Year : 2015

Development and Application of Computational Methodologies in Qualitative Modeling

CHASAPI Anastasia

UNIL | Université de Lausanne
Faculté de biologie et de médecine

SIB Swiss Institute of Bioinformatics

Development and application of computational methodologies in qualitative modeling

Thèse de doctorat en sciences de la vie (PhD)
présentée à la Faculté de biologie et de médecine
de l'Université de Lausanne
par

Anastasia Chasapi

Jury

President
Prof. Nicolas Perrin, Ecology and Evolution of Dispersal group, UNIL Lausanne

Thesis director
Prof. Ioannis Xenarios, Vital-IT/Swiss-Prot groups, Center for Integrative Genomics UNIL/ SIB Swiss Institute of Bioinformatics, Lausanne

Expert
Prof. Viesturs Simanis, Cell Cycle Control laboratory, EPFL, Lausanne

Expert
Prof. Yves Barral, Institute of Biochemistry, ETH Zurich

Expert
Dr. Frédéric Schütz, Center of Integrative Genomics/ SIB Swiss Institute of Bioinformatics, Bioinformatics Core Facility, UNIL, Lausanne

Lausanne 2015

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Ecole Doctorale

Doctorat ès sciences de la vie

Imprimatur

Vu le rapport présenté par le jury d'examen, composé de

Président	Monsieur	Prof. Nicolas **Perrin**
Directeur de thèse	Monsieur	Prof. Ioannis **Xenarios**
Experts	Monsieur	Prof. Yves **Barral**
	Monsieur	Prof. Frédéric **Schütz**
	Monsieur	Prof. Viesturs **Simanis**

le Conseil de Faculté autorise l'impression de la thèse de

Madame Anastasia Chasapi

Master of Science in Proteomics and Bioinformatics, Université de Genève

intitulée

Development and application of computational methodologies in qualitative modeling

Lausanne, le 25 mars 2015

pour La Doyenne
de la Faculté de Biologie et de Médecine

Prof. Nicolas Perrin

Table of Contents

Acknowledgements

This thesis would not have been possible without the help of many people. First, I would like to thank Prof. Ioannis Xenarios, who gave me the opportunity of realizing this thesis under his supervision and who guided me but also gave me the freedom to find my own way in this PhD. I owe my deepest gratitude to Julien Dorier, who was there for me at every step, always willing to help me.

My collaboration with Prof. Simanis and his group has been wonderful, and this is mostly due to the people, and the environment of cooperation and friendship. For that, I would like to thank all members of the group. A special thanks to Prof. Viesturs Simanis and Dr. Paulina Wachowicz for their collaboration, encouragement and support in every level, Dr. Andrea Krapp for her useful input on the model and Elena Cano for introducing me to fission yeast experimental techniques.

In the frame of our collaboration for the budding yeast project, I owe my deepest gratitude to Dr. Matteo Barberis, for this prolific and interesting collaboration, as well as his personal investment in the project.

Finally, I would like to thank my family and my dearest friends in Greece and Switzerland, whose love and patience motivates me in every step of my life.

Abbreviations

APC	Anaphase Promoting Complex
CDK	Cyclin Dependent Kinase
CFP	Cyan Fluorescent Protein
GE	Gene Expression
GFP	Green Fluorescent Protein
GRN	Gene Regulatory Network
HU	Hydroxyurea
KO	knock-out
MEN	Mitotic Exit Network
ODE	Ordinary Differential Equation
OE	over-expression
PCA	Principal Component Analysis
PKN	Prior Knowledge Network
PTM	Post Translational Modification
ROBDD	Reduced Ordered Binary Decision Diagram
SAC	Spindle Assembly Checkpoint
SBV	Systems Biology Verification
SIN	Septation Initiation Network
SPB	Spindle Pole Body
YE	Yeast Extract
YFP	Yellow Fluorescent Protein

Résumé

La présente thèse s'intitule ”Développent et Application des Méthodologies Computationnelles pour la Modélisation Qualitative”. Elle comprend tous les différents projets que j'ai entrepris en tant que doctorante. Plutôt qu'une mise en oeuvre systématique d'un cadre défini *a priori*, cette thèse devrait être considérée comme une exploration des méthodes qui peuvent nous aider à déduire le plan de processus regulatoires et de signalisation. Cette exploration a été mue par des questions biologiques concrètes, plutôt que par des investigations théoriques. Bien que tous les projets aient inclus des systèmes divergents (réseaux régulateurs de gènes du cycle cellulaire, réseaux de signalisation de cellules pulmonaires) ainsi que des organismes (levure à fission, levure bourgeonnante, rat, humain), nos objectifs étaient complémentaires et cohérents.

Le projet principal de la thèse est la modélisation du réseau de l'initiation de septation (SIN) du S.*pombe*. La cytokinèse dans la levure à fission est contrôlée par le SIN, un réseau signalant de protéines kinases qui utilise le corps à pôle-fuseau comme échafaudage. Afin de décrire le comportement qualitatif du système et prédire des comportements mutants inconnus, nous avons décidé d'adopter l'approche de la modélisation booléenne. Dans cette thèse, nous présentons la construction d'un modèle booléen étendu du SIN, comprenant la plupart des composantes et des régulateurs du SIN en tant que noeuds individuels et testable expérimentalement. Ce modèle utilise des niveaux d'activité du CDK comme noeuds de contrôle pour la simulation d'évènements du SIN à différents stades du cycle cellulaire. Ce modèle a été optimisé en utilisant des expériences d'un seul “knock-out” avec des effets phénotypiques connus comme set d'entraînement. Il a permis de prédire correctement un set d'évaluation de “knock-out” doubles. De plus, le modèle a fait des prédictions *in silico* qui ont été validées *in vivo*, permettant d'obtenir de nouvelles idées de la régulation et l'organisation hiérarchique du SIN.

Un autre projet concernant le cycle cellulaire qui fait partie de cette thèse a été la construction d'un modèle qualitatif et minimal de la réciprocité des cyclines dans la S.*cerevisiae*. Les protéines Clb dans la levure bourgeonnante présentent une activation et une dégradation caractéristique et séquentielle durant le cycle cellulaire, qu'on appelle communément les vagues des Clbs. Cet évènement est coordonné avec la courbe d'activation inverse du Sic1, qui a un rôle inhibitoire dans le système. Pour l'identification des modèles qualitatifs minimaux qui peuvent expliquer ce phénomène, nous avons sélectionné des expériences bien définies et construit tous les modèles minimaux possibles qui, une fois simulés, reproduisent les résultats attendus. Les modèles ont été filtrés en utilisant des simulations ODE qualitatives et standardisées; seules celles qui

reproduisaient le phénotype des vagues ont été gardées. L'ensemble des modèles minimaux peut être utilisé pour suggérer des relations regulatoires entre les molécules participant qui peuvent ensuite être testées expérimentalement.

Enfin, durant mon doctorat, j'ai participé au SBV Improver Challenge. Le but était de déduire des réseaux spécifiques à des espèces (humain et rat) en utilisant des données de phosphoprotéines, d'expressions des gènes et des cytokines, ainsi qu'un réseau de référence, qui était mis à disposition comme donnée préalable. Notre solution pour ce concours a pris la troisième place. L'approche utilisée est expliquée en détail dans le dernier chapitre de la thèse.

Summary

The present dissertation is entitled “Development and Application of Computational Methodologies in Qualitative Modeling”. It encompasses the diverse projects that were undertaken during my time as a PhD student. Instead of a systematic implementation of a framework defined *a priori*, this thesis should be considered as an exploration of the methods that can help us infer the blueprint of regulatory and signaling processes. This exploration was driven by concrete biological questions, rather than theoretical investigation. Even though the projects involved divergent systems (gene regulatory networks of cell cycle, signaling networks in lung cells), as well as organisms (fission yeast, budding yeast, rat, human), our goals were complementary and coherent.

The main project of the thesis is the modeling of the Septation Initiation Network (SIN) in S.*pombe*. Cytokinesis in fission yeast is controlled by the SIN, a protein kinase signaling network that uses the spindle pole body as scaffold. In order to describe the qualitative behavior of the system and predict unknown mutant behaviors we decided to adopt a Boolean modeling approach. In this thesis, we report the construction of an extended, Boolean model of the SIN, comprising most SIN components and regulators as individual, experimentally testable nodes. The model uses CDK activity levels as control nodes for the simulation of SIN related events in different stages of the cell cycle. The model was optimized using single knock-out experiments of known phenotypic effect as a training set, and was able to correctly predict a double knock-out test set. Moreover, the model has made in silico predictions that have been validated *in vivo*, providing new insights into the regulation and hierarchical organization of the SIN.

Another cell cycle related project that is part of this thesis was to create a qualitative, minimal model of cyclin interplay in S.*cerevisiae*. CLB proteins in budding yeast present a characteristic, sequential activation and decay during the cell cycle, commonly referred to as Clb waves. This event is coordinated with the inverse activation curve of Sic1, which has an inhibitory role in the system. To generate minimal qualitative models that can explain this phenomenon, we selected well-defined experiments and constructed all possible minimal models that, when simulated, reproduce the expected results. The models were filtered using standardized qualitative ODE simulations; only the ones reproducing the wave-like phenotype were kept. The set of minimal models can be used to suggest regulatory relations among the participating molecules, which will subsequently be tested experimentally.

Finally, during my PhD I participated in the SBV Improver Challenge. The goal was to infer species-specific (human and rat) networks, using phosphoprotein, gene expression and cytokine data and a reference network provided as prior knowledge. Our solution to the challenge was

selected as 3rd best performing. The approach used is explained in detail in the final chapter of the thesis.

Chapter 1

Introduction

Moving from molecular to modular

As an enthusiastic photographer I have always been fascinated by the power of perspective. Every choice you make for your picture can reveal a different aspect of your subject. Let us take a look, for example, at the starlings in picture 1.1A. The photograph must have been taken with a zoom or prime lens, allowing us to observe the birds' appearance in relative detail. We can also see that both starlings are flying, but their wings are not moving synchronously. We can continue making observations regarding each bird individually, and maybe a few for the relation between the two birds.

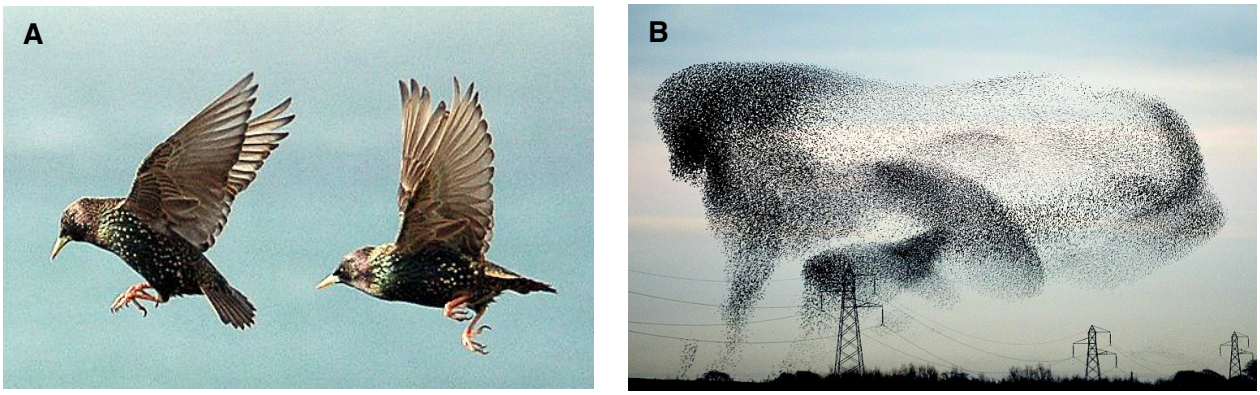

Figure 1.1: Starlings in flight. A) *Two starlings flying over Merseyside, Liverpool (photo by Steve Ward, 2013).* **B)** *Picture capturing around ten thousand starlings flying in formation over pylons in the Scottish borders near Gretna, Dumfriesshire (photo by Owen Humphreys, 2013).*

These observations are certainly of value, as they can help us understand many features of the bird. When we take off the zoom, though, a whole new level of organization and complexity is revealed, as shown in picture 1.1B. Birds interact with each other and adjust their behavior, when they are part of a flock. Those intelligent organisms, form groups with such a great level of coordination that, finally, they act like "superorganisms". Keeping the zoom on, interesting though it may be, prevents us from understanding the flock dynamics that determine each bird's behavior.

Systems biology may be thought of as a flock photographer. It shifts our focus from molecular to modular, by following a holistic, rather than reductionist approach. While the understanding of genes and proteins continues to be essential, systems biology focuses on understanding the structure and dynamics of the system as a whole. A system is more than just an assembly of genes and proteins. Therefore, its properties cannot be fully understood by merely listing the properties of its components, neither by drawing diagrams of their interconnections. Although such a diagram is an important first step, it is a static photograph, whereas what we really seek to know

are the flock patterns, why such patterns emerge, and how they can be controlled. This is analogous to drawing an exhaustive topological diagram of gene regulatory networks and their interactions. Such diagrams provide limited knowledge of how changes to one part of a system may affect other parts. To understand how a particular system functions, we must first examine how the individual components dynamically interact (Cvijovic et al, 2014; Kitano, 2002; Samaga & Klamt, 2013).

Dynamic modeling of gene regulatory networks

Discerning the blueprint of the relations among components is not a trivial task. It is usually approached with either bottom-up or top-down methodologies. The former includes data collection (literature and databases), manual curation, network reconstruction through mathematical methods, and validation of these models through experiments and literature analysis. The latter encompasses network reconstruction using 'omics' data, generated through DNA microarrays, RNA-Seq or other high-throughput genomic techniques using appropriate statistical and bioinformatics methodologies (Bruggeman & Westerhoff, 2007).

When aiming to construct dynamic models of bio-molecular networks, each case is approached individually, depending on the type and amount of knowledge and experimental data available for the specific system, as well as the size of the network (Kahlem et al, 2011; Wang et al, 2012). To account for the different quality of information that is available for a network under study, modeling formalisms of different levels of complexity have been developed over the last years, that can vary from qualitative Boolean models, to quantitative kinetic-based models (de Jong, 2002; Karlebach & Shamir, 2008; Kestler et al, 2008; Samaga & Klamt, 2013).

Quantitative modeling approaches provide detailed descriptions of the biochemical processes that are based on chemical and physical principles (Aldridge et al, 2006). Sets of ordinary differential equations (ODEs) are the most commonly used formulation, due to their well-established biophysical basis and straightforward molecular interpretation; see (Szallasi et al, 2006) for detailed review. They describe the system's evolution over time using mass-action kinetics for the rates of production and consumption of the components included in the model. The major limitation of quantitative modeling is that it requires sufficient knowledge of biological mechanisms and kinetic parameters, which limits its applicability to small, well-characterized networks; for example (Csikasz-Nagy et al, 2007; Tyson et al, 2002).

In contrast to quantitative models, qualitative approaches are mainly based on network structure and do not require kinetic information. This makes them more suitable for large-scale network analysis. There are several qualitative modeling approaches, varying in formalism and complexity. Some of the most widely known are logical modeling (Garg et al, 2012; Morris et al, 2010; Wang et al, 2012), Petri nets (Chaouiya, 2007; Hardy & Robillard, 2004) and constraint-based modeling (Feist & Palsson, 2008; Papin & Palsson, 2004). Even though these frameworks rely on network structure, rather than kinetic parameters, they enable the analysis of important functional properties of large-scale regulatory networks. For example, they include regulatory relationships among components, feedback loops and signal transfer routes. Logical models (including Boolean) and Petri nets can also derive qualitative properties of the system's dynamics by means of discrete dynamic modeling (figure 1.2). They can be used descriptively for the simulation of the steady states of the system, but they can also be used predictively to evaluate, for instance, the expected qualitative response of the system to defined perturbations.

Boolean modeling features

Boolean models integrate knowledge regarding components and interactions of biomolecular systems whose kinetic parameters have not been determined. Using the significant level of abstraction imposed by their binary definition, they manage to capture the system's dynamic repertoire and they have been successfully used in several contexts (Albert & Othmer, 2003; Azpeitia et al, 2011; Davidich & Bornholdt, 2008; Giacomantonio & Goodhill, 2010; Li et al, 2004; Li et al, 2006; Morris et al, 2010; Saez-Rodriguez et al, 2007; Samaga et al, 2009; Sanchez et al, 1997; Schlatter et al, 2012; Schlatter et al, 2009; Veliz-Cuba & Stigler, 2011). In Boolean models of Gene Regulatory Networks (GRNs), nodes can be gene products, proteins or complexes, and edges reflect the regulatory relationships between nodes. In Boolean formalism, each node is characterized by an activation state that can take the values 1 for "active" or 0 for "inactive", corresponding to the logical values TRUE and FALSE. The activation state can refer to transcription, localization, phosphorylation or other post-translational modifications (PTMs).

The input for a Boolean simulation is the GRN with all regulatory events encoded as logical rules that can be found true or not (figure 1.3). The Boolean functions (logical rules) representing the regulatory relationships of the system, can include AND, OR, NOT, IAND, XOR, XNOR (Akutsu et al, 1999; Garg et al, 2012; Garg et al, 2009; Kauffman, 1993; Kochi & Matache, 2012; Li et al, 2004). The state of each node depends on the state of the nodes regulating it, that is, the state of all the incoming edges, and the rules that govern their interaction (figure 1.2). The state of all the nodes at a given moment defines the network state. The network transitions from state to state

dictated by the underlying Boolean functions, until it reaches a steady state or a cyclic attractor (Garg et al, 2012). The possible trajectories in the state space can be represented by the state transition graph (Naldi et al, 2007; Saadatpour et al, 2011; Wang et al, 2012).

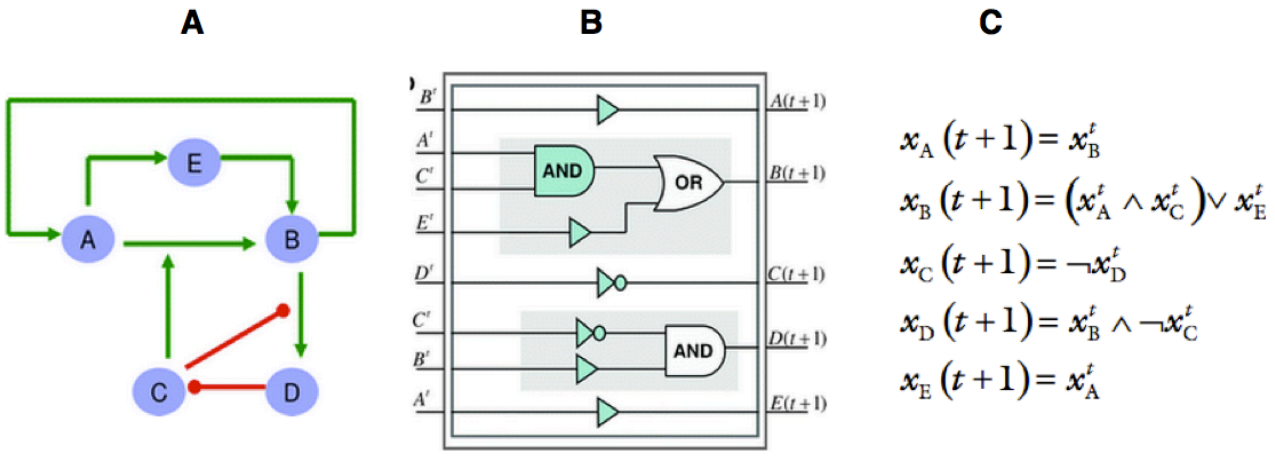

Figure 1.2: A GRN mapped to Boolean model. A) *A simple, 5 node GRN. Green arrows represent activations and red circles inhibitions. In the model there are pairwise interactions (for example "D inhibits C") and more complex regulations (for example "A together with C activate B").* **B)** *The same network, represented as Boolean functions describing the state of each node at time t+1, given the state of their regulating nodes at time t. For example, at time t+1 B will be active if at time t both A and C are active, or if E is active. In any other case, the logical rule will not be found true and B will be inactive.* **C)** *The mathematical expression of the Boolean functions. Variables* X^t_A, X^t_B, X^t_C, X^t_D *and* X^t_E *represent the expression of nodes A, B, C, D and E respectively.* $\wedge$, $\vee$ *and* $\neg$ *represent Boolean AND, OR and NOT. Source :* (Garg et al, 2012)

In Boolean modeling, time is discrete and abstract. Two main updating schemes can be used during model simulation; synchronous and asynchronous update. The former assumes that all biological events in the system have similar timescales, and all functions are updated simultaneously. In the latter, one function is updated at each time step, which can be deterministic (deterministic asynchronous) or randomly selected (stochastic asynchronous) (Chaves et al, 2005; Chaves et al, 2006). The asynchronous behavior can be controlled by setting additional rules for time delays and priorities (Garg et al, 2009). Alternatively, all possible transitions can be generated (Faure et al, 2006; Garg et al, 2008; Kahlem et al, 2011; Wang et al, 2012).

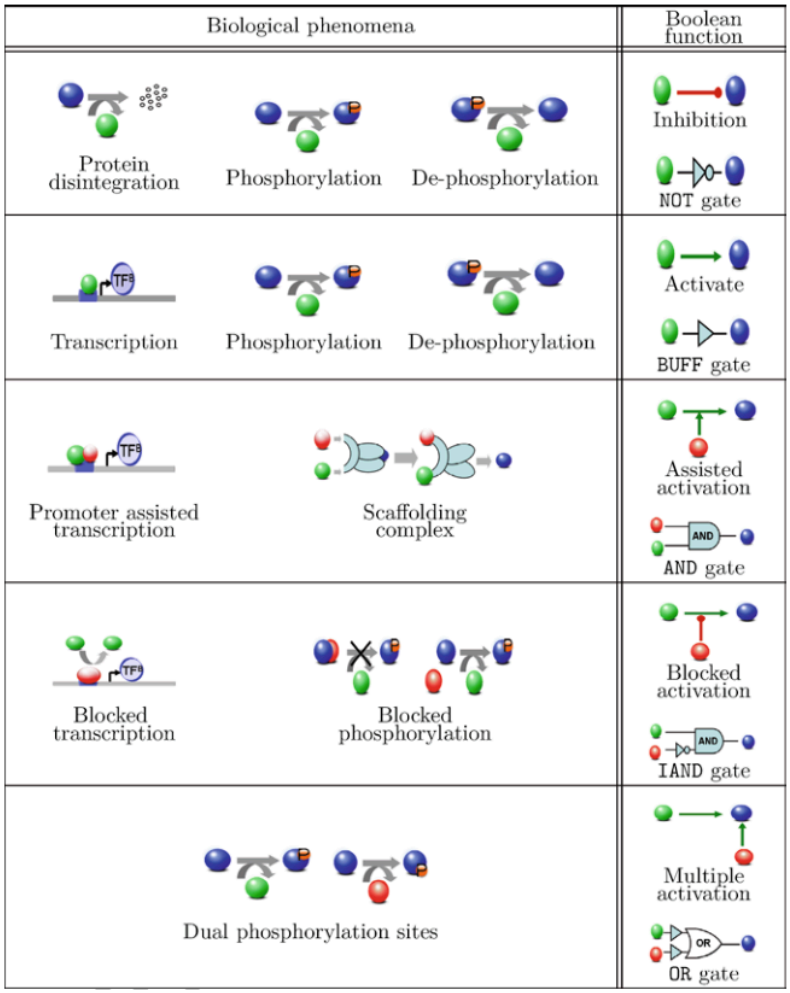

Figure 1.3: Boolean function mapping of biological phenomena in GRNs. *Examples of the encoding of biological processes to logical rules. Source: (Garg et al, 2012)*

Thesis outline

Each chapter of this dissertation is dedicated to a different biological process (gene regulation in cell cycle, signaling in lung cells) in a different organism (fission yeast, budding yeast, rat and human), based on case studies addressing specific biological questions. Despite the divergence in the methods used to approach them, all the case studies share a common goal: to derive the blueprint underlying the relations among the system's components. Boolean modeling was the principal method employed for this goal. Since each chapter discusses different biological problems, these are introduced individually *in situ*.

Chapter 2 describes the central project of my PhD; modeling the Septation Initiation Network in fission yeast. It includes an introduction on the system as well as all our efforts during these years,

successful or not. It describes the construction of the initial model, the optimization process and the change of strategy that led to the CDK-switch model of the SIN. The latter was used predictively, and the predictions were validated experimentally. **Chapter 3** focuses on the inference of a set of minimal models describing cyclin regulatory interplay in budding yeast. A description of the system is followed by the variable methods we used towards our goal and future perspectives for the project. Finally, **Chapter 4** is dedicated to the SBV Improver Challenge, where, together with colleagues at Vital-IT, I participated at the “Species Specific Network Inference” sub-challenge.

Chapter 2

An extended, Boolean model of the septation initiation network in S.pombe provides insights into its regulation

Partially based on

Chasapi A, Wachowicz P, Niknejad A, Collin P, Krapp A, Cano E, Simanis V, Xenarios I (2015) *An Extended, Boolean Model of the Septation Initiation Network in S.pombe Provides Insights into its Regulation.* PloS Comp Biol (submitted)

Introduction

Schizosaccharomyces pombe, commonly referred as fission yeast*,* has long been used as a model organism for the study of conserved, essential functions in the eukaryotic cell. It has proved highly informative in the study of the cell cycle, particularly the control of the G2/M transition. Like many somatic higher eukaryotic cells, it divides by binary fission. Cytokinesis in fission yeast is controlled by the Septation Initiation Network (SIN), a protein kinase signaling network, which uses the spindle pole body (SPB; the functional counterpart of the centrosome in yeast), as a scaffold from which to initiate signaling. Elements of the SIN's signaling architecture have been conserved throughout evolution. In *Saccharomyces cerevisiae* the corresponding pathway is known as the mitotic exit network (MEN), and controls both cytokinesis and mitotic exit. In higher eukaryotes the equivalent signaling network is the Hippo pathway, which regulates cell growth and proliferation (Bardin & Amon, 2001; Hergovich & Hemmings, 2012; Hotz & Barral, 2014; Seshan & Amon, 2004; Weiss, 2012).

The SIN comprises a group of protein kinases and their regulators that induce cytokinesis when CDK activity drops in anaphase (Chang et al, 2001; Guertin et al, 2000). Signaling failure results in multinucleated cells, as cytokinesis fails while growth and the nuclear cycle continue (Mitchison & Nurse, 1985); this is referred to as the SIN phenotype. Failure to turn off SIN signaling produces multiseptated cells that remain uncleaved and contain one or two nuclei (Minet et al, 1979). Ppc89p, Cdc11p and Sid4p form the scaffold upon which signaling proteins are assembled at the SPB (Chang & Gould, 2000; Krapp et al, 2001; Rosenberg et al, 2006; Tomlin et al, 2002). SIN signaling requires the action of three kinase-regulator modules. The kinase Cdc7p associates with the signaling GTPase Spg1p (Mehta & Gould, 2006; Schmidt et al, 1997), Sid1p associates with its regulatory subunit Cdc14p (Guertin et al, 2000; Guertin & McCollum, 2001) and the kinase Sid2p associates with its regulator Mob1p (Hou et al, 2004; Hou et al, 2000; Salimova et al, 2000). It is claimed that Cdc7p also binds directly to Cdc11p (Feoktistova et al, 2012). The nucleotide status of Spg1p is regulated by a bipartite GAP, composed of a catalytic subunit (Cdc16p), which interacts with Spg1p in the context of a scaffold, Byr4p (Furge et al, 1999; Furge et al, 1998). Etd1p regulates the nucleotide status of Spg1p, perhaps by modulating Rho1p signaling (Alcaide-Gavilan et al, 2014; Daga et al, 2005; Garcia-Cortes & McCollum, 2009; Lahoz et al, 2010). Plo1p acts upstream of the SIN (Krapp et al, 2003; Tanaka et al, 2001), and coordinates SIN activity with other mitotic events. See figure 2.1 for a diagram of SIN signaling and the participating components.

