Intracellular signaling networks, which are composed of interconnected biochemical pathways, regulate and actuate responses such as cell-cycle progression and cell migration, survival, and differentiation. Although our knowledge of the intricate biochemical mechanisms at the level of individual proteins and molecular interactions is ever expanding, those details leave us with an even murkier view of how the complex network operates as a whole. True understanding requires knowing not only what happens at the molecular level but also how these mechanisms influence the precise magnitude, timing, and spatial localization of signal transduction processes. Hence, mathematical modeling and analysis has emerged in recent years as a legitimate approach for interpreting experimental results and generating novel hypotheses for additional study and model refinement. Once conducted in isolation and scorned by most biologists, quantitative modeling has moved into the mainstream as a powerful tool for the analysis of cell signaling. In this article, the biological, chemical, and physical underpinnings of this approach are presented, as are its current applications and future challenges.

Biological research has been focused increasingly on the basis for cell regulation and function at the molecular level, and as a result, we now have a detailed map of how intracellular molecules are organized to form signal transduction pathways and cascades (Fig. 1). The “protein jungle,” as articulated by Bray a decade ago (1), is characterized by a wealth of qualitative information about the connectivity of pathways (the so-called interactome) but relatively little in the way of quantitative measurements. For example, one might wish to know how some meaningful quantity, such as the activity of a particular enzyme, changes as a function of time and stimulus in a given cell type. A quantitative assay could be developed for measuring that quantity. Understanding those kinetics fully, however, would require detailed knowledge of the underlying regulation mechanisms, of which several typically exist. Reconstituting those mechanisms in a test tube, in most cases, is prohibitively difficult, as is their systematic manipulation in the cell. Ultimately, the regulatory mechanisms would have to be characterized in terms of equilibrium and rate constants and intracellular concentrations. Finally, it would be useful to develop a correlation between the magnitude and timing of that pathway and the quality of cellular responses, such as rates of proliferation and probability of survival. Even if obtaining such a data set were feasible, it is not obvious how one would analyze it and extract useful insights and predictions from it.

This article deals with the mathematical modeling of signal transduction pathways and networks, which has emerged as a powerful tool that can aid in interpreting such quantitative data sets. In principle, quantitative models offer three advantages over the conceptual “arrow diagrams” that are encountered routinely in the signaling literature. First, a consequence of their mathematical construction is that they are precise, and the inherent assumptions may be laid out clearly. Second, to the extent that the underlying molecular biology is known, quantitative models can be mechanistic. Mechanistic models are based on established physico-chemical principles, in which case the form of the model equations is determined to a significant extent by the hypothetical mechanisms assumed. Thus, it is possible to evaluate the relative merits of different candidate mechanisms by comparison with experiment. In contrast with mechanistic models, phenomenological models are meant only to capture experimentally observed relationships in an empirical way. They naturally are less powerful, but they serve a definite and useful role and are appropriate in situations in which the underlying mechanisms are less certain. A prominent example is the sort of statistical, correlative models central to the field of bioinformatics, although such approaches...
Signal Transduction Networks: Biological Background

The molecular basis of signal transduction lies in the noncovalent interactions among cell-associated proteins, lipids, and other biomolecules, which leads to the formation of transient complexes, and the covalent modifications of those biomolecules by specific, enzyme-catalyzed reactions. Enzymatic activities, in turn, are regulated by interactions with other proteins, lipids, or small molecules/ions and by covalent modifications. Mediating all of these processes are receptors, cell-associated proteins responsible for sensing specific chemical factors (ligands), typically through noncovalent interactions at the cell surface. Receptors activated by ligation, or in some cases by mechanical forces, activate intracellular pathways and thus preside over signal transduction.

