INVITED PAPER

# Signals and Systems: Towards a Systems Biology of Signal Transduction

*Models of biological signal pathways may shed light on cellular information processing during development, physiology, disease and healing.*

By Jeremy Gunawardena

**ABSTRACT** | Systems biology seeks to understand how the behavior of cells and organisms arises from the collective interactions of their component molecules. I will discuss how signal transduction—the process by which cells sense and respond to external signals—is being reconsidered from a systems perspective. This relies on ideas and concepts from the physical sciences coupled to new experimental strategies. I will outline some of the challenges through work in our laboratory on epidermal growth factor signalling.

**KEYWORDS** | Epidermal growth factor; input/output behavior; microfluidics; multisite protein phosphorylation; signal transduction; systems biology

## I. INTRODUCTION

Systems biology is the study of how biological function emerges from the collective interactions of molecular components. It has two broad directions. The first, "omics," arose from new technologies, among which the microarray is best known [72], which allow biologists to acquire data about all, or a substantial number of, molecular components of a particular type. For the microarray, the components are expressed genes ("genomics"), while other omic technologies provide data on proteins, protein–DNA interactions, protein–protein interactions, lipids, metabolites, etc. [32]. The focus here is from pathways to whole cells and on inferring causality from statistical correlation in large data sets.

Manuscript received August 8, 2007; revised March 24, 2008. This work was supported in part by the National Science Foundation under Grants CCF-0603630 and PHY05-51164.
The author is with the Department of Systems Biology, Harvard Medical School, Boston, MA 02115 USA (e-mail: jeremy@hms.harvard.edu).

Digital Object Identifier: 10.1109/JPROC.2008.925413

The second direction has been less visible but has deeper historical roots in biochemistry and physiology [26], [33], [35], [54]. It might be called "mechanistic" systems biology. Its focus is from small subsystems to pathways. Instead of statistical data analysis, it uses mathematical models to formulate hypotheses. Instead of omic data, which is usually averaged over cell populations, single-cell analysis is often essential (see Section IV). If omics has too much data and not enough analysis, then mechanistic systems biology has not enough data and perhaps too much analysis.

In this paper, I will illustrate, from a largely mechanistic standpoint, how systems biology can be used to study signal transduction in mammalian cells. While bacteria also have signalling pathways, and those underlying bacterial chemotaxis are among the best understood of all signalling pathways, bacterial signalling has evolved to make fast decisions in the face of an unpredictable environment [6]. Pathways are short and decisions simple. The cells in a multicellular organism may have the luxury of a more predictable environment but must make more complex decisions—constructing the organism, for instance—and this is reflected in the complexity of the underlying molecular machinery. Mammalian cells are bombarded by a spectrum of chemical signals: hormones (e.g., insulin) operate on an organismal scale and neurotransmitters (e.g., serotonin) operate on the scale of synapses, while multiple cytokines (e.g., interferon) and growth factors (e.g., epidermal growth factor) participate in cell-to-cell signalling within tissues during embryonic development, immune responses, and wound healing, and in disease states like cancer.

Mammalian signal transduction is sometimes viewed as a pipe that relays information from a cell's membrane to its nucleus, where it influences gene transcription [21]. Such

 

a view fails to account for the multiple feedbacks, both positive and negative, extensive scaffolding and localization, and considerable intramolecular complexity arising from posttranslational modifications. Our central hypothesis is that, rather than being a passive pipe, a signalling pathway undertakes active information processing: the decision to influence gene transcription is the result of a calculation, whose complexity is reflected in the structure of the pathway. The approach taken to working out this calculation is based on analogy with signal processing in engineering but is also motivated by insights from other areas of biology, such as bacterial chemotaxis and neuroscience [37]. Neuroscientists, in particular, are accustomed to thinking in terms of information processing and have a long tradition of gaining deep insight from mathematical models [26], [36]. This perhaps reflects the nature of electrical signals, which can be measured with great precision, thereby focusing attention on accounting for the resulting quantitative patterns. Chemical signal processing is altogether more elusive. However, it can still be conceived of, at the level of an individual cell, as a state-dependent transformation from input signals to output responses within a given cellular context. Schematically

$$\text{INPUTS} \times \text{STATE} \times \text{CONTEXT} \longrightarrow \text{OUTPUTS}. \quad (1)$$

Such a formula serves to delineate those aspects of the system that are important and around which this paper will be organized. The INPUT signals are typically under the control of the experimenter. The STATE refers to the internal condition of the particular cell being studied, prior to stimulation by signals, and depends on the type of cell, its age and passage number, growth conditions, etc. The CONTEXT refers to the external environment in which the cell finds itself, including other cells as well as the medium in which the cells are present. The OUTPUTS are typically measures of cellular response over time (see Section IV). The complex molecular machinery within the cell implements the transformation from INPUTS to OUTPUTS. Both the STATE and the CONTEXT can conceal much biological subtlety.

Equation (1) is best suited to an experimental framework based on *in-vitro* cell culture, in which cell–cell interaction is (perhaps) less important than in an intact tissue. In a tissue, the input/output behavior of each cell may depend on the input/output behavior of its neighbors in a complex way, so that the CONTEXT becomes a dominant source of behavioral complexity. *In-vitro* cell culture has been highly successful in the search for signalling components because *in-vitro* component behavior has usually proved highly indicative of *in-vivo* physiological function. It is less clear whether the same will be true for the behavior of systems. I will gloss over these important issues here and assume that cells can be prepared in a physiologically meaningful context, within which the input/output response is reproducible. Both omics and mechanistic approaches can be taken to studying (1) (see Section VI). From a mechanistic perspective, it leads us to ask how the internal molecular mechanism within a cell implements the transformation from input signals to output responses.

The rest of this paper is organized as follows. Work in my laboratory focuses on epidermal growth factor (EGF) signalling, which I will use as the main biological example, and Section II begins with an overview of it. I then discuss in turn the input signals (Section III), the output responses (Section IV), and the internal molecular mechanisms (Section V). I review in Section VI some of the related work in signal transduction from a similar perspective. The aim is to introduce the reader to some of the biological background and to identify some of the major challenges, experimental and theoretical, in understanding cellular information processing from a systems perspective.

## II. EGF SIGNALLING

The discovery by Cohen and Levi-Montalcini that cells release chemicals (epidermal and nerve growth factors), which encourage their neighbors to grow and proliferate, opened a new window onto the complex life of metazoan cells [11], [41]. Much subsequent work has made the EGF pathway the best studied of all metazoan signalling pathways. EGF is a small peptide, of 53 amino acids in humans, which binds to receptor proteins on the outer plasma membrane of cells. Binding initiates multiple overlapping sets of internal reactions that reach out to other cellular systems, such as the cytoskeleton, as well as influence gene transcription in the cell's nucleus [91].

The EGF receptor is a member of the superfamily of receptor tyrosine kinases [22]. Such receptors have an external segment that binds ligands, linked through the membrane to a cytoplasmic segment with tyrosine kinase activity. A kinase is an enzyme that covalently modifies residues in a substrate protein by transferring a high-energy phosphate group from ATP. In mammalian cells, phosphorylation takes place predominantly on serine, threonine, and tyrosine amino-acid residues. As usual in molecular biology, a cognate group of phosphatases catalyse the opposite reaction by removing phosphate groups. This reversible posttranslational modification, discovered by Fischer and Krebs in 1955, is one of the most important of all cellular regulatory mechanisms [18], [38], and we discuss it further in Section V.

Receptors are dynamically maintained on the plasma membrane, being subject to synthesis as well as internalization, degradation, and recycling, which processes are further modulated by ligand binding [40]. Binding does not directly initiate tyrosine kinase activity. The EGF

receptor first dimerizes, then becomes catalytically active, and intermolecular phosphorylation takes place on at least nine functionally significant tyrosine residues in each monomer [73]. These act as attachment points for a range of proteins with phosphotyrosine recognition domains [75]. Proteins are now known to be composed of a succession of linked domains, each of which can independently fold into a three dimensional (tertiary) structure [61]. Domains have specific functions, such as catalytic activity or binding to molecular components, in this case, to phosphotyrosine sites on the receptor (SH2 and PTB domains). Adaptor proteins like Shc and Grb2, which have SH2 domains, accumulate on the activated EGF receptor and bring with them yet more molecular players to assemble a complex signalosome. This initiates several signalling pathways, including an important member of the family of mitogen activated protein kinase (MAPK) cascades [64]. MAPK cascades consist of three kinases, each of which activates the next; they play an important role in cellular information processing [2], [28]. The final kinase in the MAPK cascade downstream of the EGF receptor is the extracellular signal regulated kinase (Erk), which phosphorylates several transcription factors which then influence gene transcription. We will examine the response of Erk to EGF stimulation in Section IV.

