Making Sense of Complex Phenomena in Biology

Philip K. Maini, Centre for Mathematical Biology,
Mathematical Institute, 24-29 St Giles,
Oxford OX1 3LB

Abstract The remarkable advances in biotechnology over the past two decades have resulted in the generation of a huge amount of experimental data. It is now recognised that, in many cases, to extract information from this data requires the development of computational models. Models can help gain insight on various mechanisms and to be used to process outcomes of complex biological interactions. To do the latter, models must become increasingly complex and, in many cases, they also become mathematically intractable. With the vast increase in computing power these models can now be numerically solved and can be made more and more sophisticated. A number of models can now successfully reproduce detailed observed biological phenomena and make important testable predictions. This naturally raises the question of what we mean by understanding a phenomenon by modelling it computationally. This paper briefly considers some selected examples of how simple mathematical models have provided deep insights into complicated chemical and biological phenomena and addresses the issue of what role, if any, mathematics has to play in computational biology.

Keywords: Computational model, mathematical model, networks, pattern formation, robustness, microscopic to macroscopic, layered models, integration

1. Introduction

The enormous advances in molecular and cellular biology over the last two decades have led to an explosion of experimental data in the biomedical sciences. We now have the complete (or almost complete) mapping of the genome of a number of organisms and we can determine when in development certain genes are switched on; we can investigate at the molecular level complex interactions leading to cell differentiation and we can accurately follow the fate of single cells. However, we have to be careful not to fall into the practices of the nineteenth century, when biology was steeped in the mode of classification and there was a tremendous amount of list-making activity. This was recognised by D'Arcy Thompson, in his classic work "On Growth and Form", first published in 1917 (see Thompson 1992 for the abridged version). He had the vision to realise that, although simply cataloging different forms was an essential data collecting exercise, it was also vitally important to develop theories as to how certain forms arose. Only then, could one really comprehend the phenomenon under study.

Of course, the identification of a gene that causes a certain deformity, or affects an ion channel making an individual susceptible to certain diseases, has huge benefits for medicine. At the same time, one must recognise that collecting data is, in some sense, only the beginning. Knowing the spatiotemporal dynamics of a certain gene leads to the inevitable question of why that gene was switched on at that particular time and place. Genes contain the information to synthesize proteins. It is the physico-chemical interaction of proteins and cells that lead to, for example, the development of structure and form in the early embryo. Cell fate can be determined by environmental factors as cells respond to signalling cues. Therefore, a study at the molecular level alone will not help us to understand how cells

interact. Such interactions are highly nonlinear, may be nonlocal, certainly involve multiple feedback loops and may even incorporate delays. Therefore they must be couched in a language that is able to compute the results of complex interactions. At the moment, the best language we have for carrying out such calculations is mathematics. Mathematics has been extremely successful in helping us to understand physics. It is now becoming clear that mathematics, and computation, have a similar role to play in the life sciences.

Mathematics can play a number of important roles in making sense of complex phenomena. For example, in a phenomenon in which the microscopic elements are known in detail, the integration of interactions at this level to yield the observed macroscopic behaviour can be understood by capturing the essence of the whole process through focussing on the key elements, which form a small subset of the full microscopic system. Two examples of this are given in Section 2. Mathematical analysis can show that several microscopic representations can give rise to the same macroscopic behaviour (see Section 3), and that the behaviour at the macroscopic level may be greater than the sum of the individual microscopic parts (Section 4).

