The role of modelling in identifying drug targets for diseases of the cell cycle

Robert G. Clyde¹, James L. Bown¹, Ted R. Hupp², Nikolai Zhelev¹
and John W. Crawford¹,*

¹SIMBIOS, University of Abertay Dundee, Kydd Building, Bell Street, Dundee DD1 1HG, UK
²CRUK Cell Signalling Unit, University of Edinburgh, South Crewe Road, Edinburgh EH4 2XR, UK

The cell cycle is implicated in diseases that are the leading cause of mortality and morbidity in the developed world. Until recently, the search for drug targets has focused on relatively small parts of the regulatory network under the assumption that key events can be controlled by targeting single pathways. This is valid provided the impact of couplings to the wider scale context of the network can be ignored. The resulting depth of study has revealed many new insights; however, these have been won at the expense of breadth and a proper understanding of the consequences of links between the different parts of the network. Since it is now becoming clear that these early assumptions may not hold and successful treatments are likely to employ drugs that simultaneously target a number of different sites in the regulatory network, it is timely to redress this imbalance. However, the substantial increase in complexity presents new challenges and necessitates parallel theoretical and experimental approaches. We review the current status of theoretical models for the cell cycle in light of these new challenges. Many of the existing approaches are not sufficiently comprehensive to simultaneously incorporate the required extent of couplings. Where more appropriate levels of complexity are incorporated, the models are difficult to link directly to currently available data. Further progress requires a better integration of experiment and theory. New kinds of data are required that are quantitative, have a higher temporal resolution and that allow simultaneous quantitative comparison of the concentration of larger numbers of different proteins. More comprehensive models are required and must accommodate not only substantial uncertainties in the structure and kinetic parameters of the networks, but also high levels of ignorance. The most recent results relating network complexity to robustness of the dynamics provide clues that suggest progress is possible.

Keywords: cell cycle; model; mathematics; cancer; drug discovery; systems biology

1. INTRODUCTION

The cell cycle is an ordered sequence of events that involves the process by which a cell grows and then divides resulting in the production of two daughter cells that are identical to the original parent cell. The cycle may be considered in four phases commencing with the G1-phase (G1) during which the cell undergoes a series of biochemical and physiological changes, including sustained growth, all in preparation for the S-phase (S) that follows. During S, the cell copies its DNA resulting in the development of duplicate copies of each chromosome. This is followed by the G2-phase (G2), a gap phase, prior to M-phase (M), i.e. mitosis, when the cell divides, with one set of chromosomes being allocated to each daughter cell. The process by which all of this takes place is highly regulated and dependent on a pre-defined and finely tuned interplay among all of the proteins involved. The progress through the phases is governed by a group of cyclin-dependent kinases (Cdks), the principal members being Cdk4, which is active in mid-G1, Cdk2 which is active in late G1, S and M and finally Cdk1, which is active only in M. The activity of these Cdks is itself controlled by a number of mechanisms that include association with cyclins, these being cyclin D in G1, cyclin E in late G1, cyclin A in S, G2 and M, and cyclin B in M. Activity is also regulated by phosphorylation and dephosphorylation reactions as well as association with a number of inhibitors including P16, which inhibits Cdk4 and also P21 and P27, both of which can inhibit Cdk4 and Cdk2.

Received 25 May 2006
Accepted 11 July 2006
This tightly controlled Cdk activity enables the cell not only to execute the cycle correctly, but also to observe a number of control checks on its own progress. The first of these occurs in late G1 and is generally referred to as the restriction point, a point in time after which the cell becomes committed to completing its full cycle. Thereafter, checkpoint controls exist in relation to DNA damage and DNA replication, as well as spindle assembly and chromosome segregation during M-phase. Any failure by the cell to fulfill a checkpoint requirement will result in the arrest of the cell cycle followed either by abandonment of the cycle, if the restriction point has not been reached, or alternatively a delay in the cycle while the relevant error is corrected. In cases of serious error, the cell may undergo programmed cell death, a suicidal process usually referred to as apoptosis.

The importance of the cell cycle in relation to cancer is emphasized by the number of papers setting out the impact of the cell regulatory proteins on the development of this disease. Examples include Collins et al. (1997) who summarize the cell cycle describing the role of the cyclins and Cdks as well as that of the tumour suppressor Rb. Loss of function and gain of function mutations are discussed as well as their relationship to particular cancers. Also, Sandal (2002) gives a clear description of how the cyclins, Cdks, inhibitors and phosphatases combine to regulate the cell cycle as well as how the tumour suppressors P53, Rb and P19<sup>Arf</sup> keep the system in check. The relationship between mutations, along with the respective disruption in gene expression, and specific types of cancer is explained.

