A Review of Computational Models of Mammalian Cell Cycle

A. Abroudi, S. Samarasinghe and D. Kulasiri

Integrated Systems Modelling Group, Department of Informatics and Enabling Technologies, Lincoln University, Canterbury, New Zealand

Email: Ali.Abroudi@LincolnUni.ac.nz

Abstract: Cell cycle, which comprises an ordered sequence of phases (G1, S, G2, and M) that leads to growth and division of a cell, is an essential part of life and its malfunction may cause formation of tumors and cancer. Therefore, study of cell cycle system has been the topic of many computational modelling research studies. In this paper, a thorough review of all modelling methods for mammalian cell cycle and corresponding models are presented. Due to its high complexity, mammalian cell cycle system has been less modelled than other organisms, such as yeast. Majority of models have investigated various parts of mammalian cell cycle, with few covering the whole system including all the phases. There are four main modelling types (discrete, deterministic continuous, stochastic continuous, and hybrid) that enable researchers to explore and understand system properties, such as dynamics of key regulators, oscillation behaviors, feedback loops, etc. Discrete models provide an abstract view of the system where different nodes interact with each other based on discrete logic. On the other hand, continuous models usually utilize Ordinary Differential Equations (ODEs) to incorporate continuous dynamics of the system elements (i.e., protein concentrations). The effect of noise in biological systems has been modelled through stochastic models. Hybrid models combine aforementioned modelling methods to overcome limitations of individual methods. The paper covers all the above methods highlighting their strengths and weaknesses and presents some open questions as promising future prospects for modelling cell cycle.

Keywords: Mammalian cell cycle, computational modelling approaches, Ordinary Differential Equations (ODEs)
1. INTRODUCTION

Cell cycle is a precisely regulated process in which cells of living organisms undergo a sophisticated growth and division cycle which eventually leads to the production of two daughter cells. Mammalian cells are in the category of eukaryotic cells that have four phases in each cell cycle. The first phase is G1 in which a cell grows and becomes ready for DNA replication in S phase (Figure 1). The Restriction Point (R) in late G1 is considered as a point of no return because it is here that a cell is committed to either keep continuing cell cycle or exit from it (Berridge, 2012). In the second gap phase (G2), the DNA replication accuracy, environmental condition, and the volume of the cell are checked. M phase is composed of two main parts: Nuclear division and Cell division which are also called Mitosis and Cytokinesis (Cyt), respectively. Mitosis usually has five main sub-phases: Prophase (Pro), ProMetaphase (ProM), Metaphase (Meta), Anaphase (Ana), Telophase (Telo) as demonstrated in Figure 1. During Prophase, Chromatin condensation, centrosome separation, and nuclear envelope breakdown (NEB) are initiated. Next, the spindle assembly is carried out in ProMetaphase right before the chromosome alignment on the spindle plate that happens in Metaphase. Completion of spindle alignment is a signal to start the next sub-phase, Anaphase where all chromosomes are separated and moved toward the opposite spindle poles. Finally, in Telophase, the chromosomes become decondensed and two nuclei appear. Division into two distinct daughter cells is carried out in Cytokinesis phase. At the end of M phase, each of the new daughter cells has one of the newly created nuclei and its function is similar to that of the parent cell (Morgan, 2007). Although the sequence of cell cycle events has been known for decades, the underlying regulatory basis was not clear until the introduction of checkpoint concept by Hartwell and Weinert (1989). They demonstrated that there are some checkpoints which arrest cell cycle progression in the case of an incomplete event or DNA damage. This arrest provides enough time for completing the event or DNA repair process so that the fidelity of the genome is maintained (Novák et al., 2003). The two important checkpoints are G1-S and G2-M that can be seen in Figure 1.

Cyclin-Dependent Kinases (Cdks) are the core components of cell cycle control system and regulate the coordination and timing of cell cycle events. Activation of Cdks in cell cycle is mainly due to binding of cyclins. During cell cycle, the levels of Cyclin/Cdks oscillate and these oscillations drive phase transitions. Four different kinds of cyclins have crucial effects on mammalian cell cycle system (Cyclin D, Cyclin E, Cyclin A, Cyclin B). These cyclins, which bind and activate Cdks, stimulate particular cell cycle events. For example, the complex CycA/Cdk2 leads to the phosphorylation of some proteins that begin the DNA Replication in S-phase. Generally, there are three regulatory mechanisms that control Cdk activities (Berridge, 2012) as shown in Figure 2: (a) Cyclins synthesis and degradation (activates and deactivates Cdks - bottom part of Figure 2); (b) The presence of Cdk inhibitors, CKIs (inhibits Cdks - top left hand side of Figure 2); (c) Addition & removal of inhibitory phosphate group by Tyrosine kinases & phosphatases, respectively (inhibition & activation of Cdks - top right hand side of Figure 2). The detailed description of these regulators can be found in (Morgan, 2007).

