

# Deliverable 3.1

## CALAMAR (ANR-08-SYSC-003)

L. Calzone, C. Chaouiya, E. Remy, S. Soliman

February 9, 2011

### 1 Statement - Motivation

In this study, we focus on the modelling of the RB/E2F pathway, with the perspective of answering some biological questions which revolve around RB and E2F activities and their involvement in cancer.

### 2 Introduction

Among the eight known E2F transcription factors, we concentrate on three of them: E2F1, E2F2 and E2F3 are transcription factors known to be activators of their targets. They form an inactive complex with protein RB, and they are periodically expressed during the cell cycle, with a peak during transition G1/S. Although similar, they play different roles at different moments of the cell cycle. For simplicity, we choose to focus our study on E2F1 and E2F3 and not mention E2F2 yet since its role and implication in bladder tumour is not well documented. At this point, we do not make any difference between the two isoforms of E2F3, E2F3a and E2F3b. E2F3a is supposed to be an activator whereas E2F3b inhibits transcription. However, when one of them is deleted, the other isoform is capable of replacing it to a certain extent. Further investigation will be required to assess their differences.

Since the biological questions mainly revolve around RB and E2F activities, we describe how the transcription factors E2F1 and E2F3 control not only RB but also p53 pathway. Indeed, it is well known that RB and p53 are mutated in many tumours. Moreover, the two pathways are connected in various ways (for example, in RB-mutated cells, apoptosis is increased).

Using bioinformatics and systems biology approaches, we investigate the role of E2F1 and E2F3 in tumour cells, and their differences in the control of the two pathways. We built a descriptive reaction network showing their interactions with RB and p53 pathways. We then derived an influence network that served as a basis for our discrete mathematical model.

Our choice of a discrete formalism to model the network is motivated both by the lack of quantitative information and by the qualitative aspect of the biological questions that stimulate this work. We propose a logical model, which encompasses 20 components, with 4 inputs (DNADamage, TGFb, EGFR and FGFR3), and 3 outputs (Apoptosis, Proliferation and growth\_arrest). We expect that such a model will throw light on some understandable/surprising observations, that it will suggest differences between E2F1 and E2F3, and will propose ways to reestablish a lost phenotype such as apoptosis.

Some of the surprising experimental observations are,

- E2F3 amplifications are observed in bladder tumors that have RB deleted;
- loss of RB is accompanied by over-expression of p16;
- E2F3 is considered as an oncogene but not E2F1;
- CDKN2A homozygous deletions seem to frequently occur in RB potent cells. Is it always true?

We first describe the logical model and proceed with some analyses, especially related to its steady states and the corresponding phenotypes. We then describe methodological results that ease the maintenance of the two models, reaction-based and regulation-based, side by side. A conclusion summarizes the current state of this work and future perspectives.

### 3 Detailed description of the model

#### 3.1 Construction of the two maps

We identified all the major players involved in both pathways that are controlled by the same transcription factors in order to speculate on ways to re-established these lost links. Based on a thorough literature search, we built a reaction network recapitulating all that is known about the E2F activating transcription factors in response to cell receptor activation (growth factor and FGFR3), TGFb, and DNA damage (Figure 1A). The construction of the map and the choice of the different components to include in the map have been directed by the initial biological questions. As a result, the map should not be seen as the most comprehensive network linking RB/E2F and p53 pathways, but rather, it should be considered as a focused version of how the E2F transcription factors control the activity of both RB and p53 in bladder cells. The reaction network was built and annotated using CellDesigner software.

However, the lack of quantitative information did not justify the level of details we included for modelling purposes. Therefore, we simplified the

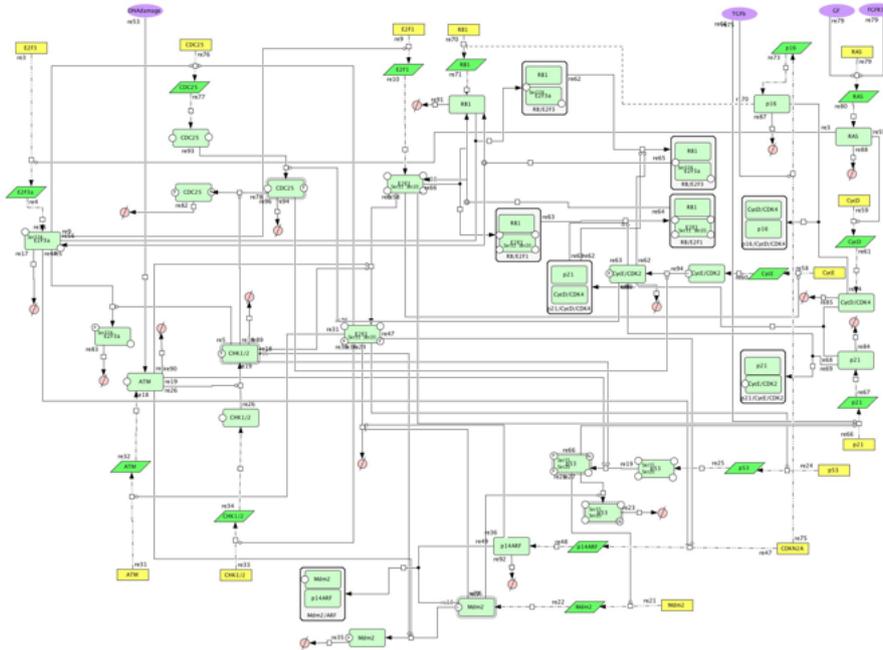


Figure 1: Reaction network of the RB/E2F map

reaction and chose to represent the same information in a more abstract way. We derived from the reaction network an influence network in which details of the states of phosphorylation, acetylation or ubiquitination are lost. Rather, the gene or the protein is either active or inactive, present or absent. Nodes (species) of the influence map are connected by positive or negative arrows (influences) depending on the type of reactions that linked them in the corresponding reaction network. For instance, RB sequesters E2F1 by forming an inactive complex in the reaction map. In the influence map, this reaction is interpreted as RB inhibiting E2F1. The influence network is this adding visual information about the activity of species that are lacking in the reaction network even though it remains less precise in terms of the nature of the interactions.