Apart from the essential SIN components described above, there are several SIN signaling regulators. SPB-associated protein kinases are important for SIN regulation. The CDK Cdc2p-Cdc13p is controlling the SIN both positively and negatively. The diverse influence of CDK depends on the stage of the cell cycle and the CDK activation levels (Cerutti & Simanis, 1999; Chang et al, 2001; Dischinger et al, 2008; Guertin et al, 2000). The ubiquitin ligase Dma1p is another SIN regulator, required to prevent septum formation if spindle function is compromised (Murone & Simanis, 1996). As part of the spindle assembly checkpoint, Dma1p binds to Sid4p and ubiquitinylates it, to prevent the recruitment of Plo1p to the SPB (figure 2.1C), thereby inhibiting SIN activity in mitotically arrested cells (Guertin et al, 2002; Johnson et al, 2012). In a subsequent paper, this is considered as a separate checkpoint, independent of the classical SAC (Johnson et al, 2013). The resetting of SIN proteins at the end of mitosis is controlled by Nuc2p, a subunit of the anaphase promoting complex APC/C (Kumada et al, 1995). This role of Nuc2p is achieved by interfering with the interaction between Spg1p and Cdc7p, maybe by stimulating the activity of Byr4p-Cdc16p (Chew & Balasubramanian, 2008), and is independent of its APC/C function.

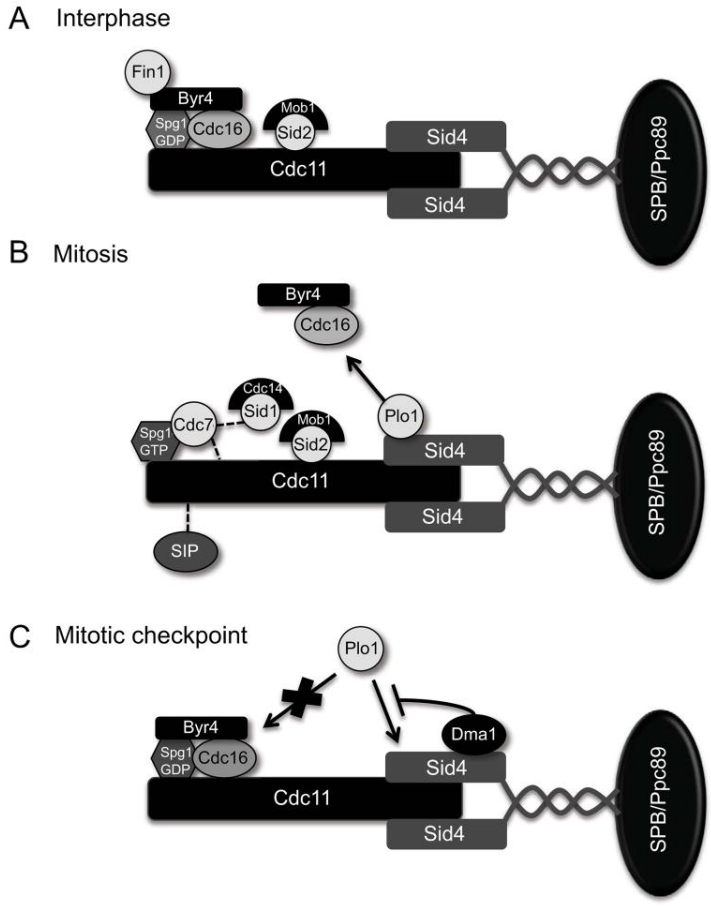

Figure 2.1: Organization of SIN components on the SPBs. *The SPB scaffold is formed by Ppc89p, Sid4p and Cdc11p.* **A)** *In interphase, Spg1p is inhibited by the Byr4p-Cdc16p complex. The inhibition is relieved in early mitosis* **(B)** *allowing the association of downstream kinases*

Cdc7p, Sid1-Cdc14p and Sid2p-Mob1p on the SPB. In case of an active mitotic checkpoint **(C)***, Dma1 inhibits SIN signaling by preventing the recruitment of Plo1p on the pole. Figure source: (Johnson et al, 2012)*

The SIN controls many aspects of cytokinesis including the assembly of the contractile ring and synthesis of the division septum (Goyal et al, 2011). Our goal is to describe the qualitative behavior of the system, investigate the role of each SIN regulator and potentially predict unknown mutant behaviors. Towards this end we decided to adopt a Boolean modeling approach. The choice of qualitative modeling was based on its suitability to simulate systems with limited kinetic data, as well as their computational efficiency, that permits large numbers of *in silico* experiments even in networks with hundreds of nodes.

Models of the fission yeast SIN have already been generated; Csikasz-Nagy et al. (2007) and Bajpai et al. (2013). In the study by Csikasz-Nagy *et al* the timing of septation in wild type and mutant cells was described using a minimal, continuous model. The SIN components were treated as two groups, the "Top of SIN" and "Bottom of SIN", with Sid1p localization to the SPB being the pivotal event that differentiates the two groups (Csikasz-Nagy et al, 2007). Subsequent analysis, (Wachowicz et al, 2015) has revealed that Sid1p is already associated with the SPB early in mitosis, suggesting that this analysis may be oversimplified. In the successive model (Bajpai et al, 2013), the asymmetric distribution of molecules at the SPBs was analyzed using a simple, non-linear model of two antagonistic molecules. The model was also extended to incorporate key regulators of the SIN (Bajpai et al, 2013).

In this work, we present an extended, Boolean model of the SIN, comprising most known SIN components and regulators as individual, experimentally testable nodes. The Boolean framework allows us to perform *in silico* knock-out and "constant activation" experiments for every combination of molecules present in the model, and to assess phenotypic predictions that could be subsequently validated experimentally. Our model provided useful insights for several aspects of SIN regulation such as the role of Fin1p, the inhibitory function of Nuc2p in interphase, as well as an *in silico,* counter-intuitive, double mutant phenotypic prediction. The model predicted that Sid4p mutant cells would septate if they express Cdc7p in high levels. The prediction has been experimentally confirmed. This work serves as a good example of the use of qualitative modeling in hypotheses generation and prediction of experimental outcomes in otherwise complicated and long experiments.

Results

Model construction through expert biocuration

For the gene regulatory network construction of the SIN we chose an expert biocuration approach (Bateman, 2010; Poux et al, 2014), taking advantage of the long-term expertise in the Swiss-Prot group. Experimentally determined interactions specific to the SIN, were retrieved, structured, curated and annotated from the literature and from available knowledge databases (for example PubMed, iHOP, UniProtKB/Swiss-Prot, ChEBI). To establish the model, we started by adding the main SIN signaling regulators such as the GTPase Spg1p, its effector kinase Cdc7p and the GAP Byr4p and Cdc16p (Furge et al, 1998; Krapp et al, 2008; Schmidt et al, 1997). We proceeded by adding the SPB scaffold for SIN, comprising Ppc89p, Sid4p and Cdc11p (Krapp et al, 2008; Rosenberg et al, 2006). Subsequently, additional regulators were added to this core unit, to complete a first working model. The collected knowledge was stored in a structure formed of pairwise interactions and regulations that include information about participating components, the origin of publications (PMID), the evidence used to evaluate the interaction was mentioned and a confidence assessment as an evidence tag from the biocurator. In detail, the database contains the following columns:

1. Node 1: The name of the first element of the interaction, the one that acts as activator or inhibitor
2. Action: A symbol characterizing the type of interaction as activation (→) or inhibition (⊣).
3. Node 2: The name of the second element of the interaction, the one that gets activated or inhibited
4. Node 1 type: The type of node 1. Can be protein, complex or miRNA
5. Node 2 type: The type of node 2. Can be protein or miRNA
6. UniProt ID 1: Reference of node 1
7. UniProt ID 2: Reference of node 2
8. PMID: Literature reference of the interaction
9. Class: A letter characterizing the confidence level of the interaction. It can be one of the following:

   Sure (S), when the interaction is confirmed or known in textbook, and/or already in the UniProt general annotation lines. Sure interactions are generally associated with many PMIDs.

Unsure (U), when the interaction is shown once and/or not confirmed by others, or when the authors are not confident about the results.

Inferred (I), when there are no results for the network in question, but the interaction was found in different cell types and/or organisms. It may also refer to cases where the information is inferred to all protein isoforms of a gene without confirmed results. Inferred interactions might be associated with more than one PMIDs.

Contradictory (C), when the interaction is based on contradictory results.

10. Evidence tag: Short extract from the publication where the interaction is mentioned.

The constructed prior knowledge network (PKN) consists of 41 nodes (gene products, proteins and complexes) and 108 directed edges (figure 2.2). The regulatory information is the result of the curation of 67 published scientific papers. The most recently published interaction contained in this model is the inhibitory regulation of CDK and Plo1p upon Byr4p recently published by (Rachfall et al, 2014).

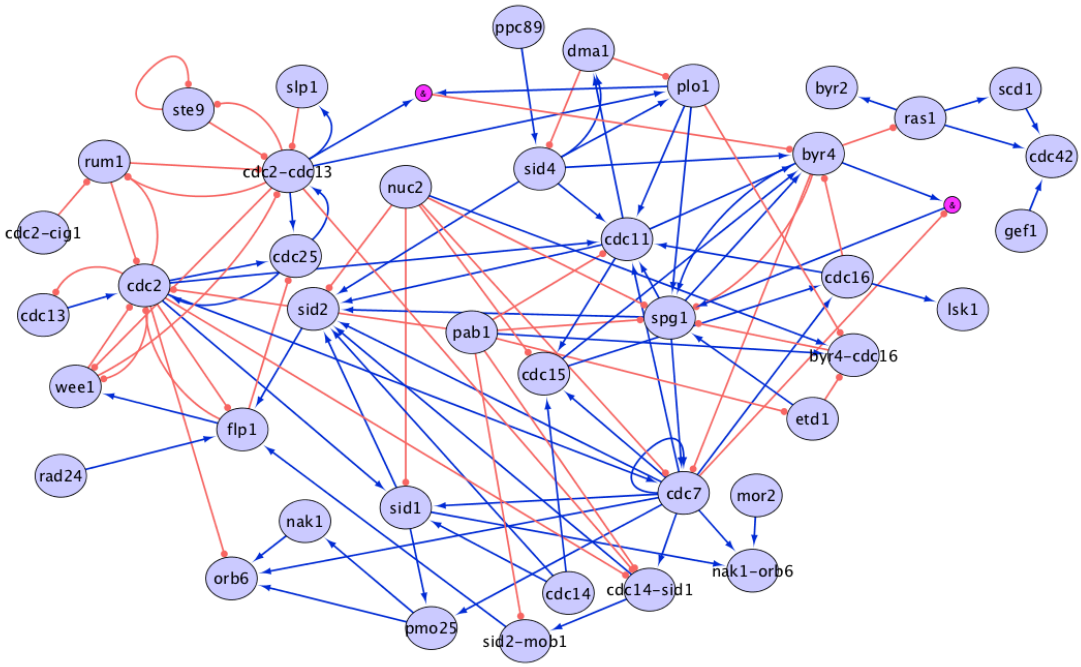

Figure 2.2: The extended Boolean SIN model. *The initial, prior knowledge network, manually re-constructed from the literature. Nodes represent proteins and complexes that take part in the regulation of the SIN. Blue arrows indicate activation events and orange circles inhibition events. More complicated logical functions are also encoded in the model, such as AND, OR and XOR regulatory gates. The pink circles, for example, are AND nodes, representing interactions like "byr4 AND NOT cdc7 activate spg1".*

The model interactions were classified as activations or inhibitions and they were represented in the network as a combination of Boolean functions including AND, OR, NOT, IAND, XOR, XNOR.

Qualitative model simulation

Despite the intensive study of the SIN over the past decades, there is little kinetic data for the protein interactions described in the literature that form the basis for our model. Obtaining such spatiotemporal data is experimentally difficult and represents one of the major challenges in systems biology research. For the simulation of the SIN model we adopted a qualitative Boolean approach, which has been successfully used in several other contexts (Albert & Othmer, 2003; Azpeitia et al, 2011; Davidich & Bornholdt, 2008; Giacomantonio & Goodhill, 2010; Li et al, 2004; Li et al, 2006; Morris et al, 2010; Saez-Rodriguez et al, 2007; Samaga et al, 2009; Sanchez et al, 1997; Schlatter et al, 2012; Schlatter et al, 2009; Veliz-Cuba & Stigler, 2011).

Asynchronous updating was chosen for the SIN model simulation in this study, since it assumes non-synchronous regulatory events, which places it closer to reality. However, the challenge in asynchronous update lies in interpreting the simulation trajectories; in stochastic asynchronous simulations, the same initial state can lead to different trajectories in the state space, due to the stochasticity of the updating scheme (Wang et al, 2012). The simulation algorithm used was based on *genysis*, a tool for synchronous and asynchronous modeling of gene regulatory networks, based on reduced ordered binary decision diagrams (ROBDDs) (Garg et al, 2008). The algorithm identifies all steady states / attractors that can be reached, by efficiently investigating all possible asynchronous state transitions. It is worth noting that for the construction and simulation of the SIN model we assumed that, for scaffold proteins, “active” (or, in Boolean terms, 1) corresponds to a state that permits the assembly of signaling complexes.

Note for the reader: To facilitate the discussion on the model’s design and results, the nomenclature for model nodes used is the protein name in lower case, without the letter p in the end; for example cdc7 instead of Cdc7p. For the interpretation parts, the appropriate nomenclature is used for gene and protein names.

Initial simulations

The asynchronous simulation of the initial model candidate resulted in 4 attractors (figure 2.3).

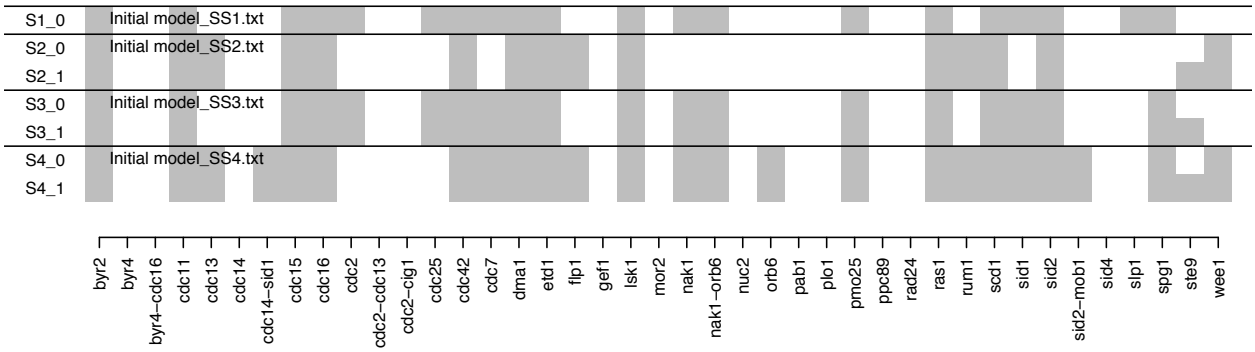

Figure 2.3: Initial model steady states and attractors. *The 4 attractors deriving from the initial simulation. The attractors are shown as horizontal lanes, separated by black lines. For each attractor, every model state is represented by a line. Grey boxes correspond to active (1) genes and white boxes to inactive (0).*

Unperturbed model attractors can provide information about the phenotypes of the system under study. In our case, we would expect the presence of a steady state (or attractor) similar to a wild type “septating” yeast, and possibly of one resembling the case when septation is not triggered. Indeed, computing the attractors there were three attractors where the main scaffold and SIN components (cdc11, spg1, sid1, cdc7) were active as expected (Bardin & Amon, 2001; Sparks et al, 1999) and one where there was a formed scaffold but inactive SIN components (attractor SS2, as represented at figure 2.3). Other participating nodes yielded unexpected results, such as sid2 and sid4 that had the same state in all attractors contrary to expectations (Bardin & Amon, 2001; Sparks et al, 1999). Specifically, sid2 appeared always active, since the only inhibitory incoming edge from Nuc2p was always inactive due to missing regulation. Sid4 remained inactive due to a positive feedback loop with its inhibitor, dma1.

Model evaluation and scoring development

In order to evaluate the model, we designed a scoring method, based on computing a similarity score between the occurring steady states and experimentally characterized gene perturbations. To score the model, we selected a test set of 7 nodes: cdc16, cdc7, spg1, sid1-cdc14, sid4, sid2-mob1 and cdc11, with activation states indicative of the expected phenotype. The selected genes and complexes have known activation states, both at the multiseptated and non-septated phenotype, as shown in table 2.1. Additionally, we selected a number of knock-out and over-expression perturbations that have a known phenotypic outcome as a way to benchmark the predicted steady state (table 2.3).

Scoring test set genes and activation states

	Multiseptation	No septation
cdc16	0	1
cdc7	1	0
spg1	1	0
sid1-cdc14	1	0
sid4	1	0
sid2-mob1	1	0
cdc11	1	0

Table 2.1: Test set gene states. *The expected activation state of all genes and complexes forming the scoring test set in the case of multiseptated and no septated phenotype.*

Gene perturbation set and phenotypic outcomes

Experiment	Phenotype
spg1 KO	no septum
spg1 OE	multiseptum
cdc16 KO	multiseptum
byr4 KO	multiseptum
byr4 OE	no septum
cdc7 KO	no septum
cdc7 OE	multiseptum

Table 2.3: Gene perturbation set. *7 gene perturbations used for model candidate scoring, and their phenotypic outcome.*

For a given model candidate, we performed the test gene perturbations *in silico*, by perturbing the gene state to remain inactive (for KO) or active (for OE) during the whole simulation. For each occurring steady state we defined a vector **x** containing the states of test genes 1,2,…,N.

$$\mathbf{x} = (x_1, x_2, ..., x_N)$$

Gene states could take the values 1 for active, 0 for inactive and 0.5 for oscillating cases that can be observed in cyclic attractors. The expected values of the test genes for each perturbation, according to prior knowledge, are given by the respective vector **y**

$$\mathbf{y} = (y_1, y_2, ..., y_N)$$

Therefore, the score s for each steady state for N genes is given by

$$s = \sum_{n=1}^{N} |x_n - y_n|$$

In cases where more than one steady state for a perturbation experiment existed, the steady state with the highest score (*smax*) was selected. The overall model score S was defined as the average of *smax* for M test set perturbations.

$$S = \frac{1}{M}\sum_{m=1}^{M} smax^{(n)}$$

A practical example of the scoring process is presented schematically in figure 2.4.

In the case of our model where 7 nodes are used as test set, the score can range from 0 to 7. A score of S=7 indicates that the model fully describes the selected perturbations, whereas a model with S=0 has no descriptive power for the test perturbations. Our initial model, as constructed from the prior knowledge body, had a score S=4.81.

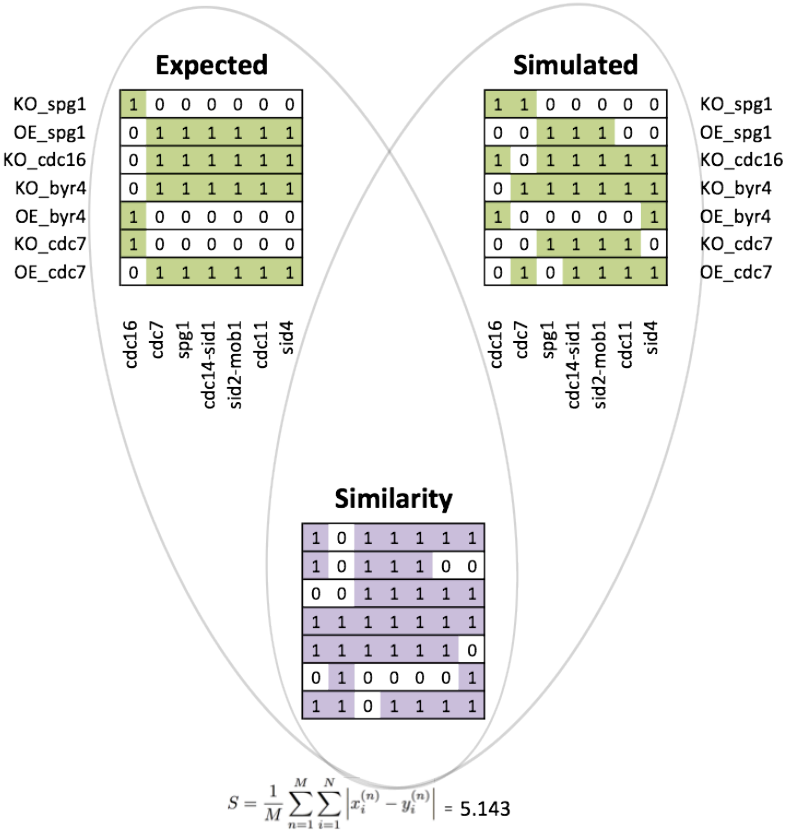

Figure 2.4: Model scoring process example. *A number of gene perturbations with known phenotypic effect were selected as test set for the model optimization. During the optimization a number of genes were tracked, chosen because their activation state was known in the perturbation test set. For each model candidate, the perturbation test set was performed in silico and the simulated results were compared to the expected. The model then received a similarity score.*

Model optimization and refinement

In order to correct the model to better describe the current state of our knowledge, we established an optimization procedure. The first step of the optimization workflow was to perturb the initial candidate model and score it. The model would then go through a refinement stage, which included several iterations of additions, modifications or deletions of regulatory edges, resulting in a number of new model candidates. Thereafter, the candidates would be perturbed and scored. Models that scored higher than the initial models were kept for further optimization iterations. This method produced a tree-like structure of model candidates, with some refinement attempts being rejected immediately while other branches were further explored.

As mentioned, during the refinement step of the optimization a number of regulatory edges were added, deleted or modified. The decision for these changes was based on a number of principles, listed below.

Feedback loops: In biology, a feedback loop is a circuit of any length that begins and ends at the same node. Feedback loops can be positive or negative, depending upon the parity of the number of negative interactions in the loop. Positive feedback loops are indicative of multistationarity and negative feedback loops of homeostasis (Plahte E., 1995) (Ferrell et al, 2011).

Identifying the feedback loops of the system is of crucial importance for the decision making during optimization. The aim is to include and maintain the biologically known feedback loops of the system throughout the optimization process. In practice, this corresponds to modifying the network topology by adjusting the regulatory links and the nodes connectivity.

An example of applying this principle at the SIN model optimization is the deletion of the sid4 inhibition edge by dma1, which was part of the initial network (as described in papers (Goyal et al, 2011; Johnson & Gould, 2011)). Dma1 used to be under constant positive regulation by cdc11, which consequently placed sid4 under constant negative regulation. Therefore, sid4 would stay in an inactive state. The problem was solved after removing the inhibition interaction.

Source nodes: There are cases where nodes act as activators/inhibitors, but their regulation is absent from the model, as they are input feeds for the system. In modeling terms, these nodes (i.e. source nodes) have only outgoing edges but no incoming. Source nodes can be a major problem, as they appear constantly inactive during the simulation, which can bias the state of every entity

under their regulation. In this case missing regulation can be added directly at the source node or the compounds under its regulation.

An example of altered information flow by source nodes at the SIN, was the case of the byr4-cdc16 complex. The complex was inhibited by etd1 which had only one incoming inhibition edge by pab1, a source node. Being a source node, pab1 remained inactive, therefore fixing etd1 in a permanently active state. Consequently, the byr4-cdc16 complex was constantly under a dominant inhibitory regulation. The problem was solved by removing the etd1⊣byr4-cdc16 inhibitory rule, an interaction that was characterized as unsure by our biocuration (Lahoz et al, 2010).

Conflicting information or low confidence level: The confidence level assignment for every interaction in the database is very helpful during optimization. Interactions that are labeled as uncertain or conflicting were selected more easily to undergo re-evaluation. Especially in the case of conflicting information, different suggested scenarios can be used as alternative regulations. For example, according to Li et al. (2000), Byr4p localization to SPBs during interphase maintains Spg1p in an inactive form (Li et al, 2000). However, the opposite statement is claimed by Furge et al. (1998), according to which in early mitosis, when Cdc16p leaves the pole (Wachowicz et al, 2015) “Byr4p appears to stabilize the GTP bound, active form of Spg1p” (Furge et al, 1998). In this case, the regulatory rules included in the model were corresponding to different stages of the cell cycle. Having both regulations during the same simulation was not meaningful, since byr4 was acting both as an activator and an inhibitor of spg1. The issue was addressed later, when we partially implemented cell cycle regulation in our model (this is described in detail later in the thesis).

Assessment of single perturbation analysis

After going through the optimization stage, we started exploring and testing the best model. First, we performed all single gene perturbations *in silico*, both KO and OE. We then used exploratory principal component analysis (PCA) on all occurring steady states, in order to observe the contributions of genes to phenotypes and the number of possible derived phenotypes. Figure 2.5 shows the PCA plot, for the first 2 principal components, which captured most of the data variation as indicated by the eigenvalues.

The PCA analysis revealed, that the steady states of all possible perturbations tend to form a small number of clusters. For each perturbation there were 2 to 4 reached steady states, out of the 2^{43} state possibilities. This observation agrees with experimental observation, where a small number

of phenotypes can be observed during septation. Moreover, most of SIN scaffold and signaling genes have a similar contribution in the steady states, as expected.