One of the major, long-term aims of cell-cycle research is to develop more effective drug targets and this review focuses on the potential impact of mathematical modelling on such development. Mathematical modelling may provide a vehicle to synthesize existing, disparate and heterogeneous datasets and manage the inherent complexity of the cell cycle. It is likely that the development of multiple drug target strategies will be rewarding and mathematics is the only realistic tool for identifying minimum toxicity regimes where the concurrent behaviour of several hundred different protein species have to be considered. Over the years, there have been various attempts to model the cell cycle and these vary in their fitness for the purpose referred to above. This review discusses these previous approaches and draws conclusions on their merits as well as identifying gaps and opportunities for the future.

2. EARLY DEVELOPMENTS

Cell-cycle modelling has tended to follow the prevailing level of the relevant biological knowledge rather than drive it. Initially, two quite different views prevailed regarding the control mechanisms underlying the cell cycle. Prior to 1980, little was known about the biochemistry or the physical processes of the proteins that regulate the cell cycle and the fact that cell reproduction followed a sequential process of growth and division had led to the idea that the cell cycle could be modelled as an oscillating system driven by changes in size and mass (Sachsemaier et al. 1972; Fantes et al. 1975; Nurse 1975). This view, taken mainly by embryologists and physiologists, and focusing specifically on rapidly dividing embryonic cells, saw the cell cycle as a biochemical machine which oscillated between two states, namely interface and mitosis. This view was generally referred to as the ‘clock’ theory.

However, geneticists focusing on somatic cells held an alternative view which saw the cell cycle as a series of independent reactions where the completion of each reaction was dependent on the previous one. This was generally referred to as the ‘domino’ theory and followed experiments by Hartwell (1978) on budding yeast and Lee & Nurse (1988) on fission yeast, which gave insight into a number of molecular mechanisms apparently linked to cell-cycle control. This had a fundamental effect on how mathematical modelling of the cell cycle would be approached in future years. While cell size and mass were still considered important, molecular interactions became the principle processes in models that linked the dynamics of particular protein concentrations to cell-cycle physiology. Initially, knowledge in terms of specific molecules and their interactions lacked detail. However, research into the various proteins and mechanisms that regulate the cell cycle has resulted in progressively better descriptions of the cell cycle and in response to this, mathematical models have become increasingly more detailed.

A unified view of the cell cycle, incorporating both the ‘domino’ and the ‘clock’ theories, was later provided by Murray & Kirschner (1989). This was based on the idea that key components of the cell cycle, both in rapidly dividing embryonic cells and in growing somatic cells, each appeared to be the products of homologous genes.

One of the earliest models depicting cell physiology in terms of molecular concentrations is due to Tyson (1991). Using differential equations, he modelled the cell-division cycle on the basis of the concentration of the proteins cdc2 and cyclin, allied to their ability to form complexes and to modify their function through phosphorylation and dephosphorylation. He demonstrated that, by varying the parameter values controlling the maximum rate of activation of the cyclin: Cdc2 complex in its active form, referred to as maturation-promoting factor (MPF), and also its rate of disassociation, the model was capable of demonstrating three separate modes of behaviour. These consisted firstly a steady state with high levels of MPF, which could be taken to represent metaphase arrest in an unfertilized egg, secondly, spontaneous oscillations representing rapid division cycles in embryonic cells and finally an excitable switch representing growth-controlled division in growing cells. During the same period, a similar model was constructed by Goldbeter (1991). Following previous work on oscillatory phenomena in biological systems (Goldbeter 1985; Goldbeter et al. 1988, 1990; Goldbeter & Dupont 1990; Goldbeter & Wolpert 1990), that included both an empirical approach and a study of the relevance of allosteric regulation, this model drew on the work that had been carried out up to that time on yeasts and embryonic cells (Murray & Kirschner 1989; Nurse 1990), and again...
modelled the cell cycle in terms of the post-translational modification of the cyclin-dependent kinase cdc2. For certain parameter sets, the model demonstrated oscillations and accorded well with the experimental data.

During the 1990s, Thron (1991, 1994, 1996, 1999) provided a series of models and analyses covering the mitotic clock, the dynamics of the cyclin B–MPF system, a bistable trigger for mitosis and a model of the activation of a cell-cycle kinase which downregulates its own inhibitor. Romond et al. (1999) considered the cell cycle in a skeletal model comprising two coupled biochemical oscillators under mutually negative control. By modelling the dynamic behaviour as a function of the strength of mutual inhibition, the model demonstrated that alternating oscillations are the result when mutual inhibition is strong and that these oscillations can be interpreted in terms of the cell in S-phase when DNA is copied and, alternately, in M-phase when the cell divides. With weak inhibition, however, the model exhibits complex dynamical behaviour including complex periodic oscillations, chaos and coexistence between multiple periodic or chaotic attractors.