The EGF receptor is one of four in the ErbB family (ErbB1), and dimerization and phosphorylation can occur between different family members [91]. ("ErbB" comes from the avian erythroblastosis virus, which acquired a truncated and oncogenic form of the EGF receptor.) EGF is also one of several ligands that can activate ErbB family members [25]. A mouse knockout of any one of the four receptors is lethal in most genetic backgrounds, indicating that they are all essential for development. Knockouts of the ligands have milder but varied phenotypes, suggesting some degree of redundancy. Sequence analysis finds one homologous ErbB receptor in both thenematode worm *C. elegans* and the fruit fly *D. melanogaster*. Two gene-duplication events between invertebrates and mammals produced the four receptor types seen in humans; and a similar, if more complex, coevolution seems to have taken place among the ligands [78]. Interestingly, duplication of the receptors has loosened their constraints: ErbB2 is incapable of ligand binding, while ErbB3 has an inactive catalytic domain. Nevertheless, both receptors are important dimerization partners.

Organismal development as well as disease states like cancer has provided some insight into the *in-vivo* role of the EGF pathway. In both *C. elegans* and *D. melanogaster*, the EGF pathway has been implicated in fine-grained spatial patterning: in delineating specific receptor cells in the rhabdomeres of the fly compound eye or in selecting individual cells in the worm vulva [19]. Omic approaches have implicated the ErbB family in several cancers, particularly those of epithelial origin, such as breast, colon, and lung cancer [29], [57]. The treatment of these solid tumors has been revolutionized by the development of drugs that target ErbB family receptors, such as monoclonal antibodies that block the extracellular domain (Herceptin, Erbitux) as well as small molecule tyrosine kinase inhibitors (Iressa, Tarceva, TykerB) [29]. Clinical trials of these have proved both exciting and frustrating [46]. They can cause dramatic improvements but only in a small subset of patients, perhaps reflecting a delicate dynamic balance between prosurvival and proapoptotic (cell death) pathways [76]. Unravelling such dynamic phenomena provides a major motivation for the kind of approach taken here.

## III. INPUT SIGNALS

Signalling systems, whether natural or man-made, are well adapted to processing certain signals and poorly adapted to processing others. The human visual system, for instance, is extraordinarily efficient at face recognition but can be unsettled by optical illusions. Cognitive psychologists have used such illusions to probe the functioning of the visual system. I take the view, in keeping with this perspective, that the molecular machinery within cells reflects the signals that the machinery was evolved to process. Hence, knowledge of these signals, and of the cellular response that they elicit, is essential to interpreting the role of the molecular machinery. Unfortunately, rather little is known about natural signals. Endocrine hormones like insulin act over long distances in the organism, and their levels in the blood stream can be measured with some precision, giving some idea of their behavior at a physiological level. Growth factors and cytokines act over shorter distances within tissues, through paracrine (affecting neighboring cells) and sometimes autocrine (affecting the same cell) signalling. While much is known about the molecular components that respond to such signals, it has proved difficult to measure the signals in their natural context. Outside of neuroscience, mammalian signal transduction has been largely the study of transduction, not signal.

One striking difference between growth factors and neurotransmitters (or hormones like insulin) is that the latter are secreted while the former are shed [71]. The secretory pathway packages neurotransmitters into vesicles within the cytoplasm, which, upon appropriate stimulation, fuse with the plasma membrane and release their contents into the extracellular space. In contrast, EGF is first constructed as part of the extracellular segment of a much larger protein, prepro-EGF of ~2000 amino acids, which is inserted into the plasma membrane. Upon appropriate stimulation, another membrane-bound protein, a protease that acts as a molecular scissors, sheds the "ectodomain" consisting of the EGF peptide by cutting it off its protein precursor. One may speculate that ectodomain shedding allows release of a substantial amount of signal, at the expense of a longer waiting time for the membrane to be resupplied with precursor, while

exocytosis provides a more continuous release of signal in smaller quanta. What effect this difference makes to information processing remains unknown.

If the natural input signals are inaccessible, cells can still be stimulated with constructed signals. Experiments of this kind laid the foundation for understanding bacterial chemotaxis [6]. A bacterium like *E. coli* moves by rotating helical flagella and can navigate towards an attractant or away from a repellent by altering the frequency of "tumbles"—uncoordinated flagellar rotation—compared to "runs"—coordinated flagellar rotation leading to straight line swimming. In classic experiments, cells were subjected to constructed temporal signals of attractant and their tumbling frequencies measured [7], [48]. It was shown that bacteria achieve high sensitivity to spatial gradients by temporal integration and that they adapt to changes across a wide range of background concentrations. Once the molecular components had been characterized through genetics and biochemistry, mathematical analysis was able to illuminate the molecular calculations responsible for chemotactic adaptation [1], [53], [92]. Such utilization of signals to probe the molecular machinery is only starting to be used in eukaryotic and mammalian signalling. For instance, Samadhani *et al.* used ultraviolet uncaging of cAMP to study individual cell variability in the chemotactic response of the slime mold *D. discoideum* [70], while Reynolds *et al.* used microspheres coated with EGF to uncover lateral propagation of EGFR activation in MCF7 cells [65]. In both cases, a constructed spatial signal was used to illuminate signalling behavior.

Microspheres and photo-uncaging, however, are unsuitable for generating complex signals. By "complex" I mean, for instance, a sequence of rectangular pulses with varying pulse width, interpulse spacing, and pulse amplitude. While it is not clear that such signals are the best choice for probing the nonlinear molecular systems within cells, they are a reasonable starting point for experimental exploration. Furthermore, exploring input/output relationships requires not only reproducible generation of complex signals but also replicated application of such signals to cells with appropriate positive and negative controls along with assays for cellular state at multiple time points. Manual bench-top methods rapidly become inadequate for such studies, and forms of parallelization and automation become essential. Microfluidic devices present an attractive option.

For biological applications, soft lithography using polydimethylsiloxane (PDMS) has considerable advantages over silicon-based microfabrication [89]. PDMS devices are easy, cheap, and fast to make, while PDMS itself is biocompatible, permeable to the $O_2$, $N_2$ and $CO_2$ required for cell culture and transparent to the wavelengths used for optical and fluoresence microscopy. Master moulds can be created by patterning photoresist on silicon using conventional photolithography. Because line widths are rarely less than 20 $\mu$m, photo masks can be made with commonly available software and high resolution printers. A PDMS layer is then spin-coated and cast from this mould. Input and output lines are punched through the PDMS, which seals efficiently around the holes. Single layers, bonded to a substrate like glass, can have passive channels and chambers. Two-layer devices allow valves to be constructed, which take advantage of the low Young's modulus of PDMS ($7.5 \times 10^5$ Pa compared to silicon's $10^{11}$ Pa) and its resulting elastomeric properties. A control channel in one layer, which crosses a flow channel in the other layer, can be pressurized so as to deform and shut the flow channel [87]. Switching rates up to 75 Hz are readily achieved. By varying the area over which the control and flow channels overlap, the pressure can be chosen so that large areas shut while small areas remain open, allowing multiplexed control and thereby the capability for large-scale integration [81]. Peristaltic pumps, which allow fluids to be moved, can be built from a sequence of valves operating suitably out of phase with each other [87]. Despite the laminar flow at low Reynolds number (around one, for typical velocities and dimensions), mixing devices can also be fabricated [56], making complex fluid handling feasible. In short, most of the requirements for the input/output experiments envisaged here are in place. Several pioneering studies have demonstrated the potential of microfluidics for undertaking a new class of biological experiments [4], [45], [51], [60].

Fig. 1 shows the schematic of a microfluidic device built by us in a joint project with Amarasinghe and Thorsen at the Massachusetts Institute of Technology. This device allows signal pulses at varying frequencies but constant amplitude (concentration) to be generated and applied to several chambers, each containing a small number of cells. This design has valves and peristaltic pumps, although some fluid movement is done by off-device pumps. The output response is measured by antibody, as discussed further in Section IV. This device is a prototype used for testing and not for data collection.

