A TOOLBOX FOR DISCRETE MODELLING OF CELL SIGNALLING DYNAMICS

Yasmin Z. Paterson[1^], *David Shorthouse*[2^], *Markus W. Pleijzier*[2^], *Nir Piterman*[3], *Claus Bendtsen*[4], *Benjamin A. Hall*[2*] & *Jasmin Fisher* [1,5*]

[1] Department of Biochemistry, University of Cambridge, Cambridge, CB2 1GA, UK

[2] MRC Cancer Unit, University of Cambridge, Cambridge, CB2 0XZ, UK

[3] Department of Informatics, University of Leicester, Leicester, LE1 7RH, UK

[4] Quantitative Biology, Discovery Sciences, IMED Biotech Unit, AstraZeneca, Cambridge, UK

[5] Microsoft Research, Cambridge, CB1 2FB, UK

[^]These authors contributed equally to this work

*Correspondence and request for materials should be addressed to B.A.H (bh418@cam.ac.uk) or J.F. (jf416@cam.ac.uk)

6435 words

ABSTRACT

In an age where the volume of data regarding biological systems exceeds our ability to analyse it, many researchers are looking towards systems biology and computational modelling to help unravel the complexities of gene and protein regulatory networks. In particular, the use of discrete modelling allows generation of signalling networks in the absence of full quantitative descriptions of systems, which are necessary for ordinary differential equation (ODE) models. In order to make such techniques more accessible to mainstream researchers, tools such as the BioModelAnalyzer (BMA) have been developed to provide a user-friendly graphical interface for discrete modelling of biological systems. Here we use the BMA to build a library of discrete target functions of known canonical molecular interactions, translated from ordinary differential equations (ODEs). We then show that these BMA target functions can be used to reconstruct complex networks, which can correctly predict many known genetic perturbations. This new library supports the accessibility ethos behind the creation of BMA, providing a toolbox for the construction of complex cell signalling models without the need for extensive experience in computer programming or mathematical modelling, and allows for construction and simulation of complex biological systems with only small amounts of quantitative data. (199 words)

AUTHOR SUMMARY

Ordinary differential equation (ODE) based models are a popular approach for modelling biological networks. A limitation of ODE models is that they require complete networks and detailed kinetic parameterisation. An alternative is the use of discrete, executable models, in which nodes are assigned discrete value ranges, and the relationship between them defined with simple mathematical operations. One tool for constructing such models is the BioModelAnalyzer (BMA), an open source and publicly available (www.biomodelanalyzer.org) software, aimed to be fully usable by researchers without extensive computational or mathematical experience. A fundamental question for executable models is whether the high level of abstraction substantially reduces expressivity relative to continuous approaches. Here, we present a canonical library of biological signalling motifs, initially defined by Tyson et al (2003), translated for the first time into the BMA. We show that; 1) these motifs are easily and fully translatable from continuous to discrete models, 2) Combining these motifs in a computationally naïve way generates a fully functional and predictive model of the yeast cell cycle.

(169 words)

INTRODUCTION

We are in an era of ever-increasing biological data. With data available from genomic studies, through to metabolomic studies, the size, scale and heterogeneity of the resources available present many triumphs in terms of advancing high-throughput technologies but also many challenges. Despite the enormous multitude of available data, our understanding of how such information encoded in a cell's genome is used to carry out the complex biological interactions found between genes and gene products is still lacking. It is therefore no surprise that a central goal of modern biology in this post-genomic era is to understand the structural and temporal nature of these control networks. Not only would this allow us to translate 'Big Data' into working models of biological systems, but also equip us with a better understanding of biological mechanisms, allowing the exploration of emergent behaviours and consequences of genomic variants, with an aim to develop real-world hypotheses for experimental validation.

If we are to meet these challenges, new tools, techniques and ways of working need to be adopted. Whilst experimental procedures using a traditional reductionist approach, focusing on the study of individual proteins or genes in isolation from other network interactions have proved useful in uncovering specific elemental functions of various cellular mechanisms, many

disease processes continue to elude us. This has fuelled the growth of new lines of scientific inquiry. The wide-ranging, vast improvements in computing power brought about at the beginning of the twenty-first century has led biologists down the path of Systems Biology as a means to organise this biological data more holistically. This strategy therefore seeks to combine traditional biological thinking with more interdisciplinary, integrated, synthetic approaches allowing for larger-scale simulations of complex systems, which could revolutionise biomedical discovery.

Computational modelling therefore presents a powerful and novel approach to combat these challenges. The application of standard mathematical modelling, such as through stochastic or ordinary differential equations (ODEs), have been faithfully reproducing the interplay between genes and proteins in small regulatory networks with relative success. Prominent examples of ODE models include that of bacterial chemotaxis [1], the lactose operon control system in *Escherichia coli* [2] and the process of X chromosome inactivation [3], the cell cycle in yeast [4], and the generation of amyloid fibrils [5].Such models employ complex kinetic equations to describe relationships between proteins or genes over time, and require highly accurate and intensive experimental data for their development as input. The complexity of such equations and experimental data required can provide a lot of dynamical detail however this complexity also begs the question of whether this approach will scale well when constructing much larger, more intricate networks in the future.

Executable modelling on the other hand, which describes biological systems as discrete systems, can provide a much simpler class of models [6]. Such models are immediately executable, allowing for much larger-scale simulations to be produced as well as providing the ability to undergo model checking and other formal verifications with ease [7–10]. One of the oldest and simplest forms of executable network models is based on Boolean states (logical models), where each node of the network represents a single gene or protein which is in one of two states: active/on (1) or inactive/off (0) [11,12]. Abstract models based on this paradigm have proved capable of forecasting dynamic processes. Clear examples include that of the control of segment polarity genes in Drosophila [13] or modelling of the neurotransmitter-signalling pathway between dopamine and glutamate receptors [14]. Yet the activities of cellular networks and signalling pathways are often subtler than this, which has resulted in various extensions being made to this model. One such refinement is *Qualitative Networks* (QNs), which uses discrete variables as opposed to Boolean states, and is able to model a much broader range of interactions by using algebraic target functions [7]. These target functions are composed of simple mathematical operations (e.g. addition, subtraction, division, multiplication) to allow for the generation of models with complex relationships between variables.

The BioModelAnalyzer (BMA) tool is a freely accessible online platform that creates QNs from user's instructions. These instructions are formed using a graphical interface, where different genes or proteins are represented by simple symbols that can be connected by inhibitory or activatory edges

negating the need for extensive experience in computer programming, logical formalisms or mathematical proofs [15]. As a result, the BMA is a highly accessible, unimposing interface that is suitable for experimental biologists, whilst still providing powerful stability checking, simulation and Linear Temporal Logic analysis abilities. Although on the surface BMA may appear to be highly abstract, elaborate biological functions can be robustly modelled such as that of *C. elegans* germline development [16], mammalian epidermis differentiation [7], gene and protein regulatory networks in chronic myeloid leukaemia (CML) [17] and acute myeloid leukaemia (AML) [18]. In the case of CML, a novel therapeutic strategy using an Imatinib and pan-Bcl2 family gene inhibitor combination has been identified, highlighting BMAs ability to work on either a hypothesis-creation or hypothesis-testing basis. Cell line specific differences in the PIM pathway were identified in the case of AML, leading to clinically relevant predictions about resistance and how to overcome it.

Although BMA provides the ability to encode complex dependencies between different genes or proteins via the use of algebraic target functions, this task can still seem quite onerous to many biologists. In 2003, Tyson, Chen and Novak [19] published a review outlining a concise mathematical vocabulary of common cellular interactions and pathways using ODEs. In their article, they identify a number of simple functional motifs, akin to electrical circuits which are found at the base of a variety of key biological processes and can be easily combined in order to model complex regulatory interactions. Here we outline a target function library which translates the ODEs outlined by Tyson et al. [19] into discrete equations encoded within nodes of a BMA model. In

order to investigate whether these target functions are capable of modelling cellular behaviours of greater complexity, we then created a BMA model of eukaryotic cell cycle regulation similar to Tyson et al [19]. *In silico* over-expression and knockouts of combinations of genes and genetic interactions were then carried out to highlight the sensitivity of our model. A key benefit of using discrete, executable modelling is that complex systems can be simulated and analysed, and experimentally testable hypotheses can be generated in the absence of large amounts of quantitative data required for ODE models.

This library of ODE translations to discrete target functions also complements the accessibility ethos behind the creation of the BMA. By providing simple building blocks that can be “plugged” into a set of specific nodes, much time and effort will be saved allowing biologists to construct elaborated valid models of biological phenomena, which can guide and direct hypotheses and ultimately drug treatments.

RESULTS

A Target Function Library Accurately Reproduces Expected Biological Behaviour in Simple Networks

We constructed QN models representing the ten major archetypal regulatory and signalling pathways. Networks were generated within the BMA, and signal/response curves compared to previous publications [19–23] for accuracy. Networks are represented by a series of nodes interconnected via activatory (i.e. generally increasing target node value), and inhibitory (generally decreasing target node value) relationships. Nodes in the system can contain values with a granularity of 5 (a range of 0-4), but are generally easily extrapolated to different system ranges. Full details are included in **Supplementary Table 1**, and all models are available in supplementary data.

1. Linear Response

A system where the signal-response is linear (i.e. an increasing signal gives a proportionally increasing response) can be accurately modelled using the default target function. A node with no specified target function will have its value calculated by:

$$average(activating\ inputs) - average(inhibiting\ inputs)$$

A schematic linear signal-response network, from Tyson et al. [19] and built within the BMA is shown in **Fig 1, A, i &ii**, with signal-response curves from both systems shown in **Fig 1, A**, **iii & iv**.

Fig 1. Comparison of Signal-Response Elements. In this illustration, the rows correspond to **(A)** linear response **(B)** hyperbolic response, **(C)** sigmoidal response, **(D)** perfect adaption, **(E)** mutual inhibition, **(F)** mutual inhibition and **(G)** homeostasis as in Tyson et al. [19] The columns correspond to **(i)** Tyson et al. [19] wiring diagrams, **(ii)** BMA wiring diagram translation, **(iii)** Tyson et al. [19] signal-response curves and **(iv)** BMA equivalent signal-response curves. Each BMA wiring diagram contains a unique set of target functions located within particular nodes of the network which can be found in **Supplementary table 1**. For most cases clear comparison between Tyson et al. [19] wiring diagrams (i) and the corresponding BMA wiring diagrams (ii) can be made. Here like in Tyson et al. [19] S indicates the input Signal and R indicates the output Response with, in our case, letters A-C representing intermediate nodes. The graphs in (iv) are derived from simulation analysis carried out in the BMA. For all cases bar (d- iv) and (g-iv) the signal is altered from 0 through to 4 directly within the S node and the output in node R recorded and subsequently plotted. For cases (d- iv) and (g-iv) a simulation is run with a set signal input of 4 as an example, and the response output from the BMA simulation plotted based on the response per calculation time step. Graphs plotted from the BMA model (iv) can then be compared to ODE counterpart (iii). In (e-iv) and (f-iv) the dashed line represents an unsteady state. In (e-iv) S_{crit}, which is denoted x in our target function (**Supplementary table 1)** represents the signal input where a switch in steady states will occur. The motifs in each case reproduce the bifurcation as expected. Similarly, in (f-

iv) S_{crit1} which is denoted y in our target function and S_{crit2} which is denoted z in our target function also correspond to the switch points in stable states.