### 3.2 Discrete modelling

Next, we translated the influence diagram into a mathematical model in order to verify the coherence of the literature-based diagram.

The model contains 14 regulatory (internal) nodes, 4 inputs (DNA damage, TGFb, EGFR and FGFR3), 3 outputs (Apoptosis, Proliferation and growth\_arrest). More specifically, the outputs correspond to possible phenotypes in response to cell signals. They correspond to:

- Proliferation: CycE/CDK2 activation is the readout for cell prolifer-

eration. When CycE/CDK complex is active, DNA replication can be initiated. The "proliferation" phenotype corresponds in reality to entry into S phase. The cell cycle could be blocked downstream.

- Apoptosis: p53-E2F1 dependent activation is the readout for apoptosis. Activation of p53 is the signal for possible apoptosis but there are events downstream that can still counteract the activation of the apoptotic signal. The model focuses on the E2F1-dependent apoptosis. E2F1 is also capable of activating p73 and launch another apoptotic program but the conditions in which this happened are not yet clear and thus, not included in the model.
- Growth arrest: p21 and RB are breaks in cell transitions. They insure that the cycle remains arrested until an external signal relieves the break. It is not rare to see apoptosis and growth arrest coexist. p21 synthesis is mediated by p53 that tries to arrest the cycle in order to repair the damage. If the damage is not repairable, apoptosis is initiated.

In the influence diagram, species is activated according to the states of its inputs and to the associated specific rules that are defined base on the known behaviours. In other words, we assign to each node (species) a logical rule that sets the conditions that will vary the value of the node, the two possible values being 0 when inactive or absent, 1 when active or present. Each node is linked to its input nodes by AND and OR gates according to what is known about the activating conditions.

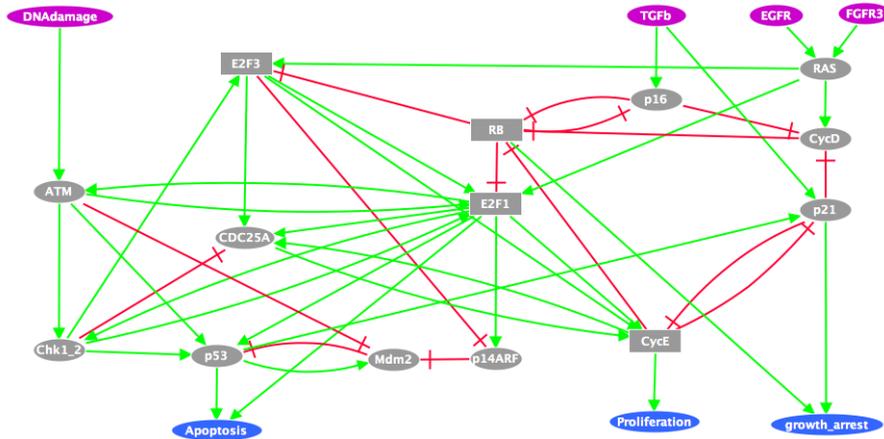


Figure 2: Influence network of the RB/E2F map (GINsim screenshot)

The content of the following tables is extracted from the GINsim annotated model, which is available in the form of a GINML archive E2F\_model.zginml.

ID	Val	Logical Term	Comment
RAS	1	$EGFR \vee FGFR3$	RAS represents the MAPK pathway activated by growth factors. RAS is often mutated in cancers. Induction of mitogens seem to activate E2F1. This would explain the role of hyperproliferative signals that lead to apoptosis through E2F1. See [Morgan 2007]
E2F1	1	$!RB:2 \wedge !(Chk1.2 \wedge ATM) \wedge (RAS \vee E2F3)$	E2F1 has two levels: 1 for cell cycle functions in G1/S, 2 for apoptotic functions. The sole presence of RB leads to the level 0. To reach level 1, E2F1 needs to be activated by RAS or E2F3, whatever the levels of ATM and Chk1.2. E2F1 reaches level 2 when E2F3 is at level 2 and in the presence of both ATM and Chk1.2. [Polager 2009], [Powers 2004], [Rogoff 2004], [Hurst, 2008], [Polager 2009], [Goberdhan 2000], [Hong 2007]
	2	$!RB:2 \wedge ATM \wedge Chk1.2 \wedge (RAS \vee E2F3)$	
E2F3	1	$!Chk1.2 \wedge RAS \wedge !RB$	E2F3 has two levels: 1 for cell cycle functions in G1/S, 2 for apoptotic functions. No distinction is made on E2F3a and E2F3b. Level 1 corresponds to normal E2F3 activation of cell cycle events. Level 2 corresponds to E2F3 apoptotic effect, which is reached in the absence of RB and presence of Chk1.2. Basal value is 1 and it decreases to 0 in the presence of RB. E2F3 transcription is induced by Myc through RAS.
	2	$Chk1.2 \wedge !RB:2$	
RB	0	$CycE \vee p16$	Rb-/- leads to apoptosis (increase of E2F3a, no change to E2F3b) Rb-/-, E2F3a-/- recovers Rb-/-, E2F3b-/- leads also to apoptosis We added the interaction to growth_arrest to ensure that in case of RB maintenance, we get growth arrest (hence outcomes are not only exclusive but one of them needs to be activated - equal to 1) [Chong 2009]
	1	$!CycE \wedge CycD \wedge !p16$	
	2	$!CycD \wedge !CycE \wedge !p16$	
p21	1	$!CycE \wedge (TGFb \vee p53)$	p21 promotes the translocation and activation of CycD/CDK complexes p21 is high in senescent cells p21 is phosphorylated by CycE/CDK2 at Ser130. But p21 can be degraded via Skp2 independently of this phosphorylation (but it's slower). [Bornstein 2003]