As in silicon chip design, microfluidic devices of this complexity require design rules, programming abstractions, and software tools, which free the designer from implementation details [83]. The device in Fig. 1 does not have an onboard mixing device and is designed to use a fixed concentration of signalling ligand, which is made available separately. Future devices will generate specified concentrations as required by undertaking their own mixing operations. Such "mixing modules" should be available to the device designer as abstract components, just as a microprocessor designer works with registers and accumulators. Looking forward, it should be possible to specify an entire biological experiment in language familiar to a biologist and to then compile this into a device design for implementing the experiment. Such technology will revolutionize our capabilities for performing complex biological experiments.

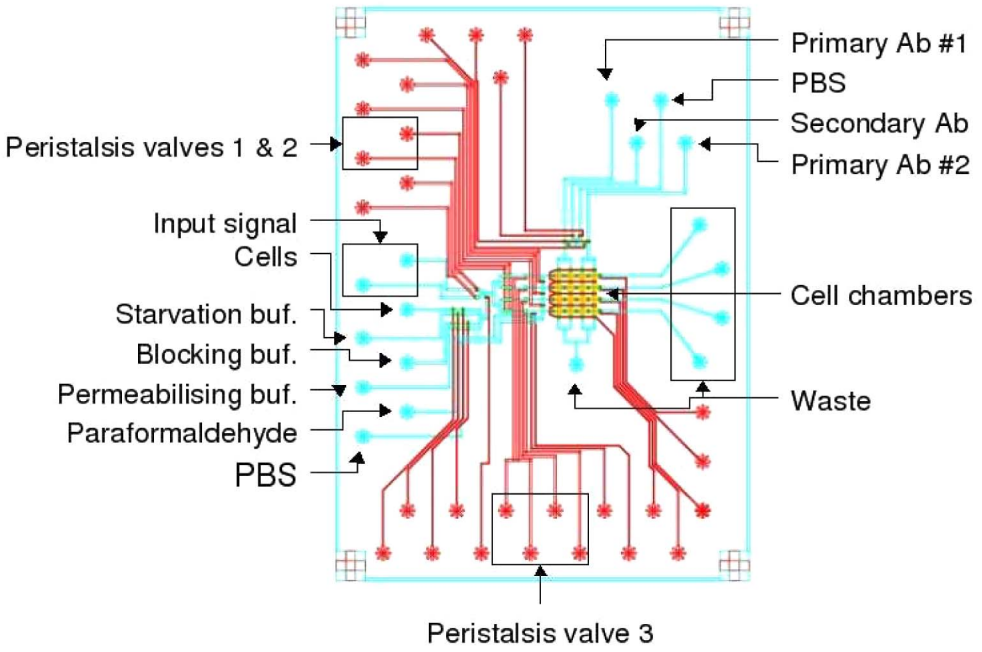


**Fig. 1.** *Prototype device for characterizing input/output behavior, showing a 4 × 4 array of cell chambers with flow channels in blue and control channels in red. This device can generate an input signal of approximately rectangular pulses at fixed amplitude (concentration), interspersed with buffer, at varying frequencies. The signal is applied from the left to all rows simultaneously, using peristaltic pumps, whose three valves are annotated. Each row can be separately fixed and permeabilized, through the flow lines on the left, allowing four time points to be obtained. Cells are normally maintained in starvation buffer for 24 hours prior to stimulation. Output is measured with antibodies, as described in Section IV. Ab = antibody, buf. = buffer, PBS = phosphate-buffered saline buffer.*

## IV. OUTPUT RESPONSES

Molecular mechanisms operate within individual cells. However, common detection techniques, such as Western blots or DNA microarrays, require a population of cells (at least $10^6$) and reflect, after suitable normalization, population averages. If the underlying population distribution shows a single mode with most of the distribution reasonably close to the mean, then an average measure may be a reasonable proxy for the distribution. On the other hand, if the distribution is multimodal, with different groups of cells having different levels of response, then the mean may give a very misleading impression of what is taking place in any individual cell.

Lahav *et al.* provide a striking demonstration of how population averaging can lead to misinterpretation of single-cell behavior [39]. Irradiation of cells with $\gamma$-rays produces a stereotyped damage response involving stabilization of the tumor suppressor p53, a key cellular transcription factor. If p53 is examined by Western blot, it exhibits a damped oscillation following DNA damage. However, using a p53-CFP fluorescent fusion reporter and time-lapse fluorescence microscopy, it was found that each cell produces a series of "digital" pulses of p53, whose heights and durations are independent of the amount of DNA damage but whose number increases with increasing damage. Because different cells generate different numbers of pulses and show variation in phase, the population average smears out the digital pulses to a damped oscillation. While signalling dynamics are not always so unusual, it is critically important to know whether or not the average response is representative of the population distribution. We will consider this for the EGF pathway to illustrate the issues that arise.

EGF stimulation has multiple effects prior to gene transcription, among which changes in the phosphorylation state of pathway components are particularly significant. The MAP kinase Erk is activated by double phosphorylation through the MAPK cascade and phosphorylates in turn several transcription factors. It serves as a convenient endpoint for the cytoplasmic portion of the EGF signalling response. The problem of measuring phosphorylation state raises a number of general issues about how molecular states can be detected and quantified.

A fluorescent fusion reporter, as used in [39], can indicate whether or not the tagged protein is present and where it is in a cell. Localization is sometimes controlled by phosphorylation state, which can then be used as a proxy for it. One major advantage of fluorescent proteins, which have revolutionized cellular imaging, is that they can be genetically encoded and introduced into cells by transfection. However, such exogenous introductions can be substantially overexpressed in comparison to their

native counterparts, particularly when coupled to strong viral promoters, and this can disturb the natural dynamics of the system being studied [5]. Chromosomal replacement of the corresponding gene under its natural promoter can avoid such difficulties, but homologous replacement of this kind is difficult to undertake in many mammalian cells. Yeast has the advantage in this respect, being much more genetically tractable.

An alternative approach is to introduce a separate sensor that can detect the protein state to be measured. This may interfere less with the system under study. Sensors based on fluorescence resonance energy transfer (FRET) have been used to detect activation of kinases [66], [86]. Since some kinases like Erk are activated by phosphorylation, this can, once again, sometimes be used as a proxy for phosphorylation state. FRET sensors are also genetically encoded. They work by exposing a short phosphorylatable peptide motif that is attractive to the kinase in question, flanked at one end by a fluorophore like YFP and at the other end by a phosphopeptide binding domain followed by a second fluorophore like CFP. If the kinase is activated, it phosphorylates the peptide. The phosphopeptide binding domain then closes on the phosphorylated residue, bringing the two fluorphopores within their Förster radius, and quantum mechanical energy transfer (FRET) takes place [63]. If the kinase is inactive, the peptide motif remains unphosphorylated, the fluorophores are far apart, and little FRET is detected. Such sensors have to be individually designed and optimized for each target kinase, carefully balancing several competing design objectives. An Erk sensor with reasonable signal/noise ratio has yet to be developed. Furthermore, such sensors couple their own dynamics into the system being studied. While this has not been seen as a problem for qualitative studies, it becomes a major issue when accurate quantification is needed. In effect, the transfer function of the sensor must be deconvoluted from the measured signal in order to uncover the actual signal [82]. A systematic strategy for designing and using quantitative real-time sensors would be of great benefit for both academic and industrial studies. Despite its importance, such an engineering challenge has failed to attract the attention of synthetic biologists [15].

While both the methods just discussed are real time and *in vivo*, they cannot be used systematically to detect phosphorylation state. At present, this can only be done with an antibody raised against the phospho-epitope. Antibodies have proved invaluable for detecting phosphorylation, although care must be taken with their specificity, particularly when the phospho-epitope contains multiple phosphorylations. The great disadvantage of antibodies is that cells must be fixed and permeabilized before the antibody can get into them and do its work. Hence different stimulations have to be applied to different cell populations. It becomes harder to deconvolute the single-cell dynamics from such data, but the population distribution can be determined.

As we see from the above discussion, there are pros and cons with each method for detecting molecular states within cells. No matter which method is used, it necessarily perturbs the state being detected (sometimes to the extent of killing the cells), and quantitative interpretation of the results can be challenging.