2. Hyperbolic Response

We generated four ways to discretely model different hyperbolic functions within the BMA. This function describes a "phosphorylation and dephosphorylation" reaction and is modelled using a three-node wiring diagram shown **in Fig 1, B, ii**. A simple function included in node A results in a hyperbolic response as a result of a linearly increasing input. Node A contains the target function:

$$ceil\left(\left(\frac{3}{2}\right)(var(signal))\right)$$

Where $var(signal)$ represents the signal received by the network. This linear approximation captures the rapid initial growth of the response, whilst the plateau is enforced by the granularity of the variable. Additional modifiers (for details see **Supplementary Table 1**) can be included to change the shape and thresholds of the response. Hyperbolic signal-response curves from Tyson et al. [19] and from within the BMA are shown in **Fig 1, B, iii & iv**.

3. Sigmoidal Response

Sigmoidal response curves represent systems that act in a switch like manner, which are reversible and increase continuously with an increasing input. Schematics for sigmoidal signal-response networks are shown in **Fig 1, C, I & ii**. The target function for Node A contains the function describing the

sigmoidal response. Multiple functions produce differing sigmoidal curves (for details see **Supplementary Table 1**). The simplest, producing a sigmoidal response from a linear signal is:

$$\left(floor\left(\frac{var(signal)}{x}\right)\right)y$$

Where x is the value at which the system switches between high and low values, and y is the upper value for the sigmoidal response. Signal-response curves are shown in **Fig 1, C, iii & iv**.

4. Perfect Adaptation Response

Adaptation is defined as “a process where a system initially responds to a stimulus, but then returns to basal or near-basal levels of activity after some period of time” [24]. Perfect adaptation is further characterised by the final response of the network returning to the exact pre-stimulus level. Perfect adaptation is used in numerous biological systems, for example, the Friedlander and Brenner [25] model of ion channel activation and inactivation. Network schematics for perfect adaptation systems can be seen in **Fig 1, D, i & ii**. Perfect adaptation is modelled with the addition of the following target function to Node C (Signal-response curves can be seen in **Fig 1, D, iii & iv**):

$$\Big(x\big(var(A)\big)\Big) - \Big(x\big(var(B)\big)\Big)$$

Where x represents the maximal height of the response before the system adapts.

5. Mutual Activation Response

Mutual activation behaviour represents irreversible cell switches, i.e. a "point-of-no-return". These discontinuous, one-way switches are typical of cell fate determination. Once a critical signal value (S_{crit}) is reached, the response immediately increases to a high level. A critical feature of mutual activation networks is that the switch is irreversible i.e. if the signal increases beyond S_{crit} and subsequently decreases, the response will not decrease. Network schematics for mutual activation systems are shown in **Fig 1, E, i & ii**. Inclusion of the following target function in Node A results in the signal-response curves shown in **Fig 1, E, iii & iv**:

$$floor\left(\frac{var(signal)}{x}\right) + var(C)$$

Where x is S_{crit} – the value at which the irreversible switch occurs.

6. Mutual Inhibition Signal Response Curve

Mutual inhibition differs from mutual activation in that these systems exhibit hysteresis; if the input decreases below a defined critical value, then the output will return to zero. Tyson et al. [19] describe this type of feedback as a "toggle switch" , where there are two defined critical values; S_{crit1} and S_{crit2}, at which point the response will shift from either upper or lower values to the opposite. This is simplified below:

$$S \geq S_{crit1} \rightarrow R = R_{max}$$

$$S \leq S_{crit2} \rightarrow R = R_{min}$$

$$S_{crit1} < S_{crit2}$$

Essentially, this works similarly to mutual activation, except if S is decreased below Scrit2 then the switch will return to the inactive state. Our model is composed of 6 nodes and is compared to the "traditional" toggle switch schematic in **Fig 1, F, ii**. B is split into two separate nodes representing active and inactive states (B_{active}, and $B_{inactive}$ respectively), and it is the interactions between these 2 states of B that give rise to hysteresis. The target function for the node representing inactive B (Node $B_{inactive}$) in a system with a granularity of 0-4 is:

$$x\left(\left(var(A) + (4 - y)\right) - var(B_{active})\right)$$

Where x represents the maximal response of the network, and y is S_{crit1}. The target function for the node representing active B (Node B_{active}) is:

$$\left(x - var(B_{inactive})\right) + (1 - z)$$

Where x represents the maximal response for the network, and z is S_{crit2}. Additionally, Node C contains the following target function:

$$x(var(A) - var(B_{active})$$

Where x represents the maximal response for the network. Signal-response curves for this network are shown in **Fig 1, E, iii & iv**.

7. Homeostasis

Homeostatic regulation involves a network where the network counteracts the activity of the stimulus such that the response is constrained to a very narrow window (in the case of our network, a single value). A schematic homeostasis network is shown in **Fig 1, F, I & ii**. In this network, the granularity of Node B is adjusted such that it is double the range of the other nodes within the system. This is required for stability of the system, as nodes can only alter by a single integer value each time step, a node of granularity 10 (0-9) will take twice as long to reach a maximal value from 0 as a node with a granularity of 5 (0-4). This temporal difference allows us to eliminate instability in the system caused by oscillatory feedback between Nodes A and B, which is present when both nodes have the same granularity. Within this network two unique target functions are required to exhibit homeostasis, for Node A:

$$var(signal) - floor\left(\frac{2}{3}var(B)\right) + 1$$

And for the node representing the system response (Node Response):

$$ceil\left(\frac{var(A)}{3}\right)$$

Signal-response curves can be seen in **Fig 1, F, iii & iv**.

8. Negative Feedback Oscillations

Negative feedback oscillations result from similar network wiring as homeostasis, with the result being a system where the response oscillates between 0 and the signal value.

A negative feedback oscillation loop is seen in **Fig 2, A**. No change in the default target functions are required to generate an oscillatory output. For these networks, an input of S will result in an oscillation that tends to between S and 0. The temporal constraints of the network however (in that each node can only update by a single integer value each step) results in cases where the oscillation will not reach the maximal value before the inhibitory portion of the network kicks in. The ultimate range of the oscillations can be tailored however, with either the addition of values to the output node (in order to adjust the oscillation range up or down), or by inserting the following formula into node A:

$$(var(signal) + x) - (var(C)) + y)$$

Where the difference between x and y changes the range of the oscillations. Additionally, the temporal properties of the system, specifically how long it takes to perform each loop, can be adjusted by the addition of more nodes to the loop, with a large number of nodes increasing the number of steps required to complete one oscillation. For an example see **Supplementary Figure** 1.

Fig 2. Comparison of Oscillatory Networks. In this illustration, the rows correspond to **(A)** negative feedback, **(B)** activator-inhibitor and **(C)** substrate-depletion oscillators as in Tyson et al. [19] The columns correspond to **(i)** Tyson et al. [19]wiring diagrams, **(ii)** BMA wiring diagram translation, **(iii)** Tyson et al. [19] signal-response curves and **(iv)** BMA equivalent signal-response curves. Each BMA wiring diagram contains a unique set of target functions located within particular nodes of the network which can be found in **Supplementary table 1**. For most cases clear comparison between Tyson et al. [19] wiring diagrams (i) and the corresponding BMA wiring diagrams (ii) can be made. Here like in Tyson et al. [19] S indicates the input Signal and R indicates the output Response with, in our case, letters A-E representing intermediate nodes. The graphs in (iv) are derived from simulation analysis carried out in the BMA. For all cases bar a simulation is run with a set signal input of 2 as an example, and the response output from the BMA simulation plotted based on the response per calculation time step and are thus not directly comparable, however clear oscillatory behaviour can still be observed.

9. Activator-Inhibitor Oscillations

The activator-inhibitor oscillation relationship is characterised by a positive and negative feedback loop within a system (shown in **Fig 2, B, i and ii**). The interactions of the two loops result in a system that oscillates between a maximal and minimal value, called a hysteresis oscillator. Including the following formula in node A results in an oscillation between the maximal and minimal values of the nodes when I = 2:

406

$$\Big(x\big(var(signal) - (var(E) + y)\big)\Big) + var(A)$$

407

408 Where x is 3 value of the nodes, and y is 0. Adjusting these values will
409 change the range of the oscillations (i.e a range of 3 or 2 is obtained by
410 reducing the value of x), and altering the value of y adjusts the start and end
411 points of the oscillation. The signal-response curves for activator-inhibitor
412 networks are shown in **Fig 2, B, iii and iv**).

413

414 **10. Substrate-Depletion Oscillations**

415

416 The substrate depletion oscillation (SDOs) is quite similar to that of negative
417 feedback. However, the number of nodes are reduced to reflect the greater
418 intimacy between enzyme-substrate reactions compared to negative feedback
419 loops. The network schematics for substrate-depletion oscillations are shown
420 in **Fig 2, C, i**. In substrate-depletion oscillations, a small signal produces a
421 small response and a large signal produces a large response. To model this,
422 the following target function is applied to Node A:

423

$$1 + \left(floor\left(\frac{var(signal)}{x}\right)\right) * \left(\left(\frac{y}{2}\right) * var(signal)\right) - \left(\frac{z}{2} * var(B)\right)$$

424

425

426 Where x is the starting point of the oscillations and y and z are the range of
427 the oscillations. Signal-response curves are presented in **Fig 2, C, ii**.

Using the BMA Target Function Library to Construct Complex Networks: Eukaryotic Cell Cycle Control

After translating these motifs that were originally defined with ODEs to BMA models and their target functions, we then sought to determine the robustness of these motifs and their target functions by modelling a complex cellular behaviour: Eukaryotic cell cycle regulation. Based on the wiring diagram presented by Tyson et al. [19], a QN model was constructed using our own BMA target function library (**Fig 3, A**). Clear descriptions of the dynamics of cell cycle regulation can be found in the following review articles: Tyson, Csikasz-Nagy & Novak (2002) [22], Tyson & Novak (2015)[20] and Hochegger, Takeda & Hunt (2008) [26].

