brought into contact with a chemical or biological substrate, and, instead of electrical signals, molecular signals are exchanged. Recorded molecular signals can be injected with additional reagents and analyzed off-line by multiple, independent techniques in parallel (e.g., fluorescence correlation spectroscopy, MALDI-MS, and fluorescence microscopy). When recombined, these analyses provide a time-resolved chemical record of a system's response to stimulation. Insulin secretion from a single murine islet of Langerhans was measured at a frequency of 0.67 Hz by using the chemistrode. This article characterizes and tests the physical principles that govern the operation of the chemistrode to enable its application to probing local dynamics of chemically responsive matter in chemistry and biology.

analysis | dispersion | flow | microscale | pulse

This article describes the "chemistrode," a droplet-based microfluidic device for manipulating and observing molecular signals with high spatial and temporal resolution. The microelectrode, voltage-clamp, and patch-clamp techniques (1) enabled stimulation and recording of electrical activity and redox-active molecules with high resolution in both space and time, revolutionizing our understanding of electroactive processes from biochemistry to neuroscience (1–3). Most biological processes, however, are fundamentally chemical rather than electrical, relying on molecular signals to orchestrate events at the correct time and location. Electrochemical approaches are widely used, but not all molecules are electrochemically active, and some electrochemically active molecules are difficult to measure selectively in complex mixtures. The grand challenge this article addresses is that of devising an analogue of the electrode that operates on molecular rather than electrical or electroactive signals.

Why could we not build such a system with today's technology? Whereas electrical signals travel through wires essentially instantly and with low losses, manipulation and transport of molecules is more challenging. First, a pulse of molecules, especially of small volume, rapidly disperses when transported through a tube by laminar flow, leading to loss of concentration and time resolution. Loss of molecules from solution by adsorption to surfaces of tubes may also occur. Therefore, methods that rely on laminar flow to transport molecular signals, such as direct sampling (4), push/pull perfusion (5), microdialysis (6), and direct microinjection, have not addressed this grand challenge. In contrast to electrical signals, molecular signals comprise multiple, often unknown, molecular species, requiring the ability to deliver multiple molecular species as stimuli and the ability to analyze a pulse of response molecules by multiple techniques. Advances in optical imaging technology, new probes and tagging methods, and photo-controllable manipulation have enabled observation and manipulation of many known molecular species, but these technologies may be time consuming to develop for each species and difficult to use for multiple or unknown species. "Biology on a chip" microfluidics technologies (7–9) can reduce dispersion by minimizing the distance that molecules are transported by the integration of a biological experiment with a specific analytical method. However, this approach requires the redevelopment and validation of the biological protocols as well as the miniaturization and integration of disparate analytical technologies. Recent advances in microfluidics have used multiphase flow to transport solutions reliably as discrete units without dilution, cross-contamination, or loss of temporal resolution (10–19).

We developed the chemistrode, a microfluidic platform that addresses this grand challenge by providing molecular stimulation and recording with high fidelity using plug-based (12) multiphase microfluidics [Fig. 1*A* and supporting information (SI) Fig. S1]. Like the electrode, the chemistrode is simply brought into contact with the surface under investigation, e.g., a cell or tissue. Instead of exchanging electrical signals, molecular signals are delivered by and captured in plugs, aqueous droplets nanoliters in volume surrounded by a fluorocarbon carrier fluid. The compartmentalization of these molecular signals eliminates dispersion and loss of sample due to surface adsorption (18).

Operation of the chemistrode relies on 9 general steps (Fig. 1 *A* and *B*): (*i*) preparation of an array of aqueous plugs containing an arbitrary sequence of stimuli (20, 21); (*ii*) delivery of the array of stimulus plugs to a hydrophilic substrate; (*iii*) coalescence of

Author contributions: D.C., W.D., Y.L., W.L., A.K., L.H.P., and R.F.I. designed research; D.C., W.D., Y.L., W.L., and F.E.M. performed research; D.C., W.D., Y.L., W.L., A.K., L.H.P., and R.F.I. analyzed data; and D.C., W.D., Y.L., W.L., and R.F.I. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]D.C., W.D., Y.L., and W.L. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: r-ismagilov@uchicago.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0807916105/DCSupplemental.

Signatures of combinatorial regulation in intrinsic biological noise

Aryeh Warmflash and Aaron R. Dinner[1]

James Franck Institute, University of Chicago, Chicago, IL 60637

Communicated by Stuart A. Rice, University of Chicago, Chicago, IL, September 19, 2008 (received for review June 10, 2008)

Gene expression is controlled by the action of transcription factors that bind to DNA and influence the rate at which a gene is transcribed. The quantitative mapping between the regulator concentrations and the output of the gene is known as the *cis*-regulatory input function (CRIF). Here, we show how the CRIF shapes the form of the joint probability distribution of molecular copy numbers of the regulators and the product of a gene. Namely, we derive a class of fluctuation-based relations that relate the moments of the distribution to the derivatives of the CRIF. These relations are useful because they enable statistics of naturally arising cell-to-cell variations in molecular copy numbers to substitute for traditional manipulations for probing regulatory mechanisms. We demonstrate that these relations can distinguish super- and subadditive gene regulatory scenarios (molecular analogs of AND and OR logic operations) in simulations that faithfully represent bacterial gene expression. Applications and extensions to other regulatory scenarios are discussed.

gene expression | mathematical modeling | noise analysis inference | flow cytometry

Transcription factors regulate the expression of genes by binding to specific sites on the DNA that are typically spatially close to, sometimes even in, sequences that code for proteins. A group of such binding sites is collectively known as a *cis*-regulatory region. When a gene processes the effects of multiple transcription factors, one can view it as performing a computation. The inputs are the occupancies of the binding sites in the *cis*-regulatory region and the output is the rate of transcription. The simplest method of describing the relationship between the transcription factors and the output is by using boolean logic (1, 2). For example, 2 activators can regulate a gene with AND logic in which both are required or with OR logic in which either transcription factor is sufficient for transcription. Knowing which mode of combinatorial regulation a gene employs can be important for determining its function in regulatory networks. For example, the coherent feed forward loop, one of the most common motifs in gene regulatory networks (3), filters noise in upstream signals differently depending on how one of the participating genes integrates its inputs (4, 5).

In general, the output of a gene will not be binary, and will depend on the concentrations of the transcription factors in a complex manner. The notion of logic operations can be generalized by introducing a continuous function that encodes the dependence of the rate of transcription on the concentrations of the inputs. Such "*cis*-regulatory input functions (CRIFs)" (5) have been evaluated experimentally for the well-studied *lac* operon. The CRIF of the wild type is complex (6, 7), but those of certain mutant operons represent molecular analogs of binary logic gates (8). In higher organisms, many genes are regulated by a large number of transcription factors, often with multiple binding sites for each factor (9–11). It is increasingly clear that the action of many such regulators depends on their molecular context (12–19). Because CRIFs are capable of encoding arbitrarily complex modes of regulation, they will likely be even more useful for describing the control of transcription in these systems than in the simpler prokaryotic ones.