We used an Applied Precision cellWoRx automated microscope to examine EGF-stimulated Erk activation in NIH 3T3 mouse fibroblasts. This epifluorescence microscope is able to automatically image a $16 \times 24$ well plate at four wavelengths, allowing a broad range of EGF concentrations, time points, and replicates to be acquired. We used primary antibody staining followed by secondary antibodies conjugated to fluorescent dyes to detect total Erk (i.e., independent of its phosphorylation state) and Erk-pp (doubly phosphorylated and thereby activated), for both of which good primary antibodies have been characterized [90]. We measured, for each cell, the per-pixel total Erk intensity and the Erk-pp intensity, summed over all pixels in the cell. The complex pipeline of fixation, antibody staining, microscopy, and image analysis may introduce significant measurement error for any particular cell, but the population distributions were reproducible and robust to controls such as fluorophore swapping.

We found that the distribution of total Erk in the population remains similar for all times and all EGF concentrations. However, it is significantly long-tailed. It is known from experiments in both *E. coli* and *S. cerevisiae* that so-called extrinsic "noise," as opposed to intrinsic noise, accounts for much of the cell-to-cell variability in protein levels [12], [16]. Intrinsic noise arises from the stochastic nature of chemical reactions: identical cells show variation in protein levels because of random differences in the timing of reaction events. Extrinsic noise arises from cell-to-cell differences in basic cellular machinery (RNA polymerases, ribosomes, etc.). As cells divide, each division introduces a (presumably) random separation of the existing components between the daughter cells, and this variation may then have further concerted effects on the activity of protein complexes (RNA polymerases, ribosomes, etc.), which contain the components. While that much is well appreciated, there is, as yet, no mathematical explanation for the long-tailed population distribution seen for total Erk. We found it to be well fitted to a log-normal, Weibull, or gamma distribution and chose the log-normal distribution for our analysis. Log-normality arises when independent processes act multiplicatively, as the normal distribution arises additively, and perhaps this provides some justification for its use here.

Antibody data are expected to reflect protein amount, but it is concentration (i.e., amount per unit volume) rather than amount that determines the rate of reaction. Cells with the same concentration of protein may have different volumes and therefore show variation in antibody staining. Since there is no easy way to determine cell volume without a confocal scan (automated confocal microscopes are now available), we normalized Erk-pp against total Erk on a per-cell basis, as previously used

in [20]. This ratio can be considered as a function of concentration rather than amount but does not require the cell volume to be known. The ratio correlated well against Erk-pp but not against total Erk, indicating that the total Erk in the denominator acted as a normalization rather than as a determining factor.

We found that the normalized Erk-pp distribution could not be well-fitted to a single log-normal distribution but became approximately bimodal in a time and dose-dependent manner. We found it could be reasonably fitted to a sum of two independent log-normals, and these smoothed distributions are shown in Fig. 2. We see from this that Erk activation is not well described by its population average. Depending on the dose and the time, cells fall into two subpopulations, which express broadly different levels of normalized Erk-pp.

We explored three hypotheses to account for this bimodality:

1) that it reflects some existing bimodality in the cell population, such as different stages of the cell cycle, which has been suggested as an important source of variation in yeast [12];
2) that it reflects the nonlinearities in the signalling pathway acting on constitutive, unimodal cell-to-cell variation in the initial concentrations of pathway components;
3) that it reflects the influence of cells upon each other through some form of intercellular communication.

We found no evidence for the first hypothesis. We believe that much of the variation can be accounted for by the second hypothesis. Surprisingly, statistical analysis of the data also revealed signs of cell-cell interaction but we have, as yet, not identified any mechanism behind this.

These findings make clear that signalling responses must be studied at the single-cell level and that it is necessary to understand both the mechanism within cells and also the causes of variation from cell to cell. Turning this around, cell-to-cell variation could prove invaluable in identifying pathway mechanisms. Suppose that variation arises largely from the second hypothesis above and that, if cell–cell interaction also plays a role, it is either marginal or can be marginalized. If the component distributions prior to stimulation can be determined empirically, then any proposed pathway mechanism would have to reproduce the dose and time-dependent response in Fig. 2 from these initial conditions. This is a stringent requirement. It is, however, the joint distribution that is needed for this calculation, and this requires simultaneous measurements of all the components. The new generation of automated microscopes possesses broader spectral capabilities, which will improve further in time, so it may soon prove possible to measure the joint distribution of several components simultaneously. Of particular interest would be components of the EGF signalosome, such as the EGF receptor family members and adaptor proteins like Grb2. Because these assemble together into a functioning complex, their joint variation from cell to cell may have a dominating influence on the variation in pathway response.

Fig. 2 shows the response of 3T3 cells to a step change in EGF concentration. The response to more complex signals will provide even more insight into pathway behavior, and this is where the microfluidics comes in. The device in Fig. 1 is designed to use antibodies to measure the response of two proteins to a complex temporal signal. Each row of cell chambers can be fixed and permeabilized independently, allowing four time points to be measured. Once all the cells are fixed, a mixture of two primary antibodies, such as goat anti-total Erk and rabbit anti-Erk-pp, is applied to columns 1 and 2, and a different set of goat and rabbit primary antibodies against a different protein is applied to columns 3 and 4. A mixture of secondary antibodies, such as mouse antigoat and mouse antirabbit IgG, each conjugated to a different dye, is then applied to all four columns, and the cells are imaged by placing the device on a suitable epi-fluorescence microscope. Two normalized measurements can be made at each time point, with two replicate chambers per measurement.

Antibody measurements on microfluidic devices are cumbersome and expensive in device area. They require lots of plumbing, and cells cannot be followed in real time, requiring many cell populations to resolve temporal

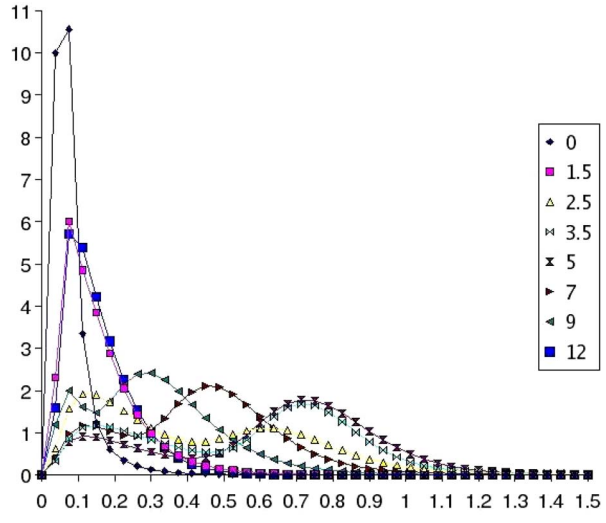


**Fig. 2.** *Distributions of activated Erk in a clonal population of NIH 3T3 cells following stimulation with 4 ng/ml EGF at the time points (in minutes) indicated in the legend. The x-axis is a measure of Erk-pp/Total Erk for each cell, obtained as the ratio of Alexa 647 (685 nm emission) to Alexa 555 (595 nm emission) intensities. Each curve is obtained by fitting the probability density function (pdf) of the raw data to a weighted sum of the pdfs of two independent log-normal distributions. The weights are $\alpha$ and $1-\alpha$, giving five fitting parameters. Note, in particular, that the area under each curve is one. Each pdf is obtained from approximately 2000 cells, pooled from two replicate wells.*

dynamics. Live-cell sensors would be more flexible but, as mentioned above, suitable sensors for phosphorylation state have yet to be developed.

## V. PATHWAY MECHANISMS

Several general mechanisms are discernible within the EGF signalling pathway. Both positive and negative feedback are present; components become localized in different compartments, including assembly of the signalosome at the plasma membrane; all components are posttranslationally modified in various ways, often on multiple sites. These mechanisms are all intertwined with each other. In this section, I will discuss one mechanism in isolation, multisite phosphorylation and dephosphorylation, and examine the range of behaviors it can exhibit. I will then speculate on how such behaviors might be exploited for cellular information processing.

It used to be thought that prokaryotes and eukaryotes differed in their phosphorylation chemistries, with eukaryotes using serine, threonine, and tyrosine (S/T/Y) residues and prokaryotes using histidine and aspartate residues in so-called two-component systems [79]. It is now understood that both mechanisms occur in all kingdoms [88]. For instance, both plants and fungi use two-component signalling. A more significant difference is the considerable expansion of multisite modification in eukaryotes as compared to prokaryotes. A recent survey of S/T/Y phosphorylation in *B. subtilis* found no more than eight sites on any single protein and only one protein with that many [47]. In contrast, the human tumor suppressor p53 has at least 16 functional sites [27], while the microtubule associated protein tau becomes hyperphosphorylated on 39–45 sites in patients with Alzheimer's disease [24].