Fig 3. Qualitative Network of Eukaryotic Cell Cycle Regulation. (A) BMA Wiring diagram. The network is constructed around a central pool of the major cell cycle regulator cyclin dependent kinase (Cdk1) and its cyclin partner (cycB). This cell cycle transitions are triggered by changes in the Cdk1-cycB activity, which is regulated by a number of different components. CKI a cyclin kinase inhibitor and Wee1 kinase subunit inactive the Cdk1-cycB complex whereas the Cdc25 phosphatase activates the complex. Cdk1-cycB activity can also be destroyed via the Anaphase-promoting complex (APC) in combination with Cdc20, which target cyclin for degradation. The activities of the Cdk1-cycB activity can then be monitored by 3 extracellular markers; G1S, G2M and MG1. **(B)** BMA simulation of Cdk1-cycB activity. The solid black line

indicates the progression of Cdk1-cycB levels through the cycle. Dotted lines and block colours represent distinct phases as determined by the key. The cycle repeats itself if growth conditions remain favourable, as is represented in this simulation.

Pool Module

This module contains the “master molecules” of the cell cycle, that being cyclin-dependent kinases (Cdks) and their cyclin partner, which as the name suggests are required in order to activate the Cdks. Our model is limited to only a single Cdk-cyclin partnership, Cdk1-cycB for simplicity. This module is fuelled by the growth of cyclin levels which we assume can have unlimited binding capacity to Cdk1. Unlike cycB, intracellular Cdk1 concentration does not fluctuate throughout the cell cycle [27], we therefore model Cdk1 as being at a constant level which can accommodate the variations in cycB levels [20].

G1/S Module

The G1/S module features mutual inhibition between Cdk1-cycB and CKI. This feedback loop is described as a “toggle switch” and is modelled using our BMA mutual inhibition target function. Here we model CKI as being present at high levels in G1 by assigning it an initial value of 10 (max based on our granularity choice). The input in this case is labelled as cyclin, which, as it increases causes an increase in bound_Cdk1 (i.e. heterodimer of CKI, Cdk1 & cycB) due to the initial high levels of CKI. As the CKI doesn’t stop cyclin accumulation and binding to Cdk molecules, the rising Cdk1-cycB levels

which are not opposed by CKIs soon tips the balance, phosphorylating the CKI and labelling them for degradation. The values chosen for the switch points can be found in **Supplementary table 2**.

G2/M Module

Following the degradation of CKI and subsequent spike in Cdk1-cycB activity the cycle enters the G2/M module. This module features both mutual activation, between Cdk1-cycB and Cdc25, and mutual inhibition between Cdk1-cycB and Wee1 [19]. The later works in a similar way to that of Ck1-cycB and CKI, with a race occurring being between Cdk1-cycB and Wee1. The Cdk-cycB and Cdc25 mutual activation interaction on the other hand is a type of positive feedback loop, where Cdc25 and Cdk1-cycB activate each other rather than inhibit each other. This is modelled using our BMA mutual inhibition target function combined with the mutual activation target function. Here we model Wee1 as being present at high levels in G2/M by assigning it an initial value of 10. The input in this case comes from the G1_Cdk1 levels, which as it increases causes an increase in phos_Cdk1 (i.e. phosphorylated form of Cdk1-cycB) due to the initial high levels of Wee1. As Wee1 does not stop cyclin accumulation and binding to Cdk molecules, the rising Cdk1-cycB levels (which are not opposed by Wee1) soon tips the balance, phosphorylating the Wee1 and marking them as inactive. Inactive Wee1 maintains active Cdc25, thus decreasing Wee1 results in an increase in Cdc25 and thus the switch like activation of Cdk1-cycB. Again, the values chosen for the switch points can be found in **Supplementary table 2**.

M/G1 Module

Once the Cdk1-cycB reaches a high level due to Cdc25 activation the cell enters mitosis. In order to exit this phase, the Cdk1-cycB activity must be destroyed and CKI levels stockpiled. This transition is aided by the Cdc20:APC complex, which itself is indirectly activated by Cdk1-cycB activity, causing degradation of CycB. This results in a substantial drop in Cdk1-cycB activity, which then allows CKI to rise again. This relationship is described as an oscillator based on a negative feedback loop, where Cdk1-cycB activates APC, which activates Cdc20, which then degrades CycB [19]. In the BMA, the negative feedback oscillators target function uses the default function. Therefore this was modelled simply by considering the whole cycle as a feedback loop by adding an inhibitory edge back to the cyclin B in order to create the desired response **(Supplementary table 2 & 3).**

Comparison to ODE Eukaryotic Cell Cycle Model Predictions

The initial conditions were set so that all nodes remained with an initial value of 0, except for CKI and Wee1 which are given an initial value of 10 (max based on our granularity choice). As Growth, Replicated_DNA, Undamaged_DNA and Aligned_Chromosomes are conditions that can be represented by a binary value, a value of 0 represents the absence of the cell phenotype, whereas a value of 1 corresponds to the presence of the phenotype. The initial values for all four of these phenotypes were therefore set to 1 to represent normal growth conditions. Simulation analysis, starting from this initial state leads to the initiation of a series of network states

(ranging between 0 and 10 based on our granularity). These steps correspond to the biological time series of protein activation and inactivation that occur during the wild-type cell cycle (**Fig 3, B**).

Similar to the Tyson et al. [19] signal-response curve, **Fig 3, B** shows the cell progresses through the cycle via a number of steady states. Firstly, at low levels of Cdk1-cycB activity the cell will remain in G1. With increased growth it will eventually pass a critical point, resulting in the irreversible disappearance of G1. As the cell moves into the S phase the level of Cdk1-cycB continues to grow until it reaches an intermediate level (3 as determined by our target function). Here in the G2 phase the cell will continue to grow until it reaches the next critical threshold, where the G2 state will disappear. This gives rise to a large spike in Cdk1-cycB activity (driving the cell into mitosis) which then decreases as cycB is degraded by APC:Cdc20, signalling cell division and resetting the system for the next round of division. One added benefit of this model is its ability to continuously cycle, as highlighted in **Fig 3, B**.

Simulation of Mutant Phenotypes Replicate Experimental Results Found in the Literature

In order to evaluate the accuracy of our model loss of function (KO) and over-expression (OP) mutations were carried out based on a sample of previous experiments found in the literature (**Table 1**). In our limited subset of mutant experiments 8 out of 9 cases were able to accurately replicate the experimental results found in the literature without making any modifications

553 to the underlying model described above beyond modelling the mutations (**Fig**
554 **4**). For instance in the case of Cdc25 OP, studies in both yeast and mice have
555 shown that over-production of Cdc25 result in premature entry into mitosis
556 due to early activation of Cdk1-cycB [28,29]. In the in silico experiment, the
557 same result can be discerned. Rather than needing 8 steps to pass through
558 G2 the Cdc25 OP, the model only takes one step. Similarly less time is spent
559 in M phase with only 1 step occurring versus 2 steps for wild types (WT). This
560 results in the mutant model undergoing each cycle in fewer calculation steps,
561 needing only 33 steps compared to the 41 needed in the WT model.
562 Descriptions of the other seven successfully reproduced experiments can be
563 found in **Table 1**. For the case concerning the CKI OP mutant, experimental
564 results were not as clearly reproduced. Experimental evidence by Moreno &
565 Nurse [30] showed that overexpression of Rum1, a fission yeast CKI, leads to
566 delays in G1, with repeated S-phase and no M-phase. This is partially
567 replicated in our mutant model, with there being a long delay in G1 phase (23
568 calculation steps compared to 14 in the WT model), as well as no M phase
569 being reached (where Cdk1-cycB hits max value of 10). The model however
570 still runs through the M/G1 phase rather than just repeating the S phase.
571
572 **Table 1:** Mutant simulations reproduce described behaviour from the
573 literature. Summary of experimental results are given in “Expected Outcome”,
574 and in silico results are given in “Model Outcome”.

Genetic Perturbation	Source	Expected Outcome	Model Output
Wee1Δ	Yeast - Nurse 1975[31] Mammal - Tominaga	Premature entry into mitosis, with long G1, short G2, but still viable	G1 same length, Short G2, cell cycles quicker

	2006[32]		
Wee1 OP	Arabidopsis – De Schutter 2007[33]	Cell cycle blocked in G2	Arrest in G2
CKIΔ	Yeast – Lengronne 2002[34]	Short G1, extended G2, increased activation of Cdks	Short G1, extended G2
CKI OP	Yeast – Moreno 1994[30]	In yeast delays G1 followed by multiple S & no M	Long delay in G1, cycles but no M phase
Wee1Δ CKIΔ	Yeast – Sveiczer 2000[35]	Cell divides very quickly, cell gets smaller with each division	Divisions occur over less time steps
Cdc25Δ	Yeast – Russell 1986[28] Mammal – Lee 2009[36]	Cell cycle blocked in G2	Arrest in G2
Wee1Δ Cdc25Δ	Yeast – Davidich 2013[37]	Cell not viable, cannot enter mitosis	Arrest in G2
Cdc25 OP	Yeast – Russell 1986[28] Mammal – Timofeev 2010[29]	Premature entry into mitosis, early activation of Cdk-cyc	Short G2, cell cycles quicker
Cdc20Δ	Yeast – Kim 1998[38] Mammal – Li 2007[39]	Lethal, cannot complete mitosis	Arrest in M

Figure 4. Mutant Phenotype Simulation Analysis. Depicts the temporal evolution of the network following perturbation of particular nodes. Each mutant perturbation can be compared to the wild type, which is listed first. Each distinct cell cycle phase is coloured coded according to the key provided. Each time step corresponds to each calculation step recoded in the BMA simulation which is exported as a CSV file.