2.1 Belousov-Zhabotinskii reaction

The phenomenon of temporal oscillations in chemical systems was first observed by Belousov in 1951 in the reaction now known as the Belousov-Zhabotinskii (BZ) reaction (for details see Field and Berger 1985). The classical BZ reaction consists of the oxidation by bromate ion in an acidic medium catalysed by metal ion oxidants, for example, the oxidation of malonic acid, in an acid medium, by bromate ions, BrO_3^-, and catalysed by cerium, which has two states Ce^{3+} and Ce^{4+}. With other metal ion catalysts and appropriate dyes, the reaction can be followed by observing changes in colour. This system is capable of producing a spectacular array of spatiotemporal dynamics, including two-dimensional target patterns and outwardly rotating spiral waves, three-dimensional scroll waves and, most recently, two-dimensional inwardly rotating spirals (Vanag and Epstein, 2001). All the steps in this reaction are still not fully determined and understood and, to date, there are of the order of about 80 reaction steps known. Detailed mathematical models have been written down for this reaction (see, for example, Field et al 1972) consisting of several coupled nonlinear ordinary differential equations. Remarkably, a vast range of the dynamics of the full reaction can be understood by a simplified model consisting of only three, coupled, nonlinear differential equations, which can be further reduced to two equations. The reduction arises due to a mixture of caricaturizing certain complex interactions and using the fact that a number of reactions operate on different time scales, so that one can use a quasi-steady-state approach to reduce some differential equations to simpler algebraic equations, allowing for the elimination of certain variables.

A phase-plane analysis of the simplifed model leads to an understanding of the essence of the pattern generator within the BZ reaction, namely the relaxation oscillator. This relies on the presence of a slow variable and a fast variable with certain characteristic dynamics (see, for example, Murray 1993). The introduction of diffusion into this model, leading to a system of coupled partial differential equations, allows for the model to capture a bewildering array of the spatiotemporal phenomena observed experimentally, such as propagating fronts, spiral waves, target patterns and toroidal scrolls.

These reduced models have proved to be an invaluable tool for the understanding of the

essential mechanisms underlying the patterning processes in the BZ reaction in the way that the study of a detailed computational model would have been impossible. With over 80 reactions and a myriad parameters (many unknown), the number of simulations required to carry out a full study would be astronomical.

2.2 Models for electrical activity

The problem of how a nerve impulse travels along an axon is central to the understanding of neural communication. The Hodgkin-Huxley model for electrical firing in the axon of the giant squid (see, for example, Cronin 1987) was a triumph of mathematical modelling in physiology and they later received the Nobel prize for their work. The model, describing the temporal dynamics of a number of key ionic species which contribute to the transmembrane potential, consists of four complicated, highly nonlinear coupled ordinary differential equations. A well studied reduction of the model, the FitzHugh-Nagumo model, is a caricature and consists of only two equations (FitzHugh 1961, Nagumo et al 1962). Again, a phase plane analysis of this model reveals the essential phenemenon of *excitability* by which a neuron "fires" and determines the kinetic properties required to exhibit this behaviour.

3. Models for aggregation in *Dictyostelium discoideum*

The amoeba *Dictyostelium discoideum* (Dd) is one of the most studied organisms in developmental biology from both experimental and theoretical aspects and serves as a model paradigm for development in higher organisms. In response to starvation conditions, these unicellular organisms chemically signal each other via cyclic AMP leading to a multicellular aggregation in which the amoebae undergo differentiation into a stalk type and a spore type. The latter can survive for many years until conditions are favourable.

Intercellular signalling in this system, which involves relay and transduction, has been widely studied and modelled. For example, the Martiel and Goldbeter model (Martiel and Goldbeter 1987) consists of nine ordinary differential equations. By exploiting the different timescales on which reactions occur, this model can be reduced to simpler two- and three- variable systems which not only capture most of the experimental behaviour, but also allow one to determine under which parameter constraints certain phenomena arise (Goldbeter 1996). This model turns out to exhibit excitable behaviour, similar in essence to that observed in electrical propagation in nerves.

Such reduced, or caricature models, can then serve as "modules" to be plugged in to behaviour at a higher level in a layered model to understand, for example, the phenomenon of cell streaming and aggregation in response to chemotactic signalling (Höfer et al 1995a,b, Höfer & Maini 1997). Assuming that the cells can be modelled as a continuum, it was shown that the resultant model could exhibit behaviour in agreement with experimental observations. Moreover, the model provided a simple (and counter-intuitive) explanation for why the speed of wave propagation slows down with increasing wave number. More sophisticated computational models, in which cells are assumed to be discrete entities, have been shown to give rise to similar behaviour (Dallon & Othmer 1997). Such detailed models can be used to compare the movement of individual cells with experimental observations and therefore allow for a degree of verification that is impossible for models at the continuum level. However, the latter are mathematically tractable and therefore can be used to determine generic behaviours.

