

Chapter XX: Computational and Mathematical Modelling of the EGF Receptor System

Colin G. Johnson¹, Emmet McIntyre¹ and William Gullick²

¹Computing Laboratory, ²Department of Biosciences, University of Kent, Canterbury,
UK

Running head: Modelling the EGF Receptor system.

* Address correspondence to Colin Johnson, PhD, Senior Lecturer, Computing
Laboratory, University of Kent, Canterbury, Kent CT2 7NF (C.G.Johnson@kent.ac.uk),
telephone +44 1227 827562.

Summary/Abstract: This chapter gives an overview of computational and mathematical modelling of the EGF receptor system. It begins with a survey of motivations for producing such models, then describes the main approaches that are taken to carrying out such modelling, *viz.* differential equations and individual-based modelling. Finally, a number of projects that applying modelling and simulation techniques to various aspects of the EGF receptor system are described.

Keywords: EGF receptor; computational models; mathematical models

1. Introduction

This paper gives an overview of computational models and simulations of the EGF receptor system; it is aimed at biologists who have no experience of such modelling and simulation techniques. It begins with a review of motivations for constructing such models, then surveys the kinds of models that can be built. Finally, a number of models of various parts of the EGF receptor system are described.

2. Why Model and Simulate?

By *modelling* we mean the construction of some computer program or mathematical description that describes some aspect of a system. *Simulation* is the running of a computer implementation of that model, i.e. setting parameters in, and the initial state of, a model, then modifying the state of that model a number of times to represent the system changing in time.

There are a number of motivations for developing such models. At the simplest level, models can be used as informal tools to develop intuitions and ideas about the functioning of a system. By attempting to build a formal model that incorporates existing knowledge about the system, the less-well understood components of the system can become clearer; furthermore conjectures can be made, and tested for plausibility, about mechanisms that might explain those components. This process is, in general, referred to as *synthetic biology*. That is, it is an attempt to gain an understanding of a system by building it.

However, such a model cannot confirm anything positive about a system. Typically, it will be used to inspire further experimental work, by providing a *prima facie* case that some experiment might produce results of interest. Another function of such a system is to demonstrate that a particular mechanism *cannot* explain a particular behaviour, by showing that an implementation of that mechanism in simulation produces different behaviour (either at a qualitative or quantitative level) to an observed system.

More formally, models can be used to integrate together a number of aspects of a system that are individually well-understood, yet where the interactions between those aspects are not. In such an approach, we build a number of computer programs or mathematical systems, each of which describes the individually well-understood subsystem and which has inputs and outputs that allow it to interact with other components of the system. Provided that such models are complete, and that their inclusion in a wider system does not produce additional effects or ill-understood non-linear interactions, such a system can produce accurate predictions about the behaviour of the system. Such an approach, however, is limited by our lack of such complete understanding of many biological systems (this is a situation which contrasts, for example, with models in physics, where many subsystems are well-understood).

More rigorous uses of modelling and simulation will attempt to combine the model with experimental or observational data. In such an approach, the model typically represents a *hypothesis* about how the system works. Typically, a hypothesis is tested by ‘bringing the data to the hypothesis’; that is, data is measured, or transformed, so that it can be directly

compared with the hypothesis. Simulation can ‘take the hypothesis’ (part of the way) to the data”. A model is constructed, based on a hypothesis about the functioning of the system, and this model is then simulated by implementing it as a computer program and measuring those aspects of the simulation that correspond to the experimental data. These measurements can be compared to the experimental or observational data.

There are a number of issues with this approach, two of which we shall explore. The first is that a typical model will have unknown parameters, which can affect both the qualitative and quantitative results that are measured. One approach to this issue is to use *parameter fitting* where the model is viewed as a parameterised space of models, and some optimization technique used to find a (heuristically) optimal setting for those parameters that maximises fit with the data. One side benefit of this is that it gives an estimate for those parameters as part of the process; however, it should be noted that for many model/dataset pairs, many different possible parameter sets can give rise to behaviour compatible with the data.