DISCUSSION

We present a library of novel Qualitative Network modules that can accurately replicate the biological behaviour of core, ubiquitous network motifs. We generate and compare our library based on biological behaviours defined previously [19], and confirm the modular nature of the library with the generation of a model for the eukaryotic cell cycle produced using motifs from the library. By simulating known genetic perturbations we further test this novel qualitative eukaryotic cell cycle model, highlighting its capacity to accurately replicate many well-known mutant phenotypes without the need for explicit parameterisation, as would generally be needed for ODE models. This study constitutes both a toolbox for biologists to construct elaborate networks with ease, but also an example of its application to a relevant biological system. The QN presented has much wider applications, with our working model having the potential to be adapted in order to provide much more dynamic details on the regulation of these core cell cycle components. Such a model could then be utilised to provide new insights into cell cycle regulation allowing the prediction of novel mutant phenotypes that have not been previously investigated. Not only could this provide a more thorough understanding of the underlying cell cycle regulatory principles, but also assist

612 in the identification of a host of mutants that contribute to cancers or other
613 pathologies, potentially allowing for the generation of novel drug therapeutics
614 [40]. Similarly, in compiling this simple, easy to use BMA target function library
615 we hope to encourage experimentalists to adopt this type of QN modelling as
616 part of mainstream biological research. This would offer a wealth of
617 advantages in terms of consolidating what is known about large networks into
618 concise descriptions, as well as by allowing the generation of novel
619 predictions about systems in the absence of large amounts of data and thus
620 help focus experimental design.
621
622 Through the construction of our BMA motif/target function library we have
623 been able to capture the dynamic behaviours of simple cell signalling
624 pathways. Although networks can be modelled using ODEs, with behaviours
625 being predicted using numerical simulations, this requires more complex and
626 harder-to-obtain biological data and may equally appear mathematically
627 complex to many biologists. As shown through the analysis of a model of
628 eukaryotic cell cycle regulation, a relatively simple QN model can capture
629 many of the advanced dynamic features of ODE models, including
630 multistability and bifurcations. Simulation analysis of the described model
631 shows strong similarities to that of the quantitative biological signal response
632 curve, first proposed by Stern & Nurse [41], which was based on the results of
633 multiple Cdk and cyclin knockout experimental studies. Like our model, they
634 described the cycle as having three distinctive phases of Cdk activity, with the
635 Cdk1-cycB levels transitioning through the cell cycle via different levels or
636 bifurcations [22,41]. These levels or bifurcations are representative of firstly a

637 stage of inactivity (G1) where Cdk activity remains low, secondly a stage of
638 moderate Cdk activity sufficient to trigger S phase, and lastly a stage of high
639 Cdk activity sufficient to initiate mitosis, all of which can be easily recognised
640 in our model simulation [41]. This ability to model varying levels of Cdk activity
641 sets our model apart from its Boolean counterparts, where only two levels of
642 detail (“on” or “off”) can be captured. Through simple manipulation of our
643 target functions we were also able to capture an extra layer of detail, by
644 allowing our model to continue cycling over unceasing divisions when
645 conditions remain favourable, a behaviour which has not always been
646 replicated in previous studies [19,37,42]. The addition of this extra layer of
647 complexity, showing sustained cell cycle oscillations, results in a model that is
648 more representative of the true clock-like oscillatory nature of the cell cycle
649 [43]. It is worth noting, however, that our model does not contain continuous,
650 biologically measurable values for components, and as such is limited in its
651 ability to interpret continuous experimental data.
652
653 As a means to further validate the model, loss of function and overexpression
654 mutants were simulated, with the simplicity and generality of the model limiting
655 the number of mutant phenotypes studies. Regardless of the simplification of
656 using discrete modelling to represent continuous protein concentrations and
657 interactions, the BMA model was capable of correctly modelling 8 out of 9
658 mutant phenotypes studied. All knockout mutations were correctly
659 reproduced, with the model capturing dynamic properties such as phase
660 length changes. For over-expression models, where the corresponding node
661 range is set to max-max, 90% of the OP mutants studied corresponded

accurately with experimental data, with CKI OP mutants being in partial agreement. This partial agreement is likely due to the minimalistic nature of our model, and could likely be overcome by using additional nodes to model the CKI interaction in more detail. Overall, the model produced using the BMA target function library accurately represents not only the WT regulation patterns of the general cell cycle control engine, but also the dynamic changes resulting from a number of mutants. This showcases our BMA target function library's ability to be easily manipulated in order to model complex networks. Of particular note is the ability of the method to accurately generate protein behaviour through the simple addition of target functions from different modules that act on the same proteins, as is the case with the Cdk1-CycB node in our QN. This ability to draw together simple motifs to create realistic and useful biological networks demonstrates the validity of the approach and the opportunities that executable modelling makes available.

686

687

688

689

690 # METHODS

691 ## ***Qualitative Networks***

692

693 Qualitative networks (QNs) are an extension of Boolean models. In Boolean
694 Networks, nodes are able to be in either and active (1), or inactive state (0),
695 and are connected via functions that describe the mathematical relationship
696 between them in an abstract way. Boolean Networks can be synchronous or
697 asynchronous, that is – they may update every node simultaneously when a
698 change is introduced in the system, or they can update in sequence from a
699 propagation point. Qualitative networks are analogous to a synchronous
700 Boolean Network, except that nodes are able to vary over a wide range of
701 discrete values (called a granularity). Simple networks may be represented as
702 Boolean, but Qualitative Networks may involve nodes with a greater range of
703 values. For example, a node may have a range of 0-2 (granularity 3), where a
704 value of 1 represents “normal activity” of an enzyme or gene product, and 0
705 and 2 represent low and high values respectively. This can be extended for
706 much larger granularities, for example 0-10, where 10 represents maximal
707 activity, and 0 represents minimal activity, with each discrete value in between
708 representing a different concentration.

709

Nodes within a Qualitative Network are associated with either activatory or inhibitory relationships. Activatory relationships generally result in a response being high when a stimulus is high, and inhibitory relationships result in a response being low when a stimulus is high. Relationships between nodes are controlled by simple mathematical functions that describe the value that a node should represent, given its current inputs, and this function is called a target function. The values of nodes within a Qualitative Network are updated simultaneously when the network is simulated, and nodes will change their values in order to reach their target function gradually – changing value by only one each step. Due to the synchronous and defined nature of Qualitative Networks, they are deterministic, and susceptible to formal verification techniques. QN's can stabilize and reach a single self-perpetuating state (called a stable point), but can also give rise to cycles and oscillations.

Models and motifs described in this document are available in supplementary information and at https://github.com/shorthouse-mrc/biomodelanalyzer_targetfunctionlibrary.

The BioModelAnalyzer (BMA) Platform

The BMA is an accessible, publicly available (www.biomodelanalyzer.org) graphical tool for discrete modelling and analysis of Qualitative Networks. The platform, with its user-friendly graphical interface, uses visual notations familiar to specialists in biology. BMA models are constructed on a gridded canvas upon which one or more cells, and cell elements (i.e. membrane receptor, cellular proteins etc.) can be placed and connected together with activatory or inhibitory links. To create a model, the user starts by dragging

and dropping a cell onto the gridded canvas. These cells have no functional role in the analysis, being purely a visual aid to assist model design clarity. Cell elements are then placed in or outside of these cells, which can represent internal proteins, external proteins or membrane bound receptors. Connections between these cell elements can then be made using activatory arrows or inhibitory bar-arrows. Each cell element can then be labelled accordingly, using the simple drop down menus and a finite value range assigned, with the BMA default being [0,1], or Boolean. This range may be altered to add different levels of concentration, for example a range of [0,2] may represent “low”, “normal” and “high” concentrations of a protein or gene. If the user does not specify a target function for a node, then the BMA assigns a default target function. The default target function assigned within the BMA is described as:

$$average(activating\ inputs) - average(inhibiting\ inputs)$$

More complex target functions can be inserted for each node manually using an autocomplete function simplifying the use of correct syntax when referencing variables or using operators.

This underlying QN can then be analysed using simulation, stability analysis or Linear Temporal Logic tools each of which is accessible using the graphical interface. Simulation analysis shows the step-by-step execution of the model starting from a set point, based on either initial values specified by the user or a randomised start point. A graphical representation of all node values as they

update over a user-defined number of time steps is produced, as well as a table of the simulation progression values, which can be exported as a CSV file for further analysis. Stability analysis can be used to test general properties of the model. If a model, given all possible starting conformations, will always result in a same self-perpetuating state, it is considered stable, and the graphical interface presents the user with the “stable values”. If stability is not achieved, however, the interface presents whether the system results in bifurcations (can potentially end in multiple states depending on the starting conformation) or oscillations (results in an infinite cycle). More advanced queries can be asked using the Linear Temporal Logic (LTL) interface, which allows the user to define simple or complex temporal logic queries with a drag and drop interface. LTL queries will return True, True sometimes, False, and False sometimes responses to queries, and the interface allows the user to see examples of systems where the behaviour occurs.

Cell Cycle Model Generation

The model was composed of 3 main modules; G1/S, G2/M and M/G1 linked to a central node representing the level of Cdk1-cycB activity throughout the cycle. Each module was represented by a different cell in the BMA and labelled accordingly. The modules themselves were comprised of 6 key components namely; Cyclin, CKI, Wee1, Cdc25, APC and Cdc20, which regulate this Cdk1-cycB activity and thus the different cell cycle transitions (**Table 2**). These 6 components, modelled in their different chemical states (phosphorylated, active, inactive etc.) thus comprise a 20 node network,

including 4 cell behaviours and 3 descriptive nodes linked by 28 interactions **(Supplementary Table 2 & 3).** Three members of the BMA target function library were combined to create the cell cycle model, with the granularity set to 11 (A range of 0-10). This granularity, which differs from the default of 5 in our target function library, was chosen to accommodate the varying levels of Cdk1-cycB activity required, and to allow for clearer analysis of mutant phenotypes. Modules were initially generated based on the wiring and target functions from the BMA target function library examples, which were linked together through appropriate nodes **(Supplementary Figure 2).** This method resulted in the creation of individual pools of Cdk1-cycB activity at the different cell cycle phases that fed into one central pool of Cdk1-cycB activity. To better represent the biological system, the model was then refined, by simply combining the Cdk1-cycB individual pool target functions into a single node via compound addition of each target function within the target function interface. To allow for multiple rounds of cell division, rather than the simulation of a single cell cycle, modification to the mutual activation target function was required. The mutual activation target function defines a one-way switch, and as such is not reversible. Here only nodes S and A (see figure 1, e, ii) and their associated target functions were used, thus allowing the cell to return from the high state achieved following the critical switch point activation.

Table 2: Cross-species nomenclature of key nodes within each module

Module	Target Function	Node	Mammalian Cells	*Xenopus* embryo	Fission Yeast	Budding Yeast	Function
G1/S	Mutual Inhibition	CKI	$p27^{Kip1}$	Xic1	Rum1	Sic1	Stoichiometric cyclin-dependent kinase inhibitor
G2/M	Mutual Inhibition	Wee1	hWee1	Xwee1	Wee1	Swe1	Inhibitory kinase that inactivate Cdk-cyclin dimer
	Mutual Activation	Cdc25	Cdc25C	Xcdc25	Cdc25	Mih1	Activatory phosphatase that activate Cdk-cyclin dimer
M/G1	Negative Feedback Oscillator	APC	APC	APC	APC	APC	Anaphase-promoting complex
		Cdc20	$p55^{Cdc}$	Fizzy	Slp1	Cdc20	Degrades cyclin in complex with APC

Knock-Out (KO) & Overexpression (OP) Analysis

In order to show if a model can faithfully reproduce known biological perturbations, loss of function and gain of function mutations can be analysed in BMA. A list of genetic perturbations curated from the literature is created and used to test the model. In the case of KO mutations, the corresponding node within the model range was set to a range of 0-0, corresponding to a permanently inactive state. OP mutations were simulated by setting the corresponding node range to max-max, (i.e. max based on the chosen granularity) simulating a permanently active state. Simulation analysis is then

carried out, and the results compared to the wild-type simulation. Differences were then compared to known biological behaviours.