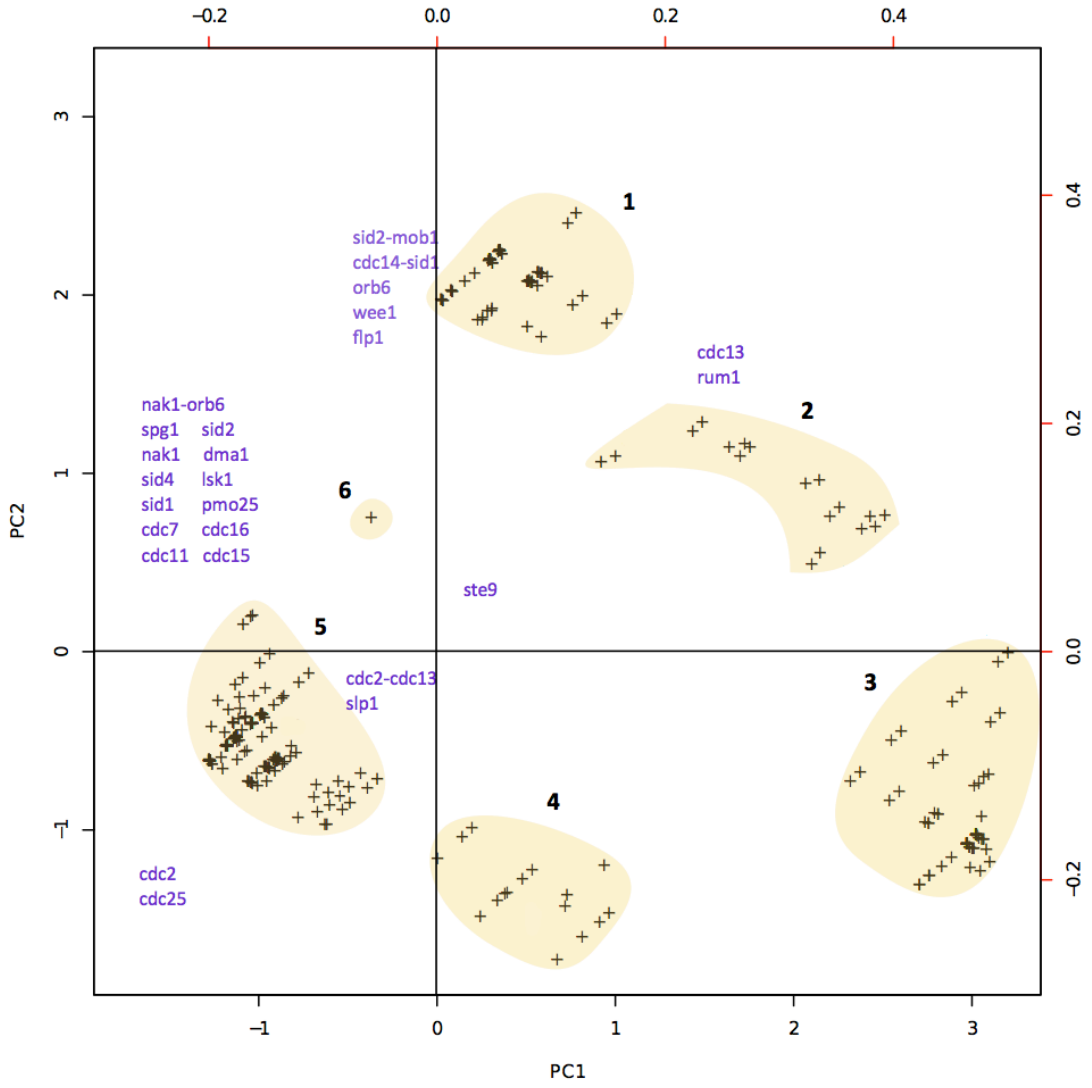

Figure 2.5: PCA analysis of all single gene perturbations. *Principal component analysis of all single gene perturbations, represented as a biplot. Each cross corresponds to a steady state of a perturbation (e.g. rum1 k.o., steady state No1). 6 main clusters are formed. Nodes with small contribution to the variation were excluded from the graph to avoid a noisy representation around 0.*

Assessment of double perturbation prediction

Experimentally performing single gene perturbations in fission yeast is time consuming and has been done by yeast biologists for many years, since most gene deletion mutants are available. However, moving towards more combinatorial, complicated experiments, as, for example, multiple perturbations, soon becomes challenging. Following the strategy of single perturbations, we

examined all possible combinations of double perturbations, KO and OE, *in silico*. The 3362 experiments performed, yielded many different steady states. To evaluate the similarity of the steady states of each experiment to the phenotypes of interest, we used the evaluation system described above, scoring the experiments both for their similarity to the multiseptated and to the non septated phenotype. Figure 2.6 shows a heatmap of all double perturbations and their similarity to the multiseptated and the non septated phenotype.

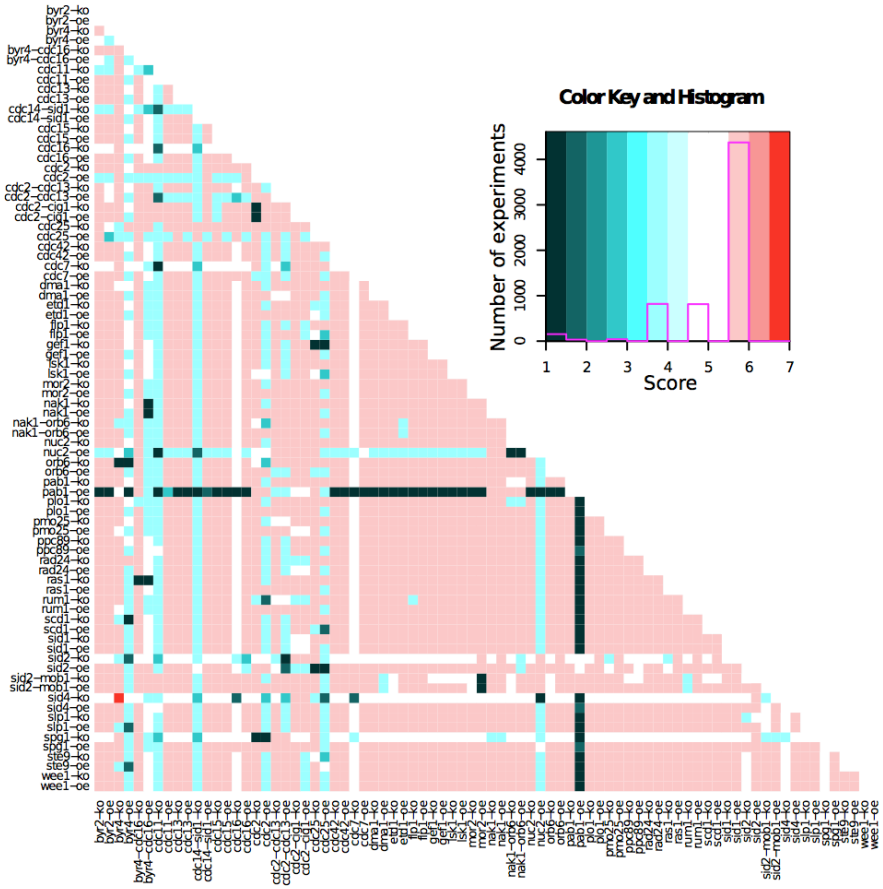

Figure 2.6: In silico predictions of double gene perturbations. *Heatmap of steady state similarity to multiseptated phenotype, for all possible combinations of in silico double perturbations. The state of 7 genes was used to score the perturbed networks. For example, in 5a, byr4 k.o. – sid4 k.o. has a score of 7, which indicates that the activation state of all test genes is in agreement with the multiseptated phenotype.*

Model optimization assessment

To assess the outcome of the model optimization process, we compared the score distribution of all double mutants to equivalent distributions where a random set of genes is selected. 1000 score sets were selected by a randomization process, having the same characteristics as our score S.

Our hypothesis was that if the distribution of random scores is similar to the distribution of S, our model will validate any random truth, which would downgrade its predictive value. Figure 2.7 shows the density distributions of 1000 random score sets, in contrast to the true score for the multiseptated phenotype. Random scores produce a normal distribution as expected whereas the distribution of true score S is significantly different.

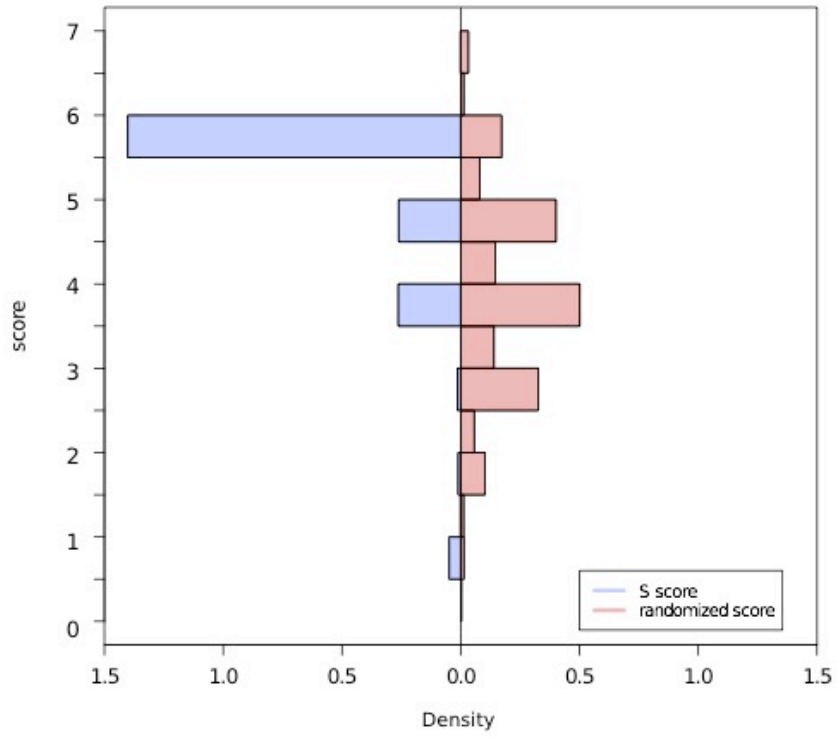

Figure 2.7: Model optimization assessment. *Score distribution of best scored steady states for multiseptated phenotype using the developed score (S) and 1000 sets of randomized scores. Each randomized set consists of the same number of genes as the original score. The distributions reveal a significant difference between the actual model score and hypothetical scores.*

Byr4 KO – *sid4* KO phenotype prediction

A counter-dogmatic prediction was generated when we performed the double perturbation *in silico* experiments, according to which in a double KO of *sid4* and *byr4* we should observe a multiseptated phenotype. The prediction was experimentally tested using 2 approaches.

Tetrads of crosses between *sid4::KAN*^{R} *pREP-sid4 ura4-D18, ade6-M210, leu1-32* and *byr4::ura4+ lys1::byr4-GFP ura4-D18, leu1-32* were dissected and the genotypes of growing colonies were assessed using replica-plating on selective media. *Byr4* and *sid4* are on different

chromosomes (I and II, respectively). Mendelian segregation predicts that 25% of progeny should be wild-type, 25% should be either single mutant and 25% should be double mutants. Since the phenotype of the double mutant was not predictable, we examined the phenotype of the colonies from non-parental ditype (NPD) tetrads, where two spores give rise to a wt colony and two should be double mutants. In all these cases, the germinated candidates of double *sid4* and *byr4* mutants (12/12 cells) were elongated and lysed, with no visible septa, suggesting that double null mutant cells do not septate.

To confirm this result, the spores of cross *sid4:: KAN^R pREP-sid4 ura4-D18, ade6-M210, leu1-32* and *byr4::ura4+ lys1::byr4-GFP ura4-D18, leu1-32* were incubated in 36°C in minimal medium (Hergovich & Hemmings) supplemented with 100mg/l adenine and leucine in order to permit the germination of *byr4*::ura4+ *sid4*+ and *byr4*::ura4+ *sid4- KAN^R* cells. Cells were incubated for 13.5h, fixed with 70% ethanol and stained for DAPI and Calcofluor for nuclear and cell wall/septum labelling, respectively. Out of 238 cells examined, 117 (49%) showed the multiseptation phenotype expected for *byr4* null *sid4*+ cells, while 121 (51%) cells showed an elongated multinucleated phenotype, without septa suggesting that double null mutant will not septate. Therefore, the observed phenotype (no septum) did not match the model's prediction (figure 2.8).

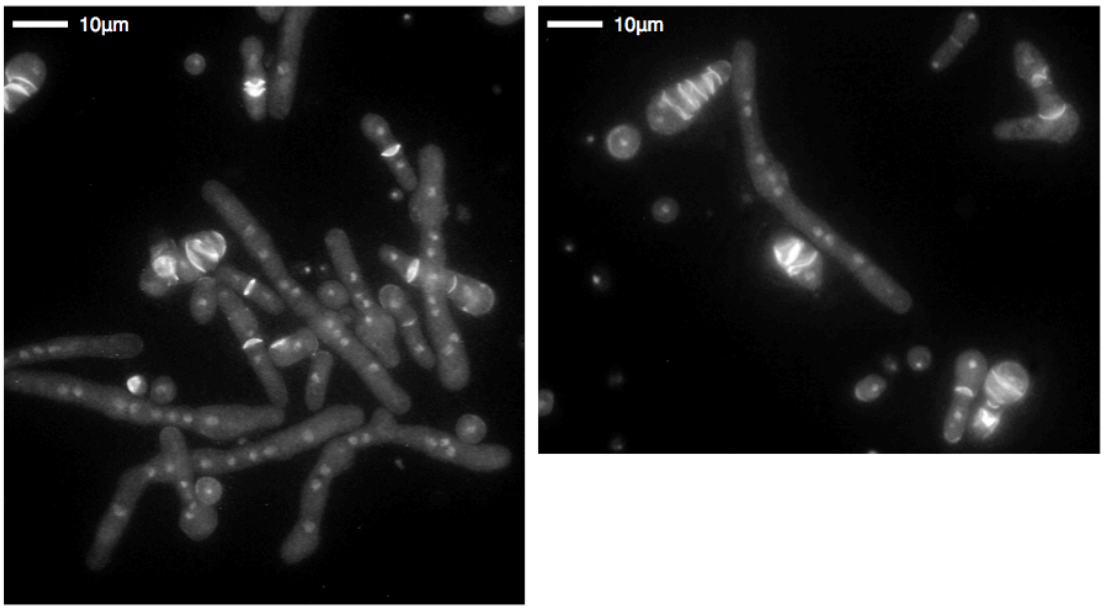

Figure 1.8: Experimental evaluation of *byr4* KO – *sid4* KO prediction. *The imaging results of germinated spores of byr4-null sid4+ and byr4-null sid4-null cells. 49% of the examined cells were multiseptated, which is typical for byr4 null sid4+ cells. The rest of the cells were elongated and multinucleated without septa, suggesting that the double null mutant does not septate.*

Even though the prediction was not successful, the experimental outcome was used to modify our scoring rules. The fact that a *byr4* KO - *sid4* KO results in a non septated phenotype was added as an additional rule in our scoring.

Change of strategy to include effective CDK regulation

During our initial modeling efforts, correctly describing CDK regulation through cell cycle was always problematic. A weakness of Boolean modeling, inherent to the definition of the method, is that there is no possibility to observe the effects of different activation levels of a protein upon the system, neither to control its activation curve during the simulation. What can be observed and adjusted are the initial conditions of the simulation, and the regulatory rules of the model. Given this limitation, we were unable to modify the regulation of Cdc2p/CDK1, in a way that is would reproduce the activity fluctuation which is important during cell cycle and for SIN related events (Stern & Nurse, 1996). The problem remained, even when we concatenated the cell cycle Boolean module that was publicly available (Davidich & Bornholdt, 2008). We, therefore, decided to change our strategy and restructure our model so that we have control over the levels of CDK during the simulation. This new approach allowed us to accurately describe SIN related events during different stages of the cell cycle and was successfully used as a predictive tool for multiple perturbation scenarios. The construction method and characteristics of the restructured model, as well as the optimization and prediction processes that followed, are described in detail below.

SIN modeling using CDK switches

Cdc2p/CDK1 influences the SIN both positively and negatively. Active Cdc2p inhibits the SIN early in mitosis; its inactivation is required for septum formation and to establish SIN protein asymmetry (Chang et al, 2001; Dischinger et al, 2008; Guertin et al, 2000; He et al, 1997; Yamano et al, 1996). Furthermore, Cdc2p and the Byr4p-Cdc16p GAP may cooperate to prevent septation in interphase (Cerutti & Simanis, 1999). However, Cdc2p and Plo1p also collaborate positively to ensure removal of Byr4p from the SPBs and facilitate SIN signaling in anaphase (Rachfall et al, 2014). Failure to increase CDK levels during early mitosis will block cytokinesis, since the cells do not enter mitosis. However, failure to decrease CDK levels through mitosis will block cytokinesis. Thus, CDK levels need to increase to permit entry into mitosis, after which cytokinesis will occur. However, this will only happen once CDK activity decreases to a very low level, and cells exit mitosis. The model must therefore accommodate these CDK-dependent regulatory events.

As a first step towards this goal, we introduced three independent nodes for CDK, representing the CDK levels before, during and after mitosis. CDK-L corresponds to the low CDK levels during interphase; these prevent re-replication of DNA, but are insufficient for entry to mitosis (Baum et al, 1997; Coudreuse & Nurse, 2010), CDK-H represents the high level of CDK activity found in early mitosis. Finally, CDK-0 represents the very low CDK activity in late mitosis as cells undergo the M-G1 transition. This multi-node representation of CDK allows us to describe the SIN-related phenotypes corresponding to several stages of the cell cycle, using the CDK nodes as inputs. For example, setting CDK-L constantly on, indicates that we are simulating the events during interphase, while, CDK-H on represents early mitosis and CDK-0 on represents late mitosis (figure 2.9).

Model refinement and simulation results

This model configuration that uses CDK levels as control nodes for the simulation of cell cycle events, allowed us to clearly define the expected steady states of the system and set our refinement strategy. First, we attributed the PKN interactions involving CDK to the correct CDK node. For example, an activation link from CDK-H was added towards Plo1p (figure 2.9), which in turn will reinforce the activity of CDK in a positive feedback loop (Grallert et al, 2013; Tanaka et al, 2001). Following the attribution step, the model was evaluated using a number of well characterized *in silico* perturbations whose phenotypic consequences are known; knock-out of *cdc11*, *spg1*, *cdc16*, *byr4*, and *cdc7*. For the evaluation, the above 5 knock-out perturbations were simulated, by setting the corresponding node to 0 throughout simulation. A fixed set of nodes, with activation states indicative of the expected phenotype was selected to score the model's ability to correctly reproduce the mutation's effects. The scoring set includes *sid4*, *cdc11*, *byr4-cdc16*, *spg1*, *cdc7*, *sid2-mob1* and *sid1-cdc14*. For each *in silico* perturbation, the resulting steady states were evaluated according to the number of the scoring set nodes that had the expected activation state.

We proceeded by refining the connections within the network. A refinement cycle consisted of altering an edge of the network, perturbing the model and evaluating the simulation outcome of the perturbations test set. The alterations could involve additions and deletions of regulatory edges, or modifications of the existing regulatory rules. The reasoning behind each change of the model's regulatory rules was based on several factors, such as the confidence level of each interaction, coupled with information from the published literature, as well as forming alternative logical rules of the given information to better represent the biological reality of the interaction. For example, "A inhibits B" can be alternatively encoded as "NOT A activates B", and is more suited for cases where the inhibition is not dominant. During this process we maintained the known, required

connections of the model and minimized the model's complexity by removing nodes that no longer served any regulatory role in the model. An example of the latter is the removal of cell cycle regulatory elements such as Cdc25p, Wee1p, Slp1p and Rum1p, which were removed after the simplification of the cell cycle representation using multi-node CDK. The final, optimized network is presented in figure 2.9. A full list of the edges comprising the final network, together with the justification for the inclusion of each edge, can be found in the appendix.

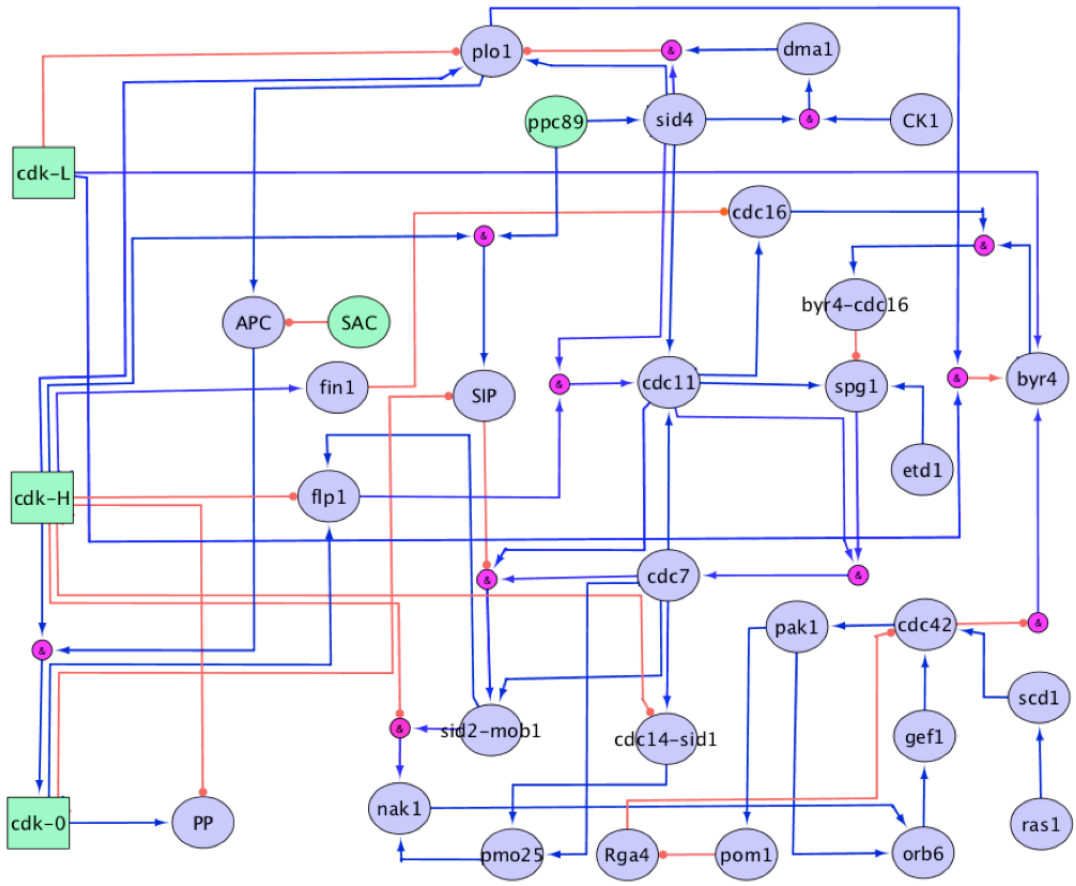

Figure 2.9: The CDK-switch model. *The restructured model, using 3-node representation of CDK activity, in its final, optimized form. Nodes in green are used as switches, and they are turned on to represent different stages of the cell cycle: interphase, early mitosis and late mitosis. SAC: spindle assembly checkpoint, APC: anaphase-promoting complex.*

The optimized model was used for *in silico* experiments in which a combination of nodes was perturbed and the phenotypic outcome in the interphase, early mitosis and late mitosis CDK-states were determined. A simulation of the wild type model, where no perturbation is introduced, is presented in figure 2.10.

To simulate interphase, CDK-L is set to 1, and Ppc89p is set to 1 as well, to permit "binding" of scaffold proteins to the SPB. In interphase, the model simulation results in a steady state where

Byr4p and Cdc16p are present and able to form the GAP complex, therefore active. The scaffold proteins Sid4p and Cdc11p are also present (therefore "active" according to our initial assumption for scaffold molecules), but no SIN signaling occurs due to the inhibitory effect of the Byr4p-Cdc16p GAP.

Early mitosis is simulated by setting CDK-H, Ppc89p and the spindle assembly checkpoint (Musacchio) to 1. Activation of the SAC inhibits the anaphase promoting complex (APC), therefore blocking the decrease of CDK levels and mitotic exit (Jia et al, 2013; Musacchio, 2011). The SIN scaffold is still formed, as expected. Cdc16p is absent from the SPBs in early mitosis, preventing formation of the GAP. This allows SIN signaling to initiate, and we observe that all the main components of the SIN are active (Plo1p, Spg1p, Cdc7p, Sid2p-Mob1p), apart from Cdc14p-Sid1p, which is inhibited by high CDK activity (Dischinger et al, 2008; Guertin et al, 2000).

Late mitosis is represented by setting CDK-0 and Ppc89p to 1 during the simulation. There are 2 resulting steady states of the system simulation. In one state, the SIN signaling scaffold is present, the Byr4p-Cdc16p complex is formed, and all SIN components, except Plo1p, are inactive. In the other state, Byr4p-Cdc16p is not active, and all proteins of the SIN scaffold and signaling including Cdc14p-Sid1p are active. Intriguingly, these resemble the asymmetric constellation of proteins observed at the old and new SPBs in late anaphase B (see (Goyal et al, 2011; Johnson et al, 2012; Simanis, 2003) for review), with the exception of Sid2p-Mob1p, which is present on both SPBs, but only active in one of the two states of the model. Setting GAP function to 0 abolishes the state that resembles the oSPB. Though it is often assumed to be the case, there is scant evidence to support the view that localization of SIN proteins to the SPB is a faithful readout of their *in vivo* activity. There is no data addressing whether Sid2p signals from one or two SPBs in late anaphase. Future experiments will investigate this. A detailed heatmap showing the activation state of all nodes of the model for all experiments presented herein can be found in the appendix.

Model evaluation by assessing experimentally validated *in silico* perturbations

The optimized model can describe the SIN related events during interphase, early and late mitosis. In order to evaluate the model's ability to describe current knowledge regarding *S. pombe* mutants, we performed a series of *in silico* knock-out and constant activation experiments mimicking those described in the literature that have an established phenotype. Figure 2.10 summarizes the steady states yielded after simulating interphase, early and late mitosis behavior of core gene mutants. Interestingly, in all the *in silico* experiments we obtained steady states where the nodes displayed, overall, the expected activation state. More specifically, *cdc11* knock-out completely blocks

septation. Both, *byr4* knock-out and *cdc16* knock-out have the same effect, which is failure to inhibit SIN signaling, and therefore SIN triggering in interphase. In a knock-out of either *spg1* or *cdc7*, signaling fails, with Spg1p still getting activated in *cdc7* deletion, indicating that Spg1p acts upstream of Cdc7p, as experimentally proven.