## Receptor–ligand interactions and receptor dynamics

Receptors are responsible for transmitting information about the external environment to the cell internum, and therefore their mechanism of activation and action must be well characterized before the details of the intracellular networks can be analyzed and modeled mathematically. In most cases, receptor activation is induced by the noncovalent binding of a specific ligand. Expression of the cognate receptor, typically existing at copy numbers of $10^7$–$10^8$ per cell, is tantamount to the ability of a cell to respond to the ligand, and thus chemical ligands signal specific cell types. Most ligands are soluble and diffusible, and such ligands are termed growth factors, cytokines, hormones, or agonists depending on the type of receptor engaged and/or the response elicited. Other ligands are associated with the extracellular matrix or other cells and mediate cell–matrix adhesion and cell–cell interactions, respectively, in addition to signaling functions. In many cases, ligation induces receptor dimerization or oligomerization as a prerequisite for signal transduction. For example, it is well established that receptor tyrosine kinases (RTKs) receptors that engage certain growth factors in a variety of cell types, must dimerize for their intrinsic enzymatic activity to phosphorylate specific tyrosine residues in the cytoplasmic portion of each receptor. The decoration of the receptor with phospho-tyrosine provides a scaffold for the recruitment of a host of signaling enzymes (4).

Receptors and receptor–ligand complexes on the cell surface are not static. Their lateral mobility in the plasma membrane enables receptor dimerization, as described above, but another critical aspect of receptor dynamics is receptor trafficking, whereby components of the plasma membrane are internalized, delivered to intracellular compartments called endosomes, and sorted either for recycling back to the cell surface or for enzymatic degradation (5). Thus, differential sorting of activated and inactive receptors provides a mechanism for regulating the number of cell surface receptors in response to chronic ligand exposure, a process called receptor downregulation, and can contribute to the clearance of external ligands over time. Typically, activated receptors are internalized at a faster rate and their intracellular sorting fate depends on the persistence of the receptor–ligand interaction in endosomes. Any realistic model of receptor-mediated signal transduction should take into consideration the processes of ligand binding, receptor dimerization or oligomerization, and receptor/ligand trafficking. Receptors that are well characterized, such as the epidermal growth factor receptor, therefore have attracted the most attention from modelers (6).

## Modular functions of signaling proteins

Signaling proteins form complexes with other molecules using conserved motifs or domains, which are modular in the sense that the domain by itself is sufficient for its function (7, 8). Protein–protein interaction domains include those responsible for binding to phosphorylated receptors and other tyrosine-phosphorylated proteins (Src homology 2 and phosphotyrosine binding domains) or to proline-rich protein

---

Figure 1 Partial representation of a signal transduction network, mediated by a cell surface receptor. The molecules are organized hierarchically (roughly, from top to bottom) according to their functions as adaptors, receptor-recruited enzymes, membrane-associated substrates, and effector kinases. Arrows represent activation mechanisms, whether by complex formation or covalent modification; “T” bars indicate negative regulators.
sequences (Src homology 3 domains); other domains are responsible for interactions with lipids (pleckstrin homology, FYVE, C1, and C2 domains). Together, these domains mediate the formation of multimolecular complexes that regulate the activities and mediate the targeting of signaling enzymes. Many signaling enzymes have a handful of such domains, whereas other signaling proteins called adaptors have no enzymatic function but possess domains that mediate the binding of enzymes to signaling complexes. From the standpoint of modeling, this complexity presents a definite challenge.

Physical organization and compartmentalization of signaling processes

Cells are not simply well-mixed reaction vessels, and cell processes are governed by physico-chemical principles, and accordingly, those principles may be used to translate known or hypothetical molecular mechanisms to mathematical equations. In this section, the general principles of chemical kinetics, mass transport, and fluid mechanics used to model chemical reaction systems are reviewed briefly.

Reaction and mass-action kinetics

In chemical kinetics, each molecule type in the system is considered a distinct species and a conservation equation is formulated to account for the change in the amount of each species per unit time. In this context, the system could be a single cell and the species could be cell-associated molecules; different states of a molecule, the simplest distinction being a two-state model (e.g., active versus inactive or phosphorylated versus unphosphorylated); are considered distinct species as are complexes of multiple molecules. The rate of each biochemical reaction or interaction is expressed in terms of the concentrations of the reacting/interacting species, defining the so-called rate law, which appears as a positive or negative term in the corresponding conservation equations depending on whether the transition generates or consumes the conserved species.

