Modelling spatio-temporal interactions within the cell

PADMINI RANGAMANI* and RAVI IYENGAR

Department of Pharmacology and Biological Chemistry, Mount Sinai School of Medicine, One Gustave L Levy Place, New York, NY 10029, USA

**Corresponding author (Fax, 212-831-0114; Email, padmini.rangamani@mssm.edu)*

Biological phenomena at the cellular level can be represented by various types of mathematical formulations. Such representations allow us to carry out numerical simulations that provide mechanistic insights into complex behaviours of biological systems and also generate hypotheses that can be experimentally tested. Currently, we are particularly interested in spatio-temporal representations of dynamic cellular phenomena and how such models can be used to understand biological specificity in functional responses. This review describes the capability and limitations of the approaches used to study spatio-temporal dynamics of cell signalling components.

[Rangamani P and Iyengar R 2006 Modelling spatio-temporal interactions within the cell; *J. Biosci.* **32** 157–167]

1. Introduction

Cell signalling pathways make up the regulatory systems of mammalian cells. With the enormous progress in biochemistry and molecular biology of cellular processes over the past four decades and the emergence of new techniques in the last few years, an immense amount of data exists on cell signalling pathways and networks. Various groups (Bhalla and Iyengar 1999; Wiley *et al* 2003; Hautaniemi *et al* 2005; Ma'ayan *et al* 2005; Kowalewski *et al* 2006; Mayawala *et al* 2006; Saha and Schaffer 2006; Singh *et al* 2006; Stamatakis and Mantzaris 2006) have organized this data into interaction networks and models. This organization makes it possible to study the signalling networks in a modelling framework and obtain non-intuitive hypotheses based on numerical simulations. In mammalian cells, signals propagate within linear pathways in a nonlinear fashion and lead to interactions between pathways. Such interactions between pathways often lead to complex nonlinear behaviour that arises from the presence of regulatory negative and positive feedback loops (Bhalla and Iyengar 1999, 2001a, b). Feedforward motifs within networks also provide the capability to have multiple modes of controls, including the ability to have redundant pathways, and the ability to mount prolonged responses to brief input signals. To study these regulatory phenomena and follow the detailed spatial and temporal dynamics can be a daunting task given the number of components involved. However, mathematical modelling of the cell signalling networks provides us with a means to overcome this barrier and get a better understanding of the information processing capability of signalling networks.

Application of mathematical modelling is not new to biological processes. Beginning from the population prey models described by the Lotka-Volterra model and the Hodgkin-Huxley model for neuron spiking (Hodgkin and Huxley 1952), many models have been proposed for various biological processes. For biochemical processes, enzyme kinetics has long been modelling by reaction engineering processes. Pharmacological systems have also been studied, using chemical kinetics and reaction engineering approaches, for ligand-receptor interactions as well as define efficacy of an antagonist or an agonist (Black and Leff 1983; Kenakin 2004; Leff *et al* 1997). From these models, it is clear that the specific formulation of the mathematical model drives the results. Cellular metabolic processes have been modelled and have provided insights into the regulatory networks (Fong *et al* 2005; Papin *et al* 2005; Thiele *et al* 2005). For cell signalling networks, Bhalla and Iyengar (1999, 2001a, b) have shown how the interaction of networks can

Keywords. Cell signalling; modelling; spatio-temporal interactions

lead to bistable behaviour and the role of positive feedback in storing information.

The temporal dynamics of signalling are modelled to provide new non-intuitive experimental hypotheses regarding the behaviour of the regulatory system. An additional benefit of modelling is to test hypotheses involving multiple variables, where experimental design may not necessarily be straightforward. In these cases, modelling can allow for identifying specific parameters that can be varied simultaneously. Identifying parameter sets that define a nondimensional parameter in the system (e.g. Damkohler number) and studying how change in this grouped parameter affects the system is one way of simultaneously changing many parameters.

Spatial dynamics of signalling is an important mechanism by which specificity is achieved. The structural organization of cells determines the mechanisms for information processing within and across cells. Because cellular interactions are determined by the spatial location of protein and their non-homogenous distribution as well as other cellular components, considering the spatial dimension becomes very important for cell signalling processes. For example, the events occurring in the spine (dendrites) of neurons and the signalling events occurring in the cell body are different. Localization of active forms of the Src kinase and RhoA have been demonstrated experimentally. As noted by others (Lemerle *et al* 2005), it becomes increasingly important to consider the spatial dimensions in modelling. The activation of a given protein in a specified location in a cell is dependent on the cell and stimulus context. For example, activated RhoA is observed in migrating mouse embryonic fibroblasts on fibronectin at the protrusions in the cell membrane (Pertz *et al* 2006). A sharp band of activity was observed in region immediately adjacent to the plasma membrane at the leading edge. In contrast, no RhoA activation was seen at the rear of the cell. Similarly, the activation of Src exhibits a spatial pattern (Wang *et al* 2005). Src regulates the integrin cytoskeleton interaction, an integral component of mechanotransduction. Local mechanical stimulation of human umbilical vein endothelial cells activates Src at the distal end and subsequently propagates Src activation along the plasma membrane in a wave-like manner. The functional outcomes of signalling in these spatially specified situations becomes critically dependent of the spatial relationships between the incoming signals, the intermediate signalling components like the small G proteins and protein kinases and the cellular machines that produce the regulated phenotypic behaviour.

Mathematical analysis provides us insight at many levels of the signalling network. Modelling allows us to generate quantitative relationships between the variables in the system. This gives us an idea of how components in the system interact. Because the input-output relationship of a given system can be measured experimentally, the mathematical model can be constrained to reflect experimentally observed results. Most of the parameters used in a model are estimates at best and guesses at worst. Once the model has been developed, it is a relatively easy to carry out parametric sensitivity analysis to study the role of different parameters. Feedback and feedforward regulation in the network by components can be identified via modelling. Responses such as hysteresis and bistability which are hard to measure experimentally can be predicted via models (Angeli *et al* 2004; Ferrell and Xiong 2001). Another major advantage of modelling is the short time frame within which hypotheses can be generated and tested.