A single substrate molecule with $n$ phosporylation sites may be in one of $2^n$ states. However, the downstream response to phosphorylation does not see a single molecule but, rather, samples the entire population. The state of the population of substrate molecules can be described by a frequency distribution, which gives the relative number of each of the $2^n$ single molecule states. This phosphoform distribution is a mathematical object whose value is determined at any time by the collective interactions of the cognate kinases and phosphatases: it is a property of a dynamical system rather than a structural property of an individual molecule. Several lines of evidence indicate that such an enhanced representation of phosphorylation state is biologically significant [77], [80]. For instance, different phosphoforms may be localized to different compartments of the cell [10]. A central question is how the kinases and phosphatases collectively regulate the phosphoform distribution. Measurement of it is challenging but in joint work with Steen at Children's Hospital, we are developing a strategy based on liquid chromatography/mass spectrometry to do so for substrates with low numbers of sites.

In the absence of experimental data, conceptual mathematical models can yield valuable insights. Consider a system in which the kinase and the phosphatase both act distributively, rather than processively, on the substrate. This means that when an enzyme molecule and a substrate molecule collide, the enzyme does at most one modification (distributivity), rather than holding on sufficiently long to do more than one modification (processivity). Examples of both mechanisms are known [28], [62], [93]. This gives rise to the kind of reaction network in Fig. 3(a), with each enzyme following a standard biochemical scheme as in Fig. 3(b). Assuming mass-action kinetics, this gives rise to a nonlinear dynamical system, which describes the time evolution of the concentrations of the various chemical species: $2^n$ phosphoforms, $2(2^n - 1)$ enzyme-substrate complexes, and two free enzymes. We introduced methods from algebraic geometry to find an analytic solution for the steady state of this system [50],

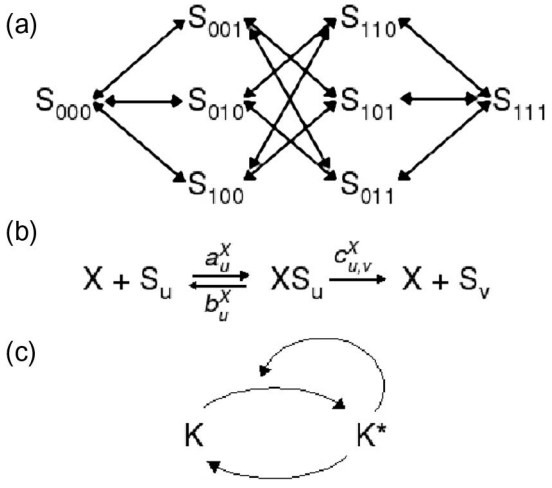


**Fig. 3.** *(a) Reaction network for multisite phosphorylation. The substrate $S$ has three phosphorylation sites. Phosphoforms are denoted $S_u$, where $u$ is a 3-bit string indicating the presence or absence (1/0, respectively) of phosphate on each site. The eight phosphoforms are intercoverted distributively by a kinase (adding a single phosphate, or bit, from left to right) and a phosphatase (removing a single phosphate, or bit, from right to left). (b) Each enzyme, denoted $X$, follows a standard biochemical scheme [13], with reversible formation of an enzyme-substrate complexes $XS_u$ and irreversible formation of product $S_v$, with $v$ having one more bit than $u$ when $X$ is the kinase and one less bit when $X$ is the phosphatase, according to the above network. The annotations over the reactions give the rate constants for mass-action kinetics ("a" for association, "b" for breakup, and "c" for catalysis). ATP, ADP, and phosphate are assumed held constant by some mechanism that is not explicitly modelled, as is the case in vivo, and their effects absorbed into the rate constants. A given substrate may yield several products, but only a single enzyme substrate complex is assumed to be formed. More general assumptions can be made without affecting our conclusions. (c) Lisman scheme for a bistable cellular memory in which kinase $K$ enhances its own phosphorylation in its active state $K^*$ [43].*

[84], [85], substantially improving upon previous models of multisite phosphorylation [17], [23], [52], [58], [68], [69]. We showed that there are values of the rate constants for which the system has $n+1$ steady states if $n$ is even and $n$ steady states if $n$ is odd. In either case there are, typically, $\lfloor (n+2)/2 \rfloor$ stable steady states, where $\lfloor - \rfloor$ is the "floor" function. In other words, the phosphoform distribution is potentially capable of complex decision making, and the capacity for this increases at least linearly with the number of sites. Bistable systems have become a common theme in systems biology [3], [31], [59], but they have always required overt positive feedback. The multisite phosphorylation system has no such overt feedback and requires only three components—kinase, phosphatase, and substrate—to encode an arbitrary amount of information.

Bistability for $n = 2$ was first shown in [52], using the Michaelis–Menten approximation. (It was subsequently claimed in [58] that no more than bistability occurs when $n > 2$ but this is incorrect, as we have just seen.) If processivity dominates over distributivity, the maximal number of stable states may be reduced; in the limiting case when all phosphates are added or removed in one collision, the system behaves as if it had only a single site and becomes monostable.

Neuroscience provides the context for understanding how this decision-making capacity may be used biologically. Individual neurons exhibit memory and learning [34], in which phosphorylation mechanisms are known to play a role. While long-term memory requires new protein synthesis to remodel synapses and make new connections, short-term memory is independent of protein synthesis and must hence be implemented by some posttranslational mechanism [34]. Crick and Lisman independently suggested the phosphorylation scheme in Fig. 3(c) as a basis for short-term neuronal memory, and Lisman showed that it exhibits bistability [14], [43]. While this scheme is similar in conception to that in Fig. 3(a), the latter does not require positive feedback as soon as $n > 1$ and may have more than two stable states. Subsequent work by Lisman and others on long-term potentiation (LTP) in hippocampal cells has implicated the calcium and calmodulin dependent protein kinase, CaMKII, a multimeric protein that is heavily expressed in synapses, as a bistable memory based on Fig. 3(c) [44]. Intriguingly, recent experimental data suggest that LTP may require a multibit capacity [42]. It remains to be seen whether the multisite phosphorylation mechanism described above is implicated in this.

Neuronal memory mechanisms were first uncovered by subjecting neurons to complex input signals, such as the learning protocols used to induce LTP (repeated high-frequency tetanic stimulation) [34]. This brings us back full circle to (1). While neuronal cells are highly specialized, they use similar components and mechanisms to all other cells, and it seems plausible that the latter may follow similar strategies in their decision making. If so, then the enhanced decision-making capacity of multisite phosphorylation found above is most likely to be revealed in the processing of sufficiently complex input signals. A simple signal, such as a step function increase of a single ligand, may not be sufficient to explore the system's broader decision-making capabilities. The characterization of input/output behaviors, along the lines suggested in this paper, thereby provides the context in which the information-processing capabilities of pathway mechanisms can be revealed and understood.

## VI. RELATED WORK ON SYSTEMS SIGNALLING

Signal transduction is of such interest from both a fundamental and a clinical perspective that many efforts are being directed towards studying it from the perspective sketched in the Introduction, using both omics and more mechanistic approaches. I cannot do justice to the field in this paper, but a few pointers may help orient the reader in the landscape.

Natarajan *et al.* have undertaken a comprehensive input/output analysis of the temporal behavior of RAW 264.7 cells, a mouse macrophage cell line, in response to 22 different ligands, applied singly and in combination [55]. The ligands were chosen to cover a cross-section of receptor types: toll-like receptors, G-protein coupled receptors, cytokine receptors, and receptor tyrosine kinases. The output responses included secreted cytokines, "second messengers" like $Ca^{2+}$ anions and cAMP, and the phosphorylation states of a panel of signalling proteins, all assayed by population averaging. Statistical analysis of this data shows substantial crosstalk between different ligands and suggests that the cell's signalling machinery is organized around a small set of core physiological responses modulated by many ligands. Using similar methods in HT29 cells, a human colon cancer cell line, but juxtaposing prodeath signals like TNF-$\alpha$ against prosurvival signals like EGF or insulin, Janes *et al.* found a similar convergence of mixed inputs upon a core "basis set" of intracellular responses [30]. The mechanistic details of how such modulation and convergence is achieved can only be schematically outlined with such omic methods and remain uncertain, as discussed above, in the absence of single-cell information.