In formulating rate laws, it is common to invoke the law of mass action (Fig. 2a). In the biochemical context, this principle applies to only two types of processes: 1) the union of two species (complex formation), the rate defined as the product of both of their concentrations and a bimolecular rate constant, and 2) a spontaneous, unimolecular transition, such as the dissociation of two species from a complex or the conformational modification of a substrate in a complex with an enzyme, which occurs with a constant mean probability per unit time that defines its rate constant. A rate process that follows mass-action kinetics often is termed an elementary reaction, although that definition technically implies specific criteria that generally are not met for reactions in liquids much less for biomolecules. The advantage of mass action kinetics is that the mathematical form of the rate law is dictated by the mechanism, in which case one needs to specify only the values of the corresponding rate constants. For complicated rate processes, a notable example being the action of highly cooperative enzymes, a more abstract rate law that does not reflect the precise mechanism but nonetheless is in quantitative agreement with experimental measurements might be assumed.

Diffusion and mass transport

Biochemical rate processes depend on local species concentrations, which are not necessarily constant. When significant concentration gradients exist, it is appropriate to include a spatial dependence in the conservation equations, which then are
Mathematical Modeling of Biological Signaling Networks

Figure 2  Mathematical modeling of cellular processes: an illustrative example. A. Consider the following mechanism for ligand/receptor dynamics: Cell surface receptors (R) are synthesized at a constant rate and bind reversibly with a ligand (L) to form a complex (C). Both free and bound receptors are internalized and later degraded, but they do so at different rates. Based on this mechanism, the law of mass action is used to construct the conservation equations for R and C.

\[
\frac{dR}{dt} = V_L - k_R R - k_{RL} LR + k_C C
\]
\[
\frac{dC}{dt} = k_{RL} LR - k_C C - k_C C
\]

B. Addition of transport effects. In the case of autocrine signaling, the cell is both the source of the secreted ligand and the responder. Spherical geometry is adopted, and a simple reaction/diffusion model is used to calculate the ligand concentration profile at steady state, assuming a dilute suspension of cells and the same receptor dynamics as in A. This profile is given by \( [L] = [L]_s \left( \frac{a}{r} \right) \), where \([L]_s\) is the value of \([L]\) at the cell surface (\(r = a\)). It is shown readily that the maximum value of \([L]_s\), achieved when no receptor binding exists, is equal to \(aV_L / D_L\), where \(V_L\) is the cell rate of ligand secretion per unit area and \(D_L\) is the diffusion coefficient of the ligand. The actual value of \([L]_s\) relative to the maximum is found to be a function of only two dimensionless variables: the ratio of receptor and ligand synthesis rates (equal to 1 for the plotted results) and a parameter that characterizes the efficiency of the receptor-mediated ligand capture (defined as \(\frac{a k_f R}{D_L (k_r + k_e)}\), where \(R_0 = V_k / k_t\); values of 100, 10, 1, and 0.1 were used here).

\[
\frac{\partial [L]}{\partial r} - \frac{D_L \partial [L]}{r^2} \left( r \frac{\partial [L]}{\partial r} \right) = 0
\]

\[-D_L \frac{\partial [L]}{\partial r} + V_c - k_{RL} [L]_s R + k_C C; \ [L]_s = 0\]

4

WILEY ENCYCLOPEDIA OF CHEMICAL BIOLOGY © 2008, John Wiley & Sons, Inc.
they evolve in a deterministic fashion, according to the aforementioned conservation equations and their initial and boundary constraints. The implementation of such models involves the solution of ordinary and partial differential equations (ODEs and PDEs, respectively). ODEs commonly are encountered in kinetic models of cell signaling systems in which the species concentrations are assumed to be spatially homogeneous within the cell/cell compartment or are averaged over the domain in an appropriate way. In kinetic models of even modest complexity, the ODEs are nonlinear and therefore must be solved using numerical methods. With the advent of efficient, implicit solution algorithms and steadily increasing computing power, handling even large systems of ODEs is straightforward and the integration of these algorithms in various software packages is widespread.

Solution of PDEs, commonly encountered in spatial modeling, is more complicated and computationally intensive by comparison, but several approximate numerical methods are available for this purpose. These include the finite difference, finite volume, and finite element methods, which vary according to their ease of implementation and applicability. The finite element method is the most complicated but generally applicable method for complex domain geometries; numerous software packages that implement this method are available, although the most powerful commercial packages are quite expensive.