Following the increase in cell cycle-related experimental data in the last two decades, a variety of computational models aimed at understanding the behavior of this complex system have been developed (Csikász-Nagy, 2009). Majority of these models have focused on yeast cell cycle. Mammalian cell cycle has been less investigated and modelled in comparison to that of other eukaryotes due to its higher complexity. Furthermore, most mammalian cell cycle models have concentrated on just a part of cell cycle, such as G1/S checkpoint. In this paper, we describe modelling approaches to mammalian cell cycle (discrete, continuous,
deterministic, stochastic and hybrid) and pros and cons of each approach in Section 2 and provide conclusions and future prospects in Section 3.

2. COMPUTATIONAL METHODS FOR MAMMALIAN CELL CYCLE MODELLING

2.1. Discrete Modelling

The first category is the group of discrete/logical models that qualitatively describe the system behavior, and mostly correspond to Boolean networks. They provide a high degree of abstraction as well as a fundamental understanding of the dynamical behavior of a system under different conditions. Such models are not only flexible but also easy to fit into cell cycle system. These models introduce the states of species in the form of active/inactive (present/absent) and construct causal relationships between them. In this way, they do not have the problem of parameter estimation that ODE models have and they in particular are useful for large scale networks. A main disadvantage of Boolean modelling is that it cannot describe intermediate states such as slow or small state changes and it can easily generate spurious results (Albert & Wang, 2009). Most Boolean models of cell cycle were developed for yeast and not many for mammalian cell cycle. In one of the most referenced Boolean models of mammalian cell cycle, Fauré et al. (2006) constructed a discrete version of an ODE model in synchronous as well as asynchronous updating schemes. They assessed the corresponding merits and limitations of these two schemes in determining the regulatory network behavior. A logical simulation software named “GINsim” was also developed by this group. However, this model does not take the important effect of some cell cycle regulatory proteins such as Cdc25, Wee1, and p16 into consideration.

Generally, for construction of a discrete model, three main steps should be followed: (1) constructing network graph; (2) determining logical nodes; (3) defining updating schemes. Logical operators, such as AND, OR, and NOT, are used to define the updating schemes. An example of a discrete graph for a small part of mammalian cell cycle, which shows the regulation of Cyclin B controlled by two ubiquitin ligases (APC/Cdc20 & APC/Cdh1), is illustrated in Figure 3 (Morgan, 2007). The logical rule for updating the activity of CycB is illustrated in Table 1. According to Figure 3 and Table 1, CycB node becomes active (state = 1) if both APC/Cdc20 AND APC/Cdh1 are absent or inactive. These ubiquitin ligases may have some effects on other elements of cell cycle, but just the effect on CycB has been highlighted to illustrate the Boolean concept. A summary of discrete models of mammalian cell cycle highlighting references and coverage (aspect of the cell cycle studied) is in Table 2.

Table 1. The updating rule for CycB activation

<table>
<thead>
<tr>
<th>The updating rule for CycB</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>CycB: APC/Cdc20 ∧ APC/Cdh1</td>
<td>To become active, CycB needs absence of two ubiquitin ligases, APC/Cdc20 AND APC/Cdh1 (Morgan, 2007).</td>
</tr>
</tbody>
</table>

Table 2. A summary of discrete models of mammalian cell cycle.

<table>
<thead>
<tr>
<th>Paper (Reference)</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huang &amp; Traub, 2000</td>
<td>G1-S-G2-M</td>
</tr>
<tr>
<td>Fauré et al., 2006</td>
<td>G1-S-G2-M</td>
</tr>
<tr>
<td>Sahin et al., 2009</td>
<td>G1-S</td>
</tr>
</tbody>
</table>

2.2. Deterministic Continuous Modelling

As discrete models are only able to present system variables discretely, continuous models have become more popular in the field of cell cycle modelling. The main approach here has been mathematical models based on Ordinary Differential Equations (ODEs) that deterministically track the exact concentration of species via continuous representation of system dynamics using a set of ODEs (Füb et al., 2005). In general, the rate of chemical reaction \( m \) in a dynamic system is expressed as a function of species \( (y_1, y_2, ..., y_k) \) that affect \( m \) and can be presented as \( r_m \):

\[
r_m = g_m(y_1, y_2, ..., y_k)
\]  

(1)