To date, the notion of a CRIF has largely been restricted to studying how the average rate of transcription depends on the average concentration of the transcription factor inputs by using bulk measurements. However, experiments performed on single cells reveal that because transcription factors are often present in low copy numbers, stochastic fluctuations in the concentrations of these molecules can have important consequences for gene regulation (20–25). Although a dialog between theory and experiment has revealed how noise enters the basic processes of transcription and translation (26, 27) and propagates through a cascade of genes (23, 28, 29), the relationship between combinatorial regulation and stochastic fluctuations has received much less attention. In particular, it is unclear how the mode of regulation encoded in the CRIF is related to the statistics that characterize the distribution of protein concentrations across a population of cells.

Our approach to this area is motivated by the fact that, at equilibrium, the fluctuations of a property of a material can be related to the response to changes in a conjugate control parameter. A well-known example is the relation between the fluctuations in the energy of a system and its heat capacity. More generally, the fluctuation–dissipation theorem relates fluctuations about equilibrium to the relaxation of the system after it has been perturbed from equilibrium (30). General fluctuation–dissipation relations for linear (or linearized) chemical systems that relate the average values of the chemical concentrations to the second moments can also be derived (31) and have been applied to biological systems (29, 32).

Here, we use a stochastic model of a gene regulated by an arbitrary number of transcription factors to show that logic operations are most closely related to the derivatives of the CRIF. In turn, the derivatives can be related to higher moments of the distribution of input and output copy numbers in analogy to molecular systems at equilibrium. These facts suggest means for inferring regulatory synergies from measurements that report on single cells. Using simulations of simple constructs with reasonable parameters, we demonstrate that the signatures of combinatorial logic should be detectable with presently available experimental methods. This is useful because proximity in DNA binding is not sufficient to infer combinatorial interactions, and they cannot be readily probed by traditional methods (e.g., knockouts) or high-throughput expression assays (e.g., microarray data). Broader implications for the work are discussed.

Results

In this section, we develop a model of gene expression and use it to derive a general expression relating the distribution of protein copy numbers to the derivatives of the CRIF. We then illustrate the utility of this relationship by considering idealized logic gates. Finally, predictions are made for an actual system to enable validation of the theory and its assumptions.

Author contributions: A.W. and A.R.D. designed research; A.W. performed research; and A.W. and A.R.D. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: dinner@uchicago.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0809314105/DCSupplemental.

Analytical distributions for stochastic gene expression

Vahid Shahrezaei[1] and Peter S. Swain[2,3]

Centre for Non-linear Dynamics, Department of Physiology, McGill University, 3655 Promenade Sir William Osler, Montreal, QC, Canada, H3G 1Y6

Edited by Charles S. Peskin, New York University, New York, NY, and approved September 5, 2008 (received for review April 22, 2008)

Gene expression is significantly stochastic making modeling of genetic networks challenging. We present an approximation that allows the calculation of not only the mean and variance, but also the distribution of protein numbers. We assume that proteins decay substantially more slowly than their mRNA and confirm that many genes satisfy this relation by using high-throughput data from budding yeast. For a two-stage model of gene expression, with transcription and translation as first-order reactions, we calculate the protein distribution for all times greater than several mRNA lifetimes and thus qualitatively predict the distribution of times for protein levels to first cross an arbitrary threshold. If in addition the fluctuates between inactive and active states, we can find the steady-state protein distribution, which can be bimodal if fluctuations of the promoter are slow. We show that our assumptions imply that protein synthesis occurs in geometrically distributed bursts and allows mRNA to be eliminated from a master equation description. In general, we find that protein distributions are asymmetric and may be poorly characterized by their mean and variance. Through maximum likelihood methods, our expressions should therefore allow more quantitative comparisons with experimental data. More generally, we introduce a technique to derive a simpler, effective dynamics for a stochastic system by eliminating a fast variable.

intrinsic noise | bursts | master equation | adiabatic approximation

Gene expression in both prokaryotes and eukaryotes is inherently stochastic (1–4). This stochasticity is both controlled and exploited by cells and, as such, must be included in models of genetic networks (5, 6). Here we will focus on describing intrinsic fluctuations, those generated by the random timing of individual chemical reactions, but extrinsic fluctuations are equally important and arise from the interactions of the system of interest with other stochastic systems in the cell or its environment (7, 8). Typically, experimental data are compared with predictions of mean behaviors and sometimes with the predicted standard deviation around this mean, because protein distributions are often difficult to derive analytically, even for models with only intrinsic fluctuations.

We will propose a general, although approximate, method for solving the master equation for models of gene expression. Our approach exploits the difference in lifetimes of mRNA and protein and is valid when the protein lifetime is greater than the mRNA lifetime. Typically, proteins exist for at least several mRNA lifetimes, and protein fluctuations are determined by only time-averaged properties of mRNA fluctuations. Following others (7, 9–11), we will use this time-averaging to simplify the mathematical description of stochastic gene expression.

For many organisms, single-cell experiments have shown that gene expression can be described by a three-stage model (3, 4, 12–14). The promoter of the gene of interest can transition between two states (10, 15–17), one active and one inactive. Such transitions could be from changes in chromatin structure, from binding and unbinding of proteins involved in transcription (3, 4, 12), or from pausing by RNA polymerase (18). Transcription can only occur if the promoter region is active. Both transcription and translation, as well as the degradation of mRNAs and proteins, are usually modeled as first-order chemical reactions (5).

By taking the limit of a large ratio of protein to mRNA lifetimes, we will study the three-stage model and a simpler two-stage version where the promoter is always active. For this two-stage model, we will derive the protein distribution as a function of time. We will derive the steady-state protein distribution for the full, three-stage model. We also include expressions for the corresponding mRNA distributions (14, 16) in the supporting information (SI) *Appendix*.

A Two-Stage Model of Gene Expression

We will first consider the model of gene expression in Fig. 1*A* (9). This model assumes that the promoter is always active and so has two stochastic variables: the number of mRNAs and the number of proteins. The probability of having m mRNAs and n proteins at time t satisfies a master equation:

$$\begin{aligned}\frac{\partial P_{m,n}}{\partial t} &= \nu_0(P_{m-1,n} - P_{m,n}) + \nu_1 m(P_{m,n-1} - P_{m,n}) \\ &\quad + d_0[(m+1)P_{m+1,n} - mP_{m,n}] \\ &\quad + d_1[(n+1)P_{m,n+1} - nP_{m,n}] \end{aligned} \qquad [1]$$

with ν_0 being the probability per unit time of transcription, ν_1 being the probability per unit time of translation, d_0 being the probability per unit time of degradation of an mRNA, and d_1 being the probability per unit time of degradation of a protein. By defining the generating function, $F(z',z)$, by $F(z',z) = \sum_{m,n} z'^m z^n P_{m,n}$, we can convert Eq. **1** into a first-order partial differential equation:

$$\frac{\partial F}{\partial \nu} - \gamma\left[b(1+u) - \frac{u}{\nu}\right]\frac{\partial F}{\partial u} + \frac{1}{\nu}\frac{\partial F}{\partial \tau} = a\frac{u}{\nu}F, \qquad [2]$$

where we have rescaled (19), with $a = \nu_0/d_1, b = \nu_1/d_0, \gamma = d_0/d_1$, and $\tau = d_1 t$, and where $u = z' - 1$ and $\nu = z - 1$.