Colman–Lerner *et al.* have used genetically encoded fluorescent reporters along with high-throughput fluorescence microscopy to study cell-to-cell variation in the yeast pheromone response pathway, a prototypical G-protein coupled receptor pathway [12]. The output responses were the activation levels of the MAP kinases, Fus3 and Kss1, measured by fluorescent reporters. No bimodality was found, and the variation in response was attributed to different cells having different capacities for transducing signals and for expressing proteins—in other words, to extrinsic differences between cells that were present prior to stimulation. This is in contrast to the bimodal variation in Erk activation described in Section IV, which was attributed to the

nonlinearities in the signalling pathway. The yeast cells were in exponential growth phase with a doubling time of around 90 min, while the mammalian 3T3 fibroblasts were serum starved prior to stimulation. Correspondingly, about half of the variation in the yeast experiment was found to be due to differences in cell cycle stage between cells, while no significant amount of the variation among the mammalian cells could be attributed to this.

The models used in both these studies are heavily abstracted. This makes it feasible to identify a few key parameters that can account for the data, but the parameters cannot be easily attributed to specific molecular components. More detailed models of signalling pathways have been built [28], and the EGF pathway, in particular, has attracted much attention from model builders [8], [9], [74]. The computational infrastructures needed for managing such complex models and the software pipelines needed for complex multidimensional data sets are just being developed [49], [67].

Paliwal *et al.* used a cleverly designed microfluidic device to subject yeast cells to pheromone concentration gradients and found bimodality in the expression of Fus1 [60]. Fus1 is a pheromone-responsive transmembrane protein that participates in cell fusion during the mating response. Its expression at the single cell level was monitored using a fluorescent fusion. The bimodality found here was attributed to bistability in the underlying molecular network on the basis of a simplified model of the feedback loops involving Fus1. The use of a microfluidic device to reproducibly control signal generation, allied to quantitative measurements and mathematical models, exemplifies the philosophy advocated here.

## VII. CONCLUSION

I have touched on some of the challenges in studying signal transduction from a systems perspective: microfluidic technology for manipulating complex signals is just being developed; measurement of cellular state necessarily alters the cellular state; cellular behavior cannot be summarized in population averages; posttranslational modification of proteins conceals much dynamical complexity. Much remains to be done.

Mathematical models will continue to play an essential role in formulating hypotheses about pathway function. Such models are necessarily phenomenological; they are not derived from quantum mechanical first principles but represent plausible hypotheses about biochemical interactions. They can be both complex, with many undetermined parameters, and contingent, being subject to alteration as assumptions change or new data become available. They present unique challenges in model building and model analysis, which have not been previously encountered in physics, mathematics, or engineering [49]. Furthermore, the dynamical systems approach taken in Section V, which is customarily used in systems biology, represents a system as a time evolution on a state space. It says nothing about signals. In contrast, the systems paradigm used in electronic engineering represents a system as a functional on a space of signals, in which state is only implicit. Equation (1) is a hybrid: inputs are signals while outputs are state. The theory of such hybrid systems has yet to be developed.

Systems biology borrows ideas from the physical sciences and presents many exciting challenges for physical scientists. Progress, however, requires a delicate integration of ideas and biological data. In the physical sciences, data are cheap and ideas are priceless. In biology, there are unsolved problems everywhere you look, and anyone can have ideas about them. Accordingly, ideas are cheap and data are priceless. These divergent cultural economies come together somewhat uneasily in systems biology. I have tried to do justice here to both. ■

## Acknowledgment

The author thanks three anonymous reviewers for their thoughtful comments, which helped improve the exposition. He also thanks members of his laboratory and their collaborators for permission to mention unpublished work, and S. Prabhakaran and T. Thorsen for their specific comments. Y. Xu is studying the distribution of Erk activation after EGF stimulation, and his data is shown in Fig. 2. Prabhakaran is examining the systems biochemistry of multisite phosphorylation in a collaboration with H. Steen at Children's Hospital, while the mathematical results in Section V are joint work with M. Thomson [84]. He is grateful to the organizers of the Biological Switches and Clocks meeting, Kavli Institute for Theoretical Physics, for providing a stimulating environment in which to begin this paper.

## REFERENCES

[1] U. Alon, M. G. Surette, N. Barkai, and S. Leibler, "Robustness in bacterial chemotaxis," *Nature*, vol. 397, pp. 168–171, 1999.

[2] C. P. Bagowski, J. Besser, C. R. Frey, and J. E. Ferrell, "The JNK cascade as a biochemical switch in mammalian cells: Ultrasensitive and all-or-none responses," *Curr. Biol.*, vol. 13, no. 4, pp. 315–320, 2003.

[3] C. P. Bagowski and J. E. Ferrell, "Bistability in the JNK cascade," *Curr. Biol.*, vol. 11, no. 15, pp. 1176–1182, 2001.

[4] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler, "Bacterial persistence as a phenotypic switch," *Science*, vol. 305, pp. 1622–1625, 2004.

[5] D. Barken, C. J. Wang, J. Kearns, A. Hoffmann, and A. Levchenko, "Comment on 'oscillations in NF $-$ $\kappa$B signalling control the dynamics of gene expression'," *Science*, vol. 308, p. 52a, 2005.

[6] H. C. Berg, *E. coli in Motion*. New York: Springer-Verlag, 2004.

[7] H. C. Berg and P. M. Tedesco, "Transient response to chemotactic stimuli in Escherichia coli," *Proc. Nat. Acad. Sci. USA*, vol. 72, pp. 3235–3239, 1975.

[8] M. R. Birtwistle, M. Hatakeyama, N. Yumoto, B. A. Ogunnaike, J. B. Hoek, and B. N. Kholodenko, "Ligand-dependent responses of the ErbB signaling network: Experimental and modeling analyses," *Mol. Syst. Biol.*, vol. 3, p. 144, 2007.

[9] M. L. Blinov, J. R. Faeder, B. Goldstein, and W. S. Hlavacek, "A network model of early events in epidermal growth factor receptor signaling that accounts for combinatorial

complexity," *Biosystems*, vol. 83, pp. 136–151, 2006.

[10] H. Cha and P. Shapiro, "Tyrosine-phosphorylated extracellular signal-regulated kinase associates with the golgi complex during G2/M phase of the cell cycle: Evidence for regulation of golgi structure," *J. Cell. Biol.*, vol. 153, pp. 1355–1367, 2001.

[11] S. Cohen, "Epidermal growth factor," in *Nobel Lectures, Physiology or Medicine 1981–1990*, T. Frängsmyr, Ed. Singapore: World Scientific, 1993.

[12] A. Colman-Lerner, A. Gordon, E. Serra, T. Chin, O. Resnekov, D. Endy, C. G. Pesce, and R. Brent, "Regulated cell-to-cell variation in a cell-fate decision system," *Nature*, vol. 437, pp. 699–706, 2005.

[13] A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, 2nd ed. London, U.K.: Portland, 1995.

[14] F. Crick, "Memory and molecular turnover," *Nature*, vol. 312, p. 101, 1984.

[15] D. A. Drubin, J. C. Way, and P. A. Silver, "Designing biological systems," *Genes Dev.*, vol. 21, pp. 242–254, 2007.

[16] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, "Stochastic gene expression in a single cell," *Science*, vol. 297, pp. 1183–1186, 2002.

[17] J. E. Ferrell and R. R. Bhatt, "Mechanistic studies of the dual phosphorylation of mitogen-activated protein kinase," *J. Biol. Chem.*, vol. 272, pp. 19 008–19 016, 1997.

[18] E. H. Fischer, "Protein phosphorylation and cellular regulation, II," in *Nobel Lectures, Physiology or Medicine 1991–1995*, N. Ringertz, Ed. Singapore: World Scientific, 1997.

[19] M. Freeman, "Cell determination strategies in the Drosophila eye," *Development*, vol. 124, pp. 261–270, 1997.

[20] A. Friedman and N. Perrimon, "A functional RNAi screen for regulators of receptor tyrosine kinase and ERK signalling," *Nature*, vol. 444, pp. 230–234, 2006.

[21] B. D. Gompaerts, L. M. Kramer, and P. E. R. Tatham, *Signal Transduction*, New York: Academic, 2002.

[22] A. Gschwind, O. M. Fischer, and A. Ullrich, "The discovery of receptor tyrosine kinases: Targets for cancer therapy," *Nat. Rev. Cancer*, vol. 4, pp. 361–370, 2004.

[23] J. Gunawardena, "Multisite protein phosphorylation makes a good threshold but can be a poor switch," *Proc. Nat. Acad. Sci. USA*, vol. 102, pp. 14 617–14 622, 2005.