Different kinetic laws are used to mathematically formulate \( r_m \). Mass Action law and Michaelis-Menten are two common kinetic laws used in mathematical models (Sauro, 2012). Mass Action Kinetic is utilized for elementary reactions with single transition step (Eq. 2) and describes the rate of reaction as in Eq. (3); therefore, change in concentration of reactants and products can be calculated by Eq. (4).

\[
dD + eE \rightarrow pP
\]  

(2)

\[
r = k \times [D]^d \times [E]^e
\]  

(3)

Figure 3. An example of a discrete Boolean graph for CycB regulation. Normal and oval arrows indicate activation and inhibition, respectively.
where $D$ and $E$ are reactants and $P$ is product; $d$, $e$, and $p$ are stoichiometric coefficients; and $k$ is rate constant (Sauro, 2012). On the other hand, Michaelis-Menten law is used when the reaction is enzymatic. The typical reaction of this type can be shown as follows:

$$S + E \xrightleftharpoons{k_s} ES \rightarrow P + E$$  \hspace{1cm} (5)

where $S$, $E$, $ES$, and $P$ are substrate, enzyme, enzyme-substrate compound, and product, respectively. $k_s$, $k_p$, and $k_c$ are rate constants. There are two assumptions to write the Michaelis-Menten kinetic: $ES$ concentration is almost constant; and the total concentration of enzyme (bounded and unbounded) ($E_{\text{total}}$) is constant. Therefore, Michaelis-Menten equation formulates the rate of enzymatic reaction or change in concentration of $P$ as hyperbolic with respect to $S$, as follows:

$$\frac{d[P]}{dt} = r_{max} \frac{[S]}{K_M + [S]}$$  \hspace{1cm} (6)

where $r_{max} = k_c x [E_{\text{total}}]$ is the maximum reaction rate and $K_M = \frac{k_p}{k_s}$ is Michaelis-Menten constant. The detailed description of other kinetic laws (such as Hill Function) can be found in (Sauro, 2012).

Due to data availability, the last decade of the 20th century has seen much growth in computational modelling of cell cycle. The first mathematical model of cell cycle was on “Modelling the cell division cycle: cdc2 and cyclin interactions” (Tyson, 1991). Comprising six species, this model could qualitatively represent cell cycle dynamics. But their model lacked a strategy to investigate the mechanism of coordination between cell cycle core components, such as Cyclin/Cdk1 and made a good comparison between mammalian and yeast cell cycle. Tyson characterized “M-phase Promoting Factor” (MPF) as the most crucial regulatory element in the cell cycle (MPF corresponds to CycB/Cdk1) which can be regulated by synthesis and degradation of cyclins, dephosphorylated by a phosphatase (Cdc25), and phosphorylated by a kinase (Wee1). Also, Novak and Tyson (Novak & Tyson, 1993) showed that “Hysteresis” exists in the MPF–Cyclin relationship and it was confirmed experimentally ten years later (Sha et al., 2003), demonstrating the value of computational modelling in biological discovery. In 2004, Novak and Tyson developed a mathematical model for the restriction point control (Novak & Tyson, 2004) where they investigated the mutations in some key elements of mammalian cell cycle, like Cyclin E and Retinoblastoma protein (Rb). They also considered the important effect of growth factor on passage through the restriction point by exploring the interactions between Growth factors and cell cycle core components, such as Cyclin/Cdk1 and made a good comparison between mammalian and yeast cell cycle dynamics. But their model lacked a strategy to investigate the mechanism of coordination between cell growth and division.

In 1999, Aguda and Tang published a detailed mathematical model to investigate the kinetic origin of restriction point in mammalian cell cycle (Aguda & Tang, 1999). In the same year, Aguda explored the effect of DNA damage on G2/M checkpoint (Aguda, 1999). Iwamoto et al. who partly utilized Aguda’s models, started to develop more complex models of mammalian cell cycle checkpoints and investigated the effect of DNA damage (following UV-irradiation) on G1/S transition in mammalian cell cycle (Iwamoto et al., 2008). This study inspired Ling et al. to investigate the robustness of G1/S checkpoint in depth. Using Type II Error, they introduced a novel approach to quantify the percentage of damaged cells passing G1/S checkpoint under different system perturbations (Ling et al., 2010). In 2011, Iwamoto et al. added G2/M checkpoint to their previous model in order to make it more comprehensive and meaningful as it then covered almost the whole cell cycle. (Iwamoto et al., 2011). This model was insightful in that it could determine cell fate based on DNA damage strength. One of the most complete models of mammalian cell cycle was constructed by Gauthier and Pohl (2011). It included both molecular and cellular systems from nucleotides and amino acids to proteins (33 cell cycle proteins) and consisted of 387 equations and around 1100 rates represented in elementary mass action kinetics. The authors presented a table to describe the corresponding location, ubiquitinator, inducer, activator, inactivator, and function of each cell cycle protein. This model presented some predictions on both molecular and system levels including the effect of cell growth (Gauthier & Pohl, 2011). A summary of ODE models of mammalian cell cycle including the corresponding reference and coverage is given in Table 3.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(Obeyesekere et al., 1995)</td>
<td>G1-S</td>
<td>(Deinuko et al., 2003)</td>
<td>G1-S</td>
<td>(Tashima et al., 2007)</td>
<td>G2-M</td>
<td>(Iwamoto et al., 2011)</td>
<td>G1-S-G2-M</td>
</tr>
</tbody>
</table>
The drawback of deterministic models such as ODE is that model complexity increases with the number of chemical reactions. Furthermore, it is hard to estimate kinetic parameters with the limited available data and it becomes even harder for large scale systems.