However, the methodology used throughout this period of early development was perhaps of more significance than the results themselves; this included the combined use of network diagrams to show protein interactions, description of the model using differential equations and the use of phase plane analysis and bifurcation theory for interpretation purposes. These have become the preferred tools of most subsequent work.

3. CELL-CYCLE MODELS BASED ON YEASTS

Perhaps, the most consistent and progressive work on cell-cycle modelling over the last decade or so has been that presented by the group headed by J. J. Tyson and B. Novak. Based on many experimental papers on the fission yeast cycle, Novak & Tyson (1995) constructed a model that simulated the results to date relating to mitotic control as well as providing a number of predictions relating to future experimental work. This was followed over the next 4 years by several papers including (Novak & Tyson 1997; Novak et al. 1997), mainly on yeasts, and looking at S-phase and checkpoint control as well as the cell cycle itself. Novak et al. (1999) produced a model focusing on mitosis and demonstrating that the cell could exit from mitosis in the absence of cyclin degradation. The methodology, including the reduction in complexity of the cell cycle into manageable parts is explained in detail in Tyson & Novak (2001), where arguments are also presented for the cell cycle taking the form of a hysteresis loop rather than a limit cycle. Cell-cycle dynamics and the methodology used in its derivation are developed in Tyson et al. (2001, 2002). While much of Tyson and Novak’s work specifically showed that the cell cycle demonstrates bistability and hysteresis. They propose that the G1-phase and the S/M-phase are the two alternative self-maintaining steady states generated as a result of mutual antagonism between the cyclin-dependent kinases and Sic1 and Hct1 which oppose the enzyme activity. These predictions were tested by Cross et al. (2002) and agreement was found in relation to predicted cyclin and Sic1 inhibitor quantities, and also with regard to the relationship between cell size and Cln3 gene dosage. Although the predictions about the interaction between Cdh1 and the G1 cyclins were found to be inaccurate, the model was capable of reproducing many of the real events of the cell cycle.

Sveiczer et al. (2000), building on previous work by Novak et al. (1997, 1999), modelled the fission yeast cell cycle using differential equations and included mass in the equations as a variable controlling the kinetics of the number of proteins involved in the regulation of the cell cycle. The topology of this model includes positive and negative feedback loops leading to the self-activation and deactivation of Cdc2. Positive feedbacks arise from Cdc2 inactivating its negative regulator as well as activating Cdc25, its positive regulator. Negative feedback is due to the reactivation of the negative regulators leading to a reduction in Cdc2 activity. In addition to modelling the cell cycle, the model identified a potential resetting mechanism for the G2/M transition when the positive feedback loops were weak. Sveiczer et al. (2002) followed this work with a stochastic model where the known asymmetry of cell division in fission yeast and differing nuclear volumes at birth were used as the basis for cyclin accumulation. Results accorded well with experimental observations. In the following year, Sveiczer et al. (2003) produced another deterministic model that was able to describe the behaviour of wild-type cells as well as a number of mutants. This model also explained why both Run1 and Ste9 are essential for blocking cells in G1 when the transcription protein Cdc10 is mutated. Chen et al. (2004) have produced further work on budding yeast which probably represents the most comprehensive model of the cell cycle to date. This substantial work demonstrates a complete methodology for modelling molecular regulatory networks. The model conforms to the phenotypes of more than 100 mutant strains and is capable of predicting further mutant phenotypes as well as predicting values for kinetic constants. The model is also shown to be extremely robust in terms of parameter variation.

4. MODELLING KEY TRANSITIONS IN THE CELL CYCLE

As biological research has continued to reveal ever more complex descriptions of cell-cycle regulation, holism has been sacrificed for increased levels of
biological detail and, in many instances, particularly in regard to mammalian cells, modelling has frequently been focused on specific key transition points within the cycle rather than on the cycle itself.

4.1. The G1/S transition
Hatzimanikatis et al. (1999) modelled this period in the cell cycle, concluding that the cell’s progress through the transition could be represented by a limit cycle. Thron (1999) produced a model that simulated the reactions between the heterodimeric kinase cyclin E : Cdk2 and the inhibitor P27. By mathematical analysis, he was able to demonstrate the conditions for a bistable biochemical system whereby cyclin E : Cdk2 switches from a low-activity state inhibited by P27 to a high-activity state resulting from the rapid degradation of P27. Qu et al. (2003a, b) modelled the G1/S transition in terms of a limit cycle resulting from the positive feedback between Cdk2 and the CKI inhibitor. However, they were also able to demonstrate that the positive feedback between cyclin E and E2F could result in a bistable state. They showed that dynamical behaviour was dependent on parameter choice and therefore ambiguous. However, the model did show that the number of phosphorylation sites activated in the case of Cdc25A, Rb and CKI, was critical to cell-cycle progression.