continued...

ID	Val	Logical Term	Comment
p16	1	$TGF\beta \wedge !RB$	p16 stands for CDKN2A or p16INK4a. In this model, it also represents p15 (CDKN2B). p16 is activated in response of oncogenic stresses. In a normal cell cycle, it is not expressed at high levels. p16 is high in senescent cells. p16 inhibits RB, presence of p16INK4 protein inversely correlated with detectable RB or cyclin D1 proteins and is not correlated with p53 mutations [Okamoto 1994] Transcriptional repression of the D-type cyclin-dependent kinase inhibitor p16 by the retinoblastoma susceptibility gene product pRb [Li 1994] In gallbladder cancer: "Inactivation of RB may stimulate cells to increase p16INK4 expression, which inhibits the activity of CDK4 and inactivates CycD1/CDK4 complex. On the contrary, pRb overexpression stimulates cells to increase p16INK4 protein loss, which enhances the activity of CDK4, thus inactivating excrescent RB." [Ma 2005]
CycD	1	$RAS \wedge !p16 \wedge !p21$	CycD stands for the CDK complexes CycD1/CDK4 or CycD1/CDK6. p16 binds to the kinase subunits of the complex CDK4 and CDK6. This binding inhibits the binding and activation of the CDK complex. CycD1 is high in senescent cells but inactive because in complex with p21 (discussions with Vjeko Dulic).
CycE	1	$!p21 \wedge CDC25A \wedge (E2F1:1 \vee (!E2F1:1 \wedge E2F3:1))$	E2F3 plays its role on CycE when E2F1 is absent (eg deleted)
Mdm2	1	$p53 \wedge !p14ARF \wedge !ATM$	p53 and Mdm2 are involved in a negative feedback loop: p53 synthesizes Mdm2 and Mdm2 ubiquitinates p53 (see link to Atlas Genetics Oncology)
p14ARF	1	$!E2F3 \wedge E2F1$	p14ARF and p16INK4a share the same gene CDKN2A. A link is not included: $ARF \rightarrow E2F1$ by degradation of E2F1 through SKP2. This inhibition will be added later. E2F1 is phosphorylated at S31 by ATM and at S364 by Chk2. These phosphorylations interfere with the ARF/SKP2- and MDM2-dependent degradation of E2F1, thus stabilizing the latter by decreasing its turnover rate (see link to Atlas Genetics Oncology)
p53	1	$ATM \wedge Chk1\_2 \wedge E2F1 \wedge !Mdm2$	Two links through miRNA exist but haven't been included in the model: $p53 \rightarrow E2F1$ : p53 inhibits the transcription of miR17 which inhibits E2F1 transcription $p53 \rightarrow !E2F3$ : p53 inhibits the transcription of miR34 which inhibits E2F3 transcription ATM and ATR phosphorylate p53 at ser15 and stabilize it [Powers 2004] Note: p53 is present in the absence of Mdm2 and presence of both Chk1_2 and ATM (response to DNA damage) or when E2F2 is at its highest apoptotic level 3 (which implies the presence of ATM and Chk1_2)
Chk1_2	1	$ATM:1 \wedge E2F1$	Expression of E2F1 leads to an increase in Chk2 mRNA and accumulation of Chk2 protein [Rogoff 2004]

continued...

ID	Val	Logical Term	Comment
CDC25A	1	$(E2F3 \vee E2F1 \vee CycE) \wedge !Chk1\_2$	<p>We assume that E2F1 or E2F3 are required for transcription and CycE is required for activation (phosphorylation)</p> <p>CDC25A is a member of the CDC25 family of phosphatases. CDC25A is required for progression from G1 to the S phase of the cell cycle. It activates the cyclin-dependent kinase CDC2 by removing two phosphate groups. CDC25A is specifically degraded in response to DNA damage, which prevents cells with chromosomal abnormalities from progressing through cell division. CDC25A is an oncogene. See link to Genecards</p> <p>CDC25 activates CDK complexes that control progression through the cell cycle. CDC25 is also involved in the DNA damage checkpoints. Regulation of CDC25 occurs through a variety of mechanisms including phosphorylation, interaction with 14.3.3 proteins, subcellular localization and protein degradation.</p> <p>Chk1 promotes degradation of CDC25A.</p> <p>CDC25C interacts more with CDC2 (CycB) whereas CDC25A and B interact with CDK2 (CycE)</p> <p>E2F1 is active during the G1 to S transition, and thus its target genes, which include regulatory elements of the cell cycle, such as CDC2, CDC25A and cyclin E, and essential components of DNA replication machinery are expressed in a cell cycle dependent manner (i.e. only in late G1 and early S phase of the cell cycle). See link to Atlas Genetics Oncology</p> <p>CDC25A is overexpressed in a variety of human cancers including breast, hepatocellular, ovarian, colorectal, esophageal, head and neck cancer and also in non-Hodgkin lymphoma [Dasgupta 2004]</p>
Apoptosis	1	$E2F1:2 \wedge p53$	
Proliferation	1	$CycE:1$	
growth_arrest	1	$p21 \vee RB$	