Modelling of signalling networks provides a way to organize the known information in a structured manner and analyse it. This process often leads to non-intuitive hypotheses and the identification of emergent properties. Combined with experiments, modelling provides us an insight into the intricacies of network behaviour and brings up new questions to be answered experimentally. With the advent of imaging technology, it is possible now more than ever to visualize a protein of interest in live cells and to study its spatial distribution. In this review, we describe some of the mathematical approaches to capture temporal and spatial dynamics of signalling networks.

2. Modelling signalling pathways

Cell signalling networks have been a subject of intense research in the past decades. Details of interactions between components in signalling pathways and information transfer in a signalling cascade are available in great detail. The mammalian cell can be viewed as a factory with different machines running multiple processes, resulting in different products. A central signalling network coordinates all of the processes in space and time. Given the complexity involved in a mammalian cell, an engineering approach to the problem would enable us to understand the problem in parts. A major challenge however, is that not all components of the machines in the factory are known. Therefore, any model should be viewed as a work in progress, constantly being revised as newer details are obtained.

With that in mind, we can proceed with our visualization of our cellular factory. In a simple representation, a cell receives a signal from its surrounding environment or from internal sources, which could be chemical or physical in nature; the signal is processed through the appropriate signalling network to lead to an output decision that regulates the cellular machines that produce the observed behaviour of the cell. This notion is simplified and considers the cell as a single entity in isolation. We can add different levels of complexity into this framework. Because the number of ongoing signalling events is large, likely to be

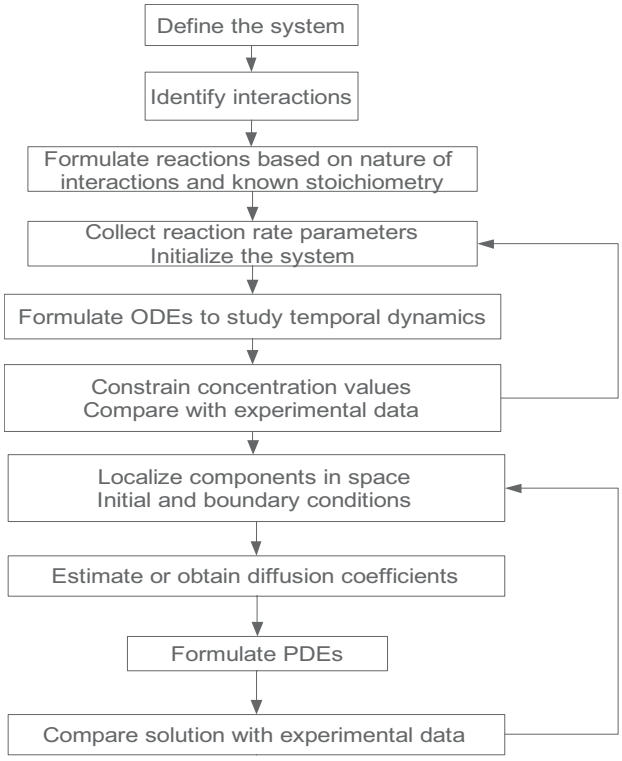

Figure 1. Flowchart for modelling. The flowchart identifies the main steps in developing mathematical models for cell signalling systems. The main steps involved include identifying direct protein-protein interactions and formulating ordinary differential equations to study temporal dynamics and then further extending the model to partial differential equations based model for study spatio-temporal dynamics. The return loops in the flow chart identify the main iterative steps in the process of model building and that validation.

on the order of ten to hundred, we choose to look at only one particular receptor and the pathway it regulates. We generalize the framework for model formulation and the mathematical technique and then proceed with a specific example. Mathematical modelling of signalling pathways in a spatio-temporal fashion can be carried out in a step-wise manner. A flowchart summarizing this process is shown in figure 1. The tasks involved in modelling can be broken up into gathering information about the system, building interaction networks and mathematical formulations, and comparing against available experimental data. The return loops in the chart indicate an iterative process. The important steps in this process are discussed in greater detail in the following sections.

2.1 *Direct interaction networks*

Cellular signalling processes are complex and involve multiple levels of regulation. More importantly, these pathways are not linear and exhibit feedback loops, thus presenting a challenge for analysis. The characteristics of the network come from the architecture of the network. Therefore, constructing an interaction map that is as accurate as possible becomes very important. It cannot be overemphasized that having a correct interaction map is the cornerstone of a biologically relevant model, since all processes downstream depend on the interactions defined. Identifying the biochemical reactions, using as many direct intracellular interactions as possible is a key to realistic modelling. This involves a thorough review of the literature to identify the direct interactions between components of the system. At this stage, one should consider building the model in a modular fashion such that each module has a distinct role or measurable outcomes. To constrain larger models, the modules within such a model should be constructed in a way such that the output from the modules can be compared to experimental data in terms of input-output relationships.

A reaction network for the activation of the transcription factor CREB in response to stimulation of the G-protein coupled receptor by a hormone is shown in figure 2. One possible way to divide the modules is based on experimentally tractable input-output relationships. The first module could be the activation of adenylyl cyclase by the dissociation of the G-protein in response to stimulation of the G-protein coupled receptor. Adenylyl cyclase catalyzes the production of cAMP from ATP. The production of cAMP in response to varying concentrations of the hormone can be compared to experimentally measured data. The second module can be the activation of PKA by changes in cellular cAMP levels resulting from the balance of activity between adenylyl cyclase and phosphodiesterase. The catalytic subunits of PKA translocate to the nucleus and phosphorylate CREB, thereby activating it. Therefore, the third module can be the temporal dynamics of CREB. Each one of these steps can be compared against experimental data and verified. It should be noted that the boundaries for computational modules and even for experimental measures of input-output relationships may or may not exist in the cell and often such boundaries are much fuzzier in the real cell as compared to experimentally defined systems. Nevertheless, the modular organization of networks is a useful feature to incorporate in computational models.