A final view of such models is that they represent *hypothesis-driven combinations of attributes*, which can be used as inputs to systems for prediction and classification problems. Typically, a statistical/computational model for prediction is produced by a *supervised learning* technique [Mitchell]. That is, we have a set of experimental or observational data, including one attribute of the system that we would like to be able to predict in the future (referred to as the *class*). For example, a medical dataset might consist of a list of patients: for each patient a list of symptoms is recorded, and an expert

diagnosis carried out. Supervised learning is any technique that takes such a dataset and produces a statistical/computational model that will make a prediction of the class; well known examples are *naïve Bayes* methods (see e.g. [2]) and *decision tree induction* [3]. In our example, the model would take a list of symptoms and make a diagnosis.

Typically, such systems work using the raw data as inputs to the *training* process, that is, the process by which a generic predictive model is adjusted to generalise from the particular set of data being used. However, in some situations, *constructed attributes* can be used: that is, attributes from the data are combined to form new data attributes [4]. Typically, such constructions are simple and based on a basic search process for useful combinations. One way to view simulations is as hypothesis-driven attribute construction methods; that is, a simulation provides a new source of data for making predictions about a model, which is based on some hypothesis about the functioning of the system. In such a situation the final test of value is simply whether the addition of the new data source from the simulation adds to the accuracy of the simulation, measured on a previously unseen set of test data.

3. Methods for Modelling and Simulation

There are a number of methods for modelling and simulating cellular systems. In this section we discuss the various methods, focusing on *differential equation-based* and *individual-based* methods.

3.1 Differential Equation Methods

One approach is to develop a set of differential equations that describe the system [5,6]. That is, the various interactions and reactions between entities in the system are described in terms of rates of exchange between different quantities (a classic example is the Michaelis-Menten equation for enzyme kinetics [7]). In such a system, when an amount of some substance is transformed in some way, the quantity of the original substance is reduced and that of the outcome of the transformation increased. So, for example, a phosphorylation event on molecule X would consist of reducing the amount of X in the system, and increasing the amount of phosphorylated-X.

This is a powerful approach to modelling the basic levels of each substance of interest, and it has an advantage over some other methods in that many methods exist to get some *analytic* understanding of the problem (i.e. to understand some general properties and overall dynamics of the system) as well as to simulate it for a particular set of parameters and initial conditions.

However, there are disadvantages to this kind of modelling. In particular, there are issues concerned with *scaling* and with representing *space*. Differential equation models provide a succinct summary of the interactions between a small number of molecule-types. However, when a system contains many types of molecules, accounting for the different types whilst retaining a comprehensible model eventually becomes intractable. In terms of spatial distribution, differential equation models are better used when dealing with a small number of components where the free-mixing assumption can be made (i.e.

that any molecule can interact with any other). In systems where genuine spatial distribution is important, this can be modelled by *partial differential equations*; however, dealing with complex interactions between different molecule types across a space is difficult, and many of the mathematical techniques for getting a qualitative understanding of the model break down in such situations.

3.2 Individual Based Methods

The second main approach to modelling is *individual-based* modelling. In such a model, each entity in the system is represented by a separate entity in the computer; this contrasts with differential equation models, which keep track of aggregate counts of objects over time. This has a number of advantages: two of particular significance are that models of systems with many different kinds of components can be readily built, and that a full spatial model can be readily incorporated.

In order to generate such a model, four aspects of the system need to be specified. Firstly, a list of the kinds of entities found in the system needs to be compiled. For a cell-biology model, these will typically be lists of molecules found in the system. Secondly, the kinds of interactions between those entities needs to be defined: most importantly, if two entities meet, do they bind? With what probability? Thirdly, the movement of the entities is defined: for example, Brownian motion, or flow through a region at a certain rate. Finally, a set of initial conditions needs to be specified.

Commonly, not all of the information required to set up such a model is known in advance. As a result, a typical “model” is not a single model, but a *parameterised space*

of possible models: i.e. there are a number of unknown parameters in the model, and setting these parameters to a particular value specifies a particular model. Sometimes, such models can be used as part of a process to *estimate* the unknown parameters. For example, a model might represent a process that is too small to observe directly experimentally; however, this process might give rise to a phenomenon that *can* be directly observed. By finding a parameter setting within that space of models that reproduces the observed behaviour, we can conclude that the parameters (which will include properties of the unobservable behaviour) are a feasible set of parameters for the real system.

Typically, this search through the parameter space will be carried out using some optimization heuristic [8], which will search for values of the parameters that maximize the fit between experimentally observable features of the system in simulation and in reality.