## 4 Validation, analysis and predictions

With the discrete model, we can propose predictions of (a minimal set of) alterations that could cause invasiveness or disappearance of needed phenotypes, such as apoptosis or that would enhance proliferation. The three outputs of the model are: Apoptosis, Proliferation and growth\_arrest:

- The phenotype "Proliferation" corresponds to the output values

$$(Apoptosis, Proliferation, growth\_arrest) = (0, 1, 0);$$

- The phenotype "Apoptosis" corresponds to the output values

$$(Apoptosis, Proliferation, growth\_arrest) = (1, 0, 1);$$

- The phenotype "Growth Arrest" corresponds to the output values

$$(Apoptosis, Proliferation, growth\_arrest) = (0, 0, 1);$$

In the wild type, the model generates 25 stable states among which the three phenotypes are represented (see Figure 3). With the following tables (Tables 1, 2 and 3), we can easily visualize the existence or the loss of these phenotypes in different mutants. To simulate a mutant, we fix the expression value of a component to its minimal or maximal value and thus delete or overexpress this component (in some cases such as for the discrete variable, we may fix it to an intermediate value). We fix a unique component (simple mutant), or two components simultaneously (double mutants). As a first exploration of the behaviour of the model, we can use this type of simulation validating the model by comparing the obtained phenotypes with known results. Then, by testing exhaustively combinations of mutants, the model could propose explanations to the disappearance of needed phenotypes, and suggest ways to reestablish them.

Phenotype	DNAdamage	TGFb	EGFR	FGFR3	RAS	E2F1	E2F3	RB	p21	p16	CycD1	CycE	Mdm2	p14ARF	p53	ATM	Chk1_2	CDC25A
proliferation	0	0	0	1	1	1	1	0	0	0	1	1	0	0	0	0	0	1
proliferation	0	0	1	*	1	1	1	0	0	0	1	1	0	0	0	0	0	1
proliferation	0	1	0	1	1	1	1	0	0	1	0	1	0	0	0	0	0	1
proliferation	0	1	1	*	1	1	1	0	0	1	0	1	0	0	0	0	0	1
growth_arrest	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
growth_arrest	0	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
growth_arrest	0	1	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0
growth_arrest	0	1	0	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0
growth_arrest	0	1	1	*	1	0	0	2	1	0	0	0	0	0	0	0	0	0
growth_arrest	0	1	1	*	1	1	1	0	1	1	0	0	0	0	0	0	0	1
growth_arrest	1	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
growth_arrest	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
growth_arrest	1	1	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0
growth_arrest	1	1	0	1	1	0	0	2	1	0	0	0	0	0	0	0	0	0
growth_arrest	1	1	1	*	1	0	0	2	1	0	0	0	0	0	0	0	0	0
apop+arrest	1	1	0	0	0	2	2	0	1	1	0	0	0	0	0	1	1	0
apop+arrest	1	1	0	1	1	2	2	0	1	1	0	0	0	0	1	1	1	0
apop+arrest	1	1	1	*	1	2	2	0	1	1	0	0	0	0	1	1	1	0

Figure 3: The 25 stable states simulated in the wild type case. The star \* indicates that the corresponding variable takes all its possible values (1 and 0 for a Boolean node)

By fixing the value of an input, we focus on a part of the state space. In Table 4, we examine the presence of stable states in each part of the space defined by the inputs, and their phenotype. This type of studies allows to give information on conditions of presence or absence of phenotypes, and is a way to treat the question of reachability of the stable states (more exactly, it gives information on non-reachability!).

## 5 Relating the regulation and reaction models

### 5.1 Not one but two models

The two previous sections mostly describe the qualitative model that was built to answer the questions delineated in introduction, however as explained in Section 3.1 in parallel with its development, the modellers maintained by hand a reaction model, using the CellDesigner software [6].

Indeed, the reaction model allows to describe precisely the mechanistic interactions when they are known and builds on the biologist-friendly graphical interface of CellDesigner. Moreover, it might permit, in a further step, to move towards a dynamical ODE-based model.

The different nature of these two models might allow to complement the above analyses with other ones based on the reactions' structure, but in order to make this step a formal relation of the two models is necessary. Moreover it might permit the discovery of differences, either due to the hand-made translation, to different modelling choices, or even to fundamental discrepancies between the two formalisms.

### 5.2 Formally relating the models

Building on deliverable 1.3 [9], we will now describe the extent of our study on the formal linkage of the two models.

#### 5.2.1 From reactions to an influence graph

Starting from the CellDesigner model `E2F_zoom_270111.xml` one can automatically extract a regulation graph corresponding to its Jacobian, even though the precise kinetics is unknown. This is a direct consequence of [3] and is implemented in the BIOCHAM software [1, 2].

However, compare figure 4 - corresponding to the current regulatory network - and Figure 5 resulting from the Jacobian of the reaction model.

It is obvious that the graphs have a completely different structure. Actually this is not unexpected since the use in the regulatory network community is to forget about linear reactions, to compact active and inactive species into a single compound with level 0 or 1 (even more if the model is multi-level), etc. They simply represent two complementary views on the same system, and we will now try to formally relate them.