Also in the same year, Qu et al. (2003a, b) presented a generic mathematical model capable, in this instance, of simulating both the G1/S and G2/M transitions. This model which demonstrated both bistability (due to the positive feedback flowing from Wee1 and Cdc25A) and limit cycle behaviour (due to the negative feedback formed by SKP2) was also able to accommodate checkpoint controls as well as the growth time involved in the cell cycle prior to duplication of DNA (referred to as the sizer phase, the period of which is determined by the birth size of the cell).

4.2. The G2/M transition
In the early 1990s, Thron (1991, 1994, 1996) constructed cell-cycle models, which were restricted to mitosis only and showed some of the conditions necessary for oscillatory and bistable behaviour. Aguda (1999) argued that the G2/M checkpoint control system consisted of a system of phosphorylation–dephosphorylation (PD) cycles, involving cyclin B : Cdc2, Cdc25, Wee1 and Myt1, which had inherent instability. It was argued that this condition could be exploited by signal-transduction pathways to delay the cell cycle. In an analysis, Aguda concluded that Cdc25 is in fact the target of the damage checkpoint pathway in G2/M, a conclusion with which many cell biologists agree. The generic model by Qu et al. (2003a, b), previously described under the G1/M transition is also applicable to G2/M.

4.3. The restriction point
Aguda & Tang (1999) constructed a model of the restriction point of the mammalian cell cycle based on the timing of a rapid increase in the level of active cyclin E : Cdk2. The sharp switching behaviour of this complex could be interpreted as emerging from a positive feedback loop between cyclin E : Cdk2 and Cdc25A as well as a mutually negative interaction between cyclin E : Cdk2 and the inhibitor P27. The timing of the restriction point was also found to depend on the level of P27 along with its affinity to bind with the enzyme. Furthermore, they were able to demonstrate that sharp switching behaviour could still occur in the absence of pRb participation, perhaps suggesting that E2F-induced expression of Myc and downgrading of P27 via the Ras pathway were also involved in defining the restriction point. A sensitivity analysis of the model was later carried out by Tashima et al. (2003) and this suggested, contrary to the generally accepted view, that the E2 : FRb pathway had little impact on switching behaviour. They concluded that further experimental work was required together with the introduction of additional pathways into the model.

Novak & Tyson (2004) modelled the restriction point in mammalian cells. Building on their previous work on yeasts, and including provision for the Rb pathway as well as cyclins D, E and A, they produced a model capable of simulating the physiological responses of cells, including the restriction point and the delayed timing of cell division, to the transient inhibition of growth through the inhibition of protein synthesis by cycloheximide. This model followed on from the experimental work of Zetterberg & Larsson (1995) who carried out experiments on cells using growth factors and cycloheximide. In this way, they located the restriction point and measured the kinetics of re-entry to the cell cycle following delay.

5. INCREASING THE COMPLEXITY OF MODELS
Many of the cell-cycle models produced to date tend, for simplicity, to avoid certain issues such as spatial considerations, changing shape and volume and the high levels of complexity involved in signal-transduction pathways. Some of the work in these areas is not necessarily aimed specifically at cell-cycle modelling; however, it seems likely that future cell-cycle models will be required to incorporate these additional complexities if they are to realistically match cell-cycle physiology.

In recent years, there has been significant progress in number of these areas.

5.1. Spatial considerations
The cell cycle culminates with division and Meinhardt & de Boer (2001) used pattern formation, resulting from an activator/depleted substrate mechanism to model the Min protein dynamics and the localization of FtsZ, the protein which, by forming a polymeric ring, determines the cell division plane in the Escherichia coli bacterium. Their model employs simulations based on a set of partial differential equations defining the concentrations of membrane bound and cytoplasmic MinE, MinD and FtsZ, to model the pole-to-pole oscillations of MinE and MinD and the localization of FtsZ. Recently, Yang et al. (2006) developed a
mathematical model which, as well as describing the cell-cycle regulatory proteins on a temporal basis also takes account of protein translocations between the cytoplasm and the nucleus. In this model, which accords well with experiments carried out by Clute & Pines (1999) on cyclin B translocation, they found that while temporal protein regulation is the primary driver of cell-cycle dynamics, including limit cycle and bistable behaviour, these dynamics are significantly modulated by the incorporation of spatial regulation in the model. Importantly, their model offers a link between cell growth and division based on protein translocation, thus removing any need for phenomenological cell growth-dependent parameters to be included in the model.