An alternative approach is to use *qualitative reasoning* methods [9,10]. This approach consists of running a simulation using qualitative features about the objects in the simulation, rather than particular values: is a quantity positive or negative, is a relationship proportional, negative-proportional, threshold, et cetera? This can give a broad understanding of a model, even in the absence of concrete parameter settings.

4. Implementing Individual Based Methods

Individual based methods are typically implemented using an object-oriented programming technique [11] such as *Java* or C++. In order to create a program in such a

language, the programmer first creates *types* of objects known as *classes*, specifying the information that is stored within an object of that class and how objects of that class interact with other objects. The simulation then progresses by the creation and interaction of individual *objects*, each of which belongs to (and has its behaviour defined by) one of the classes. There are a number of different ways in which to manage the interactions between these objects; therefore the programmer of a simulation has to make a number of choices, before writing the simulation program.

The first of these decisions is whether the model will be implemented in an *event-based* or *timestep-based* fashion. An event-based simulation [12] is one where the program maintains a list of events that change the state of the system, and the simulation is carried out by processing an event (such as an interaction between two molecules), calculating whether this generates any new events (e.g. a molecule dissociating from a complex, which might lead to a new interaction for that molecule), and then moving forward in time to the next event. This works well for systems where the “next event” can be readily calculated. However, in many biochemical models, this calculation is not easy due to processes such as Brownian motion, which can rapidly introduce a new potential interaction where there was none before. As a result, timestep-based methods, which move in a regular timestep and calculate all activity within that timestep, are commonly used in such situations.

A second decision is the level of detail that the model will use. Different questions/hypotheses will require different levels of detail in the model. Ultimately, the

model needs to be a useful abstraction from reality—incorporating those features that are needed for the question at hand, whilst ignoring features that are irrelevant. There are also practical concerns in the decision. In particular, very detailed models can be time consuming to compute (up to the point where computation might be infeasible), or else not admit the kind of analytical techniques that can be used on simpler models.

A third decision is whether the calculations will be *stochastic* (i.e. incorporating some randomness in the events) or *deterministic*. Given that all models at the cellular level will have some element of randomness in them when viewed at that level (Brownian motion and probability of two molecules binding being two examples), the stochastic modelling approach seems immediately more appropriate. However, when many objects are interacting, these individual interactions are often somewhat irrelevant. Instead, these large numbers of random events can be approximated by a deterministic rate of occurrence. Stochastic models are of most interest when the individual actions of molecules that exist in small numbers can have significant consequences for the system as a whole, as discussed by Andrews and Bray [13] and Lemerle et al. [14].

A final decision concerns how space is handled within the model. The simplest model of space is to assume that all of the components of the system interact within a single space: this is referred to as *complete mixing*. The next simplest model is that there are a number of components in the model (for example, within and outwith the cell) with some communication or exchange going on between these components, representing exchange of molecules between the domains or communication through transmembrane proteins.

Beyond this, we can develop models that have a spatial position for each component of the system: either represented as an approximation on a grid, or as a position given by decimal-number coordinates. This level of detail is important for some models (for example, studying the structure of receptor clusters or the formation of signalling complexes); however, for other models the complete mixing assumption is sufficient.

5. Examples of Simulations

Computational and mathematical models have been used for understanding a number of aspects of the EGF receptor system. Most simulations have concentrated on aspects of the intracellular signalling cascade; however, other approaches have addressed the oligomerisation behaviour on the cell surface. As note by Gullick et al. [15,16], there are three main processes in the EGF receptor system. Firstly, the liganding of the extracellular domain, secondly, the dimerisation and oligomerisation of these receptors, and finally the intracellular signalling cascade set off by this dimerisation. The majority of effort in this area has focused on the intracellular signalling cascade, using differential equation models. This is where we begin our survey. Later in this section we discuss models of the cell-surface behaviour, integrated models that examine multiple stages, and systems that introduce formal languages for the description of interactions and which make steps toward integrating models into broader systems biology projects.

5.1 Differential Equation Models of Intracellular Signalling Cascades

The largest amount of work on simulation of the EGF receptor system has focused on differential equation models of the intracellular signalling cascade. These have been surveyed by Wiley et al. [17] and Orton et al. [18].

The core of such a model is a list of the various proteins involved in the signalling process, and a list of differential equations that specify the reaction rates between these proteins. These models are then simulated by the use of a numerical method, either from a generic mathematical software package such as Mathematica [19] or Matlab [20], or by software specifically designed for sets of biochemical interactions such as Gepasi [21].