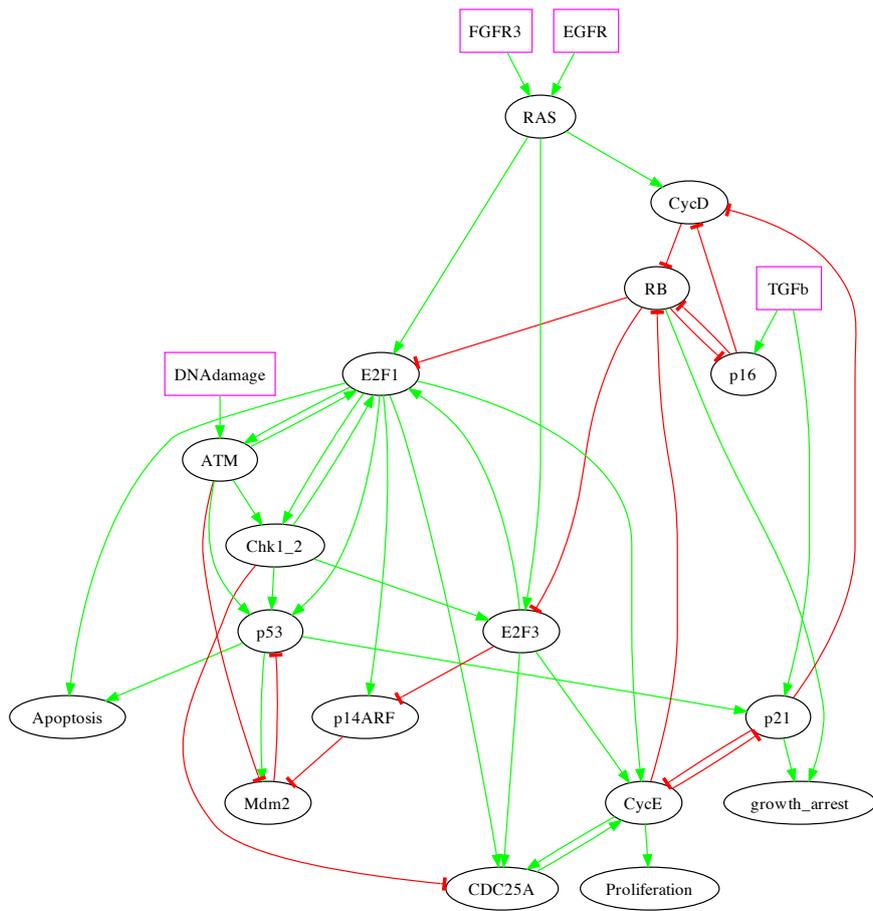


Figure 4: Automatic layout of the current GRN extracted from the zginml file

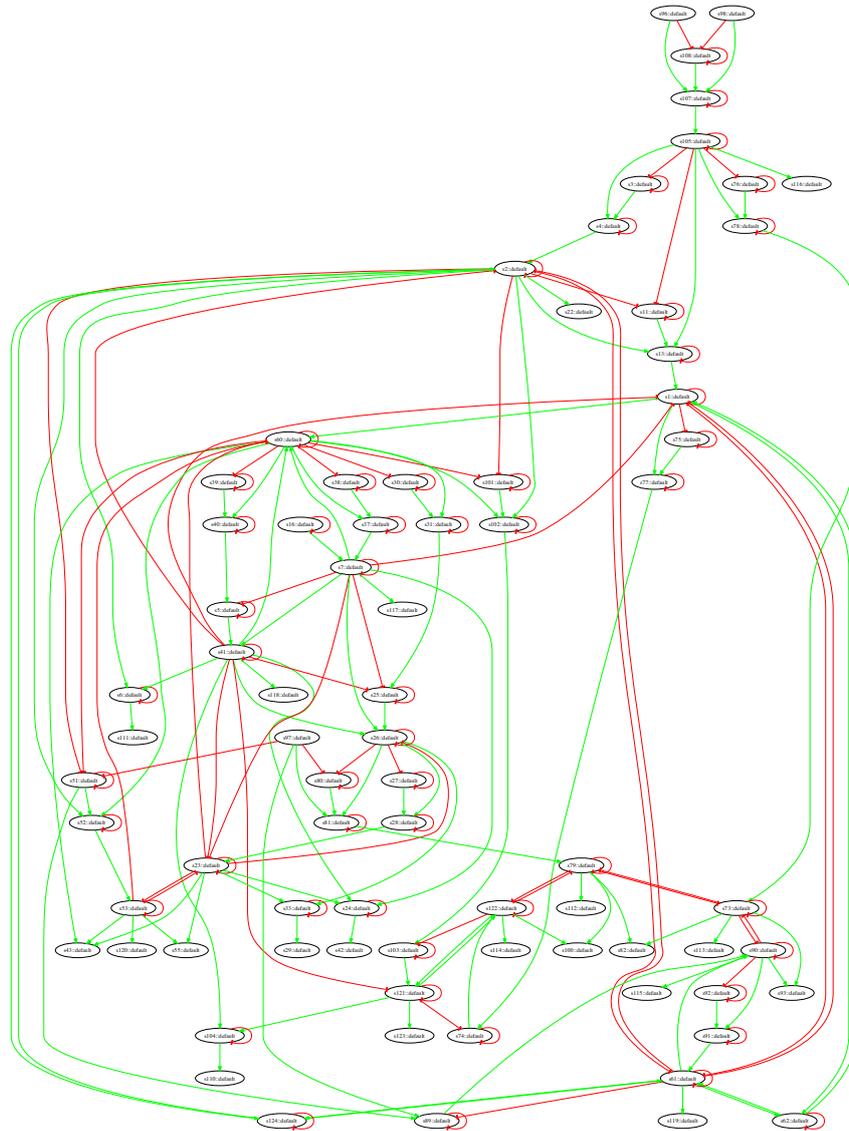


Figure 5: Automatic layout of the influences extracted by BIOCHAM from the CellDesigner file

### 5.2.2 Model reduction

Building on the reaction model transformation methods described in [7] we now propose to abstract the reaction model such that its automatically associated influence graph is as close as possible to the one obtained by modellers.