ACKNOWLEDGEMENTS

We thank the Fisher and Hall groups for useful discussions. Work was supported by Medical Research Council core funding (D.S.) and Royal Society grant UF130039 (B.A.H).

AUTHOR CONTRIBUTIONS

BH, CB, NP and JF conceived the study. DS and MP generated the target function library, YP developed the model of the cell cycle. DS and YP co-wrote the manuscript. All authors were responsible for editing of the manuscript.

COMPETING FINANCIAL INTEREST

The authors declare no competing financial interest.

REFERENCES

1. Yi TM, Huang Y, Simon MI, Doyle J. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. Proc Natl Acad Sci. 2000;97: 4649–53. doi:10.1073/pnas.97.9.4649
2. Esmaeili A, Davison T, Wu A, Alcantara J, Jacob C. PROKARYO: an illustrative and interactive computational model of the lactose operon in the bacterium Escherichia coli. BMC Bioinformatics. 2015;16: 311. doi:10.1186/s12859-015-0720-z

3. Webb C, Wang J. Cell Systems A self-enhanced transport mechanism through long noncoding RNAs for X chromosome inactivation. Nat Publ Gr. Nature Publishing Group; 2016;6: 1–12. doi:10.1038/srep31517

4. Chen KC, Calzone L, Csikasz-Nagy A, Cross FR, Novak B, Tyson JJ. Integrative analysis of cell cycle control in budding yeast. Mol Biol Cell. American Society for Cell Biology; 2004;15: 3841–62. doi:10.1091/mbc.E03-11-0794

5. Ortega F, Stott J, Visser SAG, Bendtsen C. Interplay between α-, β-, and γ-secretases determines biphasic amyloid-β protein level in the presence of a γ-secretase inhibitor. J Biol Chem. American Society for Biochemistry and Molecular Biology; 2013;288: 785–92. doi:10.1074/jbc.M112.419135

6. Fisher J, Henzinger TA. Executable cell biology. Nat Biotechnol. Nature Publishing Group; 2007;25: 1239–49. doi:10.1038/nbt1356

7. Schaub MA, Henzinger TA, Fisher J. Qualitative networks: a symbolic approach to analyze biological signaling networks. BMC Syst Biol. BioMed Central; 2007;1: 4. doi:10.1186/1752-0509-1-4

8. Clarke Edmund M., Grumberg O, Peled D. Model Checking. MIT Press; 1999.

9. Fisher J, Piterman N, Hajnal A, Henzinger TA. Predictive Modeling of Signaling Crosstalk during C. elegans Vulval Development. PLoS Comput Biol. Public Library of Science; 2007;3: e92. doi:10.1371/journal.pcbi.0030092

10. Nusser-Stein S, Beyer A, Rimann I, Adamczyk M, Piterman N, Hajnal A, et al. Cell-cycle regulation of NOTCH signaling during C. elegans vulval

development. Mol Syst Biol. European Molecular Biology Organization; 2012;8: 618. doi:10.1038/msb.2012.51

11. Kauffman SA. Metabolic stability and epigenesis in randomly constructed genetic nets. J Theor Biol. Academic Press; 1969;22: 437–467. doi:10.1016/0022-5193(69)90015-0

12. Glass L, Kauffman SA. The logical analysis of continuous, non-linear biochemical control networks. J Theor Biol. Academic Press; 1973;39: 103–129. doi:10.1016/0022-5193(73)90208-7

13. Chaves M, Albert R, Sontag ED. Robustness and fragility of Boolean models for genetic regulatory networks. J Theor Biol. 2005;235: 431–449. doi:10.1016/j.jtbi.2005.01.023

14. Gupta S, Bisht SS, Kukreti R, Jain S, Brahmachari SK. Boolean network analysis of a neurotransmitter signaling pathway. J Theor Biol. 2007;244: 463–469. doi:10.1016/j.jtbi.2006.08.014

15. Benque D, Bourton S, Cockerton C, Cook B, Fisher J, Ishtiaq S, et al. Bma: Visual Tool for Modeling and Analyzing Biological Networks. Springer, Berlin, Heidelberg; 2012. pp. 686–692. doi:10.1007/978-3-642-31424-7_50

16. Hall BA, Piterman N, Hajnal A, Fisher J. Emergent Stem Cell Homeostasis in the C. elegans Germline Is Revealed by Hybrid Modeling. Biophys J. 2015;109: 428–438. doi:10.1016/j.bpj.2015.06.007

17. Chuang R, Hall BA, Benque D, Cook B, Ishtiaq S, Piterman N, et al. Drug Target Optimization in Chronic Myeloid Leukemia Using Innovative Computational Platform. Sci Rep. 2015;5: 8190. doi:10.1038/srep08190

18. Silverbush D, Grosskurth S, Wang D, Powell F, Gottgens B, Dry J, et al.

Cell-Specific Computational Modeling of the PIM Pathway in Acute Myeloid Leukemia. Cancer Res. American Association for Cancer Research; 2017;77: 827–838. doi:10.1158/0008-5472.CAN-16-1578

19. Tyson JJ, Chen KC, Novak B. Sniffers, buzzers, toggles and blinkers: Dynamics of regulatory and signaling pathways in the cell. Current Opinion in Cell Biology. 2003. pp. 221–231. doi:10.1016/S0955-0674(03)00017-6

20. Tyson JJ, Novák B. Models in biology: lessons from modeling regulation of the eukaryotic cell cycle. BMC Biol. BMC Biology; 2015;13: 46. doi:10.1186/s12915-015-0158-9

21. Csikász-Nagy A, Battogtokh D, Chen KC, Novák B, Tyson JJ. Analysis of a generic model of eukaryotic cell-cycle regulation. Biophys J. The Biophysical Society; 2006;90: 4361–79. doi:10.1529/biophysj.106.081240

22. Tyson JJ, Csikasz-Nagy A, Novak B. The dynamics of cell cycle regulation. BioEssays. Wiley Subscription Services, Inc., A Wiley Company; 2002. pp. 1095–1109. doi:10.1002/bies.10191

23. Tyson JJ, Chen K, Novak B. Network dynamics and cell physiology. Nat Rev Mol Cell Biol. Nature Publishing Group; 2001;2: 908–916. doi:10.1038/35103078

24. Ferrell JE. Perfect and near-perfect adaptation in cell signaling. Cell Systems. 2016. pp. 62–67. doi:10.1016/j.cels.2016.02.006

25. Friedlander T, Brenner N. Adaptive response by state-dependent inactivation. Proc Natl Acad Sci. National Academy of Sciences; 2009;106: 22558–22563. doi:10.1073/pnas.0902146106

923 26. Hochegger H, Takeda S, Hunt T. Cyclin-dependent kinases and cell-
924 cycle transitions: does one fit all? Nat Rev Mol Cell Biol. 2008;9: 910–
925 916. doi:10.1038/nrm2510
926 27. Fisher DL, Nurse P. A single fission yeast mitotic cyclin B p34cdc2
927 kinase promotes both S-phase and mitosis in the absence of G1 cyclins.
928 EMBO J. European Molecular Biology Organization; 1996;15: 850–60.
929 28. Russell P, Nurse P. cdc25+ functions as an inducer in the mitotic control
930 of fission yeast. Cell. Cell Press; 1986;45: 145–153. doi:10.1016/0092-
931 8674(86)90546-5
932 29. Timofeev O, Cizmecioglu O, Settele F, Kempf T, Hoffmann I. Cdc25
933 phosphatases are required for timely assembly of CDK1-cyclin B at the
934 G2/M transition. J Biol Chem. American Society for Biochemistry and
935 Molecular Biology; 2010;285: 16978–16990.
936 doi:10.1074/jbc.M109.096552
937 30. Moreno S, Nurse P. Regulation of progression through the G1 phase of
938 the cell cycle by the rum1+ gene. Nature. 1994;367: 236–242.
939 doi:10.1016/0168-9525(94)90083-3
940 31. Nurse P. Genetic control of cell size at cell division in yeast. Nature.
941 1975;256: 547–551. doi:10.1038/256547a0
942 32. Tominaga Y, Li C, Wang R-H, Deng C-X. Murine Wee1 plays a critical
943 role in cell cycle regulation and pre-implantation stages of embryonic
944 development. Int J Biol Sci. Ivyspring International Publisher; 2006;2:
945 161–170.
946 33. De Schutter K, Joubès J, Cools T, Verkest A, Corellou F, Babiychuk E,
947 et al. Arabidopsis WEE1 kinase controls cell cycle arrest in response to

activation of the DNA integrity checkpoint. Plant Cell. American Society of Plant Biologists; 2007;19: 211–225. doi:10.1105/tpc.106.045047

34. Lengronne A, Schwob E. The yeast CDK inhibitor Sic1 prevents genomic instability by promoting replication origin licensing in late G1. Mol Cell. 2002;9: 1067–1078. doi:10.1016/S1097-2765(02)00513-0

35. Sveiczer A, Csikasz-Nagy A, Gyorffy B, Tyson JJ, Novak B. Modeling the fission yeast cell cycle: quantized cycle times in wee1- cdc25Delta mutant cells. Proc Natl Acad Sci U S A. National Academy of Sciences; 2000;97: 7865–70. doi:10.1073/pnas.97.14.7865

36. Lee G, White LS, Hurov KE, Stappenbeck TS, Piwnica-Worms H. Response of small intestinal epithelial cells to acute disruption of cell division through CDC25 deletion. Proc Natl Acad Sci U S A. National Academy of Sciences; 2009;106: 4701–4706. doi:10.1073/pnas.0900751106

37. Davidich MI, Bornholdt S. Boolean Network Model Predicts Knockout Mutant Phenotypes of Fission Yeast. Csermely P, editor. PLoS One. Public Library of Science; 2013;8: e71786. doi:10.1371/journal.pone.0071786

38. Kim SH, Lin DP, Matsumoto S, Kitazono A, Matsumoto T. Fission yeast Slp1: an effector of the Mad2-dependent spindle checkpoint. Science. 1998;279: 1045–1047. doi:10.1126/science.279.5353.1045