If the protein lifetime is much greater than the mRNA lifetime and $\gamma \gg 1$, Eq. **2** can be solved by using the method of characteristics. Let r measure the distance along a characteristic which starts at $\tau = 0$ with $u = u_0$ and $\nu = \nu_0$ for some constant u_0 and ν_0, then Eq. **2** is equivalent to (20)

$$\begin{aligned} &\frac{d\nu}{dr} = 1; && \frac{d\tau}{dr} = \frac{1}{\nu} \\ &\gamma^{-1}\frac{du}{dr} = \frac{u}{\nu} - b(1+u); && \frac{dF}{dr} = \frac{au}{\nu}F. \end{aligned} \qquad [3]$$

Consequently direct integration implies $r = \nu = \nu_0 e^{\tau}$. For $\gamma \gg 1$, $u(\nu)$ obeys (*SI Appendix*)

$$u(\nu) \simeq \left(u_0 - \frac{b\nu_0}{1-b\nu_0}\right)e^{-\gamma b(\nu-\nu_0)}\left(\frac{\nu}{\nu_0}\right)^{\gamma} + \frac{b\nu}{1-b\nu} \qquad [4]$$

or

$$u(\nu) \simeq \frac{b\nu}{1-b\nu} \qquad [5]$$

Author contributions: V.S. and P.S. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

[1]Present address: Department of Mathematics, Imperial College, London SW7 2BZ, UK.

[2]Present address: Center for Systems Biology, University of Edinburgh, Edinburgh EH9 3JU, UK.

[3]To whom correspondence should be addressed. E-mail: swain@cnd.mcgill.ca.

This article contains supporting information online at www.pnas.org/cgi/content/full/0803850105/DCSupplemental.

Small-scale copy number variation and large-scale changes in gene expression

Yuriy Mileyko[a,1], Richard I. Joh[b], and Joshua S. Weitz[a,b,1]

[a]School of Biology and [b]School of Physics, Georgia Institute of Technology, Atlanta, GA 30332

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved September 11, 2008 (received for review June 30, 2008)

The expression dynamics of interacting genes depends, in part, on the structure of regulatory networks. Genetic regulatory networks include an overrepresentation of subgraphs commonly known as network motifs. In this article, we demonstrate that gene copy number is an omnipresent parameter that can dramatically modify the dynamical function of network motifs. We consider positive feedback, bistable feedback, and toggle switch motifs and show that variation in gene copy number, on the order of a single or few copies, can lead to multiple orders of magnitude change in gene expression and, in some cases, switches in deterministic control. Further, small changes in gene copy number for a 3-gene motif with successive inhibition (the "repressilator") can lead to a qualitative switch in system behavior among oscillatory and equilibrium dynamics. In all cases, the qualitative change in expression is due to the nonlinear nature of transcriptional feedback in which duplicated motifs interact via common pools of transcription factors. We are able to implicitly determine the critical values of copy number which lead to qualitative shifts in system behavior. In some cases, we are able to solve for the sufficient condition for the existence of a bifurcation in terms of kinetic rates of transcription, translation, binding, and degradation. We discuss the relevance of our findings to ongoing efforts to link copy number variation with cell fate determination by viruses, dynamics of synthetic gene circuits, and constraints on evolutionary adaptation.

gene duplication | gene regulation | network motifs | nonlinear dynamics

Copy number variation (CNV) is an important and widespread component of within and between population genetic variation. The copy number of genes and gene fragments varies significantly over physiological to evolutionary time scales with multiple effects on phenotype. For example, CNV can cause statistically significant changes in concentrations of RNA associated with growth rate changes in bacteria (1, 2) and enzyme concentrations associated with nutrient intake in humans (3, 4). The copy number of viral genomes undergoes dynamical changes during multiple infection of bacteria by phages, leading to qualitative changes in gene regulation that may lead to alternative modes of exploitation (5, 6). The duplication of a gene can facilitate subsequent diversification—a mechanism considered to be a dominant cause of phenotypic innovation (7–10). In extreme cases, whole-genome duplications have led to lineage diversification within yeast (11). In humans, large-scale deletions and duplications of chromosomes are known to cause severe genetic disorders (12, 13) and are imputed in the onset of other diseases including cancer (14, 15). Finally, multiple studies have demonstrated that CNV in humans is far more extensive than previously believed, although its impact on phenotype is yet to be fully resolved (16–20).

Despite its ubiquity, CNV has been nearly universally overlooked in quantitative models of gene regulation. In those cases where quantitative models of CNV have been developed, the primary focus has been on changes in the copy number itself, as in the case of plasmid maintenance (21) and dynamics of transposable elements (22). In some instances, gene copy number is integrated into dynamic models of regulation to account for cell-to-cell variability of regulatory elements found on plasmids (23). More commonly, recent studies have attempted to identify statistical relations between CNV and fitness (4), protein interactions (24), or combinations of both (25). To understand the progression from CNV to changes in phenotype to changes in fitness, it seems necessary to carefully examine gene regulation itself. The dynamics of a gene regulatory network depends on network topology, the quantitative nature of feedbacks and interactions between DNA, RNA and proteins, epigenetic modifications of regulatory elements, and the biochemical state of the intracellular and surrounding environment. Additionally, as we argue here, gene regulatory dynamics can also depend sensitively on the copy number of genes and promoters. For example, in synthetically designed networks, small changes in the copy number of gene regulatory modules have been shown to lead to qualitative changes in gene expression (26, 27). In naturally occurring networks, there may be selection pressures on kinetic parameters such that normally occurring levels of copy number are far from or close to the critical threshold that would lead to a dramatic change in gene expression.

In this manuscript, we take a quantitative approach to assess when small changes in copy number can have a dramatic, nonlinear effect on gene expression. We study the effect of changing copy number within a series of small, regulatory networks commonly referred to as "network motifs" (28). These motifs are network subgraphs shown to be building blocks of complex regulatory networks (29). Increasing the number of motifs means that multiple networks are coupled together via a common pool of transcription factors. Changes in the number of promoter sites is directly linked to changes in the rate of regulated recruitment, which in turn leads to changes in translation and other transcriptional feedbacks (30). We demonstrate that small changes in gene copy number within motifs exhibiting positive and/or negative feedback can switch the network from one alternative steady state to another and switch gene expression to and from an oscillatory state. Thus, changes in copy number may act as knobs within a nonlinear dynamical system in much the same way that changes in environmental conditions can drive expression from one steady state to another (29, 31).

Results

CNV and Network Motifs. We systematically analyze the dependency of 4 network motifs, positive feedback, bistable feedback, toggle switch, and the repressilator, on the copy number, $\mathcal{N}$. The method for analyzing each of these motifs is largely the same, and illustrated in Fig. 1. Although $\mathcal{N}$ is not explicitly present in the mathematical models presented in Fig. 1 *A–D*, it factors in

Author contributions: Y.M. and J.S.W. designed research; Y.M., R.I.J., and J.S.W. performed research; and Y.M. and J.S.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. E-mail: yuriy.mileyko@biology.gatech.edu or jsweitz@gatech.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0806239105/DCSupplemental.