The main parameters in such models are rate constants for the various reactions in the system. Typically these are derived from existing experimental work; if they are missing, a *sensitivity analysis* can sometimes be performed to check whether or not the particular value of the parameter is having a significant impact on the phenomenon of interest.

A typical “experiment” using such a model will be to develop a model which introduces some new mechanism or interaction which, it is hypothesised, produces a particular experimentally-observed behaviour and therefore produces a viable hypothesis to explain the mechanism underlying that behaviour. In the remainder of this section we give a number of examples of such models.

A detailed example of such a model is given by Suresh Babu et al. [22]. This paper begins by detailing a set of differential equations that represent the various reactions in the system. At the end of this process a parameterised space of models has been created, where the parameters represent the various rate constants for the reactions in the model. They then realise a particular model by inserting rate constants found in the literature.

They then test the accuracy of the model by a number of comparisons between experimental and computational work: plotting time-courses of Raf, MEK and ERK activation levels and comparing the latter two against Western blot analyses of wet lab experiments with the same setup; studying the effects of over-expression of proteins in the model and comparisons with known experimental effects of overexpression; studying time courses of phosphorylation and dephosphorylation; and, carrying out a sensitivity analysis of the system. This work shows that an accurate model of the cascade can be produced; however, they do not apply their simulation to testing any specific new hypotheses about the functioning of the system.

One example of the application of computational methods to a specific problem in their area is the work of Brightman and Fell [23]. This paper describes a model of the MAP kinase cascade using the simulation system Gepasi [21], and applies this to form hypotheses for the difference in behaviour when the cascade is stimulated by EGF (in this case, the cascade is activated for a short time) and by NGF (in which case the cascade is stimulated for a sustained period of time). The simulation is used to narrow down where in the system a change will produce the effects seen in experimental work. In particular, it is shown that mechanisms that simply affect the intensity of signalling at the cell surface, or mechanisms that influence the phosphatase activity in the cascade are unlikely to produce the differences in effect observed in the experimental system. By contrast, their simulation of variations in the negative feedback regulation in the cascade do demonstrate a variety of differences in cascade persistence consistent with the

experimental observations. Therefore, they conclude that this final mechanism is the most likely candidate mechanism to explain the differences.

Hendriks et al. [24] also apply simulations to help make a differentiation between two competing hypotheses to explain a particular observed phenomenon. The phenomenon is the localisation of dephosphorylation activity in the ErbB-triggered signalling cascade. They simulate two hypotheses concerning this: the first, that the activity is localised in the cell surface plasma membrane; the second, that intracellular, endosomal regions are the focus for it. By comparing these simulations against experimental data, they show that the former localisation is more likely to explain the observed phenomenon.

Shvartsmann et al. [25] use a simulation to show that a proposed hypothesis is sufficient to explain an experimentally observed phenomenon. The phenomenon in question is the development of a single-peaked input into a pattern with two peaks; this is needed to show how the development of paired organs during development occurs. The computational model shows the ranges of parameters that would be required to generate the phenomenon in question: this could be seen as *refinement* of an initial qualitative hypothesis into more quantitative terms. Maly et al. [26] also carry out a simulation focused on *feasibility*. They demonstrate that a particular arrangements of feedback loops in an autocrine signalling system is capable of generating and maintaining cell polarity.

5.2 Other Modelling Methods for the Intracellular Signalling Cascade

Techniques other than differential equations have been used to model the signalling cascade. For example, Hlavacek et al. [27] have developed a system called *BioGenNet*

that is based on lists of *rewriting rules*: that is, rules that describe how parts of one structure can be transformed into another. This allows hierarchies of reaction rules to be created, rather than needing to specify each rule individually, as in the differential equation-based systems discussed above. In addition, such systems of rules permit new analytic methods such as *model checking*, which is a system for checking whether a set of rules is consistent with a formal description of how parts of a system will change with time.

Blinov et al. [28] apply similar methods to reproduce and extend the earlier model of Kholodenko et al. [29], incorporating a larger number of reactions including proteins not incorporated into the Kholodenko model.

Another method that has been used to model the signalling cascade is *Petri nets* [30]. This is a visually intuitive way of constructing and simulating systems, which can be readily visualised whilst the simulation is running.