Several models, differing in their interpretation of the relay/transduction mechanism and/or details of the chemotactic response all exhibit very similar behaviour (Dallon et al 1997). In one sense this can be thought of as a failure because modelling has been unable to distinguish between different scenarios. On the other hand, these modelling efforts illustrate that the phenomenon of Dd aggregation is very robust and has, at its heart, signal relay and chemotaxis.

4. The Turing model for pattern formation

Diffusion-driven instability was first proposed by Turing in a remarkable paper (Turing, 1952), as a mechanism for generating self-organised spatial patterns. He considered a pair of chemicals reacting in such a way that the reaction kinetics were stabilizing, leading to a temporally stable, spatially uniform steady state in chemical concentrations. As we know, diffusion is a homogenizing process. Yet combined in the appropriate way, Turing showed mathematically that these two stabilizing influences could conspire to produce an instability resulting in spatially heterogeneous chemical profiles – a spatial pattern. This is an example of an *emergent property* and led to the general patterning principle of *short-range activation, long-range inhibition* (Gierer & Meinhardt, 1972). Such patterns were later discovered in actual chemical systems and this mechanism has been proposed as a possible biological pattern generator (for a review, see Maini et al 1997, Murray 1993).

Turing's study raises a number of important points. It showed that one cannot justifably follow a purely reductionist approach, as the whole may well be greater than the sum of the parts and that one rules out, at one's peril, the possibility of counter-intuitive phenomena emerging as a consequence of collective behaviour. It also illustrates the power of the mathematical technique because, had these results been shown in a computational model without any mathematical backing, it would have been assumed that the instability which is, after all, counter-intuitive, could only have arisen due to a computational artifact. Not only did the mathematics show that the instability was a true reflection of the model behaviour, it also specified exactly the properties the underlying interactions in the system must possesss in order to exhibit the patterning phenomenon. Furthermore, mathematics served to enhance our intuitive understanding of a complex nonlinear system.

5. Discussion

For models to be useful in, for example, drug design, they must necessarily incorporate a level of detail that, on the whole, makes the model mathematically intractable. The phenomenal increase in computing power over recent years now means that very sophisticated models involving the interaction of hundreds of variables in a complex three-dimensional geometry can be solved numerically. This naturally raises a number of questions: (i) How do we validate the model? Specifically, if the model exhibits a counter-intuitive result, which is one of the most powerful uses of a model, how do we know that this is a faithful and generic outcome of the model and not simply the result of very special choice of model parameters, or an error in coding? (ii) If we take modelling to its ultimate extreme, we simply replace a biological system we do not understand by a computational model we do not understand. Although the latter is useful in that it can be used to compute the results of virtual experiments, can we say that the exercise has furthered our understanding? Moreover, since it is a model and therefore, by necessity, wrong in the strict sense of the word, how do

we know that we are justified in using the model in a particular context?

In going from the gene to the whole organism, biological systems consist of an interaction of processes operating on a wide range of spatial and temporal scales. It is impossible to compute the effects of all the interactions at any level of this spatial hierarchy, even if they were all known. The approach to be taken, therefore, must involve a large degree of caricaturizing, based on experimental experience, and reduction, based on mathematical analysis. The degree to which one simplifies a model depends very much on the question one wishes to answer. For example, to understand in detail the effect of a particular element in the transduction pathway in *Dictyostelium discoideum* will require a detailed model at that level. However, for understanding aspects of cell movement in response to the signal, it may be sufficient to consider a very simple model which represents the behaviour at the signal transduction level, allowing most of the analytical and computational effort to be spent on investigating cell movement. In this way, one can go from one spatial level to another by "modularizing" processes at one level (or layer) to be plugged in to the next level. To do this, it is vital to make sure that the appropriate approximations have been made and the correct parameter space and spatiotemporal scales are used. This comes most naturally via a mathematical treatment. Eventually, this allows for a detailed mathematical validation of the process before one begins to expand the models to make them more realistic.