Calculation of switching times in the genetic toggle switch and other bistable systems

Baruch Barzel and Ofer Biham
Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel
(Received 24 January 2008; revised manuscript received 22 July 2008; published 28 October 2008)

Genetic circuits with feedback such as the toggle switch often exhibit bistability, namely, two stable states with rare spontaneous transitions between them. These systems can be characterized by the average time between such transitions (referred to as the switching time). However, commonly used deterministic models, based on rate equations, do not account for these fluctuation-induced transitions. Stochastic methods, such as the direct integration of the master equation, do account for the transitions. However, they cannot be used to evaluate the switching time. In order to obtain the switching time, one needs to use Monte Carlo simulations. These methods require the accumulation of statistical data, which limits their accuracy. They may become infeasible when the switching time is long. Here we present an accurate and efficient method for the calculation of the switching time. The method consists of coupled recursion equations for the transition times between microscopic states of the system. Using a suitable definition of the two macroscopic bistable states (in terms of the microscopic states) and the probabilities obtained from the master equation, the method provides the switching time between the two states of the system. The method is demonstrated for the genetic toggle switch. It can be used to evaluate the switching times in a broad range of bistable and multistable systems. We also show that it is suitable for the evaluation of the oscillation periods in oscillatory systems such as the repressilator.

DOI: 10.1103/PhysRevE.78.041919 PACS number(s): 87.10.−e, 87.16.−b

I. INTRODUCTION

Multistability is common in nonlinear systems with feedback loops, particularly those in which the overall feedback throughout the loop is positive. A multistable system exhibits several macroscopic states in which it may remain for a long time. It may switch between these states either spontaneously or as a result of an external trigger. The typical time between spontaneous transitions is one of the essential features that characterize a multistable system. Examples of bistable systems appear in the context of genetic regulatory networks, which describe the interactions between genes and their products in living cells [1]. These genes regulate each other's expression using a combination of transcriptional, post-transcriptional and post-translational regulation mechanisms. Here we focus on transcriptional regulation networks, where transcription factor proteins bind to specific promoter sites on the DNA and regulate the transcription of adjacent genes [2]. In the case of positive regulation, the bound transcription factor activates the transcription while in the case of negative regulation the transcription is suppressed. Genetic networks are complex and typically include thousands of genes. In order to understand the functionality of the network in the cell, it is useful to identify functional modules and analyze each one of them separately [3–5]. Such network modules are commonly simulated using rate equations which are simple and efficient [6–9]. They consist of coupled ordinary differential equations, which account for the temporal variations in the concentrations of molecules such as messenger RNA (mRNA) and proteins. For fixed environmental conditions, these equations often approach a single steady state solution. When some control parameter is varied, a bifurcation may take place, where two distinct stable solutions emerge. Beyond the bifurcation point, the system thus exhibits bistability [10–13]. Within the rate equation approach, these two solutions are completely stable. Thus, the rate equations do not account for spontaneous transitions, which result from stochastic fluctuations. Such fluctuations are important in genetic circuits due to the fact that transcription factors and their binding sites may appear in low copy numbers [14–18].

In order to account for the spontaneous switching events one needs to use stochastic methods. These methods can be implemented either by direct integration of the master equation [19–21] or by Monte Carlo (MC) simulations [22–25]. An advantage of the direct integration of the master equation is that it provides the complete probability distribution over the microscopic states of the system. From the moments of this distribution one can obtain the averages and standard deviations of the molecular concentrations as well as the rates of biochemical reactions. However, the solution of the master equation does not provide dynamical features such as the switching times. In order to obtain the switching times one must perform MC simulations and collect a sufficient amount of statistical data. This is a difficult task because the required simulation time increases with the switching time.

The recently proposed forward flux sampling technique provides a dramatic reduction in the simulation time required for sampling these rare switching events [26]. This technique is based on a ratchetlike mechanism. The transition path between the stable states is divided into sections and the simulation is advanced across each of these sections. The statistical information is then analyzed to obtain the average switching time.

In this paper we present a different method for the calculation of the switching times in bistable and multistable systems. The method is based on a set of coupled recursion equations, which includes one equation for each microscopic state of the system. The equation associated with each of the microscopic states provides the average time it takes the system to get from this state to a predefined target state. The

The Pervasive Effects of an Antibiotic on the Human Gut Microbiota, as Revealed by Deep 16S rRNA Sequencing

Les Dethlefsen[1,2], Sue Huse[3], Mitchell L. Sogin[3], David A. Relman[1,2,4*]

1 Department of Microbiology and Immunology, Stanford University, Stanford, California, United States of America, 2 Department of Medicine, Stanford University, Stanford, California, United States of America, 3 Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, Massachusetts, United States of America, 4 Veterans Affairs Palo Alto Health Care System, Palo Alto, California, United States of America

The human intestinal microbiota is essential to the health of the host and plays a role in nutrition, development, metabolism, pathogen resistance, and regulation of immune responses. Antibiotics may disrupt these coevolved interactions, leading to acute or chronic disease in some individuals. Our understanding of antibiotic-associated disturbance of the microbiota has been limited by the poor sensitivity, inadequate resolution, and significant cost of current research methods. The use of pyrosequencing technology to generate large numbers of 16S rDNA sequence tags circumvents these limitations and has been shown to reveal previously unexplored aspects of the "rare biosphere." We investigated the distal gut bacterial communities of three healthy humans before and after treatment with ciprofloxacin, obtaining more than 7,000 full-length rRNA sequences and over 900,000 pyrosequencing reads from two hypervariable regions of the rRNA gene. A companion paper in *PLoS Genetics* (see Huse et al., doi: 10.1371/journal.pgen.1000255) shows that the taxonomic information obtained with these methods is concordant. Pyrosequencing of the V6 and V3 variable regions identified 3,300–5,700 taxa that collectively accounted for over 99% of the variable region sequence tags that could be obtained from these samples. Ciprofloxacin treatment influenced the abundance of about a third of the bacterial taxa in the gut, decreasing the taxonomic richness, diversity, and evenness of the community. However, the magnitude of this effect varied among individuals, and some taxa showed interindividual variation in the response to ciprofloxacin. While differences of community composition between individuals were the largest source of variability between samples, we found that two unrelated individuals shared a surprising degree of community similarity. In all three individuals, the taxonomic composition of the community closely resembled its pretreatment state by 4 weeks after the end of treatment, but several taxa failed to recover within 6 months. These pervasive effects of ciprofloxacin on community composition contrast with the reports by participants of normal intestinal function and with prior assumptions of only modest effects of ciprofloxacin on the intestinal microbiota. These observations support the hypothesis of functional redundancy in the human gut microbiota. The rapid return to the pretreatment community composition is indicative of factors promoting community resilience, the nature of which deserves future investigation.