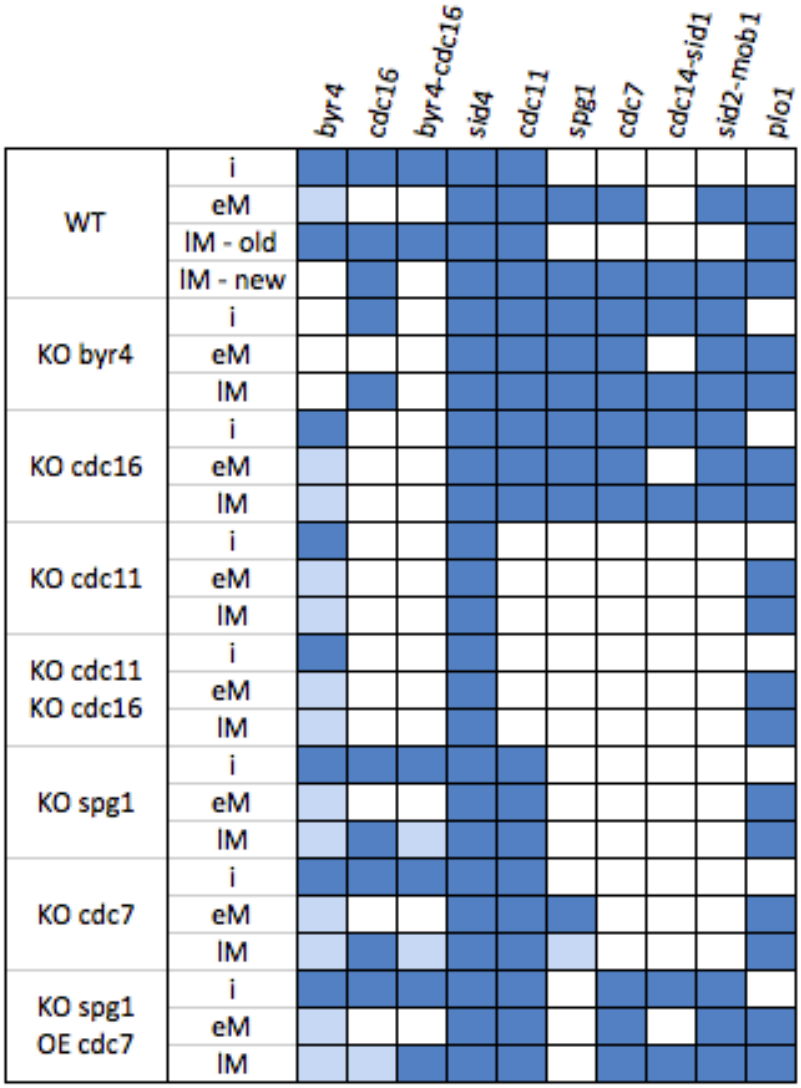

Figure 2.10: In silico steady states of the SIN, in wild type and mutated cells. *Steady states deriving from simulations performed on the final model. The boxes on the left indicate the experiments performed, which can be knock-out (KO) or over-expression (OE). When there is more than one gene in the box, it is a double perturbation. For each perturbation, 3 experiments were performed: interphase simulation (indicated as i), early mitosis (eM) and late mitosis (lM, with suffixes new and old when there are 2 resulting steady states, indicative of late mitosis asymmetry). Blue boxes correspond to active proteins, white to inactive and light blue to proteins that can be either active or inactive at the resulting steady states of the system.*

Apart from the experiments that were used as training set for the model refinement, we performed double mutant experiments that had not been used as part of the test set used to formulate the model. These experiments assess the predictive value of the model, as the *in silico* predictions are in accordance with the expected results. Specifically, the double delete of *cdc11* and *cdc16* simulation predicts that cells should not septate, as shown in figure 2.10, with supporting evidence from the literature (Marks et al, 1992). A *cdc7* over-expression in *spg1* delete will septate, in agreement with studies (Schmidt et al, 1997). Moreover, *cdc7* over-expression will produce septation in the absence of Cdc11p (figure 2.12), as confirmed by the literature (Fankhauser & Simanis, 1994). In this project, setting a node to 1 throughout the simulation has been used to simulate over-expression *in silico*, except in cases where it is known that the over-expression phenotype results from an indirect effect, such as the titration of another protein.

Other *in silico* experiments performed during the optimization provided us with insights into potential knowledge gaps regarding SIN regulation, as well as the limitations of our model. One such example was a prediction that a *byr4*-KO *sid4*-KO should septate. The relevant strains were constructed and analyzed, and the cells were found not to septate. This allowed us to refine the model, by identifying regulatory links that would permit this state to be achieved and target them as candidates for edge deletion. Moreover, the *nuc2* inhibitory links that were present in the PKN revealed our limitation of describing events that occur at the end of septation and the incomplete regulatory inputs to *cdc16* helped us discover a potential link with *fin1*. The *nuc2* and *fin1* cases are discussed in detail below.

Does Nuc2p have a role in regulating the SIN in interphase?

Increased expression of the APC/C component *nuc2* blocks septation, while incubation of *nuc2-663* at low restrictive temperature results in cutting of the cell (Hirano et al, 1988; Kumada et al, 1995). Analysis of how the SIN is reset at the end of mitosis revealed an APC/C-independent role for Nuc2p (Chew & Balasubramanian, 2008). Nuc2p interferes with formation of the Cdc7p-Spg1p complex, possibly by stimulating the GAP activity of Byr4p-Cdc16p. Since our current model does not encompass resetting of the SIN, we tested whether the inhibitory link of *nuc2* towards the Cdc7p-Spg1p complex should be included. Thereafter, we modeled the effect of inactivating Nuc2p *in silico* upon SIN behavior in interphase. The predicted outcome when including the Nuc2p inhibitory link was two steady states; one with inactive SIN and one with cells that septate in interphase.

To test whether this would be the case *in vivo*, the strain *nuc2-663 leu1-32* was arrested in S-phase by growth in medium containing 12mM hydroxyurea (HU). After 5h at 25°C, cells were shifted to 36°C to inactivate Nuc2p, and samples were fixed at hourly intervals, and stained with DAPI (to reveal DNA) and Calcofluor (to reveal the division septum). Before shift to 36°C *nuc2-663* (93%) and *nuc2*$^{+}$ cells (97%) were mononucleate with no septum, consistent with an arrest in interphase. Examination of *nuc2*$^{+}$ cells revealed that the interphase arrest was maintained for the first two hours after temperature shift (93% and 96% of cells were mononucleated without a septum, respectively). Cells then escaped from the checkpoint arrest and entered mitosis (78% and 44% of cells in interphase, at 3h and 4h, respectively). Analysis of *nuc2-663* revealed that 93% and 98% of cells remained in interphase 1h and 2h after shift, respectively. At 3h and 4h, 93 % and 97% of cells were mononucleated without a septum. Importantly, the percentage of mononucleated, septated cells was ≤4 at all the time points, which is consistent with the view that Nuc2p does not play a significant role in regulating septation in interphase, once the SIN has been reset during mitotic exit.

Fin1 over-expression may contribute to inactivation of the GAP for Spg1p at mitotic onset

Fission yeast has a single orthologue of the conserved never-in-mitosis (nimA) kinase, called *fin1* (Krien et al, 1998). Fin1p is not essential, but is important for spindle formation and regulates the affinity of Plo1p for the SPB (Grallert & Hagan, 2002). Fin1 mutant cells are delayed in the G2-M transition and Fin1p is in part regulated by Sid2p (Grallert et al, 2012). This link between *fin1* and the SIN prompted us to include *fin1* in the SIN regulatory circuit.

In the PKN of the model there were no negative regulators targeting GAP components during early mitosis, which resulted in suboptimal outcomes during the simulations of early mitosis; i.e. the simulation would produce a steady state where the GAP was still active in early mitosis. Since removal of the SIN GAP from the SPB is an early step in the activation of the SIN after entry into mitosis (Cerutti & Simanis, 1999; Li et al, 2000), we modeled whether GAP components could be regulated by *fin1*. Since Cdc16p contains several sites matching the established consensus for mammalian Nek2 (one of the orthologues of nimA), the effect of increased expression of *fin1* on Cdc16p localization was investigated. Expression of *fin1* from the medium strength *nmt-41* promoter (Basi et al, 1993) resulted in displacement of Cdc16p-YFP from SPBs in interphase cells (figure 2.11A). This required catalytically active Fin1p (figure 2.11B), and was not due to alteration of the steady state level of Cdc16p (figure 2.11C). Expression of fin1 promotes recruitment of Plo1p to the SPB in interphase cells (Grallert & Hagan, 2002); however, increased expression of

an activated allele of plo1 did not displace Cdc16p-GFP from the SPB (figure 2.11D) suggesting that its removal from the SPB is not a consequence of septation being activated by Plo1p.

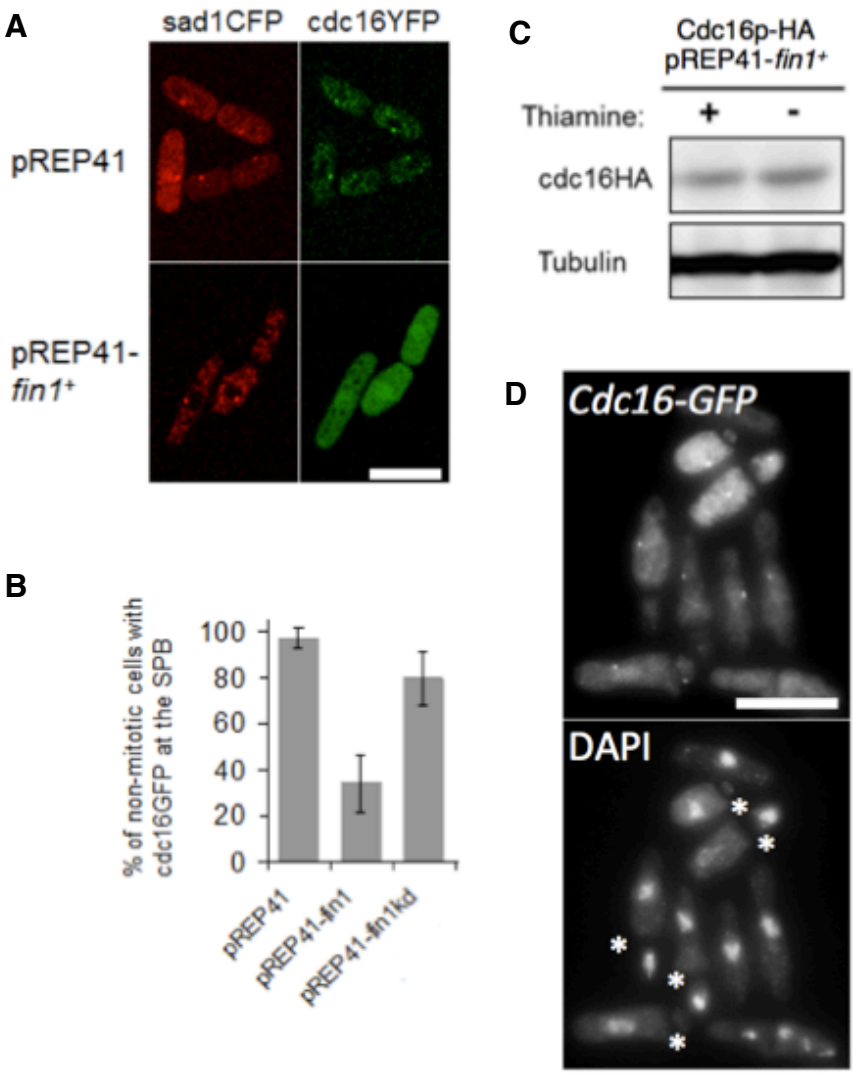

Figure 2.11: Fin1p over-expression results in Cdc16p disassociation from the SPB. (A) and (B) *Cells expressing the SPB marker sad1-CFP and cdc16-YFP were induced to express fin1 or a catalytically inactive fin1 mutant from the medium strength nmt41 promoter (Basi et al, 1993). The percentage of interphase cells retaining an SPB-associated Cdc16-YFP signal was plotted.* **(C)** *Cells bearing the cdc16-HA allele were induced to express fin1. Protein extracts were prepared 22h after induction were analyzed by western blotting using monoclonal antibody 12CA5. The anti-α-tubulin monoclonal antibody TAT-1 (Woods et al, 1989) was used as a control.* **(D)** *Cells expressing cdc16-GFP were transformed with a plasmid expressing Plo1p from the full-strength nmt1 promoter (Ohkura et al, 1995). Expression was induced for 16h at 29°C and the localization of Cdc16p-GFP was examined. The asterisks indicate septated cells; note the presence of a SPB Cdc16p-GFP signal in these cells, indicating that despite the induction of septation, Cdc16p-GFP is not displaced from the SPB. Experiments and figure produced by P. Collin.*

An unexpected prediction: cells overexpressing *cdc7* will septate in the absence of *sid4*

The final, optimized model describes the existing knowledge of the SIN, in wild type and known mutants. One of the main goals of developing this Boolean model was to use it predictively by performing *in silico* perturbations of interesting and/or experimentally challenging mutants. The regulatory relationships described in this model predict that increased expression of cdc7 should produce septation in the absence of Cdc11p and Sid4p (figure 2.12A). Previous studies have shown that *spg1* overexpression will induce septation and permit colony formation in a *cdc11* mutant (Schmidt et al, 1997), but not a *sid4* mutant (Balasubramanian et al, 1998). Moreover, increased expression of cdc7 will permit *cdc11* mutants to form colonies (Fankhauser & Simanis, 1994). In contrast to the situation with *spg1* overexpression, induction of *cdc7* expression from the very strong *nmt1* promoter in *sid4-SA1* at 36°C did not permit colony formation, but septa were formed in the cells (figure 2.12B). To test whether increased expression of *cdc7* would permit growth of a sid4 mutant, cdc7 was expressed from the ADH1 promoter, integrated at *leu1*. The *leu1::pADH1-cdc7* strain has a very high septation index at 19°C (>90%) and is barely capable of colony formation at 25°C and above (see figure 2.12D), with cells dying multiseptated at higher temperatures (see figure 2.12C). The strain *sid4-SA1 leu1::pADH1-cdc7* was capable of colony formation at 27°C and 29°C (figure 2.12D), where neither parental strain could do so. Previous studies have shown that increased expression of *cdc7* increases the level of kinase activity in immunoprecipitates of Cdc7p (Fankhauser & Simanis, 1994). This shows that septation can occur if the function of the scaffold proteins is compromised, provided the expression of Cdc7p is sufficiently elevated.

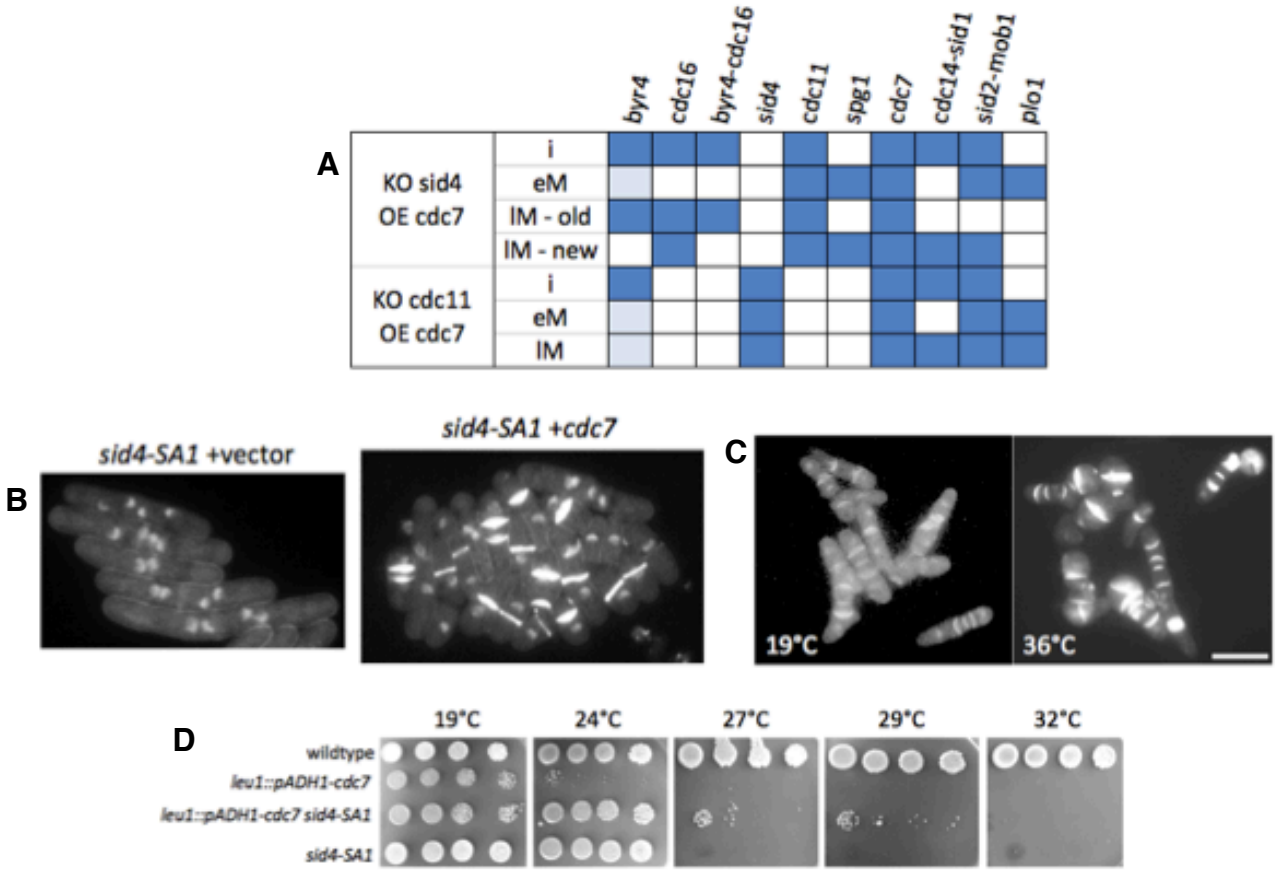

Figure 2.12: Cdc7p over-expression in a *sid4* mutant will result in septation. (A) *Steady states of in silico, double mutation experiments. The model predicts that in the absence of SIN scaffold proteins (Cdc11p or Sid4p) and over-expression of Cdc7p, the cell will septate*. **(B)** *sid4-SA1 leu1-32 was transformed with a REP1-based plasmids (Basi et al, 1993) expressing cdc7; empty vector served as a control. Cells were grown to exponential phase in EMM2 medium at 25°C containing 2mm thiamine. Expression was induced by washing with EMM2 and growth for 16h at 25°C; cells were then shifted to 36°C for 5h, fixed, and stained with DAPI and Calcofluor as described (Moreno et al, 1991). The scale bar represents 10 μm. Note that the cells carrying empty vector have become elongated and multinucleated, while 75% of cells expressing cdc7 have one or more septa.* **(C)** *The strain leu1::pADH1-cdc7 was grown to exponential phase in YE medium at 19°C. A sample was taken and cells were fixed and stained with DAPI and Calcofluor. The remainder of the culture was incubated for 5h at 36°C before fixation. Note the elevated percentage of septated cells.* **(D)** *The indicated strains were grown to exponential phase in YE medium, counted, and diluted to 10^6 ml^{-1}. 10 μl of serial 5-fold dilutions were spotted on plates, allowed to dry and then incubated at the indicated temperature until the wild-type control had formed colonies. Experiments and figure (B, C and D) produced by P. Wachowicz.*

Discussion

In this project we use qualitative Boolean modeling to represent and explore the regulatory relationships of genes participating in the Septation Initiation Network of fission yeast. Qualitative modeling is a powerful method for systems with restricted kinetic information and it is computationally efficient, allowing for thousands of *in silico* experiments in a short time, even in networks with hundreds of nodes. Moreover, it can be used predictively, to test combinations of mutations that would otherwise be time consuming, expensive and/or experimentally challenging to construct. The value of such models increases significantly when the model is coupled with *in vivo* experiments. Such experiments can be used to evaluate the regulatory rules, help the optimization procedure and test the predictions of the model (figure 2.13).

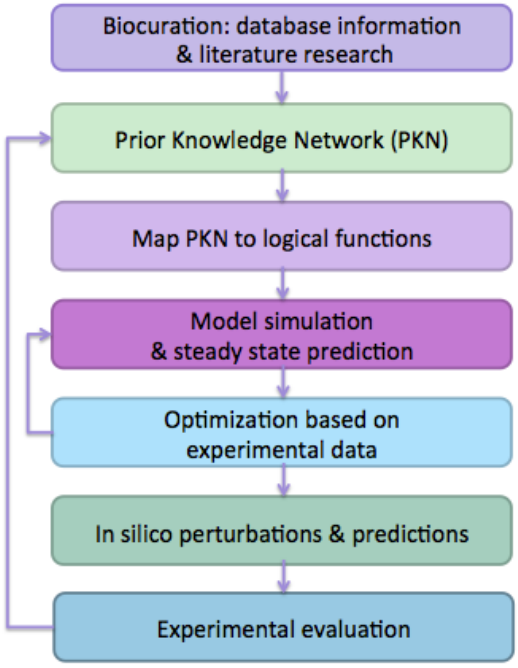

Figure 2.13: Model construction and optimization workflow. *The Prior Knowledge Network (PKN) is constructed after collecting relevant information from various sources, including network databases and literature. The PKN is translated into logical functions, describing the regulatory relations among genes products. The logical model is simulated under the preferred conditions, resulting in one or more steady states, where all the logical rules are satisfied. The model goes then through an optimization procedure, where the goal is to fit the resulting steady states with available experimental data by altering regulatory rules. The process is iterated until the simulation fits the available data. The model can then be used as a predictive tool, by performing in silico perturbations. Validation of the predictions can lead to discovery of missing regulatory links that are then added to the PKN.*

We report the construction of an extended, Boolean model of the SIN network that uses CDK levels as control nodes to simulate SIN related events in interphase, early mitosis and late mitosis. The prior knowledge network was manually curated, providing a trustworthy initial framework that could then be further optimized (figure 2.13). Information reported in literature (and used in network databases) can be conflicting, outdated, incomplete or based on *in vitro* knowledge only. Therefore, expert biocuration provides a significant advantage in order to filter the available information and construct a comprehensive network.

We optimized the model using in silico experiments with well-established outcomes based on *in vivo* data, in order to recapitulate the SIN state in different stages of the cell cycle (figure 2.13). A challenging aspect of qualitative modeling, and especially of asynchronous update, is to interpret the resulting steady states of the simulations. This is because the simulation might result in a number of steady states that are theoretically possible but never reached *in vivo*. Our approach was to use CDK levels as an initial condition for the simulation, indicating the stage of the cell cycle that the simulation corresponds, to reduce unrealistic simulation outcomes. We further restricted the simulation space by taking as a fact that the scaffold has the potential to be constructed at all times by setting the SIN-SPB linker protein Ppc89p to 1.

The optimization process under the controlled environment of CDK switches provided important insights into SIN regulation during the cell cycle. In the case of the *fin1*, the incorrect simulation results that were obtained in early mitosis helped us locate a potential missing link in the PKN. Increased expression of *fin1* removes Cdc16p from the SPB. At present we do not know whether this is by direct phosphorylation of Cdc16p or an indirect effect; this will be the subject of future analysis. However, the important point in this context is that the modeling revealed the requirement for an additional control point to turn off the GAP in early mitosis. The optimization strategy was also useful in evaluating the limitations of our model. An example of this is the role of Nuc2p in SIN regulation. In the PKN there were several inhibitory links from Nuc2p to SIN kinases, indicating the events in SIN resetting, after septation (Chew & Balasubramanian, 2008). The use of CDK switches restricts the cell cycle events that can be modeled, and our model does not presently incorporate resetting of the SIN at the M-G1 transition. Our modeling predicted that if Nuc2p continued to activate the GAP in interphase, extending the role proposed for it at the M-G1 transition (Chew & Balasubramanian, 2008), then its inactivation in post-START cells could result in septum formation; *in vivo* analysis showed this was not the case. Thus, the modeling was useful in this case, to define the possible limits of the extent of the time-window in which Nuc2p is active towards the SIN.

The great value of creating an optimized qualitative model is that it can then be used predictively to perform difficult or iconoclastic experiments *in silico*. We focused on testing whether an over-expression of SIN kinases would rescue SIN scaffold mutants. The model's prediction was that over-expression of Cdc7p in a *cdc11* or *sid4* knock-out will still septate, a prediction that was experimentally validated. Future research can investigate how signalling in such double mutant can occur. To test whether signalling is cytoplasmatic and what it depends on, future investigation can focus on tagging Cdc11p, Byr4p-Cdc16p and Plo1p in the double mutant cells. The model can be used in the future for any combination of gene mutants, and hopefully provide interesting hypotheses that can be tested experimentally.

Future directions of the project

Future studies can focus on several aspects. To begin with, the M-G1 transition can be included in the model, and it can be later extended to include the complete CDK cycle. Spatial components such as SPB localization can also be incorporated, as well as post-translational modifications of SIN proteins during the cell cycle (Bajpai et al, 2013; Feoktistova et al, 2012; Singh et al, 2011). This should allow modeling of the role of the asymmetry of SPBs with regard to SIN protein association, building upon the analysis performed by Bajpai et al., 2013. Future versions of the model can attempt to incorporate Etd1p. Finally, the same framework can be used for the construction of a SIN model for meiosis, which presents significant differences from the mitotic regulatory network. The methods that could be used to achieve these goals are explained in detail below.

CDK cycle integration in the model

For the current version of our model we attempted to simplify the CDK activity cycle and its effect on SIN regulation by setting three discrete CDK levels as control nodes; interphasic CDK, early and late mitotic CDK. This CDK switch structure proved to be a useful tool in describing interphasic, early and late mitotic events, as well as predicting unknown mutations (Chasapi et al, submitted). However, it restricted the cell cycle events that can be modeled to those three stages of the cell cycle. For example, in the PKN there were several inhibitory links from Nuc2p to SIN kinases, indicating the events in SIN resetting, after septation (Chew & Balasubramanian, 2008). However, our current model does not incorporate resetting of the SIN at the M-G1 transition. This was illustrated in practice when, during our simulations we noticed that Nuc2p continued to activate the GAP in interphase, extending the role proposed for it at the M-G1 transition (Chew & Balasubramanian, 2008), then its inactivation in post-START cells could result in septum formation;

in vivo analysis showed this was not the case. The model could, therefore, be extended to better incorporate the cell cycle event sequence.

At a first stage, a fourth CDK node could be introduced, reflecting the transition from mitosis back to interphase. The workflow to be followed for this task is already defined by our previous work; using the same PKN as the current model, we can attribute the regulatory relations that take place during the transition from mitosis to interphase to the corresponding CDK node. During this process molecules that are not present in the model so far will be added, such as the phosphatases PP1 and PP2A, and, of course, interactions relating to SIN inactivation after mitosis.

At a second stage, the cell cycle sequence can be fully incorporated to our model. This will be facilitated by the incorporation of the cell cycle Boolean module of fission yeast, by Davidich & Bornholdt (Davidich & Bornholdt, 2008). This version of the model will still use multinode CDK, but instead of CDK switches, where each simulation corresponds to a different stage of the cell cycle, it will use CDK checkpoints and during one simulation the CDK levels will change based on the activation state of the rest of the proteins in the system. Part of this link between CDK nodes is already present in our current model, even though for now it is only present for descriptive purposes. Specifically, in the current model the transition from early to late mitotic CDK is encoded as such: Activation of the spindle assembly checkpoint (Musacchio, 2011) inhibits the anaphase-promoting complex (APC), therefore blocking the decrease of CDK levels and mitotic exit (Jia et al, 2013; Musacchio, 2011).

Incorporation of spatial aspects of SIN regulation into the model

Our current Boolean model of the SIN describes certain cell cycle stages in a satisfactory manner. For example, in late mitosis the model reaches two steady states, representing the constellation of the new and old SPBs respectively. However, the model does not incorporate spatial aspects of SIN regulation for the moment. Such aspects would be the discrimination between the sequence of events in the old and new SPB and the changes complexes undergo when they associate with the SPB.

For the next version of the model the spatial aspects of SIN regulation could be implemented. This can be approached in a variety of ways such as (a) introducing multi-node configurations for proteins that display asymmetric association with the SPBs in anaphase (b) introducing multivariate nodes to simulate the effect of changes in the post-translational modifications of SIN proteins during the cell cycle (Bajpai et al, 2013; Feoktistova et al, 2012; Singh et al, 2011) and (c) a combination of the above.