### 2.3. Stochastic Continuous Modelling

The third category of models is based on the fact that the functionality of cell cycle regulatory network is often influenced by noise. Unlike deterministic models, stochastic models consider the random variations in concentration of biological species. They are suitable for cases where concentration levels of the species are lower than that which are applicable to deterministic ones (of the order of many hundreds or above). So, as concentrations become lower, the variability of concentration becomes higher and therefore the random element in chemical interactions becomes significant (Donnet & Robert, 2012). Generally, stochasticity is presented by Gillespie Stochastic Simulation Algorithm (GSSA) (Kar et al., 2009) or Stochastic Langevin Equations (SLA) (Steuer, 2004). The mathematical details of these algorithms can be found in the above references. As stochastic models take the number of molecules into account (instead of concentrations), they suffer from the problem of computational complexity. There are not many stochastic models of mammalian cell cycle (Table 4).

### 2.4. Hybrid Modelling

The last category of models is hybrid type. They integrate strengths of other methods, such as discrete models for rapid changes, continuous models for slow changes, deterministic and stochastic models for predictable and unpredictable behaviors, respectively (Kiehl et al., 2004). Majority of hybrid models have aimed to overcome the limitations of ODE-based models. Some of these limitations are: (1) problem of passing from the number of molecules to concentration levels; (2) difficulty in combining discrete state transitions and continuous dynamics which is particularly needed throughout cell cycle; (3) limited available experimental data for reaction rates, parameters, and real protein concentrations. In recent years, hybrid modelling has become popular and a few cell cycle models, but not many for mammals, have been proposed and proved to be successful due to incorporation of the best features of individual models as well as the capability to represent different time-scales. Singhania et al. (2011) developed a simplified hybrid model emphasizing the vital effect of cyclins in mammalian cell cycle regulation. The authors combined discrete (Boolean), deterministic (ODEs), and stochastic approaches, just three different cyclins (Cyclin E, A, and B) and continuous tracking of cell mass. Here, cyclins were considered as continuous variables but regulators of cyclins as discrete variables which followed Boolean logic with some time delays. However, this model is not completely autonomous and the Boolean part could not update itself based on the current system state. Noël et al. (2013) developed a hybrid model for mammalian cell cycle by combining discrete and continuous methods. They transformed an ODE model developed by Csikász-Nagy et al. (2006) into a piecewise smooth hybrid model with simplified reaction rates using an approximate hybridization scheme. Table 5 presents a summary of Hybrid models of mammalian cell cycle.

### 3. CONCLUSIONS

Mammalian cell cycle is the least studied cell cycle system experimentally or through modelling. The goal in this paper was to review computational modelling approaches for mammalian cell cycle to see how models have facilitated our understanding of it. We briefly explained the mechanism of mammalian cell cycle regulation system by highlighting the important effect of Cyclin/Cdk's as main controllers as well as their...
regulators. The role of checkpoints as the guardians of genome was also highlighted. Then, different modelling approaches (Discrete, Deterministic, Stochastic, and Hybrid) that have been used in cell cycle modelling (fewer for mammalian) were presented with their pros and cons in terms of complexity, knowledge representation, and analysis. There are many open questions that computational models can potentially answer - For example, how does a cell decide between survival and death or between temporary cell cycle arrest and permanent cell cycle arrest (senescence)? Can a cell stop cell cycle even before G1/S checkpoint if there exists a DNA damage? These and many more questions require further experimental as well as computational modelling improvements, especially for mammals. Increased understanding of different aspects of mammalian cell cycle system and its operation as whole system will greatly help future cancer treatments.

REFERENCES


659