Citation: Dethlefsen L, Huse S, Sogin ML, Relman DA (2008) The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. PLoS Biol 6(11): e280. doi:10.1371/journal.pbio.0060280

Introduction

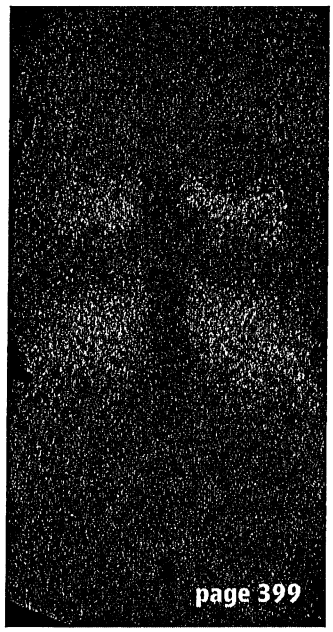
Specialized microbial communities inhabit the skin, mucosal surfaces, and gastrointestinal tract of humans (and other vertebrates) from birth until death, with by far the largest populations in the colon [1,2]. Humans rely on their native microbiota for nutrition and resistance to colonization by pathogens [3–6]; furthermore, recent discoveries have shown that symbiotic microbes make essential contributions to the development, metabolism, and immune response of the host [7–10]. Co-evolved, beneficial, human–microbe interactions can be altered by many aspects of a modern lifestyle, including urbanization, global travel, and dietary changes [1], but in particular by antibiotics [11]. The acute effects of antibiotic treatment on the native gut microbiota range from self-limiting "functional" diarrhea to life-threatening pseudomembranous colitis [12,13]. The long-term consequences of such perturbations for the human–microbial symbiosis are more difficult to discern, but chronic conditions such as asthma and atopic disease have been associated with childhood antibiotic use and an altered intestinal microbiota [14–16]. Because many chemical transformations in the gut are mediated by specific microbial populations [17], with implications for cancer [18,19] and obesity [20,21], among other conditions [22], changes in the composition of the gut microbiota could have important but undiscovered health effects. An approximate return to pretreatment conditions often (but not always) occurs within days or weeks after cessation of antibiotic treatment, as assessed by subjective judgments of bowel function and characterizations of overall

Academic Editor: Jonathan A. Eisen, University of California, Davis, United States of America

Received May 27, 2008; **Accepted** October 6, 2008; **Published** November 18, 2008

Abbreviations: AAD, antibiotic-associated diarrhea; Cp, ciprofloxacin; FDR, false discovery rate; OTU, operational taxonomic unit; PCA, principal component analysis

* To whom correspondence should be addressed. E-mail: relman@stanford.edu



INTRODUCTION

Getting Your Loops Straight

COMPLICATED BIOCHEMICAL SIGNALING PATHWAYS REGULATE THE FUNCTION OF living cells. Such regulatory networks often have "downstream" components that provide input to components that act earlier in a pathway, creating feedback loops. These feedback loops have the potential to greatly alter the properties of a pathway and how it responds to stimuli. To fully understand these regulatory systems and exploit their vast potential as targets of therapeutic strategies, we need quantitative information on the flow of signals through a pathway and on the timing and location of signaling events within cells. We need to explore the properties that determine, for example, whether a system shows a graded response to a stimulus or turns on and off like a switch, how long a pathway stays activated, whether its output oscillates, which components can be perturbed to control the output of the system, and so on. The papers assembled in this special issue and in the companion issue of *Science Signaling* highlight recent progress in tackling these challenges.

The systems-level approaches required to understand signaling networks often integrate mathematical modeling with traditional biochemical analysis. Brandman and Meyer (p. 390) describe the ways in which feedback loops allow sophisticated regulatory responses, such as adaptable sensors that respond to changes in the amplitude of an input signal rather than the absolute amount of that signal. Lewis (p. 399) summarizes recent examples in which modeling approaches allow new insights into classical problems in development. Spemann's organizer, for example, emits a gradient of signaling molecules, and recent work explains how the system adjusts when an embryo is damaged, to recreate a complete body axis. Tools that allow precise noninvasive control and monitoring of biochemical components facilitate sorting out how a system responds. As Gorostiza and Isacoff explain (p. 395), it is now possible not only to sample the output of signaling systems by monitoring the fluorescence of reporter molecules but also to engineer proteins with light-dependent isomerization switches that, when reintroduced into cells, can be controlled precisely in time and space by exposure to light.

At *Science Signaling* (see www.sciencemag.org/cellsignaling08), an original research paper by Abdi *et al.* takes strategies that engineers use to understand the vulnerability of digital circuits and applies them to biological systems to identify key elements of cellular signaling networks. In Perspectives, Dohlman discusses how scaffolding molecules can determine the graded or switchlike response of a pathway. Elston summarizes how responses to pulsatile inputs are used to understand the dynamic behavior of a signaling system in yeast. Chiang and Muir outline advances in chemical approaches providing deeper insight into signaling mechanisms.

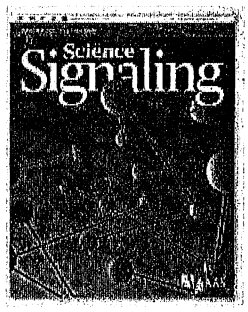
Getting your loops straight—or understanding how a complicated signaling network might respond in any given situation—does not come easily or intuitively. But the approaches mentioned here are clearly opening a large area of investigation of enormous practical potential.

– L. BRYAN RAY

Cell Signaling

CONTENTS

Reviews



Related material will be published in the 21 October issue of Science Signaling *(www.sciencesignaling.org).*

Science

REVIEW

Feedback Loops Shape Cellular Signals in Space and Time

Onn Brandman[1] and Tobias Meyer[2]

Positive and negative feedback loops are common regulatory elements in biological signaling systems. We discuss core feedback motifs that have distinct roles in shaping signaling responses in space and time. We also discuss approaches to experimentally investigate feedback loops in signaling systems.

Feedback loops are processes that connect output signals back to their inputs. The history of biological feedback goes back at least 130 years to observations by Eduard Pflüger that organs and other living systems "satisfy their own needs" (*1*). Feedback became an influential concept that led to Walter Cannon's theory of physiological homeostasis (*2*); Alan Turing's model of pattern formation (*3*); as well as investigations of metabolic end-product inhibition (*4*), metabolic oscillations (*5*), and transcriptional self-repression (*6*). Biological feedback concepts were further influenced by chemical oscillation theories (*7*) and the field of cybernetics (*8*). It has more recently become appreciated that the concept of feedback may be useful as a framework for understanding how intracellular signaling systems elicit specific cell behavior.

Mammalian species use over 3000 signaling proteins and over 15 second messengers to build hundreds of cell-specific signaling systems. Many of the signaling components have multiple upstream regulators and downstream targets, creating a web of connectivity within and between signaling pathways (*9*). The presence of multiple feedback loops in these systems (*10*) poses a challenge to understanding how receptor inputs control cellular behavior. We discuss how recurring feedback designs, or motifs (*11*), mediate biological functions such as bistability, oscillation, polarization, and robustness. Our goal was to generate a comprehensive guide for feedback in signal transduction that would also be instructive for understanding transcription networks, control of metabolism, pattern formation, the cell cycle, and the behavior of circadian oscillators.

We focus on two mammalian signaling systems: the receptor-triggered Ca^{2+} signaling system in nonexcitable cells (*12*) and the phosphoinositide 3-kinase (PI3K) signaling pathway in chemotactic neutrophils (*13*). These were chosen because of existing knowledge of feedback mechanisms that generate both simple and complex temporal and spatial signaling responses (Fig. 1). We use graphical representations of feedback motifs, with signaling components shown as vertices and directed negative and positive regulatory steps shown as arrows with and without minus symbols, respectively (Figs. 2 to 5, gray background). An arrow may consist of multiple steps so that a single positive arrow could reflect, for example, the net effect of two serial negative regulatory steps. Feedback loops are defined as paths that begin from and return to the same vertex. We also use mathematical representations with variables corresponding to vertices and equation terms corresponding to arrows (table S1). The simulated functions for each motif (Figs. 2 and 3) can be recreated by accompanying computer programs (table S2). We conclude by suggesting experimental approaches to investigate feedback loops.