5.2. Changing shape and volume

Morgan et al. (2004) addressed the problem of the changing shape and volume of the cell structure occurring during the cycle. They rightly stated that, other than early work done in the 1970s, models up to that time had generally been based on a constant volume reactor operating under steady-state conditions, conditions that were quite contrary to reality, particularly when growth was integral to the model. In their formulation, cell volume is modelled as two separate components, namely the cytoplasm and the membrane. The two volumes change at different rates, a situation that imposes periodic or oscillatory behaviour on all components within the cell. While this would obviate the need for the cell biochemistry to generate oscillations itself, it is nevertheless possible that the volume changes themselves result from the underlying biochemistry.

5.3. Signal-transduction pathways

Kholodenko (2003) considered this problem in his review of the roles of diffusion, endocytosis and molecular motors in the MAPK signal-transduction pathways. In support of previous results (Kholodenko et al. 2000a,b; Kholodenko 2002), he confirmed that membrane recruitment of specific cytoplasmic proteins could enhance receptor-induced activation of a membrane anchored target such as Ras by a factor of 1000. By analysing the spatial gradient of phosphor-proteins from the membrane inward through the cytosol, Kholodenko concluded that diffusion in the cytoplasm could not be the sole agent responsible for propagating a signal from the membrane to the nucleus. In this regard, Kholodenko concluded that endocytosis, scaffolding, molecular motors and travelling waves of phosphor-proteins may all be involved in the propagation of signals to different cell locations and that simple diffusion alone has a very limited role. If this is correct, it seems that the complexity of cell-cycle models will increase substantially in the future.

6. WHAT HAS MODELLING DELIVERED?

A number of authors have presented an overview of the modelling process, some more optimistic than others. Tyson (1999) puts a strong case for mathematical modelling, setting out the view that the cell cycle should be modelled as a series of steady states (representing the checkpoint controls) and linking cell growth to the cycle. In addition to describing mathematical modelling methods, Tyson also sets out the differing arguments for modelling the cell cycle including the limit cycle oscillator. Most importantly, Tyson stresses the need for researchers, as well as producing mechanistic data, to give due recognition to the dynamics of the system. A slightly pessimistic view of cell-cycle modelling is offered by Ingolia & Murray (2004). While identifying some of the successes, including the modelling of oscillatory behaviour and the discovery of bistability and hysteresis, the review concludes that the modelling process has added little or no understanding to the biological processes of the cell cycle. The review suggests that models need to stimulate rather than simulate experiment and points to the fact that, since mathematical models to date are intended to be interpreted qualitatively they can never be proved right or wrong at least in quantitative terms. While this may be true in general, there have nevertheless been some instances where the predictions of models have been tested. In addition to Cross et al. (as mentioned earlier), Sha et al. (2003) carried out experiments with Xenopus laevis egg extracts to test the predictions of two separate models. First, they tested the Novak Tyson model (Novak & Tyson 1993) that predicts that the cell cycle is in the form of a hysteresis loop governed by positive feedback loops between cyclin B : Cdc2 and Cdc25 and between cyclin B : Cdc2 and the inhibitors Myt1 and Wee1, together with a negative interaction between cyclin B : Cdc2 and Fizzy, a protein which activates cyclin B degradation. Second, as an alternative to this they considered the Goldbeter (1991) model where the sharp switching behaviour of cyclin B : Cdc2, which heralds entry into mitosis, is explained by a delayed negative feedback loop involving Cdc2 and Fizzy without the involvement of Cdc25 and Wee1. The experiments sought to confirm which of these models, if any, was correct and found that the levels of cyclin B displayed in various experiments was best predicted by the hysteresis loop model. Pomerening et al. (2003) carried out similar experiments and concluded that the bistable positive feedback system allied to the negative feedback loop demonstrated self-sustaining undamped spike-like oscillations.

Models have not yet reached a level of accuracy and completeness required to engage effectively with clinical research relevant to diseases of the cell cycle. The real success of cell-cycle modelling to date is rather in the methodology which has been developed as part of the modelling process, and which is now available for the future construction of full predictive models. In addition to the use of graphical networks, differential equations and bifurcation theory, there are the novel functional forms describing the regulation of cell cycle protein interactions. One of the most useful functions for mathematical modelling of biological systems was established by Goldbeter & Koshland (1981, 1984). Here, the authors were able to demonstrate that...
Role of modelling in identifying drug targets

Csikasz-Nagy

The role of modelling in identifying drug targets is to depend on the total time delay of the CFN. View is that genetic mutations are connected to cell cations that govern the events of the cell cycle. Their regulation dynamics through a succession of bifurcations, how monotonically increasing mass drives the cell budding yeast, fission yeast, frog eggs and mammals, showed, using parameters specific to the cells of a generic model of eukaryotic cell-cycle regulation and equilibrium (Kobayashi 2006), they proved that a MTN has no stable equilibrium other than periodic oscillations that depend on the total time delay of the CFN.