Schamel and Dick [31] have proposed an analogy between the signal transduction process and the Parallel Distributed Processing model used in modelling neural networks. However, this remains at the conceptual level rather than representing a way to implement simulations.

An alternative approach to modelling is given by Pawson and Linding [32]. This takes an approach sometimes known as a *synthetic biology* approach to the problem. In this

approach, signalling networks are reverse engineered from known components. By carrying out such a reconstruction, the developer of the simulation is required to think carefully about the functional role of each of the components, and therefore develops a better understanding of the role that each component plays and the possible ways in which they can interact.

6. Modelling Behaviour on the Cell Surface in the EGF Receptor System

The process of dimerisation and higher-level clustering of EGF receptors on the cell surface is the subject of a paper by Goldman et al [33]. This consists of an object-oriented individual-based model, where receptors move under Brownian motion on a model of the cell surface, are able to be liganded, and which form clusters by binding with other receptors using a probabilistic model with parameters that can be specified by the user.

A model using similar techniques has been developed to model the diffusion of ligands in the intercellular medium, and thus help to understand juxtacrine and paracrine signalling [34].

6.1 Modelling the Overall System

Recently, attempts have been made to combine models of various aspects of the system. For example, Hendriks et al. [35] have developed a differential equation model that combines a model of dimerisation of liganded receptors with a model of the consequent intracellular signalling cascade. This has been applied to model hypotheses concerning differences in the behaviour between ErbB1 receptors that are sensitive to the drug gefitinib (IRESSA), and those which are not.

6.2 Higher Level Models for Intracellular Signalling Cascades

Each piece of work described so far has consisted of a single modelling technique being applied to some particular problem. Recent papers by Calder et al. [18,36] takes a different approach. The approach taken is to describe the MAP kinase cascade in a mathematical language known as a *process algebra*. This is a formal description of the various interactions within the system. This high-level description can be automatically converted into both a deterministic, differential equation based system which can be simulated using numerical methods, *and* automatically converted into a stochastic model which can be simulated using an individual-based model. If the model is robust, both of these techniques should produce a similar outcome; however, sometimes artefacts from the particular simulation/numerical analysis method used can distort the solution.

Calder et al. [36] use a comparison between the two models, derived from the process algebra description, to show such an artefact in the earlier paper of Schoberl et al. [37], which underestimates the peak concentration of Ras-GTP in the system by a factor of two.

Descriptions such as the process algebra have two main advantages. Firstly, they can be automatically converted into simulations of different types, thus showing up problems with a particular simulation technique for a particular problem. Secondly, they have the potential advantage that models can be analysed for qualitative features, as well as being converted into executable models. Some general issues concerned with models of this kind are discussed by Kolch et al. [38].

6.3 Integration with Larger System Biology Software Systems

It has been noted by Hornberg et al. [39] that cancer is a canonical *systems biology disease*: if we want to understand cancer, we need to understand how information flows between many different parallel systems of chemical interactions. Other discussions of the impact of systems biology on signal transduction, are given by Citri and Yarden [40] and Suresh Babu et al. [41]. In recent years, attempts have been made to create software and description languages that allow the sharing and combining of models of biochemical systems. One of the most important of these languages is the *Systems Biology Markup Language* (SBML) [42]. The aim of this is to provide a common set of formal notation for the recording of diagrams of biochemical interactions, so that models can be shared between different software packages and combined into larger integrated models (for example the E-Cell project [43]).

Recently, some early efforts have been made to give an SBML description of the EGF receptor system and its associated signalling cascade [44]. A more general discussion of this kind of notation is given by Kitano [45], Blinov [28], and Cary et al. [46].

High-throughput techniques, such as microarrays, for data collection are often associated with systems biology approaches as they can provide the detailed data needed to complete a systems biology model. Studies such as that of Jones et al. [47] show how large scale protein networks can be studied and reaction rates quantified, which provide valuable input for simulations.

7 Prospects

Mathematical and computational models have proven useful in testing various hypotheses about the functioning of the EGF receptor system, and in providing a precise language for the expression of such hypotheses. In the future, we can see four new important directions for work of this type:

- The use of such methods in combination with data gained from experiment.
- The integration of these models into a wider set of tools for systems biology, leading to the integration of multiple models.
- The use of languages to describe these systems that can be realised in a number of different ways, and have a number of different analytical tools applied to them.
- The simulation of the activity of drugs on the system; and the use of computational search techniques to discover new targets for drug discovery (as illustrated by the work of Haugh et al. [48]).