Signaling with a Single Negative Feedback Loop

Negative feedback loops are found in nearly all known signaling pathways and are defined as sequential regulatory steps that feed the output signal, inverted, back to the input (Fig. 2A). Depending on its characteristics and initial conditions, a single negative feedback motif can create four distinct signaling functions: basal homeostat, output limiter, adaptation, and transient generator.

The presence of a small-amplitude negative feedback loop stabilizes the basal signaling state without preventing strong input signals from triggering maximal pathway activation (Fig. 2A, basal homeostat). In this case, small deviations of an input signal are suppressed in the output, and only large changes in the input control the output. For example, a negative feedback involving the endoplasmic reticulum (ER) Ca^{2+}-sensing protein STIM2 (stromal interaction molecule 2) (*14*) keeps basal Ca^{2+} concentration in the cytosol and in the lumen of the ER at ~50 nM and ~400 μM, respectively. STIM2 triggers influx of extracellular Ca^{2+} at ER–plasma membrane junctions in response to a reduction in ER Ca^{2+} concentration, forming a negative feedback loop (*14*). Because many cellular processes are regulated by Ca^{2+}, maintaining proper resting levels is crucial.

A different use of negative feedback is to limit maximum signaling output (Fig. 2A, limiter). Upon stimulation, the output signal rapidly increases but is attenuated once it passes a threshold. For example, receptor-triggered increases of

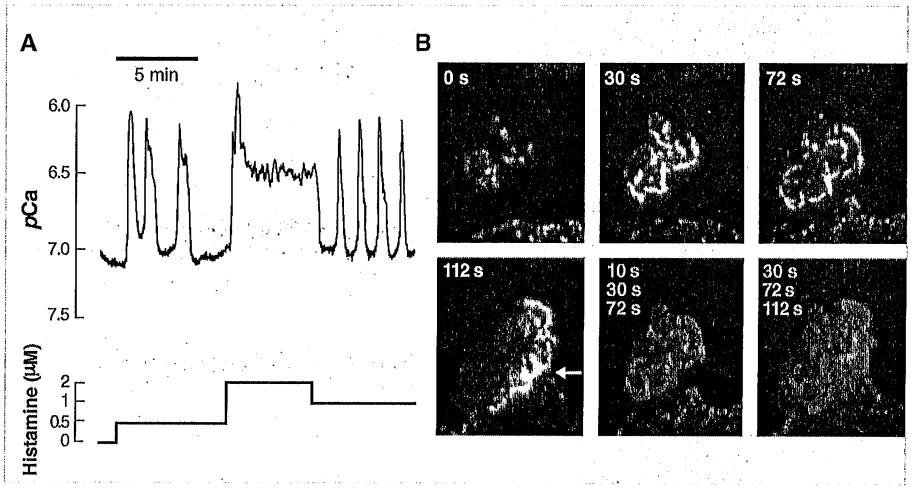


Fig. 1. Ca^{2+} and chemotaxis signaling systems exhibit complex temporal and spatial dynamics. **(A)** Histamine-triggered cytosolic Ca^{2+} signals in a single epithelial cell. The response includes spikes, oscillations, and plateaus. pCa is the negative log of the Ca^{2+} concentration (7 and 6 correspond to 10^{-7} and 10^{-6} M, respectively). [Adapted from (*28*)] **(B)** Total internal reflection fluorescence images of a neutrophil cells stimulated by chemoattractant. Hem-1 is a regulator of actin polymerization that initially concentrates in foci. It then relocalizes as part of outwardly propagating waves of actin polarization that terminate when the leading edge is reached (denoted by arrow at 112 s). The fifth and sixth images show overlays of successive Hem-1 distributions in red, blue, and green, respectively. [Adapted from (*29*)]

[1]Department of Cellular and Molecular Pharmacology, University of California-San Francisco and Howard Hughes Medical Institute, San Francisco, CA 94158, USA. E-mail: Onn.Brandman@ucsf.edu [2]Department of Chemical and Systems Biology, 318 Campus Drive, Clark Building W200, Stanford University Medical Center, Stanford, CA 94305–5174, USA. E-mail: tobias1@stanford.edu

able to many different protein classes, potentially including cytoplasmatic signaling proteins.

Conclusions

Remote and noninvasive manipulation of proteins with light provides a powerful approach for studying and reengineering signaling pathways by selectively establishing a fast and reversible remote control over specific proteins at specific locations within a cell or organism. The greatest specificity can be obtained from light-activated chemicals that are tethered to specific proteins. Naturally occurring photosensitive proteins have proven to be useful tools for experimentation and provide a major motivation to seek other naturally light-sensitive proteins. The success of synthetic PTL photoswitches introduces a tantalizing prospect of a general method that could be used to confer control of many proteins by light. The versatility of the approach comes from its superficiality. A PTL is attached to the bland surface of the protein in such a way that the covalent modification has little impact, except that presence of light of different wavelengths determines whether it presents or withdraws a ligand to and from a nearby binding site. The ligand can block or activate the site and thereby alter the function of the protein. The site can be an active site or an allosteric regulatory site. This logic should, in principle, be applicable across protein classes.

These optical approaches to the control of protein function provide an alternative to standard methods of controlling proteins through genetic or pharmacological means, which tend to be slow, difficult to reverse, and often not as specific as one would like. The contrast is particularly notable when one reflects that, in standard pharmacology, high specificity is associated with high affinity and thus with very slow dissociation and a lack of reversibility. In contrast, specificity of PTLs arises from the close relation between the geometry of the PTL and its change upon isomerization and the relative location of the attachment and ligand-binding sites. This even permits ligands of low affinity (or low specificity) to be used. The extremely rapid dissociation rates of such ligands and the subsequent fast reversal of signaling occur with rates similar to those of the fastest natural biological signals.