The particular examples considered in this article use the classical techniques of applied mathematics to help understand model behaviour. Much of the mathematical theory underlying dynamical systems and reaction-diffusion equations was motivated by problems in ecology, epidemiology, chemistry and biology. The excitment behind the Turing theory of pattern formation and other areas of nonlinear dynamics was that very simple interactions could give rise to very complex behaviour. However, it is becoming increasingly clear that often in biology very complex interactions give rise to very simple behaviours. For example, complex biochemical networks are used to produce only a limited number of outcomes (von Dassow et al 2000). This suggests that it may be the interactions, not the parameter values, that determine system behaviour and, in particular, robustness. This requires perhaps the use of topological or graph theoretical ideas as tools for investigation. Hence it is clear that it will be necessary to incorporate tools from other branches of mathematics and to develop new mathematical approaches if we are to make sense of the mechanisms underlying the complexity of biological phenomena.

Acknowledgements This paper was written while the author was a Senior Visiting Fellow at the Isaac Newton Institute for Mathematical Sciences, University of Cambridge. I would like to thank Santiago Schnell and Dr Edmund Crampin for helpful discussions.

References

Cronin J 1987 Mathematical aspects of Hodgkin-Huxley neural theory. Cambridge University Press, Cambridge

Dallon JC, Othmer HG 1997 A discrete cell model with adaptive signaling for aggregation of *Dictyostelium discoideum*. Phil Trans Roy Soc Lond B352:391-417

Dallon JC, Othmer HG, van Oss C, Panfilov A, Hogeweg P, Höfer T, Maini PK 1997 Models of *Dictyostelium discoideum* aggregation. In: Alt W, Deutsch G, Dunn G (eds) Dynamics of cell and tissue motion. Birkhauser Boston p 193-202

Field RJ, Körös E, Burger M 1972 Oscillations in chemical systems, Part 2. Thorough analysis of temporal oscillations in the bromate-cerium-malonic acid system. J Am Chem Soc 94:8649-8664

Field RJ, Burger M (eds) 1985 Oscillations and travelling waves in chemical systems. John Wiley & Sons, New York

FitzHugh R 1961 Impulses and physiological states in theoretical models of nerve membrane. Biophys J 1:445-466

Gierer A, Meinhardt H 1972 A theory of biological pattern formation. Kybernetik 12:30-39

Goldbeter A 1996 Biochemical oscillations and cellular rhythms. Cambridge University Press, Cambridge

Höfer T, Sherratt JA Maini PK 1995a *Dictyostelium discoideum*: cellular self-organization in an excitable biological medium. Proc Roy Soc Lond B259:249-257

Höfer T, Sherratt JA, Maini PK 1995b Cellular pattern formation during *Dictyostelium* aggregation. Physica D 85:425-444

Höfer T, Maini PK 1997 Streaming instability of slime mold amoebae: An analytical model. Phys Rev E 56(2):2074-2080

Martiel JL, Goldbeter A 1987 A model based on receptor desensitization for cyclic AMP signaling in *Dictyostelium* cells. Biophys J 52:807-828

Maini PK, Painter KJ, Chau, HNP 1997 Spatial pattern formation in chemical and biological systems. Faraday Transactions 93(20):3601-3610

Murray JD 1993 Mathematical biology. Springer-Verlag, Berlin, Heidelberg, New York

Nagumo JS, Arimoto S, Yoshizawa S 1962 An active pulse transmission line simulating nerve axon. Proc IRE 50:2061-2071

Thompson DW 1992 On growth and form. Cambridge University Press, Cambridge

Turing AM 1952 The chemical basis of morphogenesis. Phil Trans Roy Soc Lond B327:37-72

von Dassow G, Meir E, Munro EM, Odell GM 2000 The segment polarity network is a robust developmental module. Nature 406:188-192