- the first step is not strictly necessary, but in order to obtain easy to read models, the SBML file obtained from CellDesigner is **preprocessed** such that the IDs obtained by BIOCHAM will correspond to the names in CellDesigner annotations instead of numerical IDs.
- BIOCHAM then **deletes all linear degradations**, i.e., reactions with a single reactant, no modifier and no product.
- it would be possible to remove all compounds that only act as inputs, or only as outputs, without changing the positive and negative feedback loops of the resulting influence graph. However to abide to the choices made in this model, the strategy applied is to **delete all inputs which are marked as genes, and all outputs** (the phenotypes, like “apoptosis” do not appear in the reaction model).

All the previous steps apply quite easily but keep the modelling discrepancy between the GRN (one form with different activation levels) and the reaction model (several active or inactive forms). In the current example, all the RNAs for instance, appear in the SBML model and not in the GINML model.

A final step allows to compress species without changing the functional feedback loops<sup>1</sup> of the generated influence graph:

When some species appear only as the single output of a single rule, and as the single input of a single rule, then it is a candidate for deletion.

If the inputs of its production rule and the outputs of its degradation are distinct, one can safely merge these two reactions (see Figure 6) and delete the intermediary species. Otherwise, if its production inputs are exactly the same as its degradation outputs, we have a reversible complex and it can be deleted with its degradation (see Figure 7).

BIOCHAM can now generate the influence graph as defined in [4]. For display purposes and since negative loops required for oscillations are of length at least two [8] the negative auto-regulations are then removed, which results in Figure 8.

Formal justifications of the applied model reductions and abstractions will be detailed in the deliverables about model reduction.

---

<sup>1</sup>the feedback loops might change slightly but not the precise dynamical features like oscillations or multi-stability.