References and Notes

1. C. J. Weijer, *Science* **300**, 96 (2003).
2. C. Lutz *et al.*, *Nat. Methods* **5**, 821 (2008).
3. G. D. Housley, N. P. Raybould, P. R. Thorne, *Hear. Res.* **119**, 1 (1998).
4. J. Zhao *et al.*, *Biochemistry* **45**, 4915 (2006).
5. Y. Tong *et al.*, *J. Gen. Physiol.* **117**, 103 (2001).
6. S. Shoham, D. H. O'Connor, D. V. Sarkisov, S. S. Wang, *Nat. Methods* **2**, 837 (2005).
7. B. V. Zemelman, N. Nesnas, G. A. Lee, G. Miesenböck, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1352 (2003).
8. S. Q. Lima, G. Miesenböck, *Cell* **121**, 141 (2005).
9. M. Goeldner, R. Givens, *Dynamic Studies in Biology. Phototriggers, Photoswitches and Caged Biomolecules* (Wiley-VCH, Weinheim, Germany, 2005).
10. E. Bartels, N. H. Wassermann, B. F. Erlanger, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 1820 (1971).
11. M. Volgraf *et al.*, *J. Am. Chem. Soc.* **129**, 260 (2007).
12. G. Nagel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13940 (2003).
13. F. Zhang *et al.*, *Nat. Neurosci.* **11**, 631 (2008).
14. B. Schobert, J. K. Lanyi, *J. Biol. Chem.* **257**, 10306 (1982).
15. A. Bi *et al.*, *Neuron* **50**, 23 (2006).
16. F. Zhang *et al.*, *Nature* **446**, 633 (2007).
17. K. Deisseroth *et al.*, *J. Neurosci.* **26**, 10380 (2006).
18. L. Petreanu, D. Huber, A. Sobczyk, K. Svoboda, *Nat. Neurosci.* **10**, 663 (2007).
19. G. Nagel *et al.*, *Curr. Biol.* **15**, 2279 (2005).
20. S. Schroder-Lang *et al.*, *Nat. Methods* **4**, 39 (2007).
21. M. Bose, D. Groff, J. Xie, E. Brustad, P. G. Schultz, *J. Am. Chem. Soc.* **128**, 388 (2006).
22. A. Koçer, M. Walko, W. Meijberg, B. L. Feringa, *Science* **309**, 755 (2005).
23. K. Yoshimura, A. Batiza, C. Kung, *Biophys. J.* **80**, 2198 (2001).
24. G. A. Woolley, *Acc. Chem. Res.* **38**, 486 (2005).
25. P. Gorostiza, E. Y. Isacoff, *Mol. Biosyst.* **3**, 686 (2007).
26. P. Gorostiza, E. Y. Isacoff, *Physiology* **23**, 238 (2008).
27. I. Silman, A. Karlin, *Science* **164**, 1420 (1969).
28. H. A. Lester, M. E. Krouse, M. M. Nass, N. H. Wassermann, B. F. Erlanger, *J. Gen. Physiol.* **75**, 207 (1980).
29. D. A. Doyle *et al.*, *Science* **280**, 69 (1998).
30. R. O. Blaustein, P. A. Cole, C. Williams, C. Miller, *Nat. Struct. Biol.* **7**, 309 (2000).
31. M. Banghart, K. Borges, E. Isacoff, D. Trauner, R. H. Kramer, *Nat. Neurosci.* **7**, 1381 (2004).
32. J. J. Chambers, M. R. Banghart, D. Trauner, R. H. Kramer, *J. Neurophysiol.* **96**, 2792 (2006).
33. M. Volgraf *et al.*, *Nat. Chem. Biol.* **2**, 47 (2006).
34. P. Gorostiza *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 10865 (2007).
35. S. Szobota *et al.*, *Neuron* **54**, 535 (2007).
36. D. L. Fortin *et al.*, *Nat. Methods* **5**, 331 (2008).
37. G. C. Ellis-Davies, *Nat. Methods* **4**, 619 (2007).
38. G. Nagel *et al.*, *Biochem. Soc. Trans.* **33**, 863 (2005).

39. P.G. is supported by the Human Frontier Science Program Organization through a Career Development Award and by the European Research Council through a Starting Grant. This work was supported by the NIH Nanomedicine Development Center for the Optical Control of Biological Function (grant 5PN2EY018241).

10.1126/science.1166022

REVIEW

From Signals to Patterns: Space, Time, and Mathematics in Developmental Biology

Julian Lewis

We now have a wealth of information about the molecular signals that act on cells in embryos, but how do the control systems based on these signals generate pattern and govern the timing of developmental events? Here, I discuss four examples to show how mathematical modeling and quantitative experimentation can give some useful answers. The examples concern the Bicoid gradient in the early *Drosophila* embryo, the dorsoventral patterning of a frog embryo by bone morphogenetic protein signals, the auxin-mediated patterning of plant meristems, and the Notch-dependent somite segmentation clock.

Developmental biologists are preoccupied with patterns in space and patterns in time. What causes cells in one part of an embryo regularly to adopt one character, and those elsewhere to adopt another? What controls the timing of these patterning events as the organism develops from the egg? What halts growth when a structure has reached its proper size?

Early discussions of developmental patterning considered, in abstract terms and in total ignorance of the molecules involved, how signals exchanged between cells might drive formation of the patterns we observe. Although they lacked a concrete molecular basis, several of these speculative theories turn out to have been remarkably prescient. Moreover, right or wrong, they drive home a general lesson. To explain how embryos generate their spatial patterns and their temporal programs, we need more than lists of the molecules involved and more than simple cartoon diagrams of the qualitative control relationships between them. To make sense of the control circuitry, qualitative data and unaided intuition are not enough. Even for the simplest cases, we need mathematics and detailed measurements, just as we need mathematics and measurements to explain the orbits of the planets or the swing of a pendulum.

The necessity arises especially because of the fundamental part that feedback plays in developmental control processes, as discussed elsewhere in this issue of *Science* by Brandman and Meyer (*1*). Positive feedback can give a system a flip/flop choice between alternative steady states; it can endow a system with enduring memory of its exposure to past signals; it can generate inhomogeneity in a system that starts out spatially uniform. Negative feedback can smooth out irregularities; it can enable a system to respond to a signal more

Vertebrate Development Laboratory, Cancer Research UK London Research Institute, London WC2A 3PX, UK.

depends on its characteristics and initial conditions. Cell signaling systems are made up of multiple feedback functions that together generate a cell's input-output behavior. We propose that working models of signaling systems can be based on graphic representations that show signaling components as part of feedback motifs with assigned functions and sorted by increasing feedback time constants. If the known feedback functions are insufficient to explain a cell's input-output behavior, this analysis can guide the search to identify control loops missing from a signaling model.