Breitling and Hoeller [49] also discuss future directions for such models. They outline four main directions for future applications of modelling of the EGF system: modelling of endosomal compartmentalisation, developing more sophisticated models of the protein

interaction network, spatial modelling, and including feedback loops and crosstalk in models.

References

1. ????
2. Witten IH, Frank E. Data Mining: Practical Machine Learning Tools and Techniques. Amsterdam: Elsevier, 2005.
3. Quinlan R. C4.5: Programs for Machine Learning. San Francisco: Morgan Kaufmann, 1992.
4. Rendell L, Seshu R. Learning Hard Concepts through Constructive Induction: Framework and Rationale. Computational Intelligence, 1995; 6:247–270.
5. Fall CP, et al. **Computational Cell Biology. Berlin: Springer-Verlag, 2002.**
6. Murray JD. Mathematical Biology: I. An Introduction, 3rd ed. Berlin:Springer-Verlag, 2004.
7. Nelson DL, Cox MM. Lehninger Principles of Biochemistry. London: Palgrave-Macmillan, 2004.

8. Michalewicz Z, Fogel DB. How to Solve It: Modern Heuristics, 2nd ed. Berlin: Springer-Verlag, 2004.
9. Kuipers B. Qualitative Reasoning: Modeling and Simulation with Incomplete Knowledge. Cambridge: MIT Press, 1994.
10. Kuipers B. Qualitative simulation. In Meyers RA (Ed.), Encyclopedia of Physical Science and Technology, 2nd ed. Academic Press, 2001.
11. Johnson CG, Goldman JP, Gullick WJ. Simulating Complex intracellular processes using object-oriented computational modelling. Progress in Biophysics and Molecular Biology, 2004; 86:379–406.
12. Law AM. Simulation, Modeling and Analysis, 4th ed. New York: McGraw-Hill, 2006
13. Andrews SS, Bray D. Stochastic simulation of chemical reactions with spatial resolution and single molecule detail Physical Biology, 2005; 1:137-151.
14. Lemerle C, Di Ventura B, Serrano L. Space as the final frontier in stochastic simulations of biological systems. FEBS Letters, 2005; 559:1789–1794.
15. Gullick WJ. The type 1 growth factor receptors and their ligands considered as a complex system. Endocrine-Related Cancer, 2001; 8:75–82.

16. Bazley LA, Gullick WJ. The epidermal growth factor receptor family. *Endocrine-Related Cancer*, 2005;12(S17–S27).
17. Wiley HS, Shvartsman SY, Lauffenburger DA. Computational modeling of the EGF-receptor system: a paradigm for systems biology. *TRENDS in Cell Biology*, 2003; 13(1):43–50.
18. Orton RJ, Sturm OE, Vyshemirsky V, et al. Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway. *Biochemistry Journal*, 2005; 392:249–261.
19. Wolfram S. The *Mathematica* Book (Fifth Edition). Champaign, IL:Wolfram Media, 2004.
20. Higham DJ, Higham NJ. *Matlab Guide*. Philadelphia, Society for Industrial and Applied Mathematics, 2005.
21. Mendes P. Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3. *Trends in Biochemical Science*, 1997; 22:361–363.
22. Suresh Babu CV, Yoon S, Nam H-S, Yoom YS. Simulation and sensitivity analysis of phosphorylation of EGFR signal transduction pathway in PC12 cell model. *Systems Biology*, 2004; 1(2):213–221.

23. Brightman FA, Fell DA. Differential feedback regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signalling in PC12 cells. *FEBS Letters*, 2000; 482:169–174.

24. Hendricks BS, et al. Computational modelling of ErbB family phosphorylation dynamics in response to transforming growth factor alpha and heregulin indicates spatial compartmentalisation of phosphatase activity. *IEE Proceedings—Systems Biology*, 2006; 153(1): 22–33.

25. Shvartsman SY, Muratov CB, Lauffenburger DA. Modelling and computational analysis of EGF receptor-mediated cell communication in *Drosophila* oogenesis. *Development*, 2002; 129:2577–2589.

26. Maly IV, Wiley SH, Lauffenburger DA. Self-organizations of polarized cell signalling via autocrine circuits: computational model analysis. *Biophysical Journal*, 2004; 86:10–22.