Dividing a network into modules allows us to compare results from simulation of individual modules to behaviour observed in experiments. Often receptor stimulation at steady state produces a dose-response curve. A typical dose response curve follows a sigmoidal response that is generally given by the equation in figure 3. This type of experiment can be used to generate parameters that can be used in the model. Data comparison at different steps of the reaction network is important for validating the network and leads to iterative modelling.

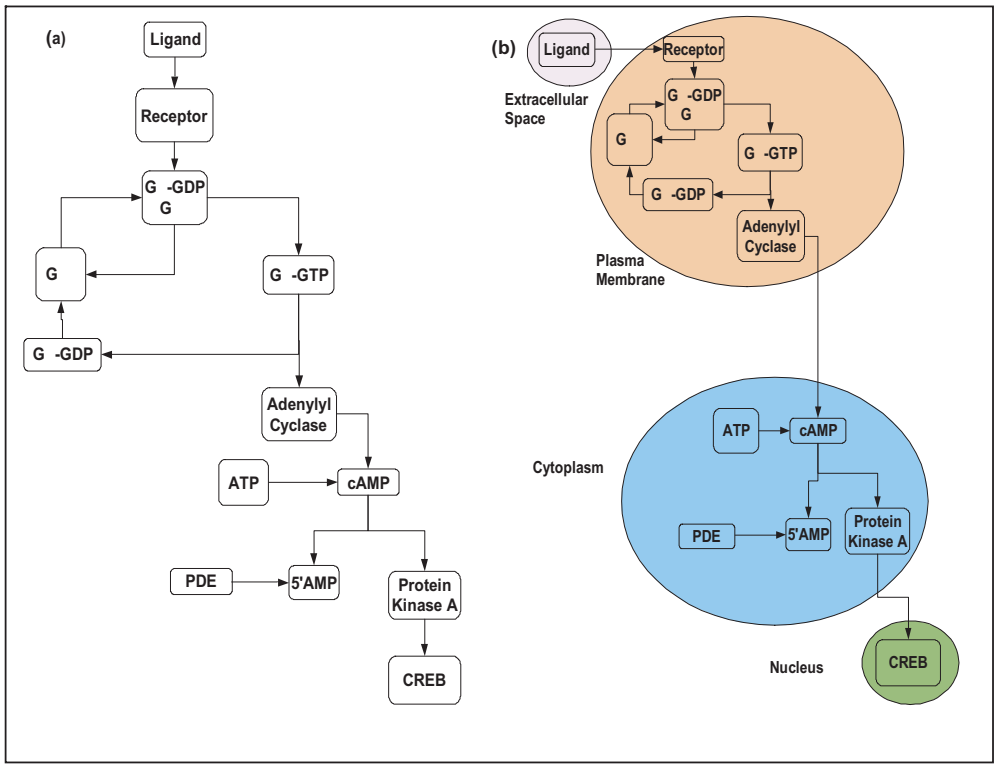

Figure 2. Signalling network of cAMP production and PKA activation by stimulation of a G-protein coupled receptor. **(a)** Interaction network showing the activation of the receptor by the ligand and subsequent activation of G-proteins leading to cAMP production. cAMP activates PKA which in turn phosphorylates the transcription factor CREB, leading to transcription. **(b)** Compartmentalization of the signalling network shown in (**a**). The ligand is in the extracellular compartment coloured pink. The membrane associated proteins like the receptor, adenylyl cyclase, and the G-proteins are in a membrane compartment coloured tan. The cytoplasmic components are place in a separate component, colored blue. The transcription factor CREB and DNA are in the nucleus coloured green.

2.2 *Chemical kinetics and temporal dynamics*

Once the interaction network is formulated, the next step is to formulate the reactions that are a part of the network. Modelling networks using chemical kinetics is a simple but powerful approach to capturing the dynamics of different concentrations. A given reaction can be written as shown in figure 4. Biochemical reactions come in two major categories, non-covalent interactions such as protein-protein interactions or nucleotide-protein interactions or enzymatic. In order to convert an interaction network into a series of biochemical reactions, we need to know the nature of the reaction. At this stage, we can introduce the protein complexes whose components have been identified by a number of experimental approaches

Once the reactions have been formulated, the reaction rate constants can be assigned. The reaction rates can be formulated as mass action kinetics where the reaction rates depend on the concentration of the reactants. The forward and backward reaction rate parameters are k_1 and k_2 for the reaction shown in figure 3. The reaction rate constants depend on temperature and the activation energy and the temperature dependence is given by the Arrhenius equation. For most mammalian systems, the temperature dependence is not an issue since reactions are set to the organismal temperature of 37°C. If the reaction is an enzyme catalyzed reaction, where there is no change in the enzyme, the rate of reaction can be formulated using Michaelis-Menten kinetics. The reaction rate can be used to formulate the ordinary differential equations that define the time and concentration dynamics of the components (see figure 4). When writing the differential equations for different species it is necessary to consider both the generation and consumption terms. This allows for maintaining mass balance for the reactants and products and maintaining consistency in the system of equations. For the example shown in figure 4, the ODEs are written following the rate laws defined by mass action kinetics. In certain cases, a certain component may be

$$y = \min + \frac{\max - \min}{1 + 10^{\log EC_{50} - x}}$$

y = response
x = concentration of stimulus
min = minimum response
max = maximal response