The methodology continues to develop and recently Csinkasz-Nagy et al. (2006) carried out an analysis using a generic model of eukaryotic cell-cycle regulation and showed, using parameters specific to the cells of budding yeast, fission yeast, frog eggs and mammals, how monotonically increasing mass drives the cell regulation dynamics through a succession of bifurcations that govern the events of the cell cycle. Their view is that genetic mutations are connected to cell phenotypes through bifurcation diagrams. On this basis, and using one and two parameter bifurcations, they show how their model can be used to explore the range of phenotypes, which can result from variation of gene expression from a deleted mutant through to some high level of overexpression.

Whether the success in methodological development can be capitalized on to build increasingly relevant and biologically useful models will depend as much on cell biology as on modelling. Provided experimental biology is strongly enough linked to theory to provide relevant and quantitative data, the theoretical methodologies are largely in place to develop models with real clinical applications.

7. LINKS TO BIOLOGY

Any cell-cycle model must be founded on the underlying biology, and in particular the various scientific papers covering the experimental cell biology that has been carried out in relation to those proteins involved in cell-cycle regulation. The scientific literature on the subject is vast, complex, subject to continual review and, at times, ambiguous. However, a degree of consensus does exist, and an initial basis for mathematical modelling can be found by referring to a generally accepted paper such as Kohn (1999) and updating it where necessary. Kohn’s paper, the intention of which is to describe protein interactions relevant to the cell cycle and relevant to DNA repair in a diagrammatic fashion, is essentially a review that selects from the vast library of biological data and then imposes structure on the data selected. Since Kohn’s original publication, as new or revised information has become available, updates and additions are clearly necessary and further reviews have become available. Excellent examples specifically useful to cell-cycle modelling include Aguda (2001) who described the pathways involved in the G1-phase of the cell cycle including activation of Cdk4/6 and activation of Cdk2 via the E2F: Rb pathway, Sears & Nevins (2002) who summarize the role of the Ras pathway in E2F and Myc transcription, and Liang & Slingerland (2003) who provide a detailed description of the role of the PI-3K pathway in cell-cycle progression. This pathway, it seems, is essential for the degradation of P21 and P27 prior to the G1/M transition via upregulation of SKP2. It also seems to have a role in the G2/M checkpoint control system by changing the transcriptional regulation of Gadd45a and modulation of Chk1 activity.

For a mathematical model of the complete cell cycle to hold credibility with the biology community, it should be of sufficient complexity to incorporate a minimum number of processes known to be involved in cell-cycle regulation, including growth and division, growth restriction, survival, programmed cell death, DNA checkpoint control and cellular damage response. To model these behaviours, it is necessary to include in the model those proteins related to the Ras and Akt pathways, proteins involved in the activation of the Myc and E2F transcription factors and all of the cyclins, Cdns, inhibitors and phosphatases involved in each of the four main phases of the cell cycle. Also to be

included are those proteins involved in survival via the PI-3K pathway, growth restriction via the Mad pathway and proteins involved in apoptosis resulting from both mitogenic and intra-cellular processes. Further, those proteins involved in DNA checkpoint and cellular damage control are required. This necessitates a minimum level model including some 150 proteins to demonstrate the regulatory processes. Much of the biological knowledge defining the interaction of these proteins is summarized in the papers referred to above but, from a modelling perspective, the knowledge is representative rather than definitive. This arises from the nature of experimental cell biology research. In relation to the various regulatory pathways to be included in any realistic model, experiments have been, and continue to be, carried out under a variety of conditions throughout the world, on many species including insects, fish, reptiles, rodents and humans, in vivo and in vitro, on many different cell lines, on different tissues and on stem, embryonic and somatic cells. Thousands of different cell types are used and while this vast range of experimentation contributes to a greater understanding of the cell-regulatory mechanisms, many of which are conserved, it is nevertheless the case that in respect of definitive modelling progress is currently restricted. While many processes are conserved among cell types, there is little relationship between the experiments for modelling purposes, since any quantitative data arising can only be considered relevant to one specific cell type. Thus, while a vast number of papers are available in relation to cell-cycle regulation, the molecular basis for whole cycle modelling remains representative.