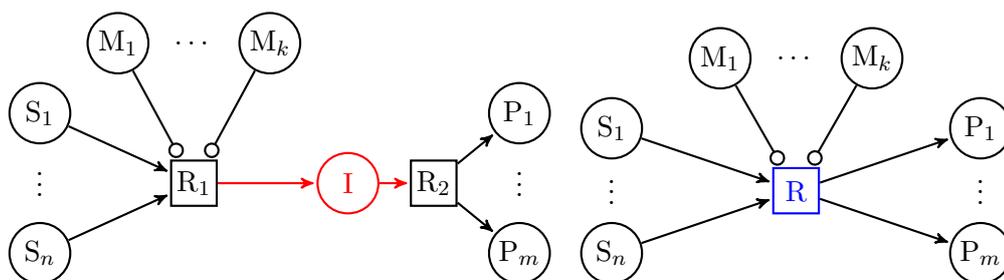


Figure 6: Example of intermediate species deletion and merge

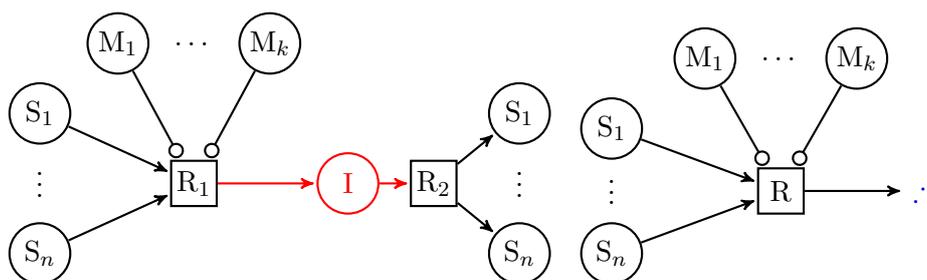


Figure 7: Example of intermediate species corresponding to a reversible complex

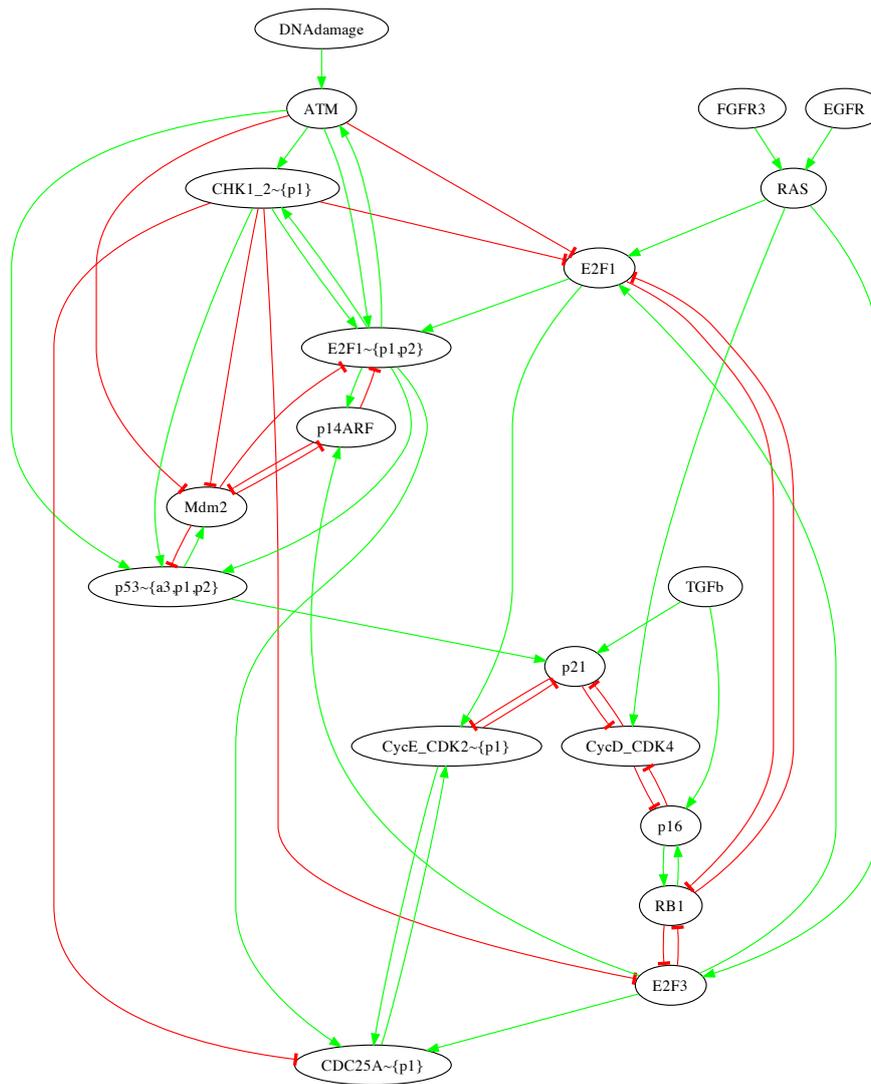


Figure 8: automatic layout of the influence graph obtained from the above procedure

### 5.3 Results

It is now easy to spot the differences. Let us review them:

- As expected the “phenotypes” do not appear in the influence graph, since they are not in the reaction graph.

They could be easily added though.

- Two regulations appear positive in the influence graph but were negative in the modellers’ GRN: E2F3 over p14ARF (re47) and p16 over RB1 (re70).

This is due to the fact that CellDesigner annotates the role of some modifiers as “inhibitor” but this information is lost in the *structure* of the SBML reaction model. It would be possible parse those annotations and to artificially add degradation reactions to the model in order to keep the correct sign of the influence, but one would then lose the mechanistic interpretation of the reaction model.

- The influence graph retains two species for E2F1, namely E2F1 and its doubly phosphorylated form. Indeed the first one is responsible for activating CycE and inhibiting RB, whereas the second form activates ATM, CDC25A, CHK1/2, p14RF and p53.

The method we described implicitly supposes a purely boolean regulation graph, but here the reactions show two different active forms, which would rather fit a multilevel GRN. Note that this corresponds to the choice that was eventually made in the logical model, and that otherwise it remains very easy to merge by hand the two active forms obtained.

- Due to their reversible complexation, the influence graph displays mutual inhibition from RB to E2F1 and E2F3, whereas the GRN shows only inhibitions from RB to the E2Fs but not the reverse.

This difference points out a modelling choice, where the reverse inhibition was not considered relevant to the dynamical behavior of the system and was thus eliminated to simplify the model.

- The exact same thing happens for Mdm2 and p14ARF: the inferred graph shows a double inhibition, whereas the modellers deemed to keep only the action of p14ARF on Mdm2 and not the reverse.
- Finally, the reaction model includes a negative action from p14ARF to E2F1 (re36) corresponding to recent findings by Zhang et al. [5]. This negative influence does not appear in the modeller’s GRN, once again by choice, since it creates a negative feedback loop that would necessitate a multilevel model to be properly incorporated.

Since the procedure is fully automatic (it can be used from the web at the following URL: <http://contraintes.inria.fr/~soliman/cd2dot.html>) it will ease the maintenance of the models, allow consistency checks in both directions, permit the use of reaction-based analyses on the model and later on help build an ODE model along the proposed GRN. It also already points out that it might be useful to consider a multilevel model of the system under study.

Note however that since the method described is purely structural and supposes a boolean (as opposed to multi-level) regulatory graph, it will not take into account features dependant on a sophisticated logical rule defining some dynamics for the logical model. Relating both dynamics is thus a further step that we might need to look into in the future

Moreover, the arbitrary choice to keep some inputs (and no outputs) was tailored to our current test-case and could be refined by a user-provided list of inputs/outputs that are meaningful to the modeller.

## 6 Conclusions and prospects

The model is still being improved. Discussions with specialists along with new publications allow its refinement. A model, by definition, is always evolving.

The next step will be to compare the predictions of the model to experimental data. We have experimental unpublished data of CGH, transcriptomics and mutations of FGFR3, RAS and p53 from 178 samples that are compared to 5 normal samples. We wish to compare the obtained results from single and double mutations to assess, among other things, the possibility of co-occurrence of alterations in invasive tissues. For instance, we could verify if two alterations occur systematically together or if the mutants (single or double) that have lost the apoptosis phenotype in the model are invasive in the experimental data.

This work is in progress.

## References<sup>2</sup>

- [1] Laurence Calzone, François Fages, and Sylvain Soliman. BIOCHAM: An environment for modeling biological systems and formalizing experimental knowledge. *Bioinformatics*, 22(14):1805–1807, 2006.
- [2] François Fages, Dragana Jovanovska, Aurélien Rizk, and Sylvain Soliman. *BIOCHAM v3.2 Reference Manual*. INRIA, 2010. <http://contraintes.inria.fr/BIOCHAM>.