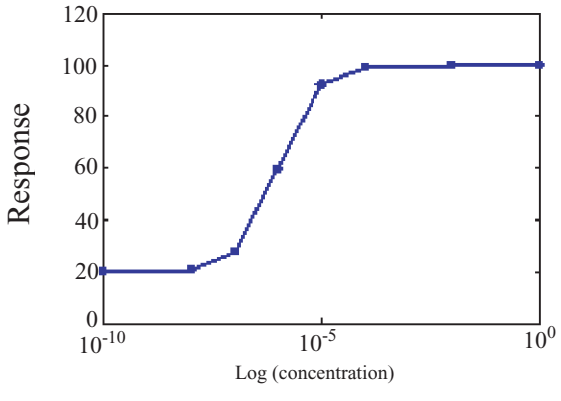

Figure 3. Dose response curves. The sigmoidal dose response curve is given by the equation relating the response (y), the dose (x) and the maximal and minimal responses. In the simple case, we do not consider the Hill slope or competition which leads to variants of this equation. For assumed values of EC50 (1 μM), maximum response (100) and minimum response (20), we show a sigmoidal dose response curve plotted for a range of concentration values.

$$A + B \underset{k_2}{\overset{k_1}{\rightleftharpoons}} C$$

Rate of forward reaction $= k_1[A][B]$
Rate of backward reaction $= k_2[C]$
Net reaction rate = forward rate – backward rate

Ordinary differential equations

$$\frac{d[A]}{dt} = -k_1.[A].[B] + k_2[C]$$
$$\frac{d[B]}{dt} = -k_1.[A].[B] + k_2[C]$$
$$\frac{d[C]}{dt} = k_1.[A].[B] - k_2[C]$$

Partial differential equations

$$\frac{\partial[A]}{\partial t} = D_A\nabla^2[A] - k_1.[A].[B] + k_2[C]$$
$$\frac{\partial[B]}{\partial t} = D_B\nabla^2[B] - k_1.[A].[B] + k_2[C]$$
$$\frac{\partial[C]}{\partial t} = D_C\nabla^2[C] + k_1.[A].[B] + k_2[C]$$

Figure 4. Chemical kinetics and differential equations. For a simple bimolecular reversible reaction, the reaction and the rates of forward reaction are formulated based on mass-action kinetics. Then the ordinary differential equations are formulated and subsequently the partial differential equations where D is the diffusion coefficient for a particular species.

treated as a 'buffered' component rather than a reactant, as its presence is necessary for the reactions to occur but its effective concentration does not change during the course of the reaction. One such molecule that is often buffered is ATP. While ATP participates in the reactions, we can treat the concentration of ATP as a constant. The reason for doing this is that in a healthy cell with ready access to nutrients, ATP levels are constant and not limiting.

The next step is to determine or estimate the reaction rate constants. Reaction rate constants are often obtained from the primary biochemical research literature. Kinetic information is recorded for many purified protein interactions. One of the areas that has lagged behind is the systematic measurement of rate constants for known biochemical reactions of a particular cell type. More recently, fluorescent imaging based techniques allow for the measurement of the kinetic and diffusion constants directly. However, since these constants are not always measured *in situ*, it is necessary to constrain the parameters with respect to observed whole cell response. In a lot of cases, where the rate constant is not immediately available, an estimate may be used, as long as observed experimental behaviour and the model behaviour are in reasonable agreement.

In order to proceed with solving the ODEs, we need to input the initial concentrations of the species. Most protein concentrations are in the range of nM to μM. The amount of a given protein depends on what protein we are looking at. For example a cytoskeletal protein like actin is present in large concentrations of a few mM. In contrast, a protein that performs a specialized function may be present (such as the β-adrenergic receptor) is present in nM or even pM quantities. Proteins that reside in the plasma membrane are better measured as number of molecules per area rather than number of molecules per volume. A good example for this is concentration of receptors. Appropriate conversion factors allow us to go back and forth between two- and

three-dimensional concentrations (i.e. surface densities and volume based concentrations).

The solution of the ODEs gives the temporal dynamics of the system. Obtaining numerical solution of ODEs is relatively straightforward by using software packages like MATLAB and Virtual Cell. Numerical solution of ODEs involves taking discrete time steps and using a numerical computation method like the Runge-Kutta method to obtain solutions with small errors over long time spans. An important aspect of solving ODEs numerically is to use a stiff solver where required. Numerical solutions of ODEs are carried out by discretizing time (using small increments of time to obtain an approximate solution). Numerical stiffness is a phenomenon exhibited by dynamical systems arising usually from discretization methods for obtaining concentration time profiles. The kinetic parameters defining the biochemical system span over three to four orders of magnitude. The discretization problem can be confounded by the large range of parameters defining the system and may give rise to numerical stiffness. In large dimensional systems, stiffness of the system can be overcome using an appropriate numerical solver. Matlab (Mathworks, Natick, MA) provides many in-built solver functions for solving stiff systems (e.g. ode15s, a variable order solver). After solving the ODEs, it is necessary to validate the temporal dynamics of the components with respect to the known dynamics. This allows us to generate input-output constraints and carry out iterative modelling.

For ODE based modelling it is also useful to carry out a parametric sensitivity analysis to identify the kinetic parameters and initial conditions that are more sensitive to small changes than others (Varma *et al* 1999). Based on the sensitivity analysis, we can vary the reaction parameters in a given range to see how the dynamics change. It is also important to carry out parametric sensitivity analysis for the initial conditions, because we cannot (or do not) measure concentrations of every protein in a pathway in an intact cell.