Regardless of the particular papers used as a basis for a model, however, it must be recognized that these biological papers rarely present information in a form suitable for mathematical modelling and, when modelling is intended, it is usual to refashion the information graphically in a form as frequently used, for example, by Hatzimanikatis et al. (1999), Tyson & Novak (2001), Tyson et al. (2003) and others referred to above. The next point to be considered is the nature of the data that can be extracted from biological papers. While qualitative data are easily available, the quantitative time-course data that would enable speedy calibration of cell-cycle models is scarce and, even if available, rarely published in the literature. The practical difficulties in obtaining absolute protein and protein complex concentrations over a time-course linked to cell-cycle physiology is one of the reasons for this. Also, the fact that this type of data is generally not required to establish conclusions from experiments means that quantification is rarely carried out. Exceptions do exist, however, and Arooz et al. (2000) and Tomasoni et al. (2003) have identified time-courses and protein concentrations that could be useful in modelling. Most common, however, are the type of experiments that seek to define relationships between specific molecular species where the results are presented in a graphical and qualitative sense using various blot analyses to support the conclusions. Taken individually, these experiments do not assist greatly in the identification of quantitative data for mathematical models.

While image analysis may be employed to approximate the relative quantity changes of individual proteins over a time-course, the experimental methods usually employed do not permit quantitative comparison between different proteins.

8. CURRENT POSITION AND THE FUTURE

From a historical viewpoint, experimental research into cell-cycle biology has not evolved according to any selection process in favour of a requirement for the production of mathematical models. While the extent of qualitative data is vast, the description of regulatory mechanisms is still representative and quantitative data is not available to any significant degree. Mathematical modelling of the cell cycle has responded to developments in biology, but the overall development of useful and applicable models has been restricted by the non-specificity of biological research allied to an absence of quantitative data. Undoubtedly, this could be redressed by better integration of theoretical and experimental research methodologies. While considerable progress has been made in modelling methodology, and the discoveries emanating from the field of experimental biology are nothing short of remarkable, it is nevertheless the fact that, as things stand at present, there are very few quantitative models available, and none that fully describe the cell’s molecular activity in a comprehensive way. Perhaps, the closest to this is the colon cancer model, produced by Gene Network Sciences of Ithaca New York, which includes some 2000 variables representing the activities of over 500 genes and proteins. This work involves a major input of data from literally thousands of sources and, being specific to a single cell line, is likely to be the first fully definitive model to include a level of complexity that matches the relevant biology. It seems that we are now at a critical point where progress in experimental cell biology research, allied to the availability of substantial computer power, means that we have reached a position where we can produce mathematical models that represent the underlying biology sufficiently to make a contribution to understanding that is at least comparable to experimentation on its own.

Mathematical models of the cell cycle have the potential to assist research into potential drug treatments for certain types of cancers. The dynamics of the networks involved is complex and highly nonlinear, and intuition alone is insufficient for a full understanding of how the various proteins interact. Accurate models depicting the protein-level dynamics of specific types of cancer cell can be used in simulation experiments to investigate the effects of potential drugs, or perhaps more importantly, combinations of drugs, on many of the proteins regulating the cell cycle. This procedure will assist in determining the best drug regimes to treat specific cancers by running parallel simulations of cancer and non-cancer cells of similar tissue type to search for ways of inducing maximum toxicity in cancer cells and minimum toxicity in non-cancer cells. Such models, however, would have to be both comprehensive and quantitatively accurate; a position which current models have not yet achieved.
Role of modelling in identifying drug targets  R. G. Clyde and others

One of the main criticisms of increasingly comprehensive quantitative models is that the demand on data is prohibitive given the current evidence of the contribution that modelling can make to the science. There would seem to be no logical end to the need to increase in complexity. As we have already discussed, there is some merit to these criticisms. However, recent studies of the behaviour of complex networks suggests that the introduction of additional complexity to models might be an essential step to reducing the demands on data, while also making models more useful. Recent work on studying the robustness of complex nonlinear networks strives to find the link between network topology and macroscopic dynamical behaviour (Albert & Barabasi 2002; Boccaletti et al. 2006). Depending on the nature of the topology of these networks, the dynamics can be surprisingly insensitive to the details of the links and the functions relating the behaviour of one node to another (Barabasi & Oltvai 2004). In the context of cell networks, this would mean that in these regimes, it may be sufficient to have qualitative constraints on couplings between proteins, and that once a certain level of complexity is incorporated, ignorance about the existence of couplings and feedbacks may have little impact on the dynamics. This would suggest an optimal level of complexity that should be incorporated in comprehensive models, and therefore that there is a logical end to the need to increase model complexity. Furthermore, early results from the study of abstract networks suggest that major modifications to the macroscopic behaviour, as might be the goal of therapies, can only be achieved by targeting several key points in the network simultaneously. This is certainly consistent with the prevailing wisdom in relation to cancer therapies. Cancer cells frequently develop mechanisms of resistance to drugs that target single pathways and to overcome this multi-drug regimes are necessary. Only by modelling the system can we determine the relevant targets. There is good evidence that cell networks have the properties suggested by these abstract models (e.g. Resendis-Antonio et al. 2005) and therefore that progress can be made (Meir et al. 2002). This suggests that an important priority for future research would be to incorporate real network topologies in the current abstract dynamical models (Boccaletti et al. 2006).