(a) Multi-node configuration

Similarly to the multi-node representation of CDK fluctuations during the cell cycle, two or more nodes can be created for other components of the SIN. This can be particularly useful in cases of proteins that participate in the asymmetry establishment of the two SPBs. There, there could be two nodes per protein, one representing its presence at the old SPB (for example Cdc7p-old) and the other representing its presence at the new SPB (for example Cdc7p-new). In fact, we can perceive this configuration as duplicating the SPBs during mitosis, same as *in vivo*. The components of the model that have an identical effect on the SPBs stay as unique nodes and the events that drive asymmetry are adjusted to regulate differently the duplicated nodes.

(a) Multivariate nodes

An alternative (or complementary) approach is to create multivariate nodes for key components of the SIN that undergo changes in the post-translational modifications during the cell cycle. A good example of this is Cdc11p, whose changes in phosphorylation levels and their effect in asymmetry establishment are demonstrated at the paper by Bajpai et al., 2013. Indeed, future work can build upon this study and extend it to other key components of the SIN. The challenge using this approach lies in the amount of information needed for the correct regulation of the transition among node levels. For example let us assume that a protein displays 4 phosphorylation levels. In a multivariate configuration this will be encoded as node levels 0,1,2,3. In order for the simulation to work successfully, the order of node level transition has to be known, as well as the regulators responsible for each transition.

Integration of Etd1p regulation in the mitotic cell cycle into the current model of the SIN

The next version of the model can attempt to incorporate Etd1p. Though its effects upon SIN signaling are evident (Alcaide-Gavilan et al, 2014), the published information does not provide a sufficiently clear, direct link to SIN components to permit its unequivocal incorporation into our current model. To address this problem, a wide exploration approach can be followed. First, probable Etd1p regulators can be identified, based on literature and molecular information. Then, a collection of regulatory relations that can exist between Etd1p and the candidate regulators will be created. The parallel work on Etd1p at a wet lab such as Simanis lab will guide the choice among regulations to be tested. The collection of candidate models with the incorporated Etd1p relations will be simulated and scored according to their ability to reproduce known *in vivo* effects. To

choose among the best-scoring candidates *in silico* perturbations of unknown effects will be performed, and the results will be evaluated with *in vivo* experiments.

Build a model of the SIN for meiosis

Our previous work involved the creation of a Boolean model of the SIN during mitosis. Employing the same techniques, a SIN model can be constructed to describe events in meiosis, where the regulatory relations of SIN proteins differ substantially. Having already established the workflow of model creation and optimization is a great asset and will facilitate the fast and effective completion of this task. Specifically, a Prior Knowledge Network will be created, specific to SIN events in meiosis, using manual biocuration. The model will be simulated with software such as BoolSim and SQUAD (Di Cara et al, 2007), and will be optimized and evaluated using a set of well characterized knock-out and constant-expression perturbations that will be divided into a test set and a training set. The meiotic model of the SIN can be a great addition to our knowledge on S.*pombe*, providing complementary information and interesting comparisons with the already existing mitotic model (Chasapi et al, submitted).

To conclude, it is worth noting that qualitative models such as the one presented here are oversimplifications of the actual regulatory processes; in our case of the regulation of the SIN. With advances in live monitoring of cell division and development of new fluorescent probes, we should be able to generate more accurate quantitative models for such a system. Our approach is nevertheless an important step towards a more comprehensive model that recapitulates known biology of the SIN and can be used as a hypothesis generator for complex experimental design.

Acknowledgements

We thank Iain Hagan (Paterson institute for Cancer Research, Manchester, UK) for the *fin1* expression plasmids and Keith Gull (Oxford, UK), for TAT-1.

Chapter 3

Minimal Boolean models reproduce the wave expression of Clb proteins in the budding yeast cell cycle

Partially based on

Linke C, Chasapi A, Gonzalez-Novo A, Klipp E, Krobitsch S, Posas F, Xenarios I, Barberis M (2015) *Fkh transcription factors control timing of Clb expression to regulate waves of mitotic cyclins.* (in preparation)

Cyclin control in budding yeast cell cycle

In *Saccharomyces cerevisiae*, commonly known as budding yeast, the coordination between cell growth and DNA replication is guaranteed by two facts: (a) the transition to S phase only occurs when cells reach a critical cell size and (b) genome duplication is limited to once per cell cycle by several mechanisms. This coordination takes place during a narrow interval in the late G1 phase known as Start and it is connected to two crucial events. On the one hand, growth-dependent accumulation of waves of G1 (Cln) and B-type (Clb) cyclins regulates the order and the timing of cell cycle phases by binding to the kinase Cdk1 (also known as Cdc28) (Andrews & Measday, 1998; Enserink & Kolodner, 2010; Fitch et al, 1992; Mendenhall & Hodge, 1998; Nasmyth, 1993; Nasmyth, 1996). On the other hand, the presence of Sic1, a potent inhibitor of Cdk1/cyclin complexes containing Clb but not Cln cyclins (Mendenhall, 1993; Schwob et al, 1994; Weinreich et al, 2001), prevents premature onset of DNA replication (Barberis et al, 2005a; Barberis et al, 2005b).

B-type cyclins Clb1–6 are expressed at different times and appear sequentially in specific cell cycle phases, resulting in a significant divergence of function (Bloom & Cross, 2007; Cross et al, 1999). Clb5,6 rise at the beginning of S phase and function primarily in the control of DNA replication (Schwob et al, 1994; Schwob & Nasmyth, 1993; Spellman et al, 1998). Clb3,4 increase in mid-S phase at about the same time as spindle pole bodies separate. Clb4 ensures proper alignment of the mitotic spindle with the cell division axis (Liakopoulos et al, 2003). However, there are still unanswered questions regarding their specific functions (Fitch et al, 1992; Richardson et al, 1992). Clb1,2 rise as mitotic spindle assembly progresses and are involved in the control of mitotic exit (Deshaies, 1997; Fitch et al, 1992; Spellman et al, 1998) (Figure 3.1A). The sequential expression curves of Clb5,6 – Clb3,4 – Clb1,2 is sometimes referred to as Clb waves; this term is used in the thesis as well (Figure 3.1B).

The regulation of active Cdk1–Clb complexes involves a combination of positive feed-forward loops – depending on the regulated transcription of CLB genes (Bloom & Cross, 2007; Fitch et al, 1992; Koch & Nasmyth, 1994) – and negative feedback loops through down-regulation of Clb levels via ubiquitin/26S proteasome pathway (Amon et al, 1994; Hochstrasser, 1995; Irniger et al, 1995; Lew & Reed, 1995; Tyers & Jorgensen, 2000).

Sic1 protein expression is limited to the M/G1 transition (Donovan et al, 1994; Mendenhall et al, 1987) and has been shown to inhibit kinase activities associated with Cdk1/Clb5, which triggers DNA replication, and Cdk1/Clb2, which triggers mitotic events (Schwob et al, 1994). These studies

highlight the fact that Sic1 strongly binds to Cdk1 its catalytic site, therefore preventing the access of substrates (Barberis et al, 2005b). Thus, Sic1 achieves two major functions in cell cycle regulation. On the one hand, it prevents premature S phase onset by inhibiting Cdk1/Clb5,6 until Sic1 is phosphorylated by Cdk1/Cln complexes and targeted for degradation via the ubiquitin-mediated proteolysis pathway (King et al, 1996; Schwob et al, 1994; Sheaff & Roberts, 1996). On the other hand, Sic1 contributes to the abolishment of Cdk1/Clb2 activity required for mitotic exit (Calzada et al, 2001; Donovan et al, 1994; Toyn et al, 1997), although its binding to the kinase does not target Clb2 for proteolysis (Amon, 1997).

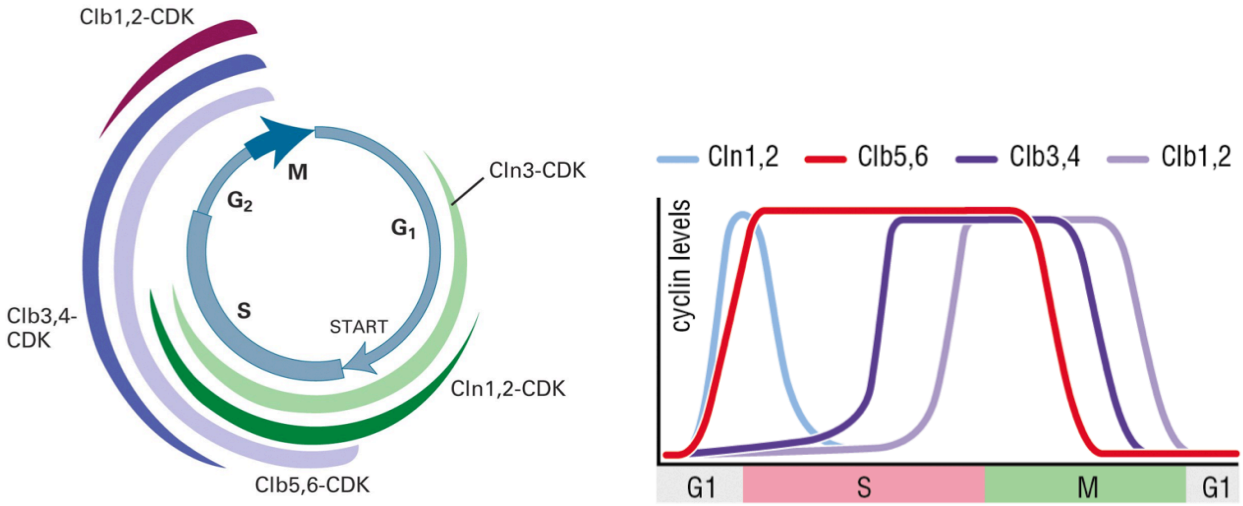

Figure 3.1: Cyclin activity during cell cycle phases A) *The sequential expression of Clb proteins during budding yeast cell cycle.* (Lodish et al, 2008). **B)** *The wave expression graph of Clb proteins. Clb5,6 gets activated first, followed by Clb3,4 and Clb1,2 (Morgan, 2007).*

The sequential activation and degradation of Clbs gives directionality to cell cycle events. Although many details of transcription of cyclin genes are well understood (Bahler, 2005; Haase & Wittenberg, 2014; Wittenberg & Reed, 2005), understanding the molecular mechanisms regulating the relative timing of waves of *CLB* activation remains a challenge.

Prior to this work, there have been several efforts on S.cerevisiae cell cycle modeling. A Boolean model comprising 11 proteins was published in 2004 (Li et al, 2004), where Sic1, Clb5 and Clb2 regulations are included, but Clb3 is absent. Other models followed top-down approaches, using biochemical kinetics and gene expression for model reconstruction (Chen et al, 2004; Chen et al, 2000; Klipp et al, 2005). Clb3 is absent from all of the models mentioned above. Some modeling efforts have been comprehensive, such as the ODE based models by Chen et al (Chen et al, 2004; Chen et al, 2000; Klipp et al, 2005), while others address specific cell-cycle phenomena, such as

the links between cell size and cycle progression (Alarcon & Tindall, 2007; Barberis et al, 2007) or DNA replication (Gidvani et al, 2012). A review on the diverse approaches that were employed by different groups can be found at (Ingalls et al, 2007).

In previous work by Barberis et al (2012), ODE modeling techniques were used to investigate whether Sic1 plays a potential role in the regulation of the oscillatory behavior of Clbs. With a combination of mathematical modeling and experimentation, it was proven that Sic1 is indeed continuously detected during the entire cell cycle, with the only exception of a temporal window during which all Clbs are present at the maximal level, and could play a role in coordinating the timing of Clbs waves (Barberis et al, 2012).

The publication of the Clb wave ODE model prompted us to form a collaboration with Matteo Barberis, aiming to investigate whether we can reproduce the Clb wave behavior using qualitative modeling frameworks. In case that could be achieved, we were interested in the type of models that would give such results. The goal, as in the case of SIN modeling in fission yeast, was not only to describe the system, but also to use the model predictively, in order to identify potential missing regulatory links.

Qualitative modeling of Clb wave behavior

To describe the Clb wave expression qualitatively, we defined the system to be modeled as a 4 node network including Clb5,6, Clb3,4, Clb1,2 and Sic1, the same proteins that had been used for the previous kinetic model (Barberis et al, 2012). For simplification reasons we named each node after one of the Clb complexes for each cell cycle stage. We started by constructing the prior knowledge network of interactions among Clb5, Clb3, Clb2 and their inhibitor Sic1 (figure 3.2A). The model was evaluated under 3 conditions: wild type simulation, knock-out of SIC1 and over-expression of SIC1. All 3 experiments have known phenotypic outcome. Specifically, in wild type we observed the sequential activation of Clb5, Clb3 and Clb2, and their sequential inactivation when the inhibitor Sic1 increases in concentration. When Sic1 activity starts to decay, the Clb molecules are reactivated, producing a cyclic behavior. SIC1 over-expression results in the same attractor, the only difference being that the Clb activity curves display delayed activation. Finally, in *sic1*Δ all the Clb proteins are active, which, in Boolean terms, would be represented by the steady state 0111 (figure 3.2B).

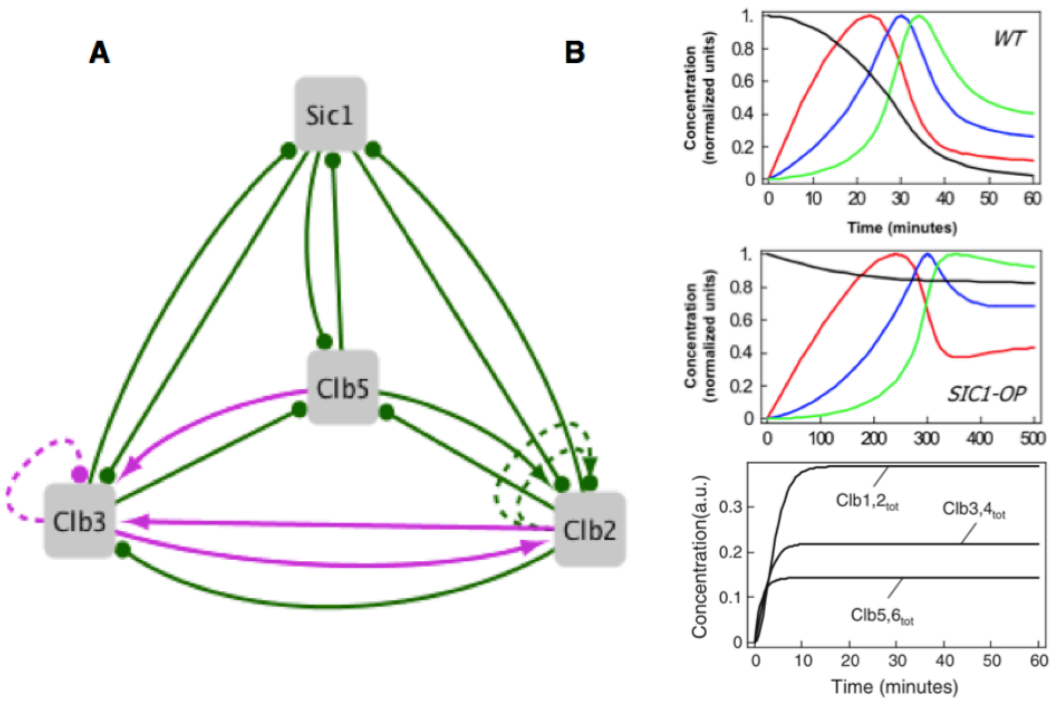

Figure 3.2: The Prior Knowledge Network of cyclin regulation A) *The minimal, prior knowledge network of Clb regulation. Green arrows represent known interactions and purple arrows represent potential, non-proven interactions.* **B)** *Expected activation profile of Clb molecules in wild type, SIC1 over-expression and sic1Δ, as simulated at (Barberis et al, 2012). The participating proteins are color coded as follows: Clb5 = red, Clb3 = blue, Clb2 = green, Sic1 = black.*

PKN-based optimization

Different versions of the PKN were simulated, by filtering the included interactions based on their level of confidence. However, none of the simulated models was able to reproduce the behavior of the 3 test experiments. The outcome was constant activation of Sic1, which suppressed all Clbs. We therefore altered our strategy, assessing whether there is a network, similar to the PKN, with the potential of reproducing the expected experimental outcomes. Two strategies were followed: (a) reducing the model and (b) investigating whether there are missing interactions.

(a) Model reduction

Starting from the initial PKN network, a genetic algorithm approach was used to test whether there is a reduced model that satisfies all 3 expected experimental outcomes. The genetic algorithm tool, developed by Julien Dorier (Vital-IT, SIB Swiss Institute of Bioinformatics), requires a set of

interactions that are used as a pool from which the model can be constructed. Additionally, a set of parameters is provided, towards which the model is optimized. In our case it corresponds to looking for an attractor with all nodes oscillating in wild type and SIC1 over-expression, and a steady state with all Clbs active in *sic1*Δ. After several iterations of the simulations it became evident that there is no model solution, based on the given interactions, that completely satisfies the requirements.

(b) Addition of possible missing regulations

The inverse approach was tried as well. First, the collection of all possible edges of the system was identified, which we will refer to as **edge pool**. The edge pool involves all possible interactions involving 1, 2, 3 or all 4 nodes acting as co-regulators (for example A→B, A&C→B, A&C&D→B, A&B&C&D→B). The interactions can be activations (→) or inhibitions (⊣), and NOT (^) is also included at the possibilities (for example ^A→B). In total, the edge pool contains 640 edges, for our minimal system of 4 nodes (Sic1, Clb5, Clb3 and Clb2). The next step was to create all PKN+1 models by adding one edge from the edge pool, as a way to investigate whether the *in silico*, unfitting results are due to a missing interaction. None of the simulated models produced, however, the expected results.

The minimal model approach

Minimal model construction rules

The failure to produce model candidates reproducing the expected results in known experiments demonstrated that when working with the minimal, 4-node system in isolation, it is difficult to explain the Clb waves profile using PKN-based networks. We therefore decided to define the minimal networks that can reproduce the experiments. In other words, we decided to create a set of networks that, with minimum number of edges, capture the known information flow of our system. Such models can indicate the type of regulatory events that are necessary for the waves and they can be used as alternative hypotheses tools, in order to explain Clbs regulation and potentially identify missing regulatory events.

The first step towards this end was to define the exact results expected from the models. Given that the experimental phenotype of wild type cells is a wave expression of the Clbs and Sic1 (figure 3.3A) we had to translate that to a Boolean attractor. The states comprising the attractor can differ according to the expression threshold set for each molecule. In order to have a comprehensive

analysis, we derived three alternative Boolean attractors for the wild type experiments (figure 3.3C). For SIC1 over-expression the expected attractor is identical to wild type, with Sic1, of course, being constantly active. For *sic1*Δ, the expected experimental result is a steady state with Sic1 inactive and all Clb molecules active (0111). An example of one of the Boolean attractors derived from the real expression curve of the Clbs is shown in Figure 3.3B.

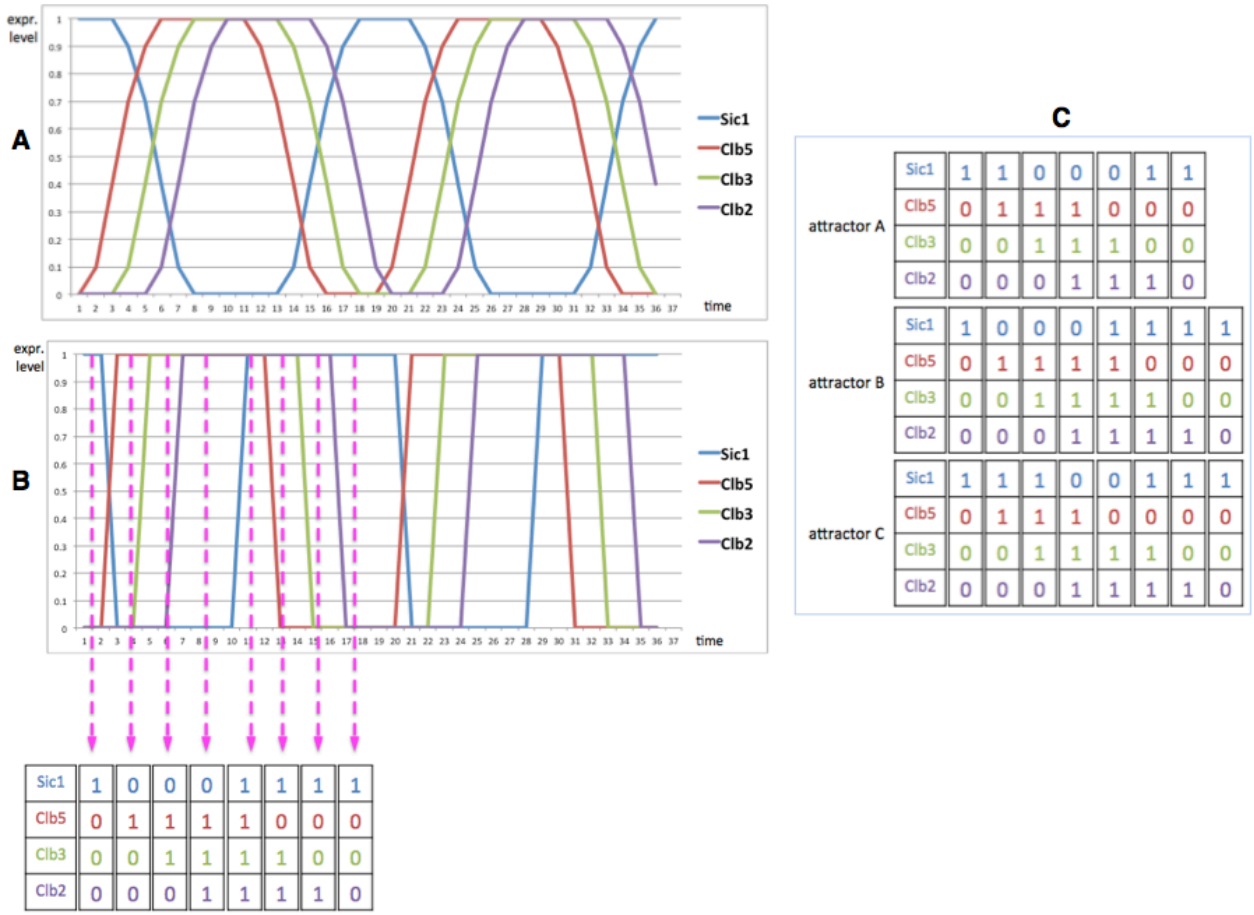

Figure 3.3: Define expected results in Boolean simulations. A) *Qualitative representation of real expression curves for Sic1, Clb5, Clb3 and Clb2 in wild type cells. The Clb molecules are activated sequentially and they start sequentially decaying in parallel to Sic1 activation.* **B)** *"Booleanized" version of the expression curves. The Boolean attractor is derived by setting activity thresholds. Each state between any activity transitions is translated to a Boolean vector.* **C)** *A list of the attractor candidates that were used for the downstream analysis. These attractors are identical for SIC1 over-expression, the only difference being that Sic1 is constantly set to 1.*

The next step was to set the construction rules for the models. Since the goal was to identify all minimal models explaining the Clb waves, our starting edge pool was the previously constructed collection of all possible edges (640 edges). Knowing that all participating genes present activity fluctuations, a rule was set that each gene should have an incoming edge. Ergo, a minimal model

should have at least 4 edges, one for the regulation of each gene. On the same note, auto-regulations were not permitted, as that would isolate the auto-regulated node from the system, the auto-regulation being the only incoming edge for that node. To summarize, the constructed models should have 4 edges, one input per node and no auto-regulations.

Filter out contradicting edges

An exhaustive model construction and testing of all models complying with the rules above would be impossible, as there are around 660 million candidate models. Therefore, we decided to follow a reconstruction approach and filter the edges that would be able to explain the expected attractors / steady states. This process would be extremely difficult to perform manually in more complex models. However, the fact that in our minimal model we only allow one edge (logical rule) to regulate each node renders the manual edge filtering feasible.

Initially, we defined the collection of possible transitions. Those included the attractors in wild type and SIC1 over-expression as well as the steady state of *sic1*Δ which can be interpreted as a constant transition 0111→0111. Then, for a given transition, we would go through the edge pool and remove the regulatory rules that would not permit it. At the end of the process the possible edge pool would be reduced to include only the regulatory rules that do not contradict any of the transitions.

Let us consider the transition 0100→0110, which is present in one of the candidate wild type attractors, as a simple example to demonstrate the edge filtering. In this vector, the proteins are represented in the following order: Sic1-Clb5-Clb3-Clb2. The transition basically constitutes the activation of Clb3, with all other proteins keeping their previous activation states. For the purposes of this example, we will only test the plausibility of 4 regulatory rules: Sic1→Clb3, ^Sic1→Clb3, Sic1&Clb5→Clb3, ^Sic1&Clb5→Clb3.

Sic1→Clb3 and ^Sic1→Clb3

The state of Clb3 at time *t*+1 is defined by the state of its regulators at time *t*. If Sic1 is the only regulator of Clb3 and acts by activating it (Sic1→Clb3), then Clb3 will be active in time *t*+1 if and only if Sic1 is active at time *t*. This, however, is not the case (Sic1 state at time *t* is 0), therefore Sic1→Clb3 cannot be the logical rule regulating Clb3. ^Sic1→Clb3 can interpreted as «the absence of Sic1 will trigger Clb3 activation». Practically, this means that for Clb3 to be active in

time *t*+1, Sic1 has to be inactive in time *t*, which is true in this case and, therefore, ^Sic1→Clb3 is not discarded.

Sic1&Clb5→Clb3 and ^Sic1&Clb5→Clb3

In cases of AND regulatory rules we have to take into account the state of all nodes participating in the regulatory rule. For the regulation to function, all states of all the participating nodes must satisfy the logical rules. At time *t*, Sic1 is inactive (0) and Clb5 is active (1). Under this setting, Clb3 switches from 0 to 1. If both Sic1 and Clb5 participate in the regulation of Clb3, the only AND rule configuration that would trigger the activation of Clb3 is the one expecting an inactive Sic1 and an active Clb5 at time *t* : ^Sic1&Clb5 → Clb3. Sic1&Clb5 → Clb3 is, of course, discarded as contradictive.

The process described above was used for all nodes in all attractor candidates that were defined initially. Only one of the 3 candidate attractors (attractor B in figure 3.3C) resulted in an edge collection that could be used to construct the minimal model candidates. In the other 2 cases, there were very few regulatory rules that did not contradict any transitions, and more importantly, not all nodes had the possibility to be regulated (all possible incoming edges had been filtered out). Without all nodes being regulated, it would have been impossible to obtain a cyclic attractor during the simulation, and, therefore, the 2 candidate attractors were rejected.

In the case of the third attractor that resulted in a permissible edge pool, the initial 640 edges were reduced to 10, and the possible minimal models to be constructed were reduced to 36. Table 3.1 shows the list of regulatory rules comprising the minimal model edge pool.