2.3 *Compartmental modelling*

The reactions in a cell take place in specific locations. For example, the receptors can be thought of as being limited to the plasma membrane and hence in a different compartment from the cytosplasmic proteins, and by the same analogy, nuclear proteins and transcription factors are located in a different compartment. One way to capture the different locations is to model the system in a compartmental model. Compartmental modelling has been used traditionally in pharmacology to study the pharmacokinetics and pharmacodynamics of a drug. Components are localized to specific compartments and the main assumption here is that in a given compartment, the reactions assigned to that compartment take place homogeneously. The compartment can be visualized as a well-mixed reactor and connections between compartments allow for means of transport across compartments. In figure 2a, the interaction network shown can be recast into a compartmental model. The ligand belongs to one compartment because it is generally present extracellularly. The G-protein coupled receptor G-protein and adenylyl cyclase are present in the membrane, therefore they can be grouped together in one compartment. Cytoplasmic components such as cAMP, ATP and protein kinase A are in a separate compartment while the transcription factor CREB is present in the nucleus, which is the fourth compartment. Such compartmentalization serves a surrogate for spatial specification. However, compartmental modelling is perhaps more realistic when one is looking for concentration dynamics of components in bulk solution, rather than spatial distribution of signalling components in highly localized spaces.

2.4 *Reaction - diffusion formulations*

Reaction diffusion equations are perhaps the most straightforward way of representing the change in concentration of a chemical species in space and time. Based on the Fick's second law of diffusion, the time and space dependence of a component can be formulated. When there is a reaction component involved, additional terms affect the time dependence of concentration as shown in figure 4.

Molecules restricted to a surface (like receptors) can be modelled with 2-dimensional diffusion, with diffusion limited to the plane. Molecules in three-dimensional (3D) spaces (either in the extracellular area or in the cytosol) can be modelled with 3D diffusion constants. Diffusion coefficients are hard to measure experimentally because of the viscous nature of the cytosol and the many possible interactions of proteins in a cell. Therefore, most diffusion constants used are estimates based on molecular weight and a value relative to known diffusion constants of some proteins.

In order to study spatio-temporal dynamics, it is necessary to identify the localization of the signalling components in space. For instance, AKAP is an anchoring protein that anchors cAMP dependent protein kinase PKA and therefore the diffusion constant of PKA can be assumed to be small when compared to other species. In contrast, integrins when bound to fibronectin exhibit clustering behaviour, indicating a high lateral diffusion constant. Therefore it is important to identify if the component is a membrane protein vs. cytoplasmic protein and subsequently identify location constraints.

Virtual Cell, a software suite optimized for the construction and analysis of spatial models computes the

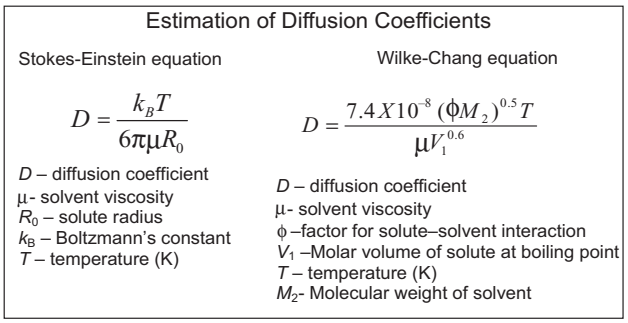

Figure 5. Estimation of diffusion coefficients. Diffusion coefficients are traditionally estimated using empirical relations. The Stokes-Einstein equation and the Wilke-Chang equations are examples of some commonly used empirical relations. Detailed explanations of the derivations of these equations and their applications are given in (Cussler 1994).

values of diffusion coefficients relative to the diffusion of calcium. The diffusion constants of proteins are usually in the range of 10^{-7} to 10^{-8} cm^2/s. This is similar to the range of diffusion in polymers. Empirical equations can be used to obtain estimates of the diffusion constant based on molecular weight, and interaction factor (e.g. Wilke-Chang equation and Stokes-Einstein equation); however these numbers are not always valid due to constraints within the cell such as molecular crowding. These empirical relations are given in figure 5. As shown, the diffusion coefficients calculated based on these relationships involves parameters like solute radius and solvent molar volumes. It is not immediately obvious how these terms can be extended to proteins and the cytoplasm. For estimated diffusion factors, appropriate correction factors must be used to account for cytoplasmic viscosity, drag and molecular crowding. These considerations significantly reduce the value of the diffusion coefficient from estimated values.

A major challenge then is to identify the diffusion constants for each protein. There is no *a priori* way to calculate the real diffusion constants; they are experimentally measured. Experimental methods such as fluorescence resonance energy transfer (FRET) and fluorescence recovery after photobleaching (FRAP) can be used to estimate diffusion constants within a cell. However, in biochemical systems because of the microscopic scale in which components interact and the large number of components present in a cellular environment, it is not always possible to accurately measure the diffusion coefficients. Diffusion constants in liquids are harder to measure because of the viscosity of the medium and the drag. An associated challenge in cellular systems is that the cytoskeletal cortex and the plasma membrane can be visualized as polymers rather than liquids. The mean displacement gives a sense of the distance traveled by a molecule of interest for a given value of time t. The mean displacement is given by $\sqrt{4Dt}$, where D is the diffusion coefficient and t is time. When looking at diffusion of a molecule in a space of known dimensions, calculating the mean displacement allows us to estimate if the reaction is going to be diffusion limited or not. For different values of diffusion constants, the mean displacement over time is shown in figure 6. Because of viscosity and drag, diffusion can theoretically be a limiting factor for reactions to occur. But as we see from figure 6, the mean displacement is of the order of μm and even for small diffusion constants, (D=1 $\mu m^2/s$) we see that the mean displacement is approximately 50 μm, and compares favourably with cellular dimensions, rendering most processes non-diffusion limiting, but rather reaction limiting.