The development of useful mathematical models is dependent on greater cooperation between cell biologists and theoreticians. This has been argued for many years without much sign of real change and it seems that significant progress will only occur once cell biology begins to deliver applications in the way that other sciences already do. Perhaps, the most important barrier to progress is the lack of a unifying conceptual framework in which the mix of experimentalists and theoreticians can operate. Setting up such a framework, along with the necessary protocols, would seem to be an essential prerequisite for full quantitative modelling to develop.

As a first step, it would be useful to know, in quantitative terms, the extent to which the level of proteins involved in the cell cycle are altered by varying anti-cancer drug regimes. All of the science necessary to deliver this knowledge already exists. Many of the regulatory proteins are known and the technology exists to carry out quantitative experiments and derive quantitative data. If these experiments were designed in conjunction with theoreticians, predictive models could certainly be constructed and validated.

It must be recognized, however, that the development of accurate and comprehensive models capable of being used for analytical purposes over a range of different cell types is a substantial task. The level of complexity is high, as indicated in Kholodenko (2003), and likely to increase in the future. It will require considerable resources in the form of teams staffed jointly by theoreticians and biologists and effective interdisciplinary teams take many years to establish.

As described earlier, the development of useful models is not directly coupled to the complexity of the underlying system. A model that represents the full cell cycle, taking account of all spatial considerations, fluid dynamics, electrical activity and every gene involved in the regulatory network is not possible, necessary or even desirable. The defining step of the worth of any model is that, through the minimal inclusion of detail required, a predictive link may be established between a known intervention process, most likely a potential drug regime, and a change in the regulatory network leading to a particular phenotype. Indeed, there may already be sufficient biological knowledge available to construct such a model. There exists a considerable bank of data and information relating to different aspects of the cell cycle and a scheme able to interrelate this information into an interdependent, dynamic model structure may be sufficient to link intervention process and phenotype. Although the majority of experimental results are published in qualitative form, this belies the fact that quantitative data is nevertheless likely to be available. It may require image analysis together with mathematical techniques involving inequalities and mathematical analysis to derive this data but it does, nevertheless, exist and could be used to model the molecular concentration dynamics of the relevant proteins. As simultaneous measurement of a large number of protein concentrations over time are extremely rare, and non-simultaneous measurements cannot be quantitatively compared, the data will be largely semi-quantitative. This means that one can infer the relative magnitude of change in the concentration of a particular protein over time, but not the absolute concentrations, nor the relative concentration of proteins that have not been simultaneously measured. Models must therefore be capable of accommodating these shortcomings.

To derive and organize this data is still a substantial task, but it is smaller than the task of repeating a large number of experiments in order to derive numerical data. There will always remain gaps and therefore future modelling must accommodate uncertainty and ignorance. Therefore, models will have to be optimized against data to infer the properties and associated parameters in regions of the network where data is missing. To this end, it is possible to employ evolutionary algorithms for parameter searching on a defined network. There is a vast knowledge base, relevant to
Role of modelling in identifying drug targets  R. G. Clyde and others 625

9. CONCLUSIONS

There is a little doubt that sufficient knowledge exists in terms of experimental results, experimental methods, mathematical methods and applied computer science to enable comprehensive quantitative models of cell-cycle pathways to be constructed. What seems to be currently lacking, however, is a unified approach to the problem and an organizational overlay that could ensure that relevant research is directed efficiently into a structured format suitable for modelling. In this regard, we may be on the verge of making progress that could result in a revolution in understanding and application. The urgent need for medical applications, particularly in regard to certain cancers such as that of the lung and skin, where new approaches are clearly needed, will almost certainly dictate this as a future requirement.

We are grateful to two anonymous referees for their comments that substantially improved an earlier draft of this paper. We also acknowledge support from the Carnegie Trust for the Universities of Scotland and an anonymous local trust.

REFERENCES

Aguda, B. D. 1999 Instabilities in phosphorylation-dephosphorylation cascades and cell cycle checkpoints. Oncogene 18, 2846–2851. (doi:10.1038/sj.onc.1202462)