Fluid mechanics and mechanical forces

Mechanical stress and strain have not been given due consideration in the broad context of cell biochemistry, yet mechanical forces can affect intracellular signaling processes in at least two distinct ways. First, at the level of macroscopic fluid flow, the conservation of fluid momentum determines the velocity field \( \mathbf{v} \), which along with active transport considerations determines the contribution of convective mass transport to the conservation of species \( i \), as discussed in the previous section. Typically, this contribution to the reaction kinetics is neglected. Second, at the microscopic level, mechanical forces might alter directly the functions of macromolecules in prescribed ways, for example by exposing a cryptic binding pocket. This mode of regulation is thought to be at the core of mechanotransduction, which refers to the ability of cells to sense and respond to applied forces (17, 18).

Mathematical Modeling Tools and Techniques

Armed with reasonably good knowledge or hypotheses of the underlying biochemistry and the ability to formulate models based on physico-chemical principles, the models must be implemented to obtain and analyze the quantitative results and to generate predictions. Except in very simple, idealized cases, this task requires the use of various numerical methods and tools. In this section, an overview of the various model types and associated methods is offered.

Continuum models

The most common approach in mechanism-based modeling is to model the system as a continuum. The underlying assumption is that the state variables of the system (e.g., species concentrations) vary continuously in time and space and that, therefore, equation and assuming that diffusion dominates over convection and \( D_i \) is constant, one obtains the common reaction–diffusion equation:

\[
\frac{\partial C_i}{\partial t} = D_i \nabla^2 C_i + \sum_j v_{ij} r_j
\]

where \( \nabla^2 \) is the Laplacian operator (sometimes written as \( \Delta \)).

Despite its lack of grounding in pure physical chemistry, which has been a source of some contention (16), Fick’s Law and the reaction–diffusion equation have been used with great success to characterize processes in the cell cytoplasm and at cell membranes. Given the empirical nature of Fick’s Law and the notable differences between molecular motions at microscopic (nm) versus macroscopic (\( \mu \)m) length scales, it is more appropriate to refer to \( D_i \) as the apparent diffusivity or macroscopic mobility coefficient when applied to cell-associated molecules. An alternative approach, from the standpoint of modeling, is to develop a microscopic description of the system using Monte Carlo-based simulations (see the section “Mathematical Modeling Tools and Techniques” below).

MATHMATICAL MODELING OF BIOLOGICAL SIGNALLING NETWORKS

Fluid mechanics and mechanical forces

Mechanical stress and strain have not been given due consideration in the broad context of cell biochemistry, yet mechanical forces can affect intracellular signaling processes in at least two distinct ways. First, at the level of macroscopic fluid flow, the conservation of fluid momentum determines the velocity field \( \mathbf{v} \), which along with active transport considerations determines the contribution of convective mass transport to the conservation of species \( i \), as discussed in the previous section. Typically, this contribution to the reaction kinetics is neglected. Second, at the microscopic level, mechanical forces might alter directly the functions of macromolecules in prescribed ways, for example by exposing a cryptic binding pocket. This mode of regulation is thought to be at the core of mechanotransduction, which refers to the ability of cells to sense and respond to applied forces (17, 18).

Mathematical Modeling Tools and Techniques

Armed with reasonably good knowledge or hypotheses of the underlying biochemistry and the ability to formulate models based on physico-chemical principles, the models must be implemented to obtain and analyze the quantitative results and to generate predictions. Except in very simple, idealized cases, this task requires the use of various numerical methods and tools. In this section, an overview of the various model types and associated methods is offered.

Continuum models

The most common approach in mechanism-based modeling is to model the system as a continuum. The underlying assumption is that the state variables of the system (e.g., species concentrations) vary continuously in time and space and that, therefore,
to probabilities of the various rate processes. This method is easy to implement and exact, but it is computationally inefficient for so-called stiff problems that possess a broad range of characteristic time scales. Approximate adaptations of the technique, using a method called tau leaping, allows more efficient computation of stiff systems albeit with a potential loss in accuracy (24, 25). Another approximate approach, valid in the limit of large numbers of molecules, is the stochastic differential equation method, which also is referred to as the chemical master equation or chemical Langevin equation approach. In this method, the probabilities of the various states being populated by different numbers of molecules evolve in time according to ODEs. Although this method is not necessarily more efficient or applicable to stiff systems, it serves as the starting point for several powerful approximations that are valid in certain limiting cases (20).