39. Li M, York JP, Zhang P. Loss of Cdc20 causes a securin-dependent metaphase arrest in two-cell mouse embryos. Mol Cell Biol. American Society for Microbiology (ASM); 2007;27: 3481–3488. doi:10.1128/MCB.02088-06

40. Sible JC, Tyson JJ. Mathematical modeling as a tool for investigating cell cycle control networks. Methods. 2007;41: 238–47. doi:10.1016/j.ymeth.2006.08.003

41. Stern B, Nurse P. A quantitative model for the cdc2 control of S phase and mitosis in fission yeast. Trends in Genetics. 1996. pp. 345–350. doi:10.1016/0168-9525(96)10036-6

42. Hong C, Lee M, Kim DD, Kim DD, Cho K-H, Shin I. A checkpoints capturing timing-robust Boolean model of the budding yeast cell cycle regulatory network. BMC Syst Biol. 2012;6: 129. doi:10.1186/1752-0509-6-129

43. Ferrell JE, Tsai TYC, Yang Q. Modeling the cell cycle: Why do certain circuits oscillate? Cell. 2011. pp. 874–885. doi:10.1016/j.cell.2011.03.006

SUPPLEMENTARY INFORMATION

Supplementary Figure 1. Adjustment of Negative Feedback Oscillatory Module. In this illustration, the rows correspond to **(A)** a one node system **(B)** a two node system and **(C)** three node system of negative feedback. The columns correspond to **(i)** the BMA wiring diagram translation and **(ii)** the BMA response curves. Each BMA wiring diagram contains a unique set of target functions located within particular nodes of the network which can be found in **Supplementary table 1**. S indicates the input Signal and R indicates the output Response with, in our case, letters A-D representing intermediate

nodes. The graphs in (ii) are derived from simulation analysis carried out in the BMA. For all cases a simulation is run with a set signal input of 2 as an example, and the response output from the BMA simulation plotted based on the response per calculation time step. Comparison of the three different node length systems highlights that with increased number of nodes there is an increased length of oscillation response, as shown in the 20 calculation steps graphed, with the three node system **(C)** carrying out only 2 full oscillations compared to the one node system **(A)** which carries out 4 oscillations in the 20 calculation steps.

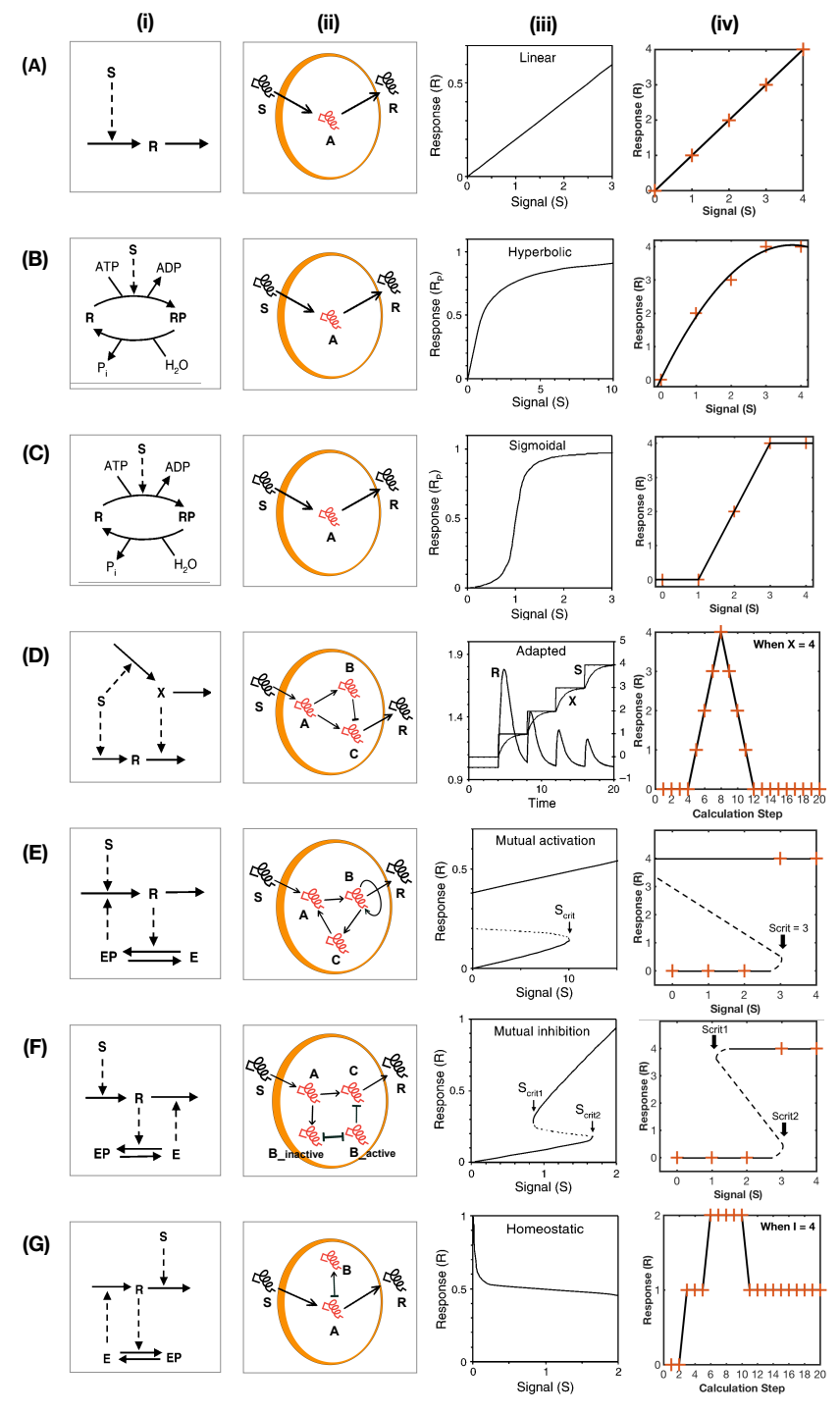
Supplementary Figure 2. Wiring Diagram. Wiring diagram composed using the network topology specified from the BMA target function library modules. Cdk-cycB activity is subdivided into individual pools which link to a central Cdk1-cycB pool. Subsequent models combine the target functions of the G1 Cdk-cycB activity and the G2-cdk-cycB activity together into one node which better represents the true biology.

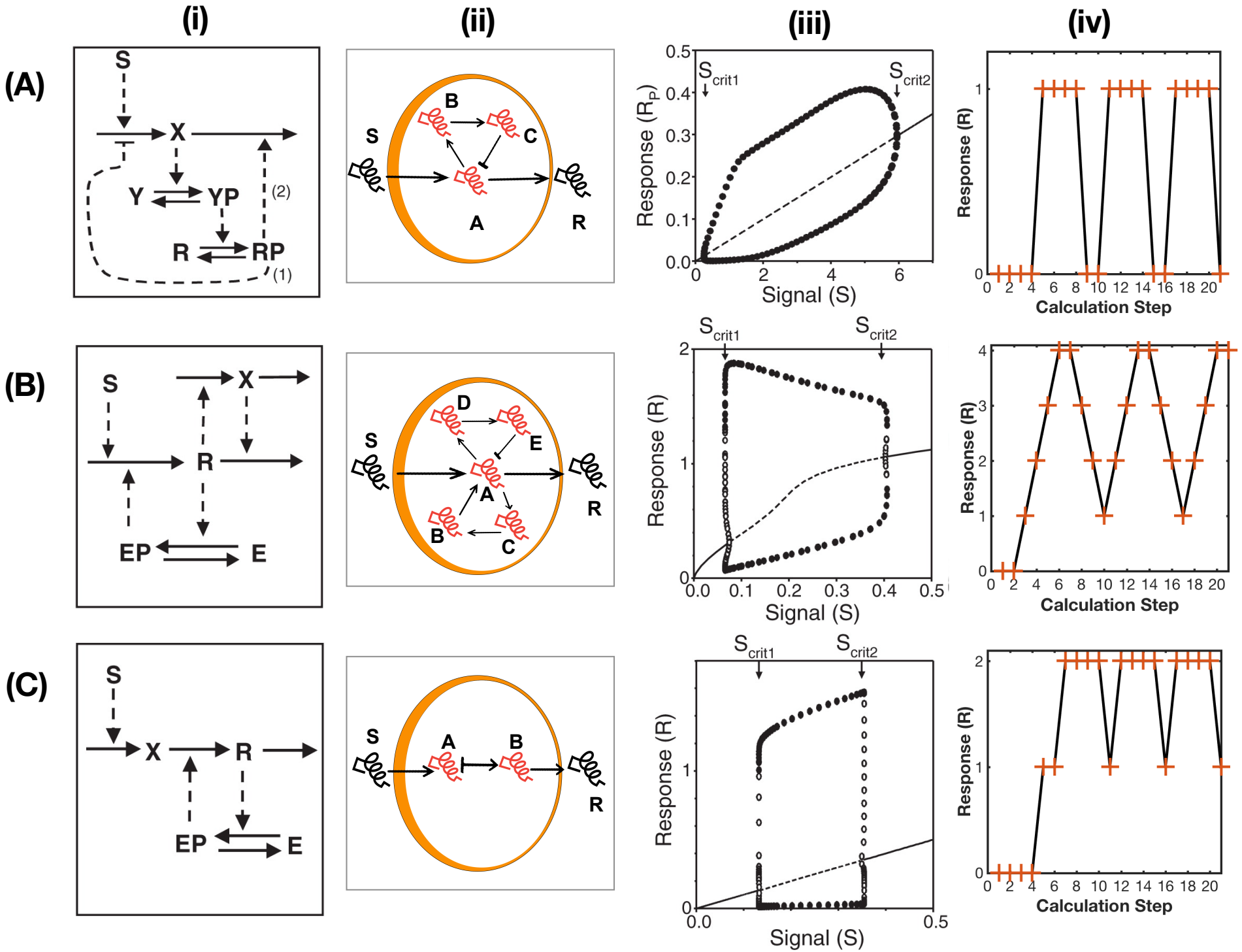
Supplemental Table 1. List of networks assessed in the manuscript. Paper reference refers to the figure in which the network occurs, included are the target functions for each node (if not the default), and comments. Included are the filenames and model names where the specific network can be found.

Supplemental Table 2. List of nodes in the network. Network ID refers to the internal label of the node. Full name is the common name found in the literature while Network Name is the name given in construction of model.