An important aspect of the spatio-temporal dynamics is the formation of patterns of signalling components within the cell. Spatial patterns can arise in systems that are reaction driven but not largely diffusion limited because of the local interactions between components. In order to understand this phenomenon in terms of the system parameters, perhaps the mathematics of spatial pattern formation (Turing mechanisms) with appropriate modifications may be applied to signalling networks. The motivation behind considering a Turing mechanism type formulation is the idea that the distribution of the cellular complexes is controlled by diffusion driven instability (Turing 1952). With this idea, we can analyse the PDEs that result from the interaction networks and obtain conditions under which one would observe the diffusion based instability. This would be a useful and important validation for the parameters (kinetic constants, diffusion coefficients and initial conditions) and to study the variation of parameters and their effect of the spatio-temporal distribution of proteins. Theoretical analysis of such diffusion-driven instability in reaction-diffusion systems allows us to compare the theoretical outcomes with experimental results in terms of observed patterns.

PDEs can be solved using finite element method using programs like FlexPDE, FEMLAB and Virtual Cell. Virtual Cell removes the burden from the user in formulating the reaction and diffusive fluxes. A standard boundary condition used is that flux across boundaries is zero or that the value of a component at a particular boundary is a constant value (figure 7); however these are formulated based on the system. The commonly used boundary conditions are summarized in figure 7. The Dirichlet boundary conditions are used when the value of a component is constant on the domain boundary. In biological systems, when modelling intracellular interactions, it is straightforward to apply the Neumann boundary conditions, namely zero flux across the membrane. This is obvious because the contents of the cell, are contained within it, and do not leave the boundary specified by the cell membrane. However, if one were modelling the formation of extracellular matrix (e.g.

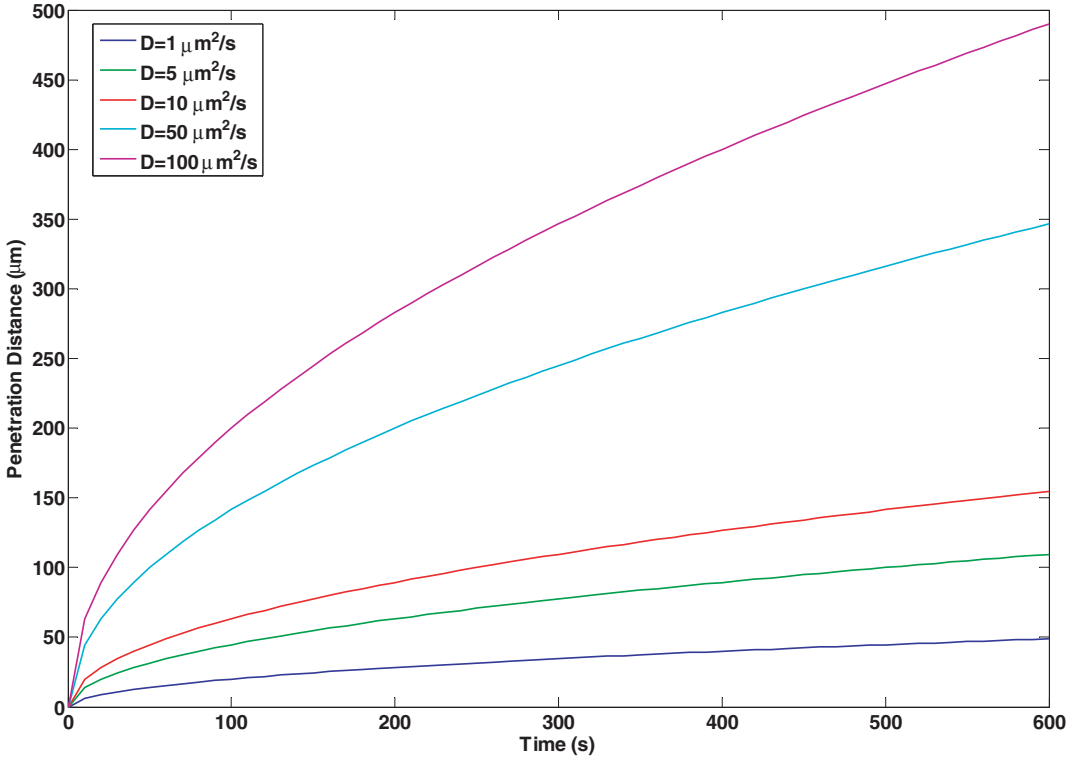

Figure 6. Mean displacement for different values of diffusion constants. The mean displacement is calculated for different values of diffusion coefficients over time. Even in the case of low diffusion constants, over 10 min, the mean displacement is around 50 μm, indicating that most biochemical reactions are not diffusion-limited.

collagen fibril), then there would be a non-zero outward flux from the cell to the extracellular space for collagen.

It has been postulated that cell surface proteins are not freely diffusible and that their localization to the local domains is regulated (Batada *et al* 2006). Diffusion coefficient of receptors and proteins are more than two orders of magnitude smaller in the membrane than in the cytoplasm. This is because diffusion in the membrane is limited to the plane comprising the membrane. For the information to be transferred and transduced from the receptor binding to the ligand in the extracellular domain to the scaffold proteins to the target enzymes, the multiple proteins need to be colocalized. The biophysical function of colocalization may be to increase coupling efficiency. If the spatial model is designed as one dependent entirely on free diffusion, the likelihood of interaction between proteins can be lower that what is actually observed. However, we do know that most interacting proteins are colocalized. Batada *et al* (2004) show that even for a large protein half-life, very few proteins (~1 in 25) interact based on free diffusion alone. Therefore, colocalization of proteins is important to increase the probability of interactions and to minimize cross-talk. The concept of colocalization must be incorporated into the initial conditions by defining initial distributions of receptors such that they can interact with the ligand and their downstream effector.