Methods for modeling spatially fluctuating reaction–diffusion systems, wherein molecules are treated as discrete particles whose coordinates are tracked and moved with time, also are available. These methods account for stochastic transitions as motivated above and are appropriate when concentration gradients are expected, as in the case of diffusion-controlled reactions. They also allow one to generate movies of the simulations to visualize the constituent processes in action. In the lattice Monte Carlo method, particles occupy nodes on a regularly spaced grid and move to adjacent nodes according to probabilities determined by the diffusion coefficient (and by long-range intermolecular potentials, if applicable). Changes in state occur probabilistically, and intermolecular interactions are subject to specific rules. For example, when the molecules are treated as point particles, interactions might occur in the next time step if the two species occupy adjacent nodes. In this limit, the method is very efficient computationally but suffers in accuracy because of its inability to resolve off-lattice events, which is especially significant when concentration gradients are steep. The Brownian dynamics or dissipative particle dynamics method is similar to the lattice Monte Carlo except that the particles adopt coordinates that are continuous throughout the domain. Thus, this method is accurate but also more computationally intensive. Both methods have been used in the context of signal transduction, particularly to study processes at cell membranes (26, 27).

Construction of biochemical networks

The multiplicity of modifications and interactions that signal- ing molecules engage in constitutes a large number of potential activity states. For instance, a protein or protein complex with 10 phosphorylation sites has \(2^{10} = 1,024\) distinct combinations of modified or unmodified sites, and this number does not even include the binding status of each modified site. This issue, termed combinatorial complexity, presents a major challenge for detailed kinetic modeling, and hence rule-based modeling methods and accompanying software tools have been developed (28–30). In this approach, the individual interactions and reactions in the network are enumerated and the specific contexts in which they are allowed to occur (rules) are specified. The network of possible species, which could number in the hundreds or thousands, is generated automatically in the form of a system of ODEs; alternatively, the rules can be used in a stochastic simulation wherein the species of the network are formed spontaneously. Efforts are underway to extend this approach to the spatial domain as well.

Applications of Mathematical Modeling in Cell Signaling

With the availability of computational tools and, more importantly, the ability to judiciously apply physico-chemical principles to the biologic realm, what systems are and have been ripe for quantitative modeling? In this section, signal transduction processes that have been modeled successfully from the standpoint of yielding biologically meaningful insights are surveyed. Limitations on the number of references preclude a comprehensive coverage of this literature.

Models of specific signaling pathways and processes

For several reasons, no pathway has received more attention from modelers recently than the mitogen-activated protein kinase (MAPK) cascade, which involves the successive activation of three enzymes (Fig. 3). MAPks are conserved in eukaryotes from yeast to man and are critical for cell functions that range from proliferation to differentiation to stress responses, among others. MAPKs are considered master integrators of upstream signals and master controllers of transcription factors and other downstream effectors. Aside from their obvious importance in cell biology, MAPK cascades have achieved “über-pathway” status among modelers because of their potentially interesting dynamical properties. Building on pioneering work by Goldbeter and Koshland (31), who examined a hypothetical sequence of reversible enzymatic steps, the first model of a MAPK cascade was offered by Huang and Ferrell (32), who showed that the pathway was capable of a sensitive, switch-like relationship between the input (upstream of the cascade) and output (activation of MAPK); the basis for this response was the assumed biochemical mechanism, whereby MAPK and the upstream kinase are each activated by dual phosphorylation steps in a nonpro- cessive fashion. Building on this finding, more recent modeling efforts have focused on the consequences of the negative feedback regulation of the cascade (33–35) and the possibility that MAPK cascades operate as bistable switches, wherein two stable states (low and high MAPK activation) can be achieved at the same input strength (36, 37). Because of these modeling efforts, a solid understanding exists of what the cascade is capable of and how it apparently is modulated. Other models have been used to study the dynamics of MAPK cascades in the context of a full pathway, initiated by the activation of a specific receptor (38, 39). This approach is equally insightful albeit less general.