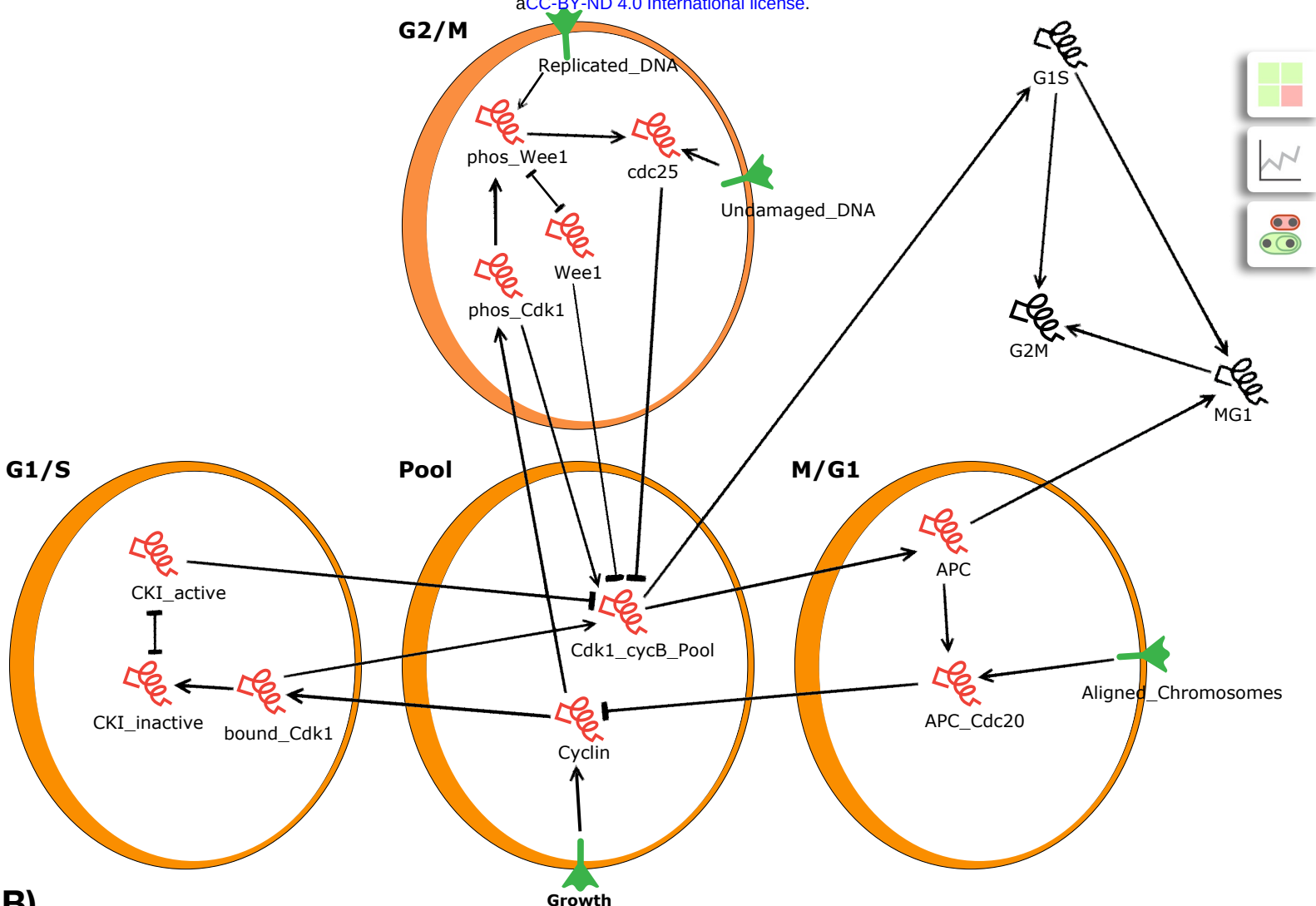
Supplemental Table 3. List of nodes in the network. Network ID refers to the internal label of the node. Full name is the common name found in the literature while Network Name is the name given in construction of model. The target function located in each node is found under the "Target Function" heading.

Model Files. We have additionally included all model files in .json format within an enclosed .zip file, they are also available at https://github.com/shorthouse-mrc/biomodelanalyzer_targetfunctionlibrary. Importing any file into the BMA will load the model and allow manipulation and simulation/stability analysis. Each file is named explicitly in Supplementary Table 1, with some files containing multiple models, which are referenced independently.

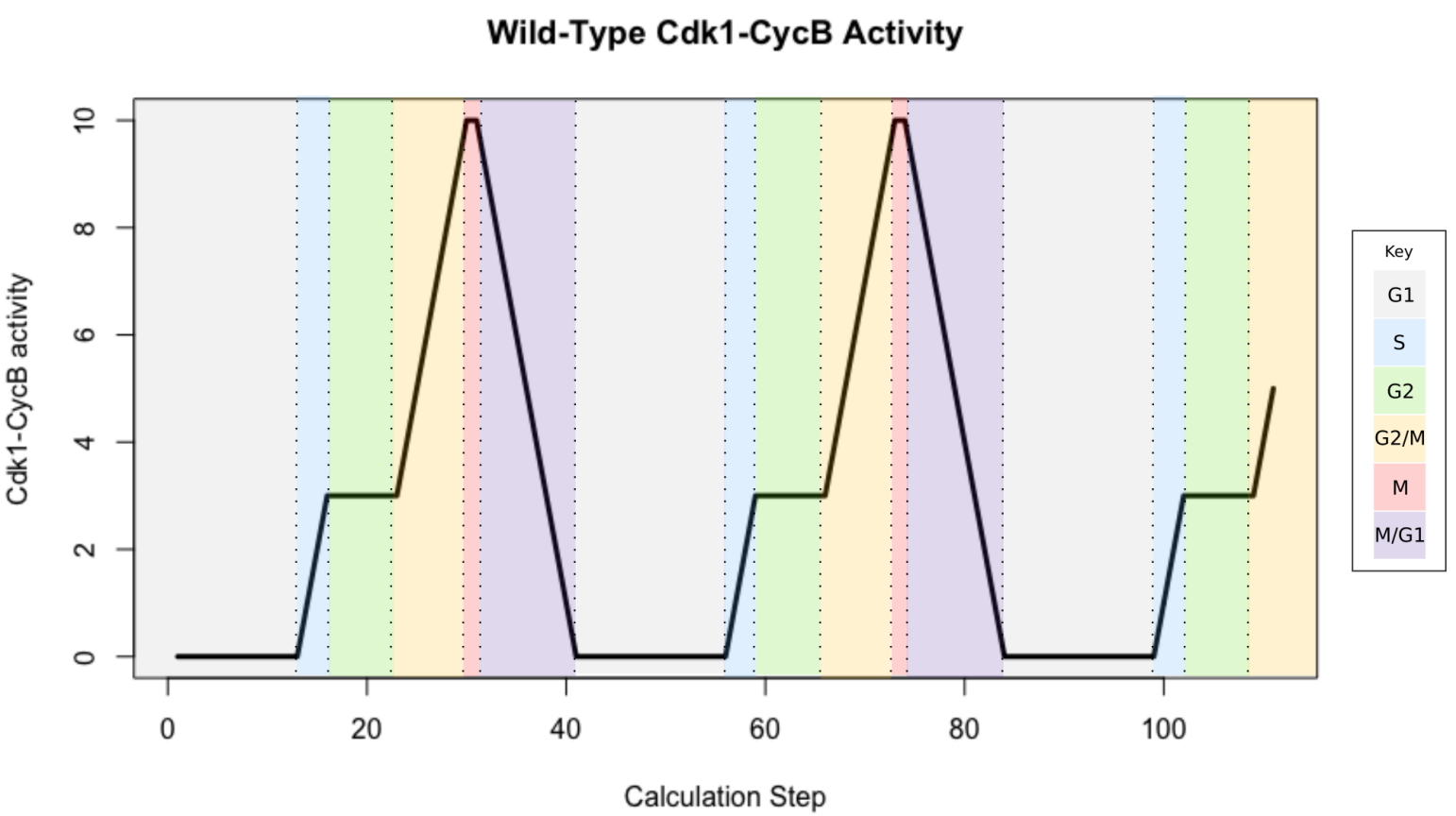

(i)
(ii)
(iii)
(iv)
(A)
S
R
A
Linear
Response (R)
Signal (S)
(B)
ATP
ADP
RP
P_i
H_2O
Hyperbolic
Response (R_p)
(C)
Sigmoidal
(D)
X
B
C
Adapted
Time
When X = 4
Calculation Step
(E)
EP
E
Mutual activation
S_{crit}
Scrit = 3
(F)
C
B_inactive
B_active
Mutual inhibition
S_{crit1}
S_{crit2}
Scrit1
Scrit2
(G)
Homeostatic
When I = 4


(i)
(ii)
(iii)
(iv)
(A)
S
X
Y
YP
(2)
R
RP
(1)
B
C
A
R
Response (R_P)
S_{crit1}
S_{crit2}
Signal (S)
Response (R)
Calculation Step
(B)
D
E
(C)
EP
E

A)



B)



Mutation	Time Step	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61
	Wild Type	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	3	3	3	3	3	3	3	4	5	6	7	8	9	10	10	9	8	7	6	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	3	3
	Wee1Δ	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	9	10	10	9	8	7	6	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	9	10	10	9
	Wee1 OP	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	CKI Δ	0	0	0	0	1	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	5	6	7	8	9	10	10	9	8	7	6	5	4	3	2	1	0	0	0	0	0	0	0	1	2	3	3	3	3	3	3	3	3	3	3	3	3
	CKI OP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	7	6	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	CKI & Wee1Δ	0	0	0	0	1	2	3	3	3	3	3	3	3	4	5	6	7	8	9	10	10	9	8	7	6	5	4	3	2	1	0	0	0	0	0	0	0	1	2	3	3	3	3	3	3	3	4	5	6	7	8	9	10	10	9	8	7	6	5	4	3
	Cdc25 Δ	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	Cdc25 & Wee1Δ	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	Cdc25 OP	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	9	10	9	8	7	6	5	4	3	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	4	5	6	7	8	9	10	9	8	7	6
	Cdc20 Δ	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	3	3	3	3	3	3	3	4	5	6	7	8	9	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	No Growth	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Unreplicated DNA	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	Damaged DNA	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	Misaligned Chromosomes	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	3	3	3	3	3	3	3	3	4	5	6	7	8	9	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10