[24] D. P. Hanger, H. L. Byers, S. Wray, K. Y. Leung, M. J. Saxton, A. Seereeram, C. H. Reynolds, M. A. Ward, and B. H. Anderton, "Novel phosphorylation sites in tau from Alzheimer brain support a role for case in kinase 1 in disease pathogenesis," *J. Biol. Chem.*, vol. 282, pp. 23 645–23 654, 2007.

[25] R. C. Harris, E. Chung, and R. J. Coffey, "EGF receptor ligands," *Exp. Cell Res.*, vol. 284, pp. 2–13, 2003.

[26] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *J. Physiol.*, vol. 117, pp. 500–544, 1942.

[27] C. I. Holmberg, S. E. F. Tran, J. E. Eriksson, and L. Sistonen, "Multisite phosphorylation provides sophisticated regulation of transcription factors," *Trends Biochem. Sci.*, vol. 27, pp. 619–627, 2002.

[28] C.-Y. F. Huang and J. E. Ferrell, "Ultrasensitivity in the mitogen-activated protein kinase cascade," *Proc. Nat. Acad. Sci. USA*, vol. 93, pp. 10 078–10 083, 1996.

[29] N. E. Hynes and H. A. Lane, "ErbB receptors and cancer: The complexity of targeted inhibitors," *Nat. Rev. Cancer*, vol. 5, pp. 341–354, 2005.

[30] K. A. Janes, J. G. Albeck, S. Gaudet, P. K. Sorger, D. A. Lauffenburger, and M. B. Yaffe, "A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis," *Science*, vol. 310, pp. 1646–1653, 2005.

[31] R. J. Johnston, S. Chang, J. F. Etchberger, C. O. Ortiz, and O. Hobert, "MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision," *Proc. Nat. Acad. Sci. USA*, vol. 102, pp. 12 449–12 454, 2005.

[32] A. R. Joyce and B. O. Palsson, "The model organism as a system: Integrating 'omics' data sets," *Nat. Rev. Mol. Cell Bio.*, vol. 7, pp. 198–210, 2006.

[33] H. Kacser and J. A. Burns, "The control of flux," *Biochem. Soc. Trans.*, vol. 23, pp. 341–66, 1995, reprint of 1973 paper in *Symp. Soc. Exp. Biol.*

[34] E. Kandel, "The molecular biology of memory storage: A dialog between genes and synapses," in *Nobel Lectures, Physiology and Medicine 1996–2000*, H. Jörnvall, Ed. Singapore: World Scientific, 2003.

[35] M. Kirschner, "The meaning of systems biology," *Cell*, vol. 121, pp. 503–504, 2005.

[36] C. Koch, *Biophysics of Computation: Information Processing in Single Neurons*. Oxford, U.K.: Oxford Univ. Press, 1999.

[37] D. E. Koshland, "Bacterial chemotaxis in relation to neurobiology," *Ann. Rev. Neurosci.*, vol. 3, pp. 43–75, 1980.

[38] E. G. Krebs, "Protein phosphorylation and cellular regulation, I," in *Nobel Lectures, Physiology or Medicine 1991–1995*, N. Ringertz, Ed. Singapore: World Scientific, 1997.

[39] G. Lahav, N. Rosenfeld, A. Sigal, N. Geva-Zatorsky, A. J. Levine, M. B. Elowitz, and U. Alon, "Dynamics of the p53-mdm2 feedback loop in individual cells," *Nature Genetics*, vol. 36, pp. 147–150, 2004.

[40] D. A. Lauffenburger and J. Linderman, *Receptors: Models for Binding, Trafficking and Signaling*. Oxford, U.K.: Oxford Univ. Press, 1995.

[41] R. Levi-Montalcini, "The nerve growth factor: Thirty five years later," in *Nobel Lectures, Physiology or Medicine 1981–1990*, T. Frängsmyr, Ed. Singapore: World Scientific, 1993.

[42] J. Lisman and S. Raghavachari, "A unified model of the presynaptic and postsynaptic changes during LTP at CA1 synapses," *Sci. STKE*, vol. 356, 2006, 10.1126/stke.3562006re11.

[43] J. E. Lisman, "A mechanism for memory storage insensitive to molecular turnover: A bistable autophosphorylating kinase," *Proc. Nat. Acad. Sci. USA*, vol. 82, pp. 3055–3057, 1985.

[44] J. E. Lisman, H. Schulman, and H. Kline, "The molecular basis of CaMKII function in synaptic and behavioural memory," *Nat. Rev. Neurosci.*, vol. 3, pp. 175–190, 2002.

[45] E. M. Lucchetta, J. H. Lee, L. A. Fu, N. H. Patel, and R. F. Ismagilov, "Dynamics of Drosophila embryonic patterning perturbed in space and time using microfluidics," *Nature*, vol. 434, pp. 1134–1138, 2005.

[46] T. J. Lynch, D. W. Bell, R. Sordella, S. Gurubhagvatula, R. A. Okimoto, B. W. Brannigan, P. L. Harris, S. M. Haserlat, J. G. Supko, F. G. Haluska, D. N. Louis, D. C. Christiani, J. Settleman, and D. A. Haber, "Activating mutations in the epidermal growth factor receptor underlying responsivness of non-small-cell-lung cancer to Gefinitib," *N. Engl. J. Med.*, vol. 350, pp. 2129–2139, 2004.

[47] B. Macek, I. Mijakovic, J. V. Olsen, F. Gnad, C. Kumar, P. R. Jensen, and M. Mann, "The serine/threonine/tyrosine phosphoproteome of the model bacterium Bacillus subtilis," *Mol. Cell. Proteomics*, vol. 6, pp. 697–707, 2007.

[48] R. M. Macnab and D. E. Koshland, "The gradient-sensing mechanism in bacterial chemotaxis," *Proc. Nat. Acad. Sci. USA*, vol. 69, pp. 2509–2512, 1972.

[49] A. Mallavarapu, M. Thomson, B. Ullian, and J. Gunawardena, "Programming with models: Modularity and abstraction provide powerful capabilities for systems biology," *J. R. Soc. Interface*, 2008, to appear.

[50] A. Manrai and J. Gunawardena, "The geometry of multisite phosphorylation," submitted for publication.

[51] Y. Marcy, C. Ouverney, E. M. Bik, T. Lösekann, N. Ivanova, H. G. Martin, E. Szeto, D. Platt, P. Hugenholtz, D. A. Relman, and S. R. Quake, "Dissecting biological 'dark matter' with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth," *Proc. Nat. Acad. Sci. USA*, vol. 104, pp. 11 889–11 894, 2007.

[52] N. I. Markevich, J. B. Hoek, and B. N. Kholodenko, "Signalling switches and bistability arising from multisite phosphorylation in protein kinase cascades," *J. Cell Biol.*, vol. 164, pp. 353–359, 2004.

[53] B. A. Mello and Y. Tu, "Perfect and near-perfect adaptation in a model of bacterial chemotaxis," *Biophys. J.*, vol. 84, pp. 2943–2956, 2003.

[54] J. Monod and F. Jacob, "General conclusions: Teleonomic mechanisms in cellular metabolism, growth and differentiation," in *Proc. Cold Spring Harbor Symp. Quant. Biol.*, 1961, vol. 26, pp. 389–401.

[55] M. Natarajan, K.-M. Lin, R. C. Hsueh, P. C. Sternweis, and R. Ranganathan, "A global analysis of cross-talk in a mammalian cellular signalling network," *Nat. Cell Biol.*, vol. 8, pp. 571–580, 2006.

[56] K. P. Nichols, J. R. Ferullo, and A. J. Baeumner, "Recirculating, passive micromixer with a novel sawtooth structure," *Lab Chip*, vol. 6, pp. 242–246, 2006.

[57] N. Normano, A. D. Luca, C. Bianco, L. Strizi, M. Mancino, M. R. Maiello, A. Carotenuto, G. D. Feo, F. Caponigro, and D. S. Salomon, "Epidermal growth factor recpetor (EGFR) signaling in cancer," *Gene*, vol. 366, pp. 2–16, 2006.

[58] F. Ortega, J. L. Garcés, F. Mas, B. N. Kholodenko, and M. Cascante, "Bistability from double phosphorylation in signal transduction," *FEBS J.*, vol. 273, pp. 3915–3926, 2006.

[59] E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. van Oudenaarden, "Multistability in the lactose utilization network of Escherichia coli," *Nature*, vol. 427, pp. 737–740, 2004.