2.5 *Model robustness*

A model is often built to obtain systems level understanding and generation of new hypotheses regarding functional regulation. The model depends on many parameters that are not known or cannot be measured experimentally. In fact, most reaction rate parameters are measured for biochemically purified proteins and not *in situ*. The model outcome depends largely on the parameters used. As mentioned before, it is crucial to carry out sensitivity analysis for the various parameters and generate sensitivity indices (Varma *et al* 1999). It is also important to run the model for different values of the parameters. A criterion for a robust model is that it displays mathematically consistent behaviour – non negative and determined value of concentrations. Negative concentrations and indeterminate values (division by zero) are indicators of mathematical inconsistency and the model formulation should be checked. Once the mathematical robustness has been established, the model must be constrained to biological relevant ranges, by comparing with experimental data at every stage possible.

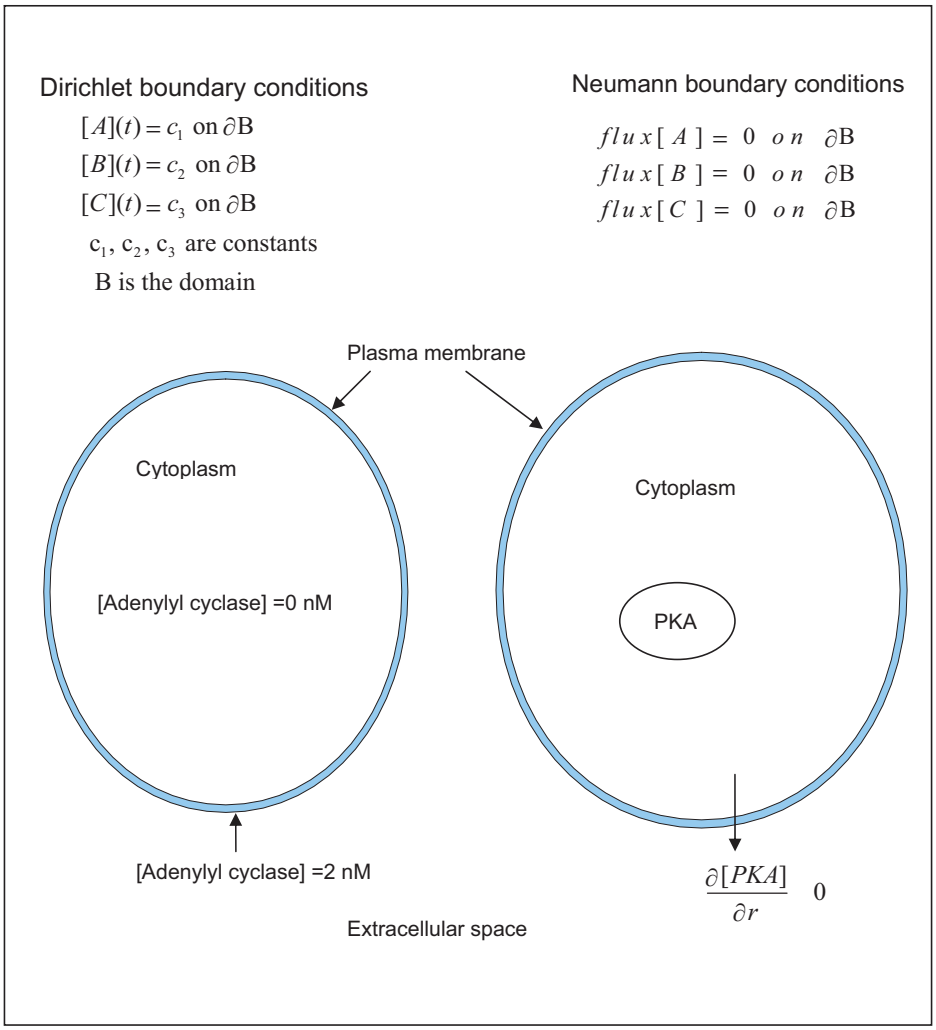

Figure 7. Boundary condition for PDEs. The Dirichlet boundary condition states that the value of a component is a constant on the domain boundary while the Neumann condition states that the flux of a component across the boundary is zero. Using the Dirichlet conditions we can define the concentration of adenylyl cyclase to be zero in the cytoplasm but a constant value in the membrane. Using Neumann condition, we can define the flux of PKA across the cell membrane is zero.

Then the model can be tested for different conditions and analysed for non-intuitive hypotheses.

3. Challenges for modelling

The main validation for models is comparison with experimental data. When we generate a model, for every reaction that we add to the model, we are introducing at least two or three new species. Nominally, we also introduce a forward and backward reaction rate parameter. In order to compare the dynamics of many concentrations, we need as much experimental data. Because we do not have all the experimental data, it is hard to fully validate the models. Another major challenge for modelling cell signalling systems is the number of parameters that are not known. For every reaction that is formulated, one or more kinetic parameters are introduced. As we introduce more components and their intermediates, we add progressively increasing non-linear levels of complexity to the problem. The increasing size of these models poses enormous problems of how the model can be constrained to realistically represent the boundaries within which biological systems operate. This is a challenge that we need to deal with for each model before extensive numerical simulations are conducted. Once the model is built and we are ready to simulate and obtain numerical results, we are faced with computational challenges. Finite element models and packages are often used to obtain numerical solutions.

Thus far, we have considered deterministic modelling. However, some processes are better modelled using stochastic methods. This is the case where the number of molecules is less, or the process is driven by the random distribution of events. For temporal dynamics Gillespie's algorithm (Gillespie 1977) is efficient for computing the probability of reactions occurring in the next time step. Spatial distributions can be evaluated using random walks and probabilities of molecular reactions. Such representations are going to be necessary to consider many types of interactions in all regions of the cell, the plasma membrane, the cytoplasm and nucleus, where dynamic co-localization becomes an important regulator of information flow.