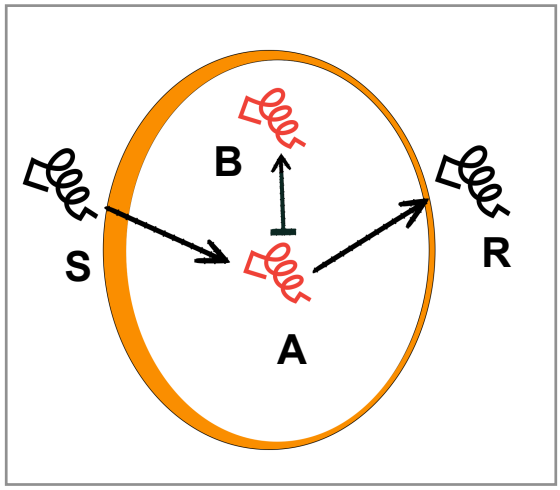
Key:

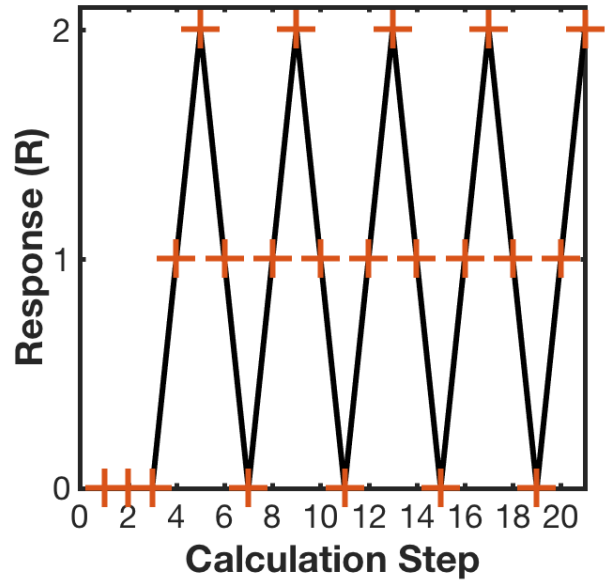
G1	S	G2	G2/M	M	M/G1

(i)

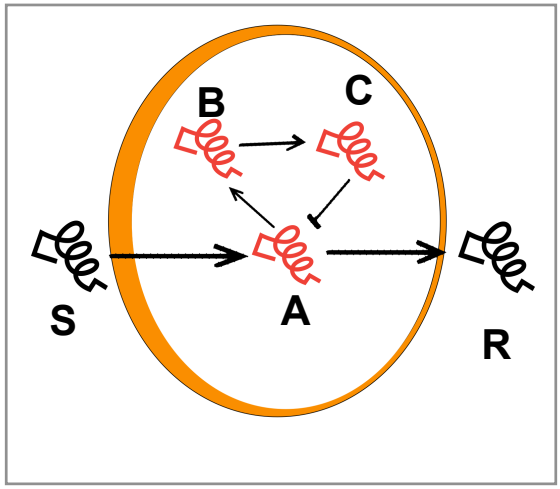
(ii)

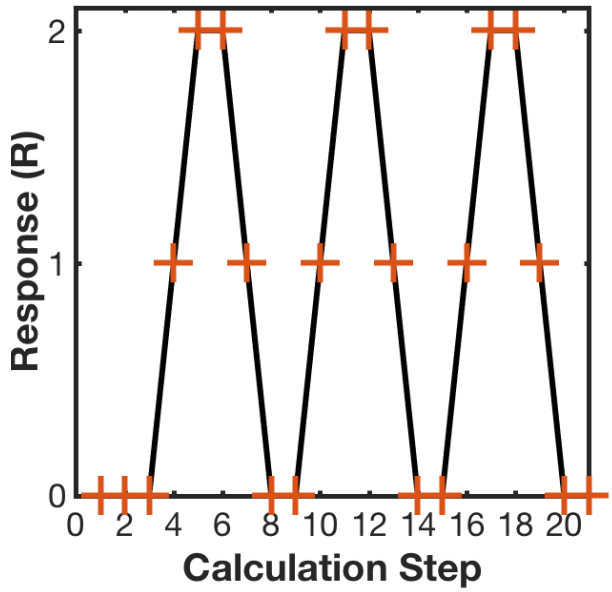
(A)


B
S
R
A

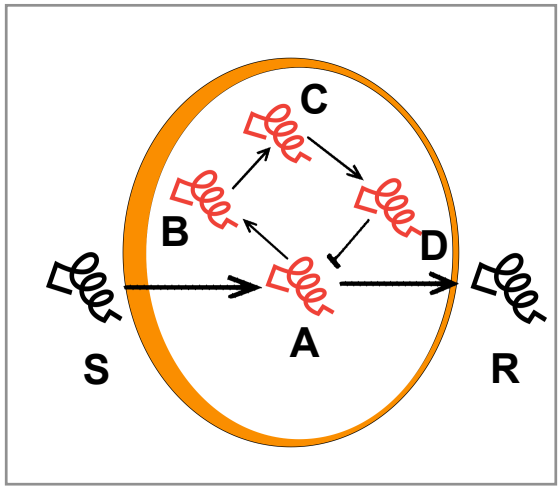

Response (R)
Calculation Step
0
1
2
0 2 4 6 8 10 12 14 16 18 20

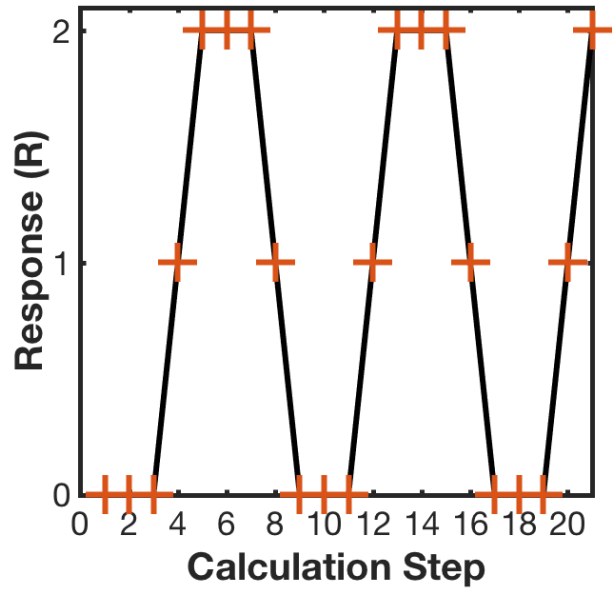
(B)

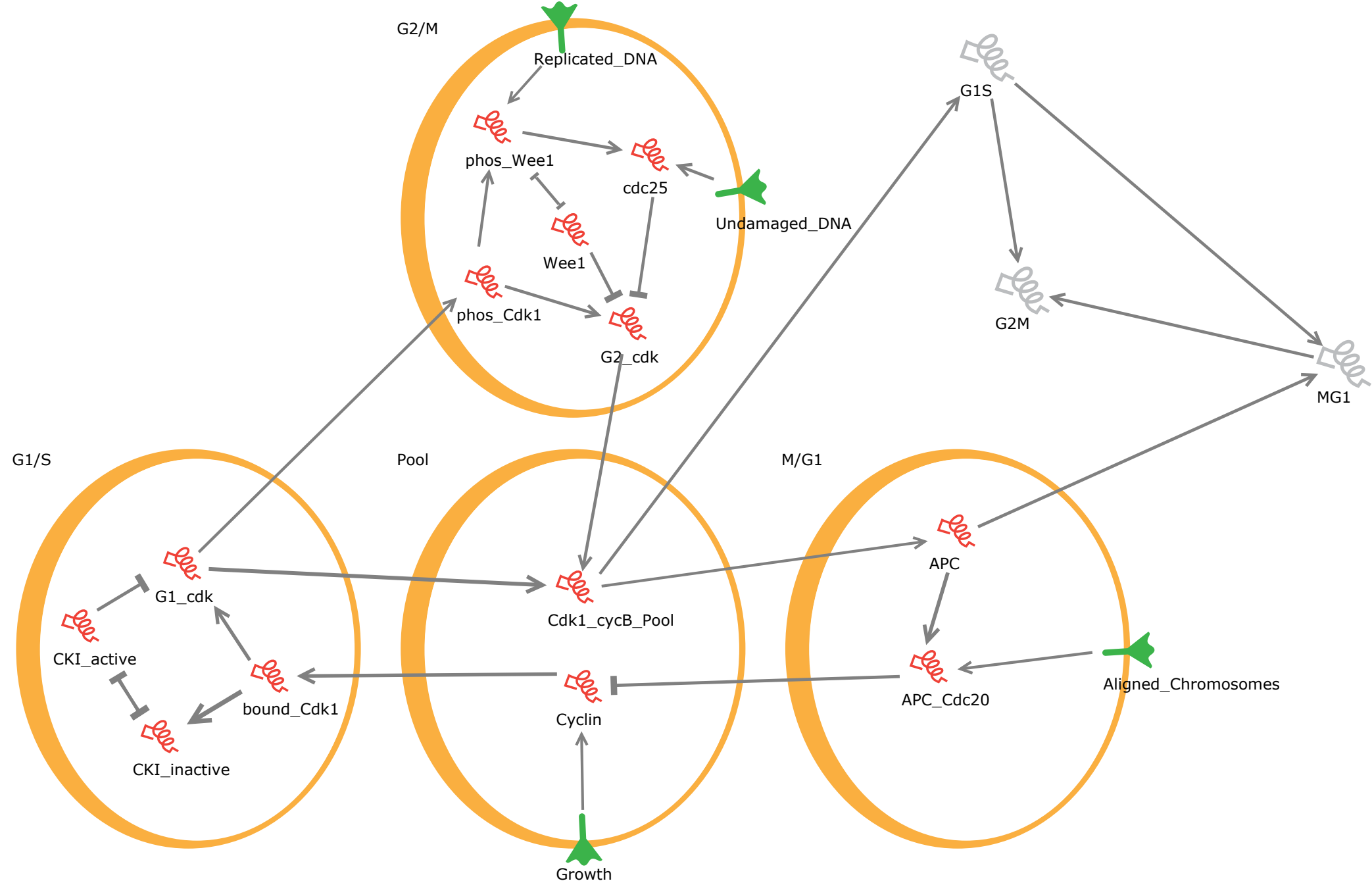

B
C
S
A
R


Response (R)
Calculation Step
0
1
2
0 2 4 6 8 10 12 14 16 18 20

(C)


C
B
D
S
A
R


Response (R)
Calculation Step
0
1
2
0 2 4 6 8 10 12 14 16 18 20


G2/M
Replicated_DNA
phos_Wee1
cdc25
Undamaged_DNA
Wee1
phos_Cdk1
G2_cdk
G1S
G2M
MG1
G1/S
Pool
M/G1
G1_cdk
CKI_active
bound_Cdk1
CKI_inactive
Cdk1_cycB_Pool
Cyclin
Growth
APC
APC_Cdc20
Aligned_Chromosomes