27. Hlavacek WS, Faeder JR, Blinov ML, et al. Rules for modelling signal-transduction systems. *Science's STKE*, 2006:

28. Blinov ML, Yang J, Faeder JR, Hlavacek W. Depicting signaling cascades. *Nature Biotechnology*, 2006; 24(2):137–138.

29. Kholodenko BN, Demin OV, Moehren G, Hoek JB. Quantification of short term signalling by the epidermal growth factor receptor. *Journal of Biological Chemistry*, 1999; 274:30169–30181.
30. Lee D-Y, Zimmer R, Lee SY, Park S. Colored Petri net modeling and simulation of signal transduction pathways. *Metabolic Engineering*, 2006; 8:112–122.
31. Schamel WWA, Dick TP. Signal transduction: specificity of growth factors explained by parallel distributed processing. *Medical Hypotheses*, 1996; 47:249–255.
32. Pawson T, Linding R. Synthetic Modular Systems—reverse engineering of signal transduction. *FEBS Letters*, 2005; 579:1808–1814.
33. Goldman JP, Gullick WJ, Johnson CG. Individual-based simulation of the clustering behaviour of epidermal growth factor receptors. *Scientific Programming*, 2004; 12(1):25–43
34. Walker D, Wood S, Southgate J, Holcombe M, Smallwood R. An integrated agent-mathematical model of the effect of intercellular signalling via the epidermal growth factor receptor on cell proliferations. *Journal of Theoretical Biology*, 2006; 242(3):774–789.

35. Hendricks BS, et al. Decreased internalisation of ErbB1 mutants in lung cancer is linked with a mechanism conferring sensitivity to gefitinib. IEE Proceedings—Systems Biology, 2006; 153(6):457–466.

36. Calder M, Duguid A, Gilmore S, Hillston J. Stronger computational modelling of signalling pathways using both continuous and discrete-state methods. In Priami C (Ed.), Proceedings of the Fourth International Workshop on Computational Methods in Systems Biology. Berlin: Springer-Verlag (Lecture Notes in Computer Science vol. 4210), 2006.

37. Schoeberl B, Eichler-Jonsson C, Gilles ED, Muller G. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalised EGF receptors. Nature Biotechnology, 2002; 20:370–275.

38. Kolch W, Calder M, Gilbert D. When kinases meet mathematics: the systems biology of MAPK signalling. FEBS Letters, 2005; 579:1891–1895.

39. Hornberg JJ, Bruggeman FJ, Westerhoff HV, Lankelma J. Cancer: a systems biology disease. BioSystems, 2005; 83:81–90.

40. Citri A Yarden Y. EGF-ERBB signalling: towards the systems level. Nature Reviews Molecular Cell Biology, 2006; 7:505–515.

41. Suresh Babu CV, Song EJ, Yoo YS. Modeling and simulation in signal transduction pathways: a systems biology approach. *Biochemie*, 2006; 88(3–4):277–83.
42. Hucka M, Finney A, Sauro HM, et al. The Systems Biology Markup Language (SBML): A Medium for Representation and Exchange of Biochemical Network Models. *Bioinformatics*, 2003; 19(4): 524-531.
43. Tomita M, Hasimoto K, Takahashi K, et al. E-CELL: software environment for whole-cell simulation. *Bioinformatics*, 1999; 15(1):72-84.
- 44. Oda K, et al. A comprehensive pathway map of epidermal growth factor receptor signaling. *Molecular Systems Biology*, 2005; doi:10.1038/msb4100014.**
45. Kitano H, Funahashi A, Matsuoka Y, Oda K. Using process diagrams for the graphical representation of biochemical networks. *Nature Biotechnology*, 2005; 23:961–966.
46. Cary MP, Bader GD, Sander C. Pathway information for systems biology. *FEBS Letters*, 2005; 579:1815–1820.
47. Jones RB, Gordus A, Krall JA, MacBeath G. A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature*, 2006; 439(7073):168–174.

48. Haugh JM, Wells A, Lauffenburger DA. Mathematical modeling of epidermal growth factor receptor signalling through the phospholipase pathway: mechanistic insights and predictions for molecular interventions. *Biotechnology and Bioengineering*, 2000; 70(2):225–238.

49. Breitling R, Hoeller D. Current challenges in quantitative modelling of epidermal growth factor receptor signalling. *FEBS Letters*, 2005; 579:6289–6294.