[60] S. Paliwal, P. A. Iglesias, K. Campbell, Z. Hilioti, and A. Levchenko, "MAPK-mediated bimodal gene expression and adaptive gradient sensing in yeast," *Nature*, vol. 446, pp. 46–51, 2007.

[61] T. Pawson and P. Nash, "Assembly of cell regulatory systems through protein

interaction domains," *Science*, vol. 300, pp. 445–452, 2003.

[62] P. Pellicena and W. T. Miller, "Processive phosphorylation of p130Cas by Src depends on SH3-polyproline interactions," *J. Biol. Chem.*, vol. 276, pp. 28 190–28 196, 2001.

[63] D. W. Piston and G. J. Kremers, "Fluorescent protein FRET: The good, the bad and the ugly," *Trends Biochem. Sci.*, vol. 32, pp. 407–414, 2007.

[64] M. Raman, W. Chen, and M. H. Cobb, "Differential regulation and properties of MAPKs," *Oncogene*, vol. 26, pp. 3100–3112, 2007.

[65] A. R. Reynolds, C. Tischer, P. J. Verveer, O. Rocks, and P. I. H. Bastiaens, "EGFR activation coupled to inhibition of tyrosine phosphatases causes lateral signal propagation," *Nature Cell Biol.*, vol. 5, pp. 447–453, 2003.

[66] D. M. Rothman, M. D. Shults, and B. Imperiali, "Chemical approaches for investigating phosphorylation in signal transduction networks," *Trends Cell Biol.*, vol. 15, pp. 502–510, 2005.

[67] J. Saez-Rodriguez, A. Goldsipe, J. Muhlich, L. G. Alexopoulos, B. Millard, D. A. Lauffenburger, and P. K. Sorger, "Flexible informatics for linking experimental data to mathematical models via DataRail," *Bioinformatics*, vol. 24, pp. 840–847, 2008.

[68] C. Salazar and T. Höfer, "Allosteric regulation of the transcription factor NFAT1 by multiple phosphorylation sites: A mathematical analysis," *J. Mol. Biol.*, vol. 327, pp. 31–45, 2003.

[69] C. Salazar and T. Höfer, "Versatile regulation of multisite protein phosphorylation by the order of phosphate processing and protein-protein interactions," *FEBS J.*, vol. 274, pp. 1046–1060, 2007.

[70] A. Samadhani, J. Mettetal, and A. van Oudenaarden, "Cellular asymmetry and individuality in directional sensing," *Proc. Nat. Acad. Sci. USA*, vol. 103, pp. 11549–11554, 2006.

[71] M. P. Sanderson, P. J. Dempsey, and A. J. Dunbar, "Control of ErbB signaling through metalloprotease mediated ectodomain shedding of EGF-like factors," *Growth Factors*, vol. 24, pp. 121–136, 2006.

[72] M. Schena, D. Shalon, R. W. Davis, and P. O. Brown, "Quantitative monitoring of gene expression patterns with a complementary DNA microarray," *Science*, vol. 270, pp. 467–470, 1995.

[73] J. Schlessinger, "Ligand-induced, receptor-mediated dimerization and activation of EGF receptor," *Cell*, vol. 110, pp. 669–672, 2002.

[74] B. Schoeberl, C. Eichler-Jonsson, E. D. Gilles, and G. Müller, "Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors," *Nat. Biotechnol.*, vol. 20, pp. 370–375, 2002.

[75] W. X. Schulze, L. Deng, and M. Mann, "Phosphotyrosine interactome of the ErbB-receptor kinase family," *Mol. Syst. Biol.*, vol. msb4100012, pp. E1–E13, 2005.

[76] S. V. Sharma, M. A. Fischbach, D. A. Haber, and J. Settleman, "'Oncogenic shock': Explaining oncogene addiction through differential signal attenuation," *Clin. Cancer Res.*, vol. 12, pp. 4392s–4395s, 2006.

[77] M. Springer, D. D. Wykoff, N. Miller, and E. K. O'Shea, "Partially phosphorylated Pho4 activates transcription of a subset of phosphate-responsive genes," *PLOS Biol.*, vol. 1, pp. 261–270, 2003.

[78] R. A. Stein and J. V. Staros, "Evolutionary analysis of the ErbB receptor and ligand families," *J. Mol. Evol.*, vol. 50, pp. 397–412, 2000.

[79] A. M. Stock, V. L. Robinson, and P. N. Goudreau, "Two component signal transduction," *Annu. Rev. Biochem.*, vol. 69, pp. 183–215, 2000.

[80] C. Sweeney and K. L. Carraway, "Ligand discrimination by ErbB receptors: Differential signaling through differential phosphorylation site usage," *Oncogene*, vol. 19, pp. 5568–5573, 2000.

[81] T. Thorsen, S. J. Maerkl, and S. R. Quake, "Microfluidic large-scale integration Science, vol. 298, pp. 580–584, 2002.

[82] L. H. Tay, O. Griesbeck, and D. T. Yue, "Live-cell transforms between $Ca^{2+}$ transients and FRET responses a troponin-based $Ca^{2+}$ sensor," *Biophys. J.*, vol. 93, pp. 4031–4040, 2007.

[83] W. Thies, J. P. Urbanski, T. Thorsen, and S. Amarasinghe, "Abstraction layers for scalable microfluidic biocomputing," in *Proc. 2007 Nat. Comput.*

[84] M. Thomson and J. Gunawardena, "A new method of symbolic parameter analysis reveals unlimited multistability in multisite phosphorylation systems," submitted for publication.

[85] M. Thomson and J. Gunawardena, "The steady states of a multisite kinase, phosphatase, substrate system form a rational projective algebraic curve," in preparation.

[86] A. Ting, K. H. Kain, R. L. Klemke, and R. Y. Tsien, "Genetically encoded fluorescent reporters of protein tyrosine kinase activities in living cells," *Proc. Nat. Acad. Sci. USA*, vol. 98, pp. 15003–15008, 2001.

[87] M. A. Unger, H.-P. Chou, T. Thorsen, A. Scherer, and S. R. Quake, "Monolithic microfabricated valves and pumps by multilayer soft lithography," *Science*, vol. 288, pp. 113–116, 2000.

[88] A. H. West and A. M. Stock, "Histidine kinases and response regulator proteins in two-component signaling systems," *Trends Biochem. Sci.*, vol. 26, pp. 369–376, 2001.

[89] G. M. Whitesides, E. Ostuni, S. Takayama, X. Jiang, and D. E. Ingber, "Soft lithography in biology and biochemistry," *Annu. Rev. Biomed. Eng.*, vol. 3, pp. 335–373, 2001.

[90] Z. Yao, Y. Dolginov, T. Hanoch, Y. Yung, G. Ridner, Z. Lando, D. Zharhary, and R. Seger, "Detection of partially phosphorylated forms of ERK by monoclonal antibodies reveals spatial regulation of ERK activity by phosphatases," *FEBS Lett.*, vol. 18, pp. 37–42, 2000.

[91] Y. Yarden and M. X. Sliwkowski, "Untangling the ErbB signalling network," *Nat. Rev. Mol. Cell Biol.*, vol. 2, pp. 127–137, 2001.

[92] T.-M. Yi, Y. Huang, M. I. Simon, and J. Doyle, "Robust perfect adaptation in bacterial chemotaxis through integral feedback control," *Proc. Nat. Acad. Sci. USA*, vol. 97, pp. 4649–4653, 2000.

[93] Y. Zhao and Z.-Y. Zhang, "The mechanism of dephosphorylation of extracellular signal-regulated kinase 2 by mitogen-activated protein kinsae phosphatase 3," *J. Biol. Chem.*, vol. 276, pp. 32382–32391, 2001.

## ABOUT THE AUTHOR

**Jeremy Gunawardena** received the Ph.D. degree in pure mathematics from Trinity College, Cambridge, U.K.

He is Director of the Virtual Cell Program in the Department of Systems Biology, Harvard Medical School, Cambridge, MA. Prior to returning to academia, he spent 14 years with Hewlett-Packard (HP) Research, becoming Director of HP's Basic Research Institute in the Mathematical Sciences. He has held faculty and visiting appointments in mathematics and computer science at the University of Chicago, Stanford University, the Massachusetts Institute of Technology, Cambridge University, and the École Normale Supérieure. He was a Member of the Council of EPSRC, the United Kingdom's engineering and physical sciences funding agency. His current research interests lie in mammalian signal transduction, approached through a combination of experiment, computation, and theory.

Dr. Gunawardena received a Research Fellowship from Trinity College.