4. Perspective

In this review we have considered the various types of deterministic modelling approaches separately. These distinctions arise from separate development of mathematical formalisms and their use in naturally occurring physical or engineered systems that can be adequately represented for numerical simulation by one approach. In contrast, biological systems even at the cellular level display a dazzling array of complex behaviours. When such behaviours are parsed in modules that can be represented mathematically so as to carry out numerical simulations, it becomes quickly apparent that different modules may require different types of models: ODE vs. PDE vs. stochastic models. How these models with differing representations for various modules can be seamlessly integrated into higher order models that represent relationships in biological functions across scales remains a challenge for the future.

References

Angeli D, Ferrell J E Jr and Sontag E D 2004 Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems; *Proc. Natl. Acad. Sci. USA* **101** 1822–1827

Batada N N, Shepp L A and Siegmund D O 2004 Stochastic model of protein-protein interaction: why signalling proteins need to be colocalized; *Proc. Natl. Acad. Sci. USA* **101** 6445–6449

Batada N N, Shepp L A, Siegmund D O and Levitt M 2006 Spatial Regulation and the Rate of Signal Transduction Activation; *PLoS Comput. Biol.* **2** e44

Bhalla U S and Iyengar R 1999 Emergent properties of networks of biological signalling pathways; *Science* **283** 381–387

Bhalla U S and Iyengar R 2001a Functional modules in biological signalling networks; *Novartis Found. Symp.* **239** 4–13; discussion 13–15, 45–51

Bhalla, U S and Iyengar R 2001b Robustness of the bistable behaviour of a biological signalling feedback loop; *Chaos* **11** 221–226

Black J W and Leff P 1983 Operational models of pharmacological agonism; *Proc R Soc London B Biol. Sci.* **220** 141–162

Cussler E L 1994 *Diffusion, mass transfer in fluid systems* (Cambridge, UK: Cambridge University Press)

Ferrell J E and Xiong W 2001 Bistability in cell signalling: How to make continuous processes discontinuous, and reversible processes irreversible; *Chaos* **11** 227–236

Fong S S, Burgard A P, Herring C D, Knight E M, Blattner F R, Maranas C D and Palsson B O 2005 In silico design and adaptive evolution of *Escherichia coli* for production of lactic acid; *Biotechnol. Bioeng.* **91** 643–648

Gillespie D T 1977 Exact Stochastic Simulation of Coupled Chemical Reactions; *J. Phys. Chem.* **81** 2340–2361

Hautaniemi S, Kharait S, Iwabu A, Wells A and Lauffenburger D A 2005 Modelling of signal-response cascades using decision tree analysis. *Bioinformatics* **21** 2027–2035

Hodgkin A L and Huxley A F 1952 A quantitative description of membrane current and its application to conduction and excitation in nerve; *J. Physiol.* **117** 500–544

Kenakin T 2004 Efficacy as a vector: the relative prevalence and paucity of inverse agonism; *Mol. Pharmacol.* **65** 2–11

Kowalewski J M, Uhlen P, Kitano H and Brismar H 2006 Modelling the impact of store-operated Ca(2+) entry on intracellular Ca(2+) oscillations; *Math Biosci*. (in press)

Leff P, Scaramellini C, Law C and McKechnie K 1997 A three-state receptor model of agonist action. *Trends Pharmacol. Sci.* **18** 355–362

Lemerle C, Di Ventura B and Serrano L 2005 Space as the final frontier in stochastic simulations of biological systems; *FEBS Lett.* **579** 1789–1794

Ma'ayan A, Jenkins S L, Neves S, Hasseldine A, Grace E, Dubin-Thaler B, Eungdamrong N J, Weng G, Ram P T, Rice J J, Kershenbaum A, Stolovitzky G A, Blitzer R D and Iyengar R 2005 Formation of regulatory patterns during signal propagation in a Mammalian cellular network; *Science* **309** 1078–83

Mayawala K, Vlachos D G and Edwards J S 2006 Spatial modelling of dimerization reaction dynamics in the plasma membrane: Monte Carlo vs. continuum differential equations; *Biophys. Chem.* **121** 194–208

Papin J A, Hunter T, Palsson B O and Subramaniam S 2005 Reconstruction of cellular signalling networks and analysis of their properties; *Nat. Rev. Mol. Cell Biol.* **6** 99–111

Pertz O, Hodgson L, Klemke R L and Hahn K M 2006 Spatiotemporal dynamics of RhoA activity in migrating cells; *Nature* (*London*) **440** 1069–1072

Saha K and Schaffer D V 2006 Signal dynamics in Sonic hedgehog tissue patterning; *Development* **133** 889–900

Singh A, Jayaraman A and Hahn J 2006 Modelling regulatory mechanisms in IL-6 signal transduction in hepatocytes; *Biotechnol Bioeng*. (in press)

Stamatakis M and Mantzaris N V 2006 Modelling of ATP-mediated signal transduction and wave propagation in astrocytic cellular networks; *J. Theor. Biol*. (in press)

Thiele I, Vo T D, Price N D and Palsson B O 2005 Expanded metabolic reconstruction of *Helicobacter pylori* (iIT341 GSM/

GPR): an *in silico* genome-scale characterization of single- and double-deletion mutants; *J. Bacteriol.* **187** 5818–5830

Turing A 1952 The Chemical Basis of Morphogenesis. *Philos. Trans. R. Soc. Series B.* **237** 37–72

Varma A, Morbidelli M and Wu H 1999 *Parametric sensitivity in chemical systems* (Cambridge, UK: Cambridge University Press)

Wang Y, Botvinick E L, Zhao Y, Berns M W, Usami S, Tsien R Y and Chien S 2005 Visualizing the mechanical activation of Src; *Nature* (*London*) **434** 1040–1045

Wiley H S, Shvartsman S Y and Lauffenburger D A 2003 Computational modelling of the EGF-receptor system: a paradigm for systems biology; *Trends Cell Biol.* **13** 43–50

*e*Publication: 5 December 2006