Clb2	→	Sic1
Clb2 & Clb3	→	Sic1
^ Clb2	⊣	Sic1
Sic1 & Clb2	⊣	Clb5
Sic1 & Clb3 & Clb2	⊣	Clb5
Clb5	→	Clb3
^ Clb5 & Sic1	⊣	Clb3
Clb3	→	Clb2
^ Clb3	⊣	Clb2
Sic1 & ^ Clb3	⊣	Clb2

Table 3.1: Minimal model edge pool. *A list of the regulatory rules not contradicting any of the transitions for wild type, sic1Δ and SIC1 over-expression attractors / steady states. This edge pool is the result of filtering the attractor candidate B (see figure 3.3C). 3 logical rules can explain the regulation of Sic1 and Clb2, and 2 the regulation of Clb5 and Clb3.*

Capturing the wave phenotype with minimal models

The 36 model candidates were simulated using BoolSim and, all of them reproduced the Boolean attractors expected for wild type and SIC1 over-expression, as well as the steady state expected at *sic1*Δ. Next, we used SQUAD (Di Cara et al, 2007) to further filter the model candidates and only keep the ones that reproduce the waves during qualitative ODE simulations. SQUAD is a dynamic simulation software that uses a standardized qualitative dynamic approach. It works by first identifying the steady states of the network using Boolean modeling. A system of standardized ordinary differential equations is then used to simulate the dynamic behavior of the network in time (Di Cara et al, 2007).

When we simulated the 36 candidate models with SQUAD, only 6 of them reproduced the wave curves that are experimentally observed (figure 3.4A). Furthermore, these 6 models reproduced not only the expected attractors, but also the slight delay in activation of Clbs relative to wild type that is observed in SIC1 over-expression (Barberis et al, 2012). The latter can be observed in figure 3.4B, where SQUAD simulations of one of the minimal model candidates are shown.

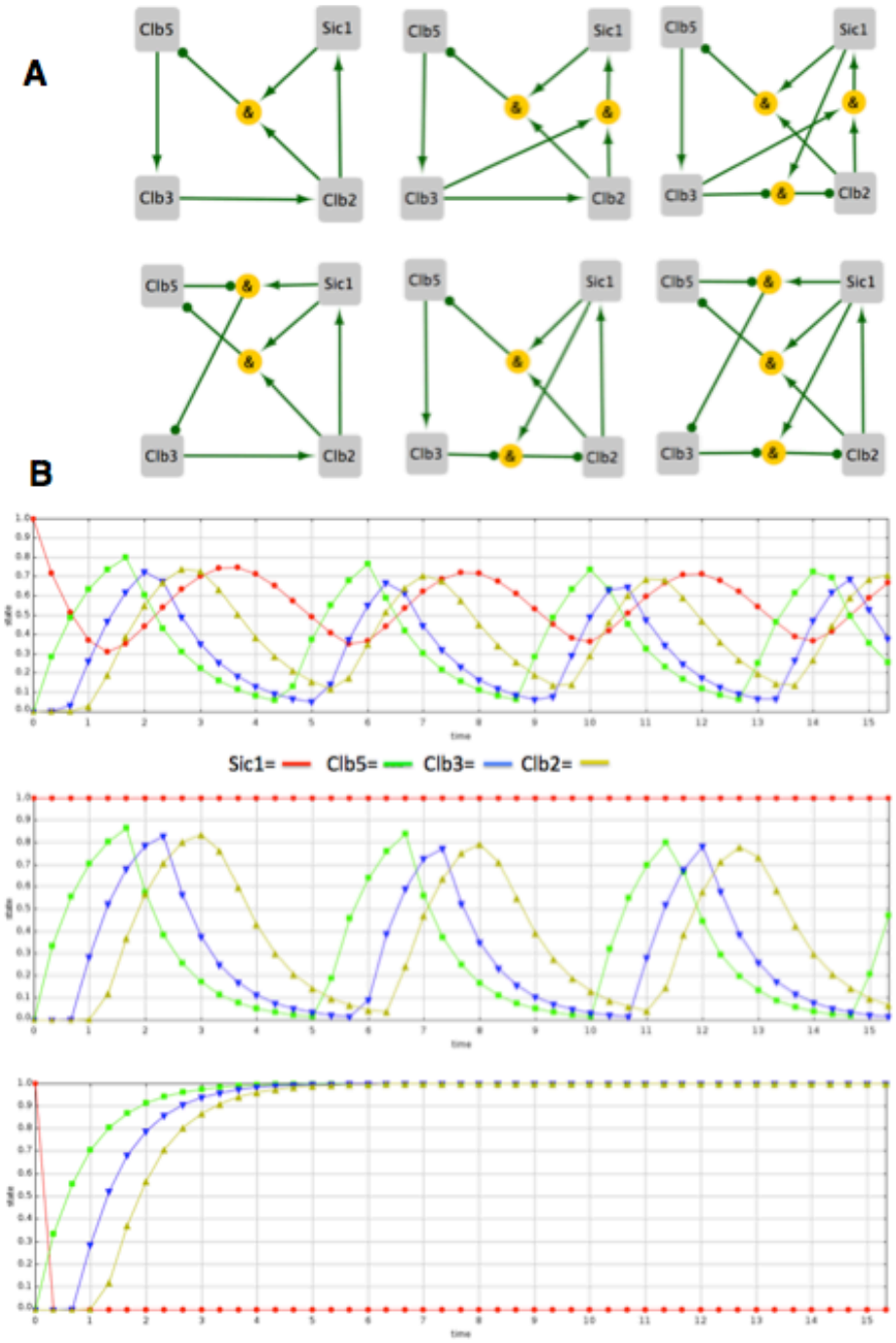

Figure 3.4: Minimal model candidates and simulation results. A) *The 6 minimal models that satisfy all experiments, both with Boolean simulations and with standardized ODE simulations. Arrows indicate activations and circles inhibitions. Many interactions appear in more than one model, with Sic1&Clb2⊣Clb5 being the only shared interaction among all models.* **B)** *SQUAD simulations of the simplest model, shown at the top right frame. From top to bottom are plotted the simulations of wild type, SIC1 over-expression and sic1Δ. Notice that by time t=15 wild type cells have entered the fourth cycle, whereas in SIC over-expression we observe a delay, with cells having just competed the third cycle. These results agree with published literature* (Barberis et al, 2012).

Result verification using MaBoSS

To independently validate our results we decided to use MaBoSS, which is a software that applies Markov processes onto Boolean networks. MaBoSS describes heterogeneous cell population behavior using continuous and discrete Markov processes (Stoll et al, 2012). It uses a specific language to associate transition rates to each node. Given some initial conditions, MaBoSS applies a Monte-Carlo kinetic algorithm to the network to produce time trajectories. Time evolution of probabilities for each network state are then estimated (Stoll et al, 2012). Since MaBoSS describes a cell population rather than single cells, the probabilities of all states converge to a constant when time goes to infinity. The Boolean feedback loop depends only on the topology of the regulations and not on transition rates or time; this is the kind of cycle we observe, for example, with our models when using BoolSim. However, these cycles cannot be linked perfectly to periodic behavior of instantaneous probabilities that MaBoSS uses, because the set of instantaneous probabilities cannot be perfectly periodic. They can display a damped oscillating behavior, or none at at all. In our case, we observe a damped oscillation (figure 3.5A). In this simulation, the probability distributions of all network states tend to a stationary distribution. In this case, we can use the entropy and transition entropy measures that are also provided in the MaBoSS simulation results, and can help characterize cyclic stationary distributions. As defined at the software publication (Stoll et al, 2012):

*“ The **Entropy (H)** measures the disorder of the system. Maximum entropy means that all states have the same probability; H=0 means that one of the states has a probability of one (i.e. steady state). “*

*“ The **Transition Entropy (TH)** characterizes the system at the level of a single trajectory. For each state S, there exists a set of possible transitions. TH(S)=0 means that there is no transition from S to any other state. The TH for all trajectories is a way to measure how deterministic the dynamics is. If the transition entropy is always zero, the system can only make a transition to a given state. “*

To identify a cyclic stationary distribution (meaning a stationary distribution that ‘’hides’’ a cyclic attractor), the entropy H has to be non-zero and the transition entropy TH has to be 0. This is indeed the case with our simulation (figure 3.5B).

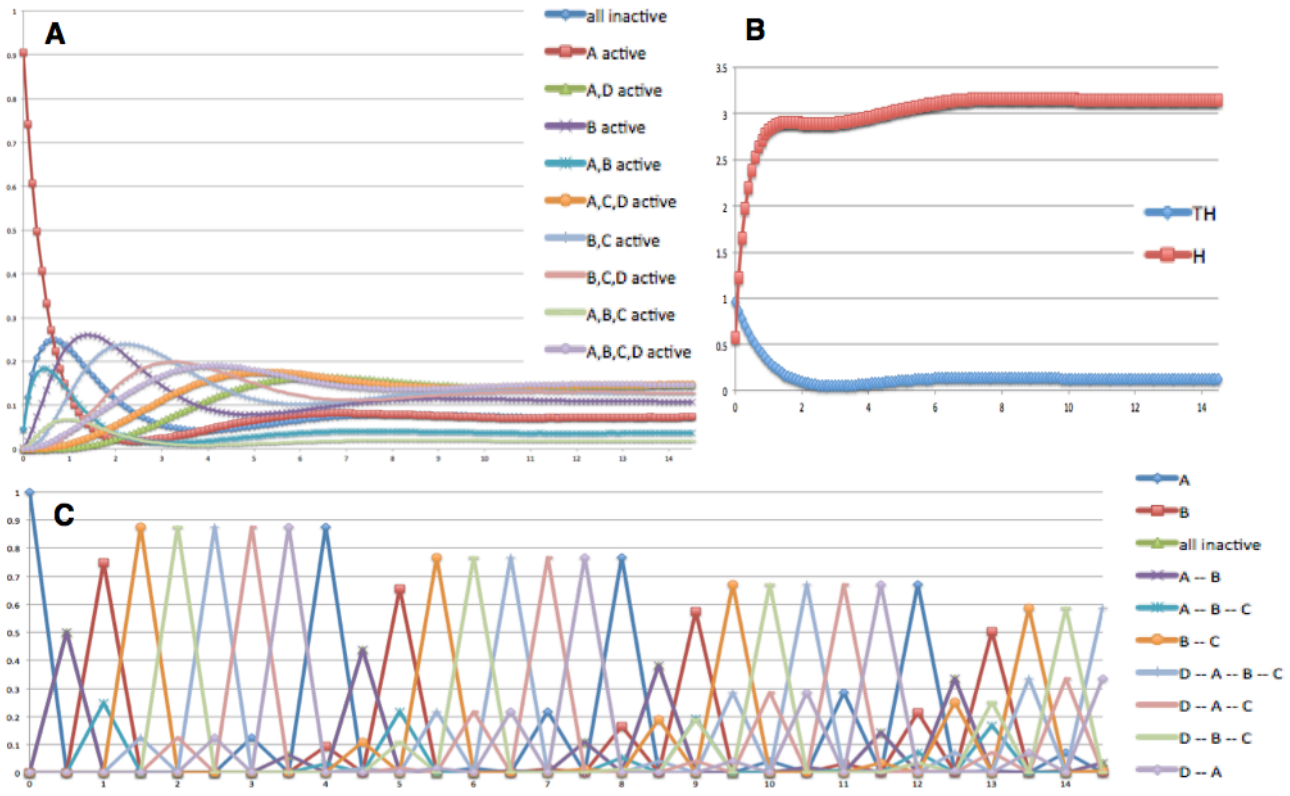

Figure 3.5: MaBoSS simulations of a minimal model candidate. A) *Standard continuous Markov process simulation. The probability distributions for network states tend to stationary distribution. However, during the first cell cycle we observe that the sequence of states corresponds to the wild type attractor we obtain with BoolSim and SQUAD.* **B)** *Entropy H and transition entropy TH distributions. H is non-zero and TH is almost 0, which indicated that the stationary distribution shown in panel A is "hiding" a cyclic attractor.* **C)** *Discrete time simulation with MaBoSS. Even though the state probabilities reduce with time, we can clearly observe the cyclic behavior of the system.*

To keep the cyclic behavior of the probabilities, we can use discrete time modeling with MaBoSS. The results obtained with this approach are shown in figure 3.5C. In this case, the cyclic behavior of the model is clearly maintained, with a much slower decrease through time due to the stochastic events. Here, we clearly see the state transition replicating the results we obtain using BoolSim and SQUAD.

To sum up, both continuous and discrete MaBoSS simulations validate the results we obtain with BoolSim and SQUAD. In the case of continuous time simulation we obtain an attractor with a state probability graph similar to the state transition graph we obtain with BoolSim/SQUAD. Given that

MaBoSS is a stochastic, probability-based software that focuses on cell populations (instead of single cell simulations in the case of SQUAD), we end up in a damped oscillation where each state has similar probabilities. However, looking at the asymptotic behavior of entropy (not converging to 0) and transition entropy (converging to 0) we can characterize this stationary distribution as cyclic stationary distribution.

Discussion

In this project we make use of qualitative methods to evaluate whether we can characterize the budding yeast cyclin wave behavior with such a level of abstraction. For this purpose we constructed the possible model configuration, starting from the experimental facts that had to be described. The 6 minimal models derived from Boolean and Standardized dynamic simulation analysis, compose an interesting collection that can be used as an important tool for research around Clb regulation.

It is important to note that the minimal models capture the information flow necessary for the system to present the expected phenotype, rather than direct biochemical relations. Many regulatory processes that would require an extension of the model have been omitted for the time being, such as the transcription and degradation regulation of the cyclins. It was in our interest, however, to try to understand the mechanisms regulating the relative timing of Clb waves, rather than the detailed activation and degradation processes. Therefore, we focused on the interplay among cyclins and their inhibitor Sic1.

Our future work will focus on choosing the best candidate among the minimal models. Since for the construction of the models we only used Sic1 perturbations, we will now perform *in silico* perturbations of the cyclin molecules and test the predictions in the lab. This will be achieved in collaboration with M. Barberis lab where they will evaluate the plausibility of the models' wiring and eventually choose the best fitting model.

Chapter 4

SBV Improver Challenge: Species Translation Sub-challenge 4: Species Specific Network Inference A solution using treatment-specific network inference

Partially based on the supplementary material of

Bilal E, Sakellaropoulos T, Participants C, Melas IN, Messinis DE, Belcastro V, Rhrissorrakrai K, Meyer P, Norel R, Iskandar A, Blaese E, Rice JJ, Peitsch MC, Hoeng J, Stolovitzky G, Alexopoulos LG, Poussin C (2014) *A crowd-sourcing approach for the construction of species-specific cell signaling networks*. Bioinformatics

Introduction

Sbv IMPROVER stands for Systems Biology Verification combined with Industrial Methodology for Process Verification in Research. This approach aims to provide a measure of quality control of industrial research and development by verifying the methods used. The scope of sbv IMPROVER is the verification of methods and concepts in systems biology research using crowd-sourcing challenges. A complex research program is typically built upon research projects (consisting of "building blocks") that synergistically support each other towards a final goal. Each building block is formulated as a standalone challenge of a complex workflow. It has a defined input that results in a defined output (Meyer et al, 2011; Meyer et al, 2012).

Species Translation Challenge: Species specific network inference

Systems biology emphasizes in studying relationships and connectivity between the components of complex systems. As such, pathway diagrams are the primary representation of complex biological systems, and the construction of accurate and complete pathway maps is an on-going challenge in the field. The two main approaches that have been taken to build pathway maps are knowledge driven and data driven. The knowledge-driven approach uses *a priori* data, often curated from the literature, to define entities (nodes) and connections (edges) that can be assembled into network diagrams. In contrast, the data-driven approach seeks to infer the connection based on inference from large dataset using methods such as regression analysis and Bayesian probabilistic models. Combining disparate data types in pathway maps is a useful way of synthesizing such diverse knowledge into a consistent and unified view of a complex biological system. In addition, knowledge-driven approaches are often used to construct the scaffold network that can be augmented and refined using data-driven approaches.

The aim of the Species Translation Challenge was to identify rules which map measurements derived from systematic perturbations in one species to another species, quantify the translatability between species and understand the limitation of species translatability. The Species Translation Challenge addressed the translatability of findings between rat and human model systems. The four Sub Challenges addressed different aspects of this problem.

The goal of the SBV Improver sub-challenge 4 was to infer human and rat networks given phosphoprotein, gene expression and cytokine data and a reference network provided as prior knowledge. Participants were required to use network inference to add or remove edges from the reference map in order to produce specific rat and human networks (Figure 4.1)

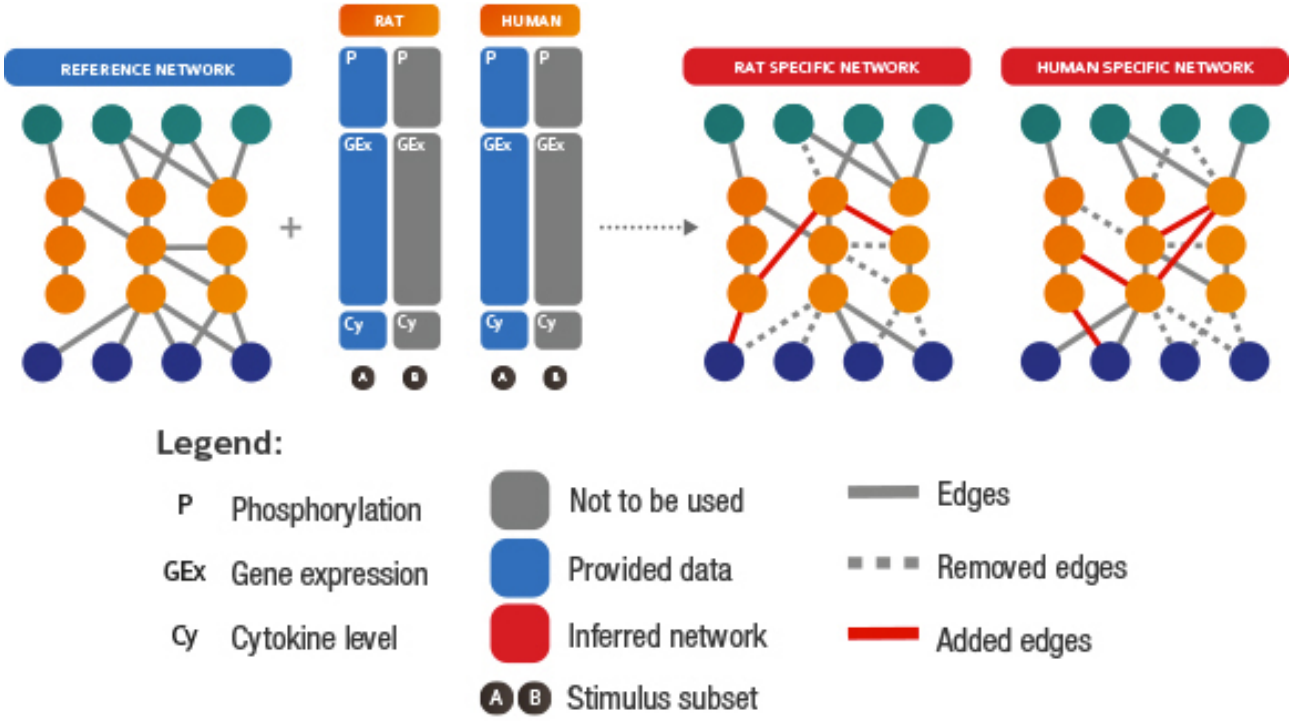

Figure 4.1: The sub-challenge 4 – Species specific network inference. *In sub-challenge 4, a Reference Network was provided to participants. Participants were asked to construct human- and rat-specific networks given the omics data provided. Participants would have to use network inference to add or remove edges from the Reference Network based on Phosphoprotein (P), Gene Expression (GEx) and Cytokine (Cy) data in training sets for rat and human. Only human and rat data from subset A were to be used to allow for proper comparability between the respective networks. Figure source: SBV Improver website (sbvimprover.com/challenge-2/sub-challenge-4-species-network-inference).*

Produced Datasets

Normal bronchial epithelial cells from human (NHBE) and rat (NRBE) were exposed, in parallel, to 52 various stimulus types or to control medium (DME: Dulbecco's Modified Eagle's Medium corresponding to standard cell culture medium for those cells). Cells were collected and lysed at different time points depending on the type of measurements done with the samples (5 and 25 minutes for phosphoproteins; 6 hours for gene expression). The cell supernatant was collected at 24 hours for cytokine level measurement. The exposure of cells to each stimulus was performed in triplicate, and in quintuplicate and sextuplicate for the DME control. Two experiments were performed to collect all samples. Experiment 1 contained 3/4 of the stimuli (with respective

controls) and experiment 2 contained the remaining 1/4 of the stimuli with respective controls. More information on the data set used can be found at (Poussin et al, 2014).

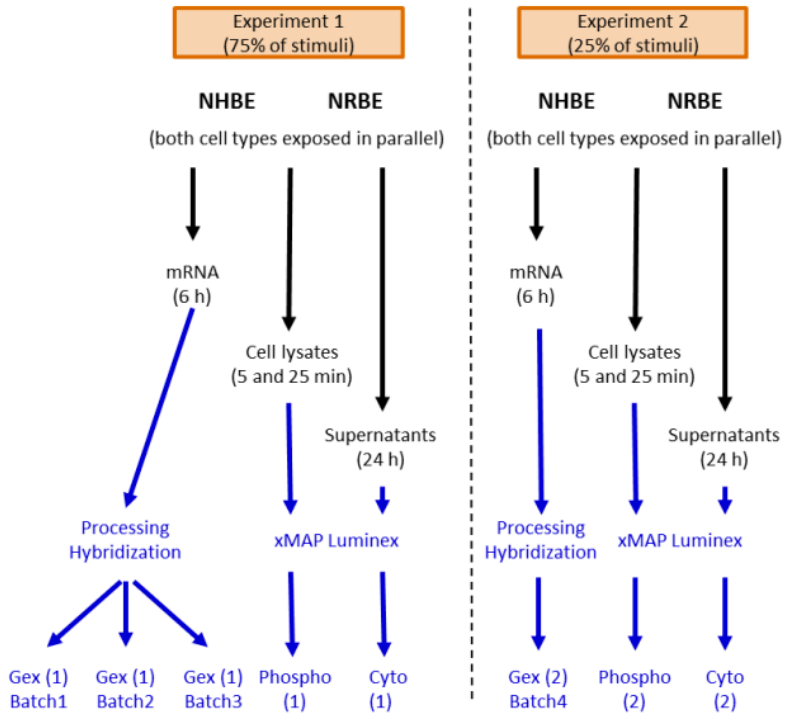

Figure 4.2: Challenge data generation scheme. *Schema summarizing the experiments performed to generate the complete dataset for the Species Translation challenge (mRNA: messenger RNA; NHBE: normal human bronchial epithelial cells; NRBE: normal rat bronchial epithelial; GEx: gene expression; Phospho: phosphorylation; Cyto: cytokine). Figure source: SBV Improver, Challenge documentation, Detailed Information about the Data.*

Results

For the Species Specific Network Inference Sub-Challenge of SBV Improver, we decided to conduct a treatment-based analysis. The goal was to identify the genes that are differentially expressed in each treatment, and create treatment specific networks, based on the provided reference network (figure 4.3). The ensemble of all treatment specific networks for a species constitutes the species-specific network. All scripts of this work were developed in R, and several packages were used, such as limma and igraph.

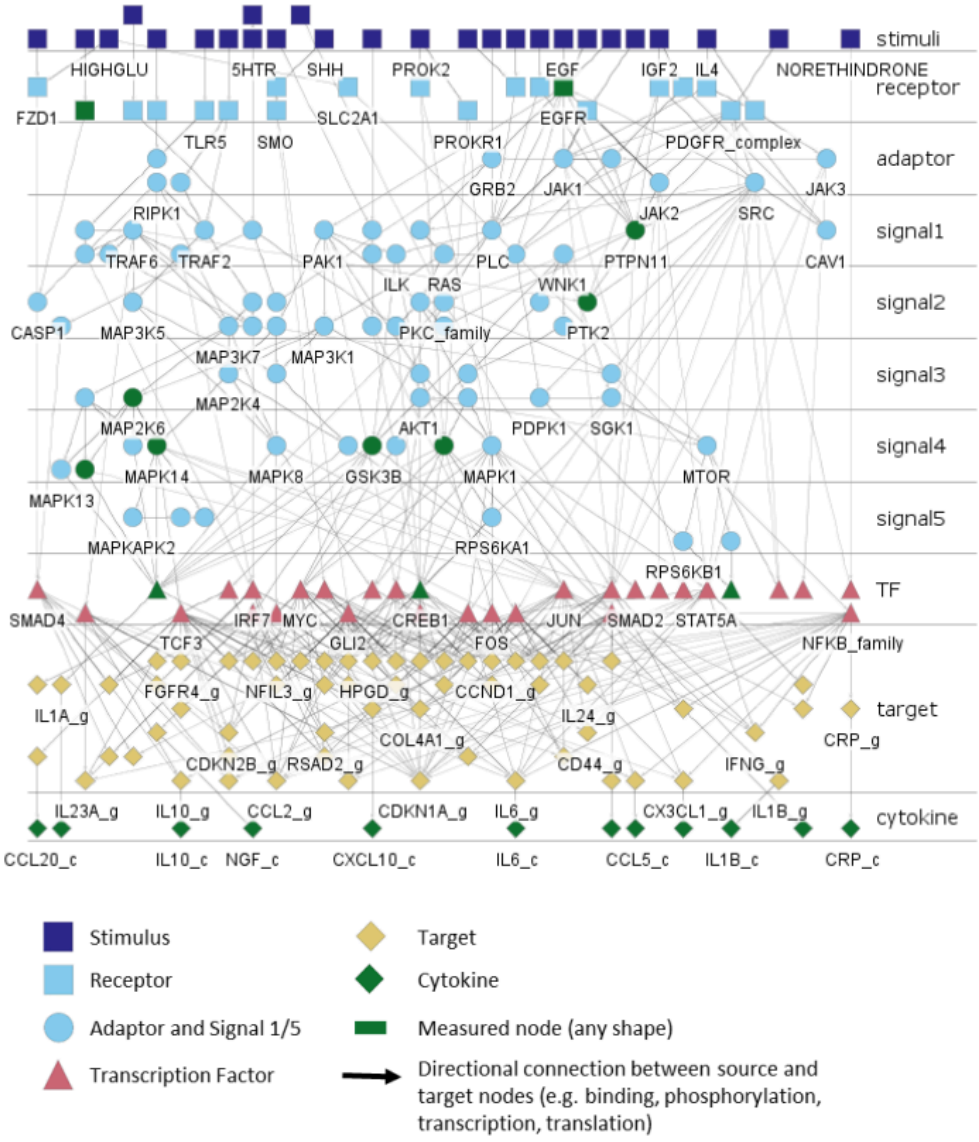

Figure 4.3: Challenge reference network. *Snapshot of the Reference Network visualized in Cytoscape. Figure source: SBV Improver, Challenge documentation, Detailed Information about the Data.*

Data Mapping

The first step of the analysis was to identify the genes that are present in the reference network but do not appear with the same name in the gene expression data. The NCBI gene database was queried for those genes and the gene expression datasets were scanned against the alternative names retrieved. In case there was more than one matching measurement in the data, the most variant gene would be selected. All cytokine and phosphoprotein names were mapped to their gene names using UniProtKB. Thereafter, all data files were filtered to contain only elements found in the reference network.