---

<sup>2</sup>references in Section3.2 are given in the table documenting the model (pages 3.2 to 3.2)

- [3] François Fages and Sylvain Soliman. Abstract interpretation and types for systems biology. *Theoretical Computer Science*, 403(1):52–70, 2008.
- [4] François Fages and Sylvain Soliman. From reaction models to influence graphs and back: a theorem. In *Proceedings of Formal Methods in Systems Biology FMSB'08*, number 5054 in Lecture Notes in Computer Science. Springer-Verlag, February 2008.
- [5] Akira Funahashi, Yukiko Matsuoka, Akiya Jouraku, Mineo Morohashi, Norihiro Kikuchi, and Hiroaki Kitano. Celldesigner 3.5: A versatile modeling tool for biochemical networks. *Proceedings of the IEEE*, 96(8):1254–1265, August 2008.
- [6] Steven Gay, Sylvain Soliman, and François Fages. A graphical method for reducing and relating models in systems biology. *Bioinformatics*, 26(18):i575–i581, 2010. special issue ECCB'10.
- [7] El Houssine Snoussi. Necessary conditions for multistationarity and stable periodicity. *J Biol Syst* 6:3?9,1998.
- [8] Sylvain Soliman, Claudine Chaouiya, Grégory Batt, François Fages, Élisabeth Remy, Franck Pommereau, and Laurence Calzone. Modelling molecular networks: relationships between different formalisms and levels of details. Research Report RR-7221, INRIA, 2010.
- [9] Zhang HJ, Li WJ, Gu YY, Li SY, An GS, Ni JH, Jia HT. p14ARF interacts with E2F factors to form p14ARF-E2F/partner-DNA complexes repressing E2F-dependent transcription. *J Cell Biochem*. 109(4):693-701, 2010.

mutant	# ss	Loss of Apopt.	Loss of Prolif.	Loss of Gr. Arrest
Wild	(25)			
RB- RB==1 RB+		•	•	
RAS- RAS+			•	
Mdm2- Mdm2+		•		
ATM- ATM+		•	•	
Chk1/2- Chk1/2+		•	•	
p53- p53+		•		
Cdc25- Cdc25+			•	
p14ARF- p14ARF+				
p21- p21+		? A	•	
CycD- CycD+				
p16- p16+		•		
CycE- CycE+		•	•	•
E2F1- E2F1+		•		
E2F3- E2F3==1 E2F3+	(18)			

Table 1: List of single mutations. A green (*resp.* red or black) dot indicates that the considered mutant has no stable state with the Apoptosis (*resp.* Proliferation or Growth arrest) phenotype.

mutant1	mutant2	# ss	Loss of Apopt.	Loss of Prolif.	Loss of Gr. Arrest
E2F3-	p14ARF-				
	RB+		●	●	
	RB-				
	p53-		●		
	p53+				
	p21-		? A		
	p21+			●	
	E2F1-		●	●	
	E2F1+				
E2F3+	RB-				
	RB==1				
	RB+		●	●	
	RAS-				
	RAS+				
	Mdm2-				
	Mdm2+		●		
	ATM-		●		
	ATM+				
	Chk1/2-		●		
	Chk1/2+			●	
	p53-		●		
	p53+				
	Cdc25-			●	
	Cdc25+				
	p14ARF-				
	p14ARF+				
	p21-		?A		
	p21+			●	
	CycD-				
	CycD+				
	p16-		●		
	p16+				
	CycE-			●	
	CycE+		●		●
	FGFR3-				
	FGFR3+				
	E2F1-		●		
	E2F1==1		●		
	E2F1+				

Table 2: List of double mutations. The legend is the same as the single mutation table

mutant1	mutant2	# ss	Loss of Apopt.	Loss of Prolif.	Loss of Gr. Arrest	Presence of Anormal state
E2F1+	RB-					
	RB+			•		P+G
	RAS-					
	RAS+					
	Mdm2-					
	Mdm2+		•			T
	ATM-		•			
	ATM+			•	•	
	Chk1/2-		•			
	Chk1/2+			•		T
	p53-		•			T
	p53+			•	•	A+P
	Cdc25-			•		T
	Cdc25+					A+P
	p14ARF-					
	p14ARF+					
	p21-				•	A
	p21+			•		
	CycD-					
	CycD+					
	p16-					
	p16+					
	CycE-			•		T
	CycE+		•		•	A+P
	FGFR3-					
	FGFR3+					
E2F1-	RB-		•			T
	p53-	(11)	•			
	p53+	(13)	•			
Chk1/2+	p14ARF-	(16)		•		T
p53+	p21-	(17)	? A			T

Table 3: List of double mutants involving E2F1

Input	Wild			RB del		
	# ss	# Prolif.	# Apopt.	# ss	# Prolif.	# Apopt.
DNA dam=0	15	6	0	11	6	0
DNA dam=1	10	0	4	10	0	8
TGFb=0	5	3	0	9	3	4
TGFb=1	20	3	4	12	3	4
EGFR=0	13	2	2	11	2	4
EGFR=1	12	4	2	10	4	4
FGFR3=0	13	2	2	11	2	4
FGFR3=1	12	4	2	10	4	4
FGFR3=*	6	2	1	5	2	2

Table 4: Number of stable states and their distribution through the 3 phenotypes. Each line of the tabular corresponds to a constraint on an input. For "FGFR3 = \*", are counted all the stable states which do not depend on the value of input FGFR3