References and Notes

1. E. F. W. Pflüger, *Pflügers Arch.* **XV**, 57 (1877).
2. W. B. Cannon, *Physiol. Rev.* **IX**, 399 (1929).
3. A. M. Turing, *Philos. Trans. R. Soc. London Ser. B* **237**, 37 (1952).
4. H. E. Umbarger, *Science* **123**, 848 (1956).
5. B. Chance, R. W. Estabrook, A. Ghosh, *Proc. Natl. Acad. Sci. U.S.A.* **51**, 1244 (1964).
6. J. Monod, F. Jacob, *Cold Spring Harbor Symp. Quant. Biol.* **26**, 389 (1961).
7. I. Prigogine, *Thermodynamics of Irreversible Processes* (Interscience, New York, ed. 2, 1961).
8. N. Wiener, *Cybernetics: Or the Control and Communication in the Animal and the Machine* (MIT Press, Cambridge, MA, 1948).
9. J. D. Jordan, E. M. Landau, R. Iyengar, *Cell* **103**, 193 (2000).
10. M. Freeman, *Nature* **408**, 313 (2000).
11. U. Alon, *Nat. Rev. Genet.* **8**, 450 (2007).
12. D. E. Clapham, *Cell* **131**, 1047 (2007).
13. L. Stephens, L. Milne, P. Hawkins, *Curr. Biol.* **18**, R485 (2008).
14. O. Brandman, J. Liou, W.S. Park, T Meyer, *Cell* **131**, 1327 (2007).
15. D. Nicholls, K. Akerman, *Biochim. Biophys. Acta* **683**, 57 (1982).
16. U. Alon, M. G. Surette, N. Barkai, S. Leibler, *Nature* **397**, 168 (1999).
17. M. E. Burns, D. A. Baylor, *Annu. Rev. Neurosci.* **24**, 779 (2001).
18. S. H. Zigmond, S. J. Sullivan, *J. Cell Biol.* **82**, 517 (1979).
19. S. M. DeWire, S. Ahn, R. J. Lefkowitz, S. K. Shenoy, *Annu. Rev. Physiol.* **69**, 483 (2007).
20. D. M. Bautista, M. Hoth, R. S. Lewis, *J. Physiol.* **541**, 877 (2002).
21. A. Lotka, *J. Phys. Chem.* **14**, 271 (1910).
22. I. Bezprozvanny, J. Watras, B. E. Ehrlich, *Nature* **351**, 751 (1991).
23. D. O. Mak, S. McBride, J. K. Foskett, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 15821 (1998).
24. D. E. Koshland Jr., A. Goldbeter, J. B. Stock, *Science* **217**, 220 (1982).
25. T. Meyer, L. Stryer, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5051 (1988).
26. H. Meinhardt, A. Gierer, *Bioessays* **22**, 753 (2000).
27. W. Xiong, J. E. Ferrell Jr., *Nature* **426**, 460 (2003).
28. R. Jacob, J. E. Merritt, T. J. Hallam, T. J. Rink, *Nature* **335**, 40 (1988).
29. O. D. Weiner, W. A. Marganski, L. F. Wu, S. J. Altschuler, M. W. Kirschner, *PLoS Biol.* **5**, e221 (2007).
30. M. B. Elowitz, S. Leibler, *Nature* **403**, 335 (2000).
31. T. Y. Tsai *et al.*, *Science* **321**, 126 (2008).
32. I. Parker, Y. Yao, *Proc. Biol. Sci.* **246**, 269 (1991).
33. G. J. Augustine, E. Neher, *Curr. Opin. Neurobiol.* **2**, 302 (1992).
34. T. Meyer, L. Stryer, *Annu. Rev. Biophys. Biophys. Chem.* **20**, 153 (1991).
35. A. H. Cornell-Bell, S. M. Finkbeiner, M. S. Cooper, S. J. Smith, *Science* **247**, 470 (1990).
36. J. Sai *et al.*, *J. Biol. Chem.*, in press.
37. M. Fivaz *et al.*, *Curr. Biol.* **18**, 44 (2008).
38. J. K. Foskett *et al.*, *Physiol. Rev.* **87**, 593 (2007).
39. O. Brandman, J. E. Ferrell Jr., R. Li, T. Meyer, *Science* **310**, 496 (2005).
40. A. Politi, L. D. Gaspers, A. P. Thomas, T. Hofer, *Biophys. J.* **90**, 3120 (2006).
41. J. E. Ferrell Jr., *Curr. Biol.* **18**, R244 (2008).
42. A. T. Harootunian, J. P. Kao, S. Paranjape, R. Y. Tsien, *Science* **251**, 75 (1991).
43. J. Liou *et al.*, *Curr. Biol.* **15**, 1235 (2005).
44. R. S. Lewis, *Annu. Rev. Immunol.* **19**, 497 (2001).
45. C. Arrieumerlou, T. Meyer, *Dev. Cell* **8**, 215 (2005).
46. P. A. Iglesias, P. N. Devreotes, *Curr. Opin. Cell Biol.* **20**, 35 (2008).
47. J. W. Walker, A. V. Somlyo, Y. E. Goldman, A. P. Somlyo, D. R. Trentham, *Nature* **327**, 249 (1987).
48. O. D. Weiner *et al.*, *Nat. Cell Biol.* **4**, 509 (2002).
49. T. Inoue, T. Meyer, *PLoS One* **3**, e3068 (2008).

50. We thank M. Hammer, J. Ferrell, M. Covert, A. Salmeen, Y. Brandman, and P. Vitorino for helpful discussion and suggestions and NIH grants R01GM063702, R01GM030179, and R01MH064801 for funding the work.

Supporting Online Material

www.sciencemag.org/cgi/content/full/322/5900/390/DC1
Tables S1 and S2

10.1126/science.1160617

REVIEW

Optical Switches for Remote and Noninvasive Control of Cell Signaling

Pau Gorostiza[1] and Ehud Y. Isacoff[2,3]*

Although the identity and interactions of signaling proteins have been studied in great detail, the complexity of signaling networks cannot be fully understood without elucidating the timing and location of activity of individual proteins. To do this, one needs a means for detecting and controlling specific signaling events. An attractive approach is to use light, both to report on and control signaling proteins in cells, because light can probe cells in real time with minimal damage. Although optical detection of signaling events has been successful for some time, the development of the means for optical control has accelerated only recently. Of particular interest is the development of chemically engineered proteins that are directly sensitive to light.

Signaling proteins operate in complex networks in cells. The networks are wired into long serial chains, and these chains are arrayed in numerous parallel pathways that diverge from common inputs, converge onto intermediate nodes, and diverge again to many different effectors. Signals from the external world that are detected at the cell membrane are transmitted in the plane of the membrane and through the cytoplasm, with feedback and feedforward loops onto organelles and the nucleus. The upshot of this complex connectivity is the control of outputs as diverse as membrane transport, cell metabolism, protein translation, cell shape and migration, gene transcription, cell cycle, and cell survival. The shear number of signaling proteins and complexity of their connectivity is staggering, and the depictions in textbooks and on glossy posters from chemical companies are as dense and as difficult to decipher as spirographs.

Optical detection of cellular signaling has been accomplished, first with chemical dyes and more recently with fluorescent protein sensors to measure ions, membrane potential, second messengers, enzyme activity, protein/protein interaction, and the structural changes in proteins that underlie their functional transitions (*1*). Methods for remote control of cell signaling have seen explosive growth. Early methods used caged compounds, which release a signaling molecule when exposed to an intense pulse of light. "Reversibly caged" photochromic ligands, whose ability to function can be toggled on and off, were also developed. Most recently, efforts here have led to the engineering of light-sensitivity into proteins by the attachment of photoswitched tethered ligands (PTLs) that turn the function of the protein on and off in response to light. Although demonstrated in only a limited number of cases, this form of protein engineering appears likely to be broadly applicable across protein classes, and it complements recently discovered naturally light-sensitive ion channels and pumps. We focus here on the optical methods of remote control of protein function for elucidating signaling circuits inside cells and in cell circuits.

The Power of Light

To manipulate cell signaling, it is essential to control the function of key proteins by means that differ from, and thus will not cross-react with,

[1]Institució Catalana de Recerca i Estudis Avançats and Institut de Bioenginyeria de Catalunya, Parc Científic de Barcelona, Edifici Hèlix. C/Baldiri Reixac 15, Barcelona 08028, Spain. [2]Department of Molecular and Cell Biology, University of California at Berkeley, 271 Life Sciences Addition, Berkeley, CA 94720, USA. [3]Divisions of Material and Physical Bioscience, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed. E-mail: ehud@berkeley.edu