CD120	IKB_family
PI3K	IKK_family
PKA	NFKB_family
PLC	PGR_family
SOS	PKC_family
AC_family	PDGFR_family
ADR_family	FGFR_family

Table 4.1: Name inconsistencies in data. *List of reference network nodes that were not represented in the GEx data with the same name and were retrieved by public database query.*

Data Normalization

The workflow in its entirety was performed separately for the human and rat measurements. During the first step of normalization, each batch was treated separately. The average control (DME) value of all repetitions from the specific batch was calculated for each gene. The average values vector was then subtracted from the treatment expression values belonging to the same batch. No averaging was performed for treatment repetitions.

For the second normalization step all data were merged. For each repetition of each treatment the z score was calculated. The same process was followed for the phosphoprotein and cytokine datasets.

Treatment Specific Differential Expression

To define treatment specific, differentially expressed genes, a moderated t-test for each treatment against the remaining treatments and controls was performed , and genes with p value < 0.05 were selected as differentially expressed. This process was repeated in all datasets (gene expression, phosphoproteins, cytokines). ANOVA analysis of the gene expression dataset revealed a number of genes that were not found with the t-tests. We assumed that this is a result of treatments that have very similar expression patterns but their effect is underestimated when one of the treatments is used in the "control set" for the t-test against the other. We used hierarchical clustering on the treatments, and divided them to clusters to test this hypothesis. We performed again the t-test for each treatment against the rest, excluding members of the same cluster. Indeed, this method yielded additional differentially expressed genes that were added to the results.

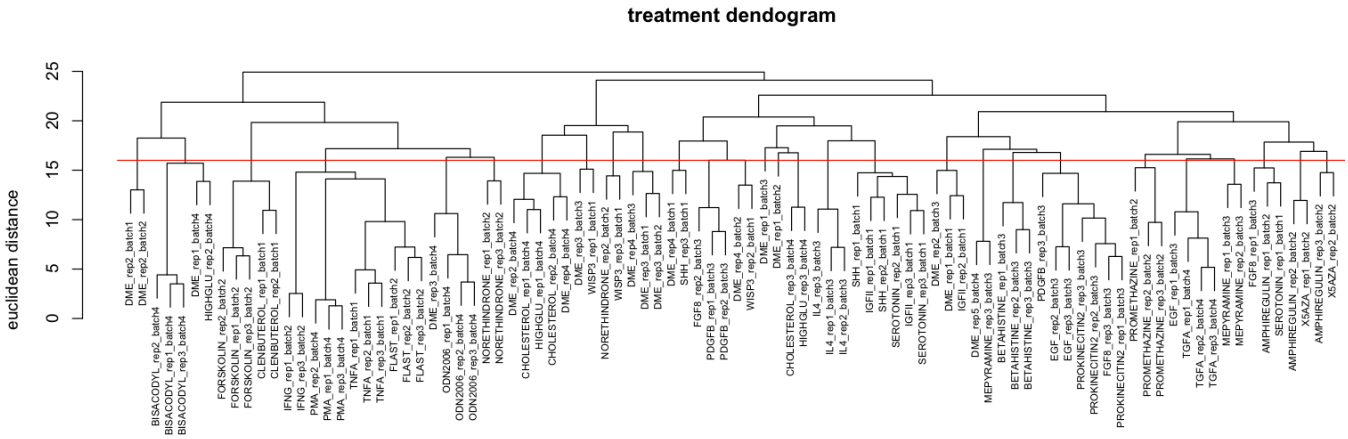

Figure 4.4: Clustering of treatments. *Hierarchical clustering of treatments in human, based on euclidean distance. The red line corresponds to the clustering cutoff, same for rat and human.*

Treatment specific network inference

At this stage we had a list of highlighted (i.e. differentially expressed) genes for each treatment, derived from t-tests performed in gene expression, phosphoproteins, cytokines and GEx clustered datasets. For the network inference of each stimulus, we evaluated the reachability of every highlighted gene pair in the reference network. Having as a principle that every highlighted gene should be reachable from the stimulus node, we added missing edges according to the following rules:

1. Every node should be reachable from the stimulus node
2. If not, test if the node is connected to other upstream, highlighted nodes
3. If it is not connect node to all highlighted nodes of the level above
4. If none, reveal the paths of all upstream nodes to all downstream nodes, and connect the node to all nodes of the level above.

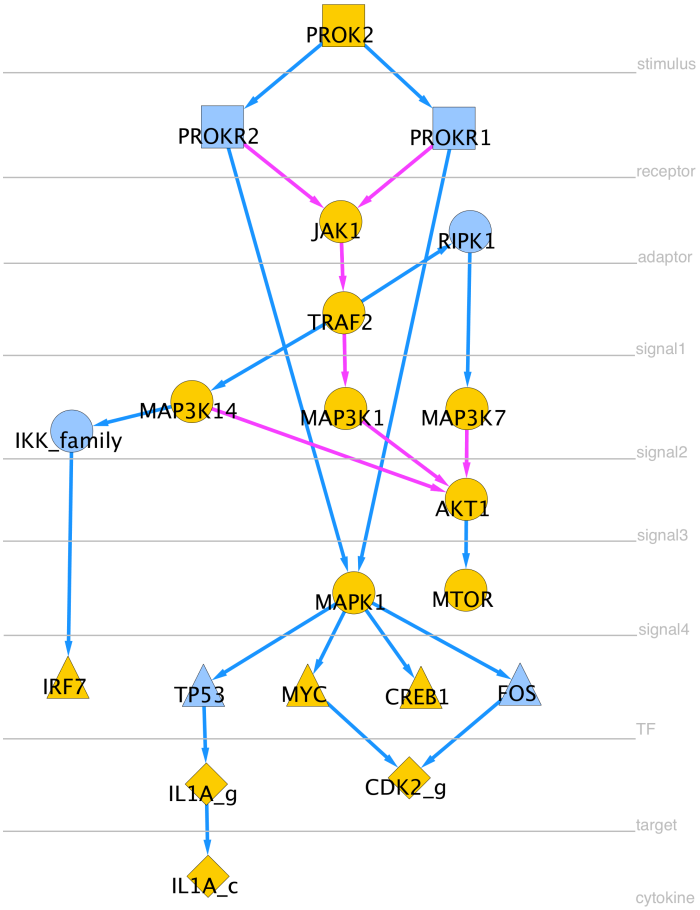

Figure 4.5: Treatment specific network inference. *Adding missing links for treatment specific network inference, the case of human PROKINECITIN2. Yellow nodes are differentially expressed for PROK2 and blue nodes are not, but are used for path reconstruction. Blue links represent already existing paths and pink links are the ones added. In the case of JAK1, there were no highlighted nodes on the level above. Therefore, all paths that connect PROK2 to downstream nodes were revealed and JAK1 was connected to the nodes of the level above, i.e. PROKR1 and PROKR2. For the connection of TRAF2 no path revealing was necessary since JAK1, which is on the level above, was already highlighted. The same principles are applied for the rest of the added edges.*

Species specific network inference

The added edges of all treatments were appended to the reference network. For every highlighted gene pair of each stimulus, all shortest paths were calculated. Every time an edge was found in the path, the edge received +1 point (which corresponds to number of visits). The final, species-specific network was derived by keeping only the visited edges of the extended reference network. A presentation of the challenge meta-analysis, the evaluation methods as well as the approaches followed by the best scoring teams including our approach, can be found at (Bilal et al, 2014). Our

team scored 3rd in the challenge. The best performing teams were announced at the sbv IMPROVER Symposium 2013 in Greece and the identity of the best performing teams for each challenge was published in Nature (Volume 503 Number 7476, 21 November 2013, Naturejobs page 12).

Acknowledgements

I would like to thank the members of Vital-IT that participated with me in this challenge: Mark Ibberson, Leonore Wigger, Nicolas Guex, Julien Dorier and Ioannis Xenarios.

For the Septation Initiation project, as well as the Clb interplay project, a wide range of tools and scripts were used throughout the workflow.

Septation Initiation Project Scripts

For the workflow stages of “model simulation and steady state prediction” as well as “optimization based on experimental data” (figure 2.13), several scripts were developed using R and Groovy programming languages. The produced scripts served as parts of the standard analysis workflow as well as specific efforts such as the scoring randomization (figure 2.7) and they covered the following functionalities:

- construct experimental tree and input for simulation
- structure and parse results
- score model candidates / selected experiments
- analyze / visualize results and scoring
- scoring randomization

The developed scripts were used for the initial stages of the Clb interplay project as well.

Modeling software used

BoolSim

BoolSim was the main modeling software used in this thesis, for the identification of Boolean attractors and *in silico* perturbation experiments. Gene regulatory networks can be modeled efficiently using reduced ordered binary decision diagrams (ROBDDs or in short BDDs). BDDs are directed acyclic graphs that can represent large Boolean functions in a space efficient manner, and are computationally suitable for complex Boolean operations (e.g. logical AND, OR, etc.) and set operations (e.g. Union, Intersection, etc.). To map gene regulatory networks on BDDs, the first step is to transform networks into Boolean functions, which represent the dynamics of a model. All the operations that can be performed on Boolean functions can also be performed on their

corresponding BDD representations (Garg et al, 2008). BoolSim uses this strategy to efficiently compute cyclic attractory of large networks that are not feasible using other existing software.

SQUAD

SQUAD was mainly used in the Clb interplay project, for all continuous simulations of Clb sequential expression. It was also used in SIN analysis for exploratory purposes, in order to observe the nodes that did not change activation state during simulation (with the initial, non controlled modeling strategy), indicating problemating wiring.

SQUAD is a software for the dynamic simulation of signaling networks using the standardized qualitative dynamical systems approach. SQUAD converts the network into a discrete dynamical system, and it uses a binary decision diagram algorithm to identify all the steady states of the system. Then, the software creates a continuous dynamical system and localizes its steady states which are located near the steady states of the discrete system. The software permits to make simulations on the continuous system, allowing for the modification of several parameters. Importantly, SQUAD includes a framework for perturbing networks in a manner similar to what is performed in experimental laboratory protocols, for example by activating receptors or knocking out molecular components (Di Cara et al, 2007).

MaBoSS

MaBoSS was used in the Clb interplay project, for the verification of the results obtained using BoolSim and SQUAD, by another software.

MaBoSS is a modeling framework based on a qualitative approach that is intrinsically continuous in time. The algorithm fills the gap between qualitative and quantitative modeling. It is based on continuous time Markov process applied on a Boolean state space. In order to describe the temporal evolution of the biological process to be modeled, the transition rates for each node are sexplicitely specified. For that purpose, MaBoSS uses a generalization of Boolean equations. Mathematically, this approach can be translated in a set of ordinary differential equations on probability distributions. The software is developed in C++ , and is able to simulate such a system by applying Kinetic Monte-Carlo (or Gillespie algorithm) on the Boolean state space. This software, parallelized and optimized, computes the temporal evolution of probability distributions and estimates stationary distributions (Stoll et al, 2012).

Concluding Remarks

Computational models of biological processes have gained an important role in systems biology, not because they are correct, but because they can be useful. Indeed, models can only approximate reality. Yet the simplicity of these models makes them useful anyway, often more useful in day-to-day life than more complicated models that better fit the data. The idea that models never fit all the data was nicely summarized by the statistician George Box in his famous phrase, “all models are wrong, some models are useful” (Box, 1979). Systems biologist Arthur Lander, in his thought-provoking article “The edges of understanding” further advocates this idea, by insisting on “utility before validation” (Lander, 2010).

In the present thesis, Boolean modeling has been used as core approach, in order to infer the regulatory relations and systems’ dynamics in several biological contexts. Even though Boolean models have limited capacity to describe quantitative characteristics of dynamic systems, they reveal considerable dynamic richness and were proven effective in describing qualitative behaviors of biological systems (Albert & Thakar, 2014). The fact that Boolean models do not require the knowledge of kinetic parameters makes them a practical choice for systems where these parameters have not been measured. The success of Boolean models illustrates that, in at least a subset of biological systems, the organization of network structure plays a more important role than kinetic details of individual interactions.

An advantage of working on diverse biological problems is that the toolset and methods developed for a project are directly available to be applied in other contexts. For example, budding yeast modeling was the most recent project I undertook during my PhD. The minimal model strategy developed in this project could be very helpful for the continuation of the SIN modeling research. In fact, applying the technique to subsets of the S.*pombe* model, such as, for example, to the Byr4p, Cdc16p, Spg1p and Etd1p minimal node set, could potentially facilitate the deciphering of Etd1p regulatory relation to the system.

Despite the diversity of the biological questions described in this thesis, many are the modeling challenges that were not addressed or deeply investigated. To begin with, an important field of research in computational modeling is reverse engineering of regulatory networks. This is achieved by identifying the underlying network on the basis of correlated molecular behavior observed in genome-wide "omics" studies (top-down approach). Even though there are some success stories using this approach (see, for example, (Davidson, 2002 #223)) there is still extensive research on

model inference. The aim, in this case, is to identify the real underlying network that results in the observed phenotypes. However, the "gene-system-phenotype" puzzle is not an easy one to solve. Many networks can theoretically generate the same phenotype. This happens, for example, in sister cells with same phenotype but different transcriptomes / proteomes / metabolomes, as well as in a single cell where transcriptomes / proteomes / metabolomes change over time but the phenotype remains the same. An old but always relevant illustration by Conrad Hal Waddington exemplifies this point (figure 5.1). In his illustration we can see a network of interactions putting specific tensions to the system, defining the system state trajectories and, ultimately, the cell phenotype. Still, any other combination of interaction network with the same tensions would result in the same phenotype.

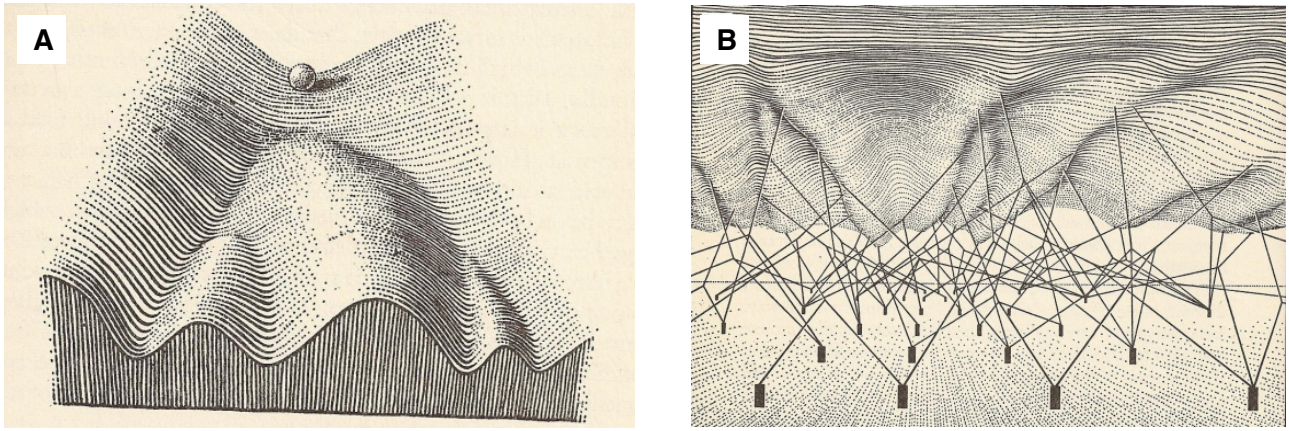

Figure 5.1: A) *A depiction of the epigenetic landscape. The ball represents a cell, and the bifurcating system of valleys represents the trajectories in state space.* **B)** *Behind the scenes of Waddington's landscape. The valleys are formed by tension on cables (interactions) that are attached to pegs (genes). Source: (Waddington, 1957)*

Another important issue in modeling, and a big bet for future research, is the problem of scales. Biological systems consist of many spatial and temporal scales, each rich and complex. The levels of biological organization range from molecules to genes, proteins, individual biological cells, tissues, organs, and up to organisms that interact with the environment (Southern et al, 2008). Associated with this spatially based organization are the temporal scales of biological processes that range from microsecond (10^{-6} s) for molecular interactions to 80 years (10^{9} s) for the average human life expectancy (Walker & Southgate, 2009). This enormous diversity of biological events' scales is one of the greatest challenges for computational modeling. Indeed, successful physiological analyses require an understanding of the functional interactions between the key components of cells, organs, and systems, as well as how these interactions change in disease states (Noble, 2002). Current model frames are restricted in magnitude of description, temporally as well as spatially. There have been some multi-scaling modeling efforts in the past (Dada &

Mendes, 2011; Southern et al, 2008). The resulting complex integrated models, however, are difficult to solve numerically even with the incredible advances in computational power. The major challenge in multi-scale modeling is how to couple together the various models that are available at different spatial and temporal scales. Even if all the models of a biological system at different scales can easily be integrated, the task of simulating the integrated models is computationally very expensive. High performance computing, parallel processing, grid computing and similar technologies will undoubtedly help in the future, but the most important step will be to develop adequate software in order to take full advantage of the hardware.

The future of biological research lies in dealing with complexity. As *lex parsimoniae* rarely applies, we will have to ascertain the central factors of each process and focus our attention on them, in order to gain better understanding. Computational modeling will most definitely play a significant role in this process. Asking the right questions will help us define the adequate assumptions that will lead to the answers we are looking for. And then, who knows, maybe we will be able to understand the flock and create even more majestic patterns.

- **Chasapi A**, Wachowicz P, Niknejad A, Collin P, Krapp A, Cano E, Simanis V, Xenarios I (2015) *An Extended, Boolean Model of the Septation Initiation Network in S.pombe Provides Insights into its Regulation.* PloS Comp Biol (submitted)

- Wachowicz P, **Chasapi A**, Krapp A, Cano Del Rosario E, Schmitter D, Sage D, Unser M, Xenarios I, Rougemont J, Simanis V (2014) Analysis of S. pombe SIN protein SPB-association reveals two genetically separable states of the SIN. *Journal of cell science*

- Linke C, **Chasapi A**, Gonzalez-Novo A, Klipp E, Krobitsch S, Posas F, Xenarios I, Barberis M (2015) *Fkh transcription factors controls timing of Clb expression to regulate waves of mitotic cyclins.* (in preparation)

- Schmitter D, Wachowicz P, Sage D, **Chasapi A**, Xenarios I, Simanis, Unser M (2013) A 2D/3D image analysis system to track fluorescently labeled structures in rod-shaped cells: application to measure spindle pole asymmetry during mitosis. *Cell division*

- SBV Improver Species Translation Challenge. Best performing team announcement published in Nature (Volume 503 Number 7476, 21 November 2013, Naturejobs page 12).

- Naldi A, Monteiro PT, Mussel C, the Consortium for Logical M, Tools, Kestler HA, Thieffry D, Xenarios I, Saez-Rodriguez J, Helikar T, Chaouiya C (2015) Cooperative development of logical modelling standards and tools with CoLoMoTo. *Bioinformatics*

- Bilal E, Sakellaropoulos T, Participants C, Melas IN, Messinis DE, Belcastro V, Rhrissorrakrai K, Meyer P, Norel R, Iskandar A, Blaese E, Rice JJ, Peitsch MC, Hoeng J, Stolovitzky G, Alexopoulos LG, Poussin C (2014) A crowd-sourcing approach for the construction of species-specific cell signaling networks. *Bioinformatics* (supplementary material)

- Poster presentation in ECCB'12 Conference (9-12/9/2012) in Basel, Switzerland. "*Modeling the Septation Initiation Network in fission yeast*". **Chasapi A**, Wachowicz P, Niknejad A, Schmitter D, Cerutti L, Dorier J, Sage D, Simanis V, Xenarios I.

- Poster presentation in ISMB/ECCB 2013 Conference (19-23/7/2013) in Berlin, Germany. "*Boolean modeling of the Septation Initiation Network in fission yeast*". **Chasapi A**, Wachowicz P, Schmitter D, Niknejad A, Sage D, Simanis V, Xenarios I.

- Poster presentation in SBV Improver Symposium (29-31/10/2013) in Lagonissi, Greece. "*Treatment-based networks for species specific network inference*". **Chasapi A**, Wigger L, Dorier J, Xenarios I, Ibberson M, Guex N

- Poster presentation in Biocuration 2014 Conference (6-9/4/2014) in Toronto, Canada. "*Modeling and Simulation: The vital role of biocurated models*". **Chasapi A**, Niknejad A, Dorier J, Crespo I, Naldi A, Bridge A, Ibberson M, Xenarios I.

- Poster presentation in SB@NL Symposium (15-16/12/2014) in Maastricht, Netherlands. "Boolean Modeling of the Septation Initiation in fission yeast". **Chasapi A**, Wachowich P, Niknejad A, Krapp A, Cano E, Simanis V, Xenarios I.

Bibliography

Akutsu T, Miyano S, Kuhara S (1999) Identification of genetic networks from a small number of gene expression patterns under the Boolean network model. *Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing***:** 17-28

Alarcon T, Tindall MJ (2007) Modelling cell growth and its modulation of the G1/S transition. *Bulletin of mathematical biology* **69:** 197-214

Albert R, Othmer HG (2003) The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in Drosophila melanogaster. *Journal of theoretical biology* **223:** 1-18

Albert R, Thakar J (2014) Boolean modeling: a logic-based dynamic approach for understanding signaling and regulatory networks and for making useful predictions. *Wiley interdisciplinary reviews Systems biology and medicine* **6:** 353-369

Alcaide-Gavilan M, Lahoz A, Daga RR, Jimenez J (2014) Feedback regulation of SIN by Etd1 and Rho1 in fission yeast. *Genetics* **196:** 455-470

Aldridge BB, Burke JM, Lauffenburger DA, Sorger PK (2006) Physicochemical modelling of cell signalling pathways. *Nature cell biology* **8:** 1195-1203

Amon A (1997) Regulation of B-type cyclin proteolysis by Cdc28-associated kinases in budding yeast. *The EMBO journal* **16:** 2693-2702

Amon A, Irniger S, Nasmyth K (1994) Closing the cell cycle circle in yeast: G2 cyclin proteolysis initiated at mitosis persists until the activation of G1 cyclins in the next cycle. *Cell* **77:** 1037-1050

Andrews B, Measday V (1998) The cyclin family of budding yeast: abundant use of a good idea. *Trends in genetics : TIG* **14:** 66-72

Azpeitia E, Benitez M, Padilla-Longoria P, Espinosa-Soto C, Alvarez-Buylla ER (2011) Dynamic network-based epistasis analysis: boolean examples. *Frontiers in plant science* **2:** 92

Bahler J (2005) Cell-cycle control of gene expression in budding and fission yeast. *Annual review of genetics* **39:** 69-94

Bajpai A, Feoktistova A, Chen JS, McCollum D, Sato M, Carazo-Salas RE, Gould KL, Csikasz-Nagy A (2013) Dynamics of SIN asymmetry establishment. *PLoS computational biology* **9:** e1003147

Balasubramanian MK, McCollum D, Chang L, Wong KC, Naqvi NI, He X, Sazer S, Gould KL (1998) Isolation and characterization of new fission yeast cytokinesis mutants. *Genetics* **149:** 1265-1275

Barberis M, De Gioia L, Ruzzene M, Sarno S, Coccetti P, Fantucci P, Vanoni M, Alberghina L (2005a) The yeast cyclin-dependent kinase inhibitor Sic1 and mammalian p27Kip1 are functional homologues with a structurally conserved inhibitory domain. *The Biochemical journal* **387:** 639-647

Barberis M, Klipp E, Vanoni M, Alberghina L (2007) Cell size at S phase initiation: an emergent property of the G1/S network. *PLoS computational biology* **3:** e64

Barberis M, Linke C, Adrover MA, Gonzalez-Novo A, Lehrach H, Krobitsch S, Posas F, Klipp E (2012) Sic1 plays a role in timing and oscillatory behaviour of B-type cyclins. *Biotechnology advances* **30:** 108-130

Barberis M, Pagano MA, Gioia LD, Marin O, Vanoni M, Pinna LA, Alberghina L (2005b) CK2 regulates in vitro the activity of the yeast cyclin-dependent kinase inhibitor Sic1. *Biochemical and biophysical research communications* **336:** 1040-1048

Bardin AJ, Amon A (2001) Men and sin: what's the difference? *Nature reviews Molecular cell biology* **2:** 815-826

Basi G, Schmid E, Maundrell K (1993) TATA box mutations in the Schizosaccharomyces pombe nmt1 promoter affect transcription efficiency but not the transcription start point or thiamine repressibility. *Gene* **123:** 131-136

Bateman A (2010) Curators of the world unite: the International Society of Biocuration. *Bioinformatics* **26:** 991

Baum B, Wuarin J, Nurse P (1997) Control of S-phase periodic transcription in the fission yeast mitotic cycle. *The EMBO journal* **16:** 4676-4688

Bilal E, Sakellaropoulos T, Participants C, Melas IN, Messinis DE, Belcastro V, Rhrissorrakrai K, Meyer P, Norel R, Iskandar A, Blaese E, Rice JJ, Peitsch MC, Hoeng J, Stolovitzky G, Alexopoulos LG, Poussin C (2014) A crowd-sourcing approach for the construction of species-specific cell signaling networks. *Bioinformatics*

Bloom J, Cross FR (2007) Multiple levels of cyclin specificity in cell-cycle control. *Nature reviews Molecular cell biology* **8:** 149-160

Box GEP (1979) Robustness in Statistics. In *Robustness in the strategy of scientific model building*, Launer RL, G.N. W (eds), pp 201–236. Academic Press

Bruggeman FJ, Westerhoff HV (2007) The nature of systems biology. *Trends in microbiology* **15:** 45-50

Calzada A, Sacristan M, Sanchez E, Bueno A (2001) Cdc6 cooperates with Sic1 and Hct1 to inactivate mitotic cyclin-dependent kinases. *Nature* **412:** 355-358

Cerutti L, Simanis V (1999) Asymmetry of the spindle pole bodies and spg1p GAP segregation during mitosis in fission yeast. *Journal of cell science* **112 (Pt 14):** 2313-2321

Chang L, Gould KL (2000) Sid4p is required to localize components of the septation initiation pathway to the spindle pole body in fission yeast. *Proceedings of the National Academy of Sciences of the United States of America* **97:** 5249-5254

Chang L, Morrell JL, Feoktistova A, Gould KL (2001) Study of cyclin proteolysis in anaphase-promoting complex (APC) mutant cells reveals the requirement for APC function in the final steps of the fission yeast septation initiation network. *Molecular and cellular biology* **21:** 6681-6694

Chaouiya C (2007) Petri net modelling of biological networks. *Briefings in bioinformatics* **8:** 210-219

Chaves M, Albert R, Sontag ED (2005) Robustness and fragility of Boolean models for genetic regulatory networks. *Journal of theoretical biology* **235:** 431-449

Chaves M, Sontag ED, Albert R (2006) Methods of robustness analysis for Boolean models of gene control networks. *Systems biology* **153:** 154-167