Another proving ground for mathematical modeling in cell signaling has been the dynamics of intracellular calcium. Spurred by the nonlinear nature of the calcium release mechanisms, which exhibit cooperativity and both positive and negative feedbacks, and by the availability of fluorescent dyes for...
Prospects and Challenges

Quantitative models of signal transduction processes, in conjunction with quantitative experimentation, are being used to evaluate biochemical mechanisms, predict the outcomes of novel experiments, and generate nonintuitive insights and hypotheses warranting additional study. One challenge we now face is how best to integrate such models to analyze complex intracellular and cell–cell communication systems. Although it is envisioned that quantitative models of cell physiology ultimately will progress in lock step with our knowledge of biochemical mechanisms, several hurdles loom on the horizon. Arguably the most important of these hurdles is the specification of parameter values and parameter estimation from data sets. With a few notable exceptions, mining the literature for “known” parameters is a dicey prospect, and this practice should be used only to provide reasonable, order-of-magnitude guesses for parameter values. Estimating parameters in their particular intracellular context would be ideal, and algorithms for doing so are applied readily, but a significant amount of quantitative data is needed. This exercise is only tractable for models with a modest number of parameters, and hence the challenge arises: a very detailed signaling network model might require hundreds of free parameters. At the pathway/network level, the formulation of simplified, coarse-grain models with lumped parameters offers a reasonable compromise. Even as experimental techniques for quantifying cellular processes become increasingly refined, the formulation of a useful model will continue to rely on biological and mathematical intuition, which, at its core, is the art of modeling.

Pathway models related to specific cell responses

Signaling pathways drive phenotypic responses, but the precise details of how this occurs at the molecular level remain elusive. Consider cell migration. From the signaling literature, we know which receptor-mediated pathways are more or less important in controlling migration, and from the literature on cytokine- and adhesion dynamics we have a good understanding of how migration is coordinated, but currently the interface between those two fields is poorly understood by comparison. To move forward with a mathematical model capable of yielding mechanistic insights, the prudent course of action is to focus on the molecular details of the upstream signaling and use a coarse-grained, phenomenological model of the cell response or vice versa. In the context of cell migration and chemotaxis signaling, both approaches have been adopted in recent models (66–68). In other cell response–specific models, understandably more attention is paid to the execution of the response and less to the dynamics of the upstream regulation, as in the modeling of the cell cycle (69) and programmed cell death (70, 71). It is expected that such models will be refined as the details of how upstream signaling pathways regulate the execution of cell responses are revealed, and quantitative modeling and analysis could play a key role in elucidating those mechanisms.

Crosstalk in signal transduction networks

It has long been appreciated that signaling pathways/cascades do not operate in isolation. As considered theoretically by Bray (63), the use of multiple intermediates in signaling pathways affords more opportunities for regulation, often from parallel pathways (crosstalk). In fact, most pathways as classically defined are simply dominant routes of regulation embedded in larger networks of interactions in which proteins may interact with and/or modify multiple substrates (branch points) and receive regulatory inputs from multiple molecular partners (convergence points). Although experimental characterization of this important layer of complexity remains on the horizon, several groups have begun collecting data and building models of more complete signal transduction networks (38, 64, 65).

Mathematical Modeling of Biological Signaling Networks

Figure 3 The MAPK cascade. A signaling cascade generally refers to a series of enzyme modification processes, as in the activation of extracellular signal-regulated kinase (ERK) systems, which are mammalian mitogen-activated protein kinases (MAPks). The first kinase, Raf, is activated (indicated by an asterisk) by numerous inputs, allowing it to phosphorylate MEK on two sites; ERK is phosphorylated dually in a similar fashion by MEK. Each phosphorylation event is thought to require a separate encounter between enzyme and substrate, which gives rise to interesting dynamical properties. Not depicted here, but equally important, are the phosphatases that catalyze the reverse reaction.
References