Chen KC, Calzone L, Csikasz-Nagy A, Cross FR, Novak B, Tyson JJ (2004) Integrative analysis of cell cycle control in budding yeast. *Molecular biology of the cell* **15:** 3841-3862

Chen KC, Csikasz-Nagy A, Gyorffy B, Val J, Novak B, Tyson JJ (2000) Kinetic analysis of a molecular model of the budding yeast cell cycle. *Molecular biology of the cell* **11:** 369-391

Chew TG, Balasubramanian MK (2008) Nuc2p, a subunit of the anaphase-promoting complex, inhibits septation initiation network following cytokinesis in fission yeast. *PLoS genetics* **4:** e17

Coudreuse D, Nurse P (2010) Driving the cell cycle with a minimal CDK control network. *Nature* **468:** 1074-1079

Cross FR, Yuste-Rojas M, Gray S, Jacobson MD (1999) Specialization and targeting of B-type cyclins. *Molecular cell* **4:** 11-19

Csikasz-Nagy A, Kapuy O, Gyorffy B, Tyson JJ, Novak B (2007) Modeling the septation initiation network (SIN) in fission yeast cells. *Current genetics* **51:** 245-255

Cvijovic M, Almquist J, Hagmar J, Hohmann S, Kaltenbach HM, Klipp E, Krantz M, Mendes P, Nelander S, Nielsen J, Pagnani A, Przulj N, Raue A, Stelling J, Stoma S, Tobin F, Wodke JA, Zecchina R, Jirstrand M (2014) Bridging the gaps in systems biology. *Molecular genetics and genomics : MGG* **289:** 727-734

Dada JO, Mendes P (2011) Multi-scale modelling and simulation in systems biology. *Integrative biology : quantitative biosciences from nano to macro* **3:** 86-96

Daga RR, Lahoz A, Munoz MJ, Moreno S, Jimenez J (2005) Etd1p is a novel protein that links the SIN cascade with cytokinesis. *The EMBO journal* **24:** 2436-2446

Davidich MI, Bornholdt S (2008) Boolean network model predicts cell cycle sequence of fission yeast. *PloS one* **3:** e1672

de Jong H (2002) Modeling and simulation of genetic regulatory systems: a literature review. *Journal of computational biology : a journal of computational molecular cell biology* **9:** 67-103

Deshaies RJ (1997) Phosphorylation and proteolysis: partners in the regulation of cell division in budding yeast. *Current opinion in genetics & development* **7:** 7-16

Di Cara A, Garg A, De Micheli G, Xenarios I, Mendoza L (2007) Dynamic simulation of regulatory networks using SQUAD. *BMC bioinformatics* **8:** 462

Dischinger S, Krapp A, Xie L, Paulson JR, Simanis V (2008) Chemical genetic analysis of the regulatory role of Cdc2p in the S. pombe septation initiation network. *Journal of cell science* **121:** 843-853

Donovan JD, Toyn JH, Johnson AL, Johnston LH (1994) P40SDB25, a putative CDK inhibitor, has a role in the M/G1 transition in Saccharomyces cerevisiae. *Genes & development* **8:** 1640-1653

Enserink JM, Kolodner RD (2010) An overview of Cdk1-controlled targets and processes. *Cell division* **5:** 11

Fankhauser C, Simanis V (1994) The cdc7 protein kinase is a dosage dependent regulator of septum formation in fission yeast. *The EMBO journal* **13:** 3011-3019

Faure A, Naldi A, Chaouiya C, Thieffry D (2006) Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics* **22:** e124-131

Feist AM, Palsson BO (2008) The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli. *Nature biotechnology* **26:** 659-667

Feoktistova A, Morrell-Falvey J, Chen JS, Singh NS, Balasubramanian MK, Gould KL (2012) The fission yeast septation initiation network (SIN) kinase, Sid2, is required for SIN asymmetry and regulates the SIN scaffold, Cdc11. *Molecular biology of the cell* **23:** 1636-1645

Ferrell JE, Jr., Tsai TY, Yang Q (2011) Modeling the cell cycle: why do certain circuits oscillate? *Cell* **144:** 874-885

Fitch I, Dahmann C, Surana U, Amon A, Nasmyth K, Goetsch L, Byers B, Futcher B (1992) Characterization of four B-type cyclin genes of the budding yeast Saccharomyces cerevisiae. *Molecular biology of the cell* **3:** 805-818

Furge KA, Cheng QC, Jwa M, Shin S, Song K, Albright CF (1999) Regions of Byr4, a regulator of septation in fission yeast, that bind Spg1 or Cdc16 and form a two-component GTPase-activating protein with Cdc16. *The Journal of biological chemistry* **274:** 11339-11343

Furge KA, Wong K, Armstrong J, Balasubramanian M, Albright CF (1998) Byr4 and Cdc16 form a two-component GTPase-activating protein for the Spg1 GTPase that controls septation in fission yeast. *Current biology : CB* **8:** 947-954

Garcia-Cortes JC, McCollum D (2009) Proper timing of cytokinesis is regulated by Schizosaccharomyces pombe Etd1. *The Journal of cell biology* **186:** 739-753

Garg A, Di Cara A, Xenarios I, Mendoza L, De Micheli G (2008) Synchronous versus asynchronous modeling of gene regulatory networks. *Bioinformatics* **24:** 1917-1925

Garg A, Mohanram K, De Micheli G, Xenarios I (2012) Implicit methods for qualitative modeling of gene regulatory networks. *Methods Mol Biol* **786:** 397-443

Garg A, Mohanram K, Di Cara A, De Micheli G, Xenarios I (2009) Modeling stochasticity and robustness in gene regulatory networks. *Bioinformatics* **25:** i101-109

Giacomantonio CE, Goodhill GJ (2010) A Boolean model of the gene regulatory network underlying Mammalian cortical area development. *PLoS computational biology* **6**

Gidvani RD, Sudmant P, Li G, DaSilva LF, McConkey BJ, Duncker BP, Ingalls BP (2012) A quantitative model of the initiation of DNA replication in Saccharomyces cerevisiae predicts the effects of system perturbations. *BMC systems biology* **6:** 78

Goyal A, Takaine M, Simanis V, Nakano K (2011) Dividing the spoils of growth and the cell cycle: The fission yeast as a model for the study of cytokinesis. *Cytoskeleton (Hoboken)* **68:** 69-88

Grallert A, Connolly Y, Smith DL, Simanis V, Hagan IM (2012) The S. pombe cytokinesis NDR kinase Sid2 activates Fin1 NIMA kinase to control mitotic commitment through Pom1/Wee1. *Nature cell biology* **14:** 738-745

Grallert A, Hagan IM (2002) Schizosaccharomyces pombe NIMA-related kinase, Fin1, regulates spindle formation and an affinity of Polo for the SPB. *The EMBO journal* **21:** 3096-3107

Grallert A, Patel A, Tallada VA, Chan KY, Bagley S, Krapp A, Simanis V, Hagan IM (2013) Centrosomal MPF triggers the mitotic and morphogenetic switches of fission yeast. *Nature cell biology* **15:** 88-95

Guertin DA, Chang L, Irshad F, Gould KL, McCollum D (2000) The role of the sid1p kinase and cdc14p in regulating the onset of cytokinesis in fission yeast. *The EMBO journal* **19:** 1803-1815

Guertin DA, McCollum D (2001) Interaction between the noncatalytic region of Sid1p kinase and Cdc14p is required for full catalytic activity and localization of Sid1p. *The Journal of biological chemistry* **276:** 28185-28189

Guertin DA, Venkatram S, Gould KL, McCollum D (2002) Dma1 prevents mitotic exit and cytokinesis by inhibiting the septation initiation network (SIN). *Developmental cell* **3:** 779-790

Haase SB, Wittenberg C (2014) Topology and control of the cell-cycle-regulated transcriptional circuitry. *Genetics* **196:** 65-90

Hardy S, Robillard PN (2004) Modeling and simulation of molecular biology systems using petri nets: modeling goals of various approaches. *Journal of bioinformatics and computational biology* **2:** 595-613

He X, Patterson TE, Sazer S (1997) The Schizosaccharomyces pombe spindle checkpoint protein mad2p blocks anaphase and genetically interacts with the anaphase-promoting complex. *Proc Natl Acad Sci U S A* **94:** 7965-7970

Hergovich A, Hemmings BA (2012) Hippo signalling in the G2/M cell cycle phase: lessons learned from the yeast MEN and SIN pathways. *Seminars in cell & developmental biology* **23:** 794-802

Hirano T, Hiraoka Y, Yanagida M (1988) A temperature-sensitive mutation of the Schizosaccharomyces pombe gene nuc2+ that encodes a nuclear scaffold-like protein blocks spindle elongation in mitotic anaphase. *The Journal of cell biology* **106:** 1171-1183

Hochstrasser M (1995) Ubiquitin, proteasomes, and the regulation of intracellular protein degradation. *Current opinion in cell biology* **7:** 215-223

Hotz M, Barral Y (2014) The Mitotic Exit Network: new turns on old pathways. *Trends in cell biology* **24:** 145-152

Hou MC, Guertin DA, McCollum D (2004) Initiation of cytokinesis is controlled through multiple modes of regulation of the Sid2p-Mob1p kinase complex. *Molecular and cellular biology* **24:** 3262-3276

Hou MC, Salek J, McCollum D (2000) Mob1p interacts with the Sid2p kinase and is required for cytokinesis in fission yeast. *Current biology : CB* **10:** 619-622

Ingalls BP, Duncker BP, Kim DR, McConkey BJ (2007) Systems level modeling of the cell cycle using budding yeast. *Cancer informatics* **3:** 357-370

Irniger S, Piatti S, Michaelis C, Nasmyth K (1995) Genes involved in sister chromatid separation are needed for B-type cyclin proteolysis in budding yeast. *Cell* **81:** 269-278

Jia L, Kim S, Yu H (2013) Tracking spindle checkpoint signals from kinetochores to APC/C. *Trends in biochemical sciences* **38:** 302-311

Johnson AE, Chen JS, Gould KL (2013) CK1 is required for a mitotic checkpoint that delays cytokinesis. *Current biology : CB* **23:** 1920-1926

Johnson AE, Gould KL (2011) Dma1 ubiquitinates the SIN scaffold, Sid4, to impede the mitotic localization of Plo1 kinase. *The EMBO journal* **30:** 341-354

Johnson AE, McCollum D, Gould KL (2012) Polar opposites: Fine-tuning cytokinesis through SIN asymmetry. *Cytoskeleton (Hoboken)* **69:** 686-699

Kahlem P, DiCara A, Durot M, Hancock JM, Klipp E, Schachter V, Segal E, Xenarios I, Birney E, Mendoza L (2011) Strengths and Weaknesses of Selected Modeling Methods Used in Systems Biology. In *Systems and Computational Biology* N.S. Y (ed), Vol. Bioinformatics and Computational Modeling. InTech

Karlebach G, Shamir R (2008) Modelling and analysis of gene regulatory networks. *Nature reviews Molecular cell biology* **9:** 770-780

Kauffman SA (1993) *The origins of order : self-organization and selection in evolution*, New York ; Oxford: Oxford University Press.

Kestler HA, Wawra C, Kracher B, Kuhl M (2008) Network modeling of signal transduction: establishing the global view. *BioEssays : news and reviews in molecular, cellular and developmental biology* **30:** 1110-1125

King RW, Deshaies RJ, Peters JM, Kirschner MW (1996) How proteolysis drives the cell cycle. *Science* **274:** 1652-1659

Kitano H (2002) Systems biology: a brief overview. *Science* **295:** 1662-1664

Klipp E, Nordlander B, Kruger R, Gennemark P, Hohmann S (2005) Integrative model of the response of yeast to osmotic shock. *Nature biotechnology* **23:** 975-982

Koch C, Nasmyth K (1994) Cell cycle regulated transcription in yeast. *Current opinion in cell biology* **6:** 451-459

Kochi N, Matache MT (2012) Mean-field Boolean network model of a signal transduction network. *Bio Systems* **108:** 14-27

Krapp A, Cano E, Simanis V (2003) Mitotic hyperphosphorylation of the fission yeast SIN scaffold protein cdc11p is regulated by the protein kinase cdc7p. *Current biology : CB* **13:** 168-172

Krapp A, Collin P, Cano Del Rosario E, Simanis V (2008) Homoeostasis between the GTPase Spg1p and its GAP in the regulation of cytokinesis in S. pombe. *Journal of cell science* **121:** 601-608

Krapp A, Schmidt S, Cano E, Simanis V (2001) S. pombe cdc11p, together with sid4p, provides an anchor for septation initiation network proteins on the spindle pole body. *Current biology : CB* **11:** 1559-1568

Krien MJ, Bugg SJ, Palatsides M, Asouline G, Morimyo M, O'Connell MJ (1998) A NIMA homologue promotes chromatin condensation in fission yeast. *Journal of cell science* **111 (Pt 7):** 967-976

Kumada K, Su S, Yanagida M, Toda T (1995) Fission yeast TPR-family protein nuc2 is required for G1-arrest upon nitrogen starvation and is an inhibitor of septum formation. *Journal of cell science* **108 (Pt 3):** 895-905

Lahoz A, Alcaide-Gavilan M, Daga RR, Jimenez J (2010) Antagonistic roles of PP2A-Pab1 and Etd1 in the control of cytokinesis in fission yeast. *Genetics* **186:** 1261-1270

Lander AD (2010) The edges of understanding. *BMC biology* **8:** 40

Lew DJ, Reed SI (1995) A cell cycle checkpoint monitors cell morphogenesis in budding yeast. *The Journal of cell biology* **129:** 739-749

Li C, Furge KA, Cheng QC, Albright CF (2000) Byr4 localizes to spindle-pole bodies in a cell cycle-regulated manner to control Cdc7 localization and septation in fission yeast. *The Journal of biological chemistry* **275:** 14381-14387

Li F, Long T, Lu Y, Ouyang Q, Tang C (2004) The yeast cell-cycle network is robustly designed. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 4781-4786

Li S, Assmann SM, Albert R (2006) Predicting essential components of signal transduction networks: a dynamic model of guard cell abscisic acid signaling. *PLoS biology* **4:** e312

Liakopoulos D, Kusch J, Grava S, Vogel J, Barral Y (2003) Asymmetric loading of Kar9 onto spindle poles and microtubules ensures proper spindle alignment. *Cell* **112:** 561-574

Lodish H, Berk A, Zipursky SL, Baltimore D, Darnell J (2008) Molecular Cell Biology. 6th edition edn, p 973. W. H. Freeman

Marks J, Fankhauser C, Simanis V (1992) Genetic interactions in the control of septation in Schizosaccharomyces pombe. *Journal of cell science* **101 (Pt 4):** 801-808

Mehta S, Gould KL (2006) Identification of functional domains within the septation initiation network kinase, Cdc7. *The Journal of biological chemistry* **281:** 9935-9941

Mendenhall MD (1993) An inhibitor of p34CDC28 protein kinase activity from Saccharomyces cerevisiae. *Science* **259:** 216-219

Mendenhall MD, Hodge AE (1998) Regulation of Cdc28 cyclin-dependent protein kinase activity during the cell cycle of the yeast Saccharomyces cerevisiae. *Microbiology and molecular biology reviews : MMBR* **62:** 1191-1243

Mendenhall MD, Jones CA, Reed SI (1987) Dual regulation of the yeast CDC28-p40 protein kinase complex: cell cycle, pheromone, and nutrient limitation effects. *Cell* **50:** 927-935

Meyer P, Alexopoulos LG, Bonk T, Califano A, Cho CR, de la Fuente A, de Graaf D, Hartemink AJ, Hoeng J, Ivanov NV, Koeppl H, Linding R, Marbach D, Norel R, Peitsch MC, Rice JJ, Royyuru A,

Schacherer F, Sprengel J, Stolle K, Vitkup D, Stolovitzky G (2011) Verification of systems biology research in the age of collaborative competition. *Nature biotechnology* **29:** 811-815

Meyer P, Hoeng J, Rice JJ, Norel R, Sprengel J, Stolle K, Bonk T, Corthesy S, Royyuru A, Peitsch MC, Stolovitzky G (2012) Industrial methodology for process verification in research (IMPROVER): toward systems biology verification. *Bioinformatics* **28:** 1193-1201

Minet M, Nurse P, Thuriaux P, Mitchison JM (1979) Uncontrolled septation in a cell division cycle mutant of the fission yeast Schizosaccharomyces pombe. *Journal of bacteriology* **137:** 440-446

Mitchison JM, Nurse P (1985) Growth in cell length in the fission yeast Schizosaccharomyces pombe. *Journal of cell science* **75:** 357-376

Moreno S, Klar A, Nurse P (1991) Molecular genetic analysis of fission yeast Schizosaccharomyces pombe. *Methods in enzymology* **194:** 795-823

Morgan DO (2007) The Cell Cycle: Principles of Control. p 210. London: New Science Press

Morris MK, Saez-Rodriguez J, Sorger PK, Lauffenburger DA (2010) Logic-based models for the analysis of cell signaling networks. *Biochemistry* **49:** 3216-3224

Murone M, Simanis V (1996) The fission yeast dma1 gene is a component of the spindle assembly checkpoint, required to prevent septum formation and premature exit from mitosis if spindle function is compromised. *The EMBO journal* **15:** 6605-6616

Musacchio A (2011) Spindle assembly checkpoint: the third decade. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* **366:** 3595-3604

Naldi A, Thieffry D, Chaouiya C (2007) Decision Diagrams for the Representation and Analysis of Logical Models of Genetic Networks. In *Computational Methods in Systems Biology*, Calder M, Gilmore S (eds), pp 233-247. Springer Berlin Heidelberg

Nasmyth K (1993) Control of the yeast cell cycle by the Cdc28 protein kinase. *Current opinion in cell biology* **5:** 166-179

Nasmyth K (1996) At the heart of the budding yeast cell cycle. *Trends in genetics : TIG* **12:** 405-412

Noble D (2002) Modeling the heart--from genes to cells to the whole organ. *Science* **295:** 1678-1682

Ohkura H, Hagan IM, Glover DM (1995) The conserved Schizosaccharomyces pombe kinase plo1, required to form a bipolar spindle, the actin ring, and septum, can drive septum formation in G1 and G2 cells. *Genes & development* **9:** 1059-1073

Papin JA, Palsson BO (2004) Topological analysis of mass-balanced signaling networks: a framework to obtain network properties including crosstalk. *Journal of theoretical biology* **227:** 283-297

Plahte E. MT, Omholt S. W. (1995) Feedback loops, stability and miltistationarity in dynamical systems. *J Biol Systems* **3:** 409-413

Poussin C, Mathis C, Alexopoulos LG, Messinis DE, Dulize HJ, Belcastro V, Melas IN, Sakellaropoulos T, Rhissorrakrai K, Bilal E, Meyer P, Talikka M, Boue S, Norel R, Rice JJ,

Stolovitzky G, Ivanov NV, Peitsch MC, Hoeng J (2014) The species translation challenge - A systems biology perspective on human and rat bronchial epithelial cells. *Scientific Data* **1**

Poux S, Magrane M, Arighi CN, Bridge A, O'Donovan C, Laiho K (2014) Expert curation in UniProtKB: a case study on dealing with conflicting and erroneous data. *Database : the journal of biological databases and curation* **2014:** bau016

Rachfall N, Johnson AE, Mehta S, Chen JS, Gould KL (2014) Cdk1 Promotes Cytokinesis in Fission Yeast through Activation of the Septation Initiation Network. *Molecular biology of the cell*

Richardson H, Lew DJ, Henze M, Sugimoto K, Reed SI (1992) Cyclin-B homologs in Saccharomyces cerevisiae function in S phase and in G2. *Genes & development* **6:** 2021-2034

Rosenberg JA, Tomlin GC, McDonald WH, Snydsman BE, Muller EG, Yates JR, 3rd, Gould KL (2006) Ppc89 links multiple proteins, including the septation initiation network, to the core of the fission yeast spindle-pole body. *Molecular biology of the cell* **17:** 3793-3805

Saadatpour A, Wang RS, Liao A, Liu X, Loughran TP, Albert I, Albert R (2011) Dynamical and structural analysis of a T cell survival network identifies novel candidate therapeutic targets for large granular lymphocyte leukemia. *PLoS computational biology* **7:** e1002267

Saez-Rodriguez J, Simeoni L, Lindquist JA, Hemenway R, Bommhardt U, Arndt B, Haus UU, Weismantel R, Gilles ED, Klamt S, Schraven B (2007) A logical model provides insights into T cell receptor signaling. *PLoS computational biology* **3:** e163

Salimova E, Sohrmann M, Fournier N, Simanis V (2000) The S. pombe orthologue of the S. cerevisiae mob1 gene is essential and functions in signalling the onset of septum formation. *Journal of cell science* **113 (Pt 10):** 1695-1704

Samaga R, Klamt S (2013) Modeling approaches for qualitative and semi-quantitative analysis of cellular signaling networks. *Cell communication and signaling : CCS* **11:** 43

Samaga R, Saez-Rodriguez J, Alexopoulos LG, Sorger PK, Klamt S (2009) The logic of EGFR/ErbB signaling: theoretical properties and analysis of high-throughput data. *PLoS computational biology* **5:** e1000438

Sanchez L, van Helden J, Thieffry D (1997) Establishement of the dorso-ventral pattern during embryonic development of drosophila melanogasater: a logical analysis. *Journal of theoretical biology* **189:** 377-389

Schlatter R, Philippi N, Wangorsch G, Pick R, Sawodny O, Borner C, Timmer J, Ederer M, Dandekar T (2012) Integration of Boolean models exemplified on hepatocyte signal transduction. *Briefings in bioinformatics* **13:** 365-376

Schlatter R, Schmich K, Avalos Vizcarra I, Scheurich P, Sauter T, Borner C, Ederer M, Merfort I, Sawodny O (2009) ON/OFF and beyond--a boolean model of apoptosis. *PLoS computational biology* **5:** e1000595

Schmidt S, Sohrmann M, Hofmann K, Woollard A, Simanis V (1997) The Spg1p GTPase is an essential, dosage-dependent inducer of septum formation in Schizosaccharomyces pombe. *Genes & development* **11:** 1519-1534

Schwob E, Bohm T, Mendenhall MD, Nasmyth K (1994) The B-type cyclin kinase inhibitor p40SIC1 controls the G1 to S transition in S. cerevisiae. *Cell* **79:** 233-244

Schwob E, Nasmyth K (1993) CLB5 and CLB6, a new pair of B cyclins involved in DNA replication in Saccharomyces cerevisiae. *Genes & development* **7:** 1160-1175

Seshan A, Amon A (2004) Linked for life: temporal and spatial coordination of late mitotic events. *Current opinion in cell biology* **16:** 41-48

Sheaff RJ, Roberts JM (1996) End of the line: proteolytic degradation of cyclin-dependent kinase inhibitors. *Chemistry & biology* **3:** 869-873

Simanis V (2003) The mitotic exit and septation initiation networks. *Journal of cell science* **116:** 4261-4262

Singh NS, Shao N, McLean JR, Sevugan M, Ren L, Chew TG, Bimbo A, Sharma R, Tang X, Gould KL, Balasubramanian MK (2011) SIN-inhibitory phosphatase complex promotes Cdc11p dephosphorylation and propagates SIN asymmetry in fission yeast. *Current biology : CB* **21:** 1968-1978

Southern J, Pitt-Francis J, Whiteley J, Stokeley D, Kobashi H, Nobes R, Kadooka Y, Gavaghan D (2008) Multi-scale computational modelling in biology and physiology. *Progress in biophysics and molecular biology* **96:** 60-89

Sparks CA, Morphew M, McCollum D (1999) Sid2p, a spindle pole body kinase that regulates the onset of cytokinesis. *The Journal of cell biology* **146:** 777-790

Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, Brown PO, Botstein D, Futcher B (1998) Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. *Molecular biology of the cell* **9:** 3273-3297

Stern B, Nurse P (1996) A quantitative model for the cdc2 control of S phase and mitosis in fission yeast. *Trends in genetics : TIG* **12:** 345-350

Stoll G, Viara E, Barillot E, Calzone L (2012) Continuous time Boolean modeling for biological signaling: application of Gillespie algorithm. *BMC systems biology* **6:** 116

Szallasi Z, Stelling J, Periwal V. (2006) Systems Modeling in Cell Biology, From Concepts to Nuts and Bolts. Cambridge, MA: MIT Press.

Tanaka K, Petersen J, MacIver F, Mulvihill DP, Glover DM, Hagan IM (2001) The role of Plo1 kinase in mitotic commitment and septation in Schizosaccharomyces pombe. *The EMBO journal* **20:** 1259-1270

Tomlin GC, Morrell JL, Gould KL (2002) The spindle pole body protein Cdc11p links Sid4p to the fission yeast septation initiation network. *Molecular biology of the cell* **13:** 1203-1214

Toyn JH, Johnson AL, Donovan JD, Toone WM, Johnston LH (1997) The Swi5 transcription factor of Saccharomyces cerevisiae has a role in exit from mitosis through induction of the cdk-inhibitor Sic1 in telophase. *Genetics* **145:** 85-96

Tyers M, Jorgensen P (2000) Proteolysis and the cell cycle: with this RING I do thee destroy. *Current opinion in genetics & development* **10:** 54-64

Tyson JJ, Csikasz-Nagy A, Novak B (2002) The dynamics of cell cycle regulation. *BioEssays : news and reviews in molecular, cellular and developmental biology* **24:** 1095-1109

Veliz-Cuba A, Stigler B (2011) Boolean models can explain bistability in the lac operon. *Journal of computational biology : a journal of computational molecular cell biology* **18:** 783-794

Wachowicz P, Chasapi A, Krapp A, Cano Del Rosario E, Schmitter D, Sage D, Unser M, Xenarios I, Rougemont J, Simanis V (2015) Analysis of S. pombe SIN protein SPB-association reveals two genetically separable states of the SIN. *Journal of cell science*

Waddington CH (1957) A Discussion of Some Aspects of Theoretical Biology. In *The Strategy of the Genes*.

Walker DC, Southgate J (2009) The virtual cell--a candidate co-ordinator for 'middle-out' modelling of biological systems. *Briefings in bioinformatics* **10:** 450-461

Wang RS, Saadatpour A, Albert R (2012) Boolean modeling in systems biology: an overview of methodology and applications. *Physical biology* **9:** 055001

Weinreich M, Liang C, Chen HH, Stillman B (2001) Binding of cyclin-dependent kinases to ORC and Cdc6p regulates the chromosome replication cycle. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 11211-11217

Weiss EL (2012) Mitotic exit and separation of mother and daughter cells. *Genetics* **192:** 1165-1202

Wittenberg C, Reed SI (2005) Cell cycle-dependent transcription in yeast: promoters, transcription factors, and transcriptomes. *Oncogene* **24:** 2746-2755

Woods A, Sherwin T, Sasse R, MacRae TH, Baines AJ, Gull K (1989) Definition of individual components within the cytoskeleton of Trypanosoma brucei by a library of monoclonal antibodies. *Journal of cell science* **93 (Pt 3):** 491-500

Yamano H, Gannon J, Hunt T (1996) The role of proteolysis in cell cycle progression in Schizosaccharomyces pombe. *EMBO J* **15:** 5268-5279