# Cell-based multi-scale modeling for systems and synthetic biology : from stochastic gene expression in single cells to spatially organized cell populations

François Bertaux

École Doctorale: Science Mathématiques de Paris Centre

THÈSE DE DOCTORAT

présentée par

**François Bertaux**

---

# *Cell-based multi-scale modeling for systems and synthetic biology: from stochastic gene expression in single cells to spatially organized cell populations*

---

Soutenue le 15 juin 2016 devant le jury composé de :

| | | |
|---|---|---|
| M^me^ Alessandra Carbone | Université Pierre et Marie Curie | président |
| M. Mark Chaplain | University of St-Andrews | rapporteur |
| M. Paul Macklin | University of Southern California | rapporteur |
| M. Emmanuel Barillot | Institut Curie | examinateur |
| M. Jérémie Roux | Université de Nice | examinateur |
| M. Dirk Drasdo | Inria Paris | directeur |
| M. Gregory Batt | Inria Paris | directeur |

# Abstract

Cell-intrinsic, non-environmental sources of cell-to-cell variability, such as stochastic gene expression, are increasingly recognized to play an important role in the dynamics of tissues, tumors, microbial communities… However, they are usually ignored or oversimplified in theoretical models of cell populations.

In this thesis, we propose a *cell-based* (each cell is represented individually), *multi-scale* (cellular decisions are controlled by biochemical reaction pathways simulated in each cell) approach to model the dynamics of cell populations. The main novelty compared to traditional approaches is that the fluctuations of protein levels driven by stochastic gene expression are systematically accounted for (i.e., for every protein in the modeled pathways). This enables to investigate the joint effect of cell-intrinsic and environmental sources of cell-to-cell variability on cell population dynamics. Central to our approach is a parsimonious and principled parameterization strategy for stochastic gene expression models.

The approach is applied on two case studies. First, it is used to investigate the resistance of HeLa cells to the anti-cancer agent TRAIL, which can induce apoptosis specifically in cancer cells. A single-cell model of TRAIL-induced apoptosis is constructed and compared to existing quantitative, single-cell experimental data. The model explains fractional killing and correctly predicts transient cell fate inheritance and reversible resistance, two observed properties that are out of reach of previous models of TRAIL-induced apoptosis, which do not capture the dynamics of cell-to-cell variability. In a second step, we integrate this model into multi-cellular simulations to study TRAIL resistance in virtual scenarios constructed to help bridging the gap between standard *in-vitro* assays and the response of *in-vivo* tumors. More precisely, we consider the long-term response of multi-cellular spheroids to repeated TRAIL treatments. Analysis of model simulations points to a novel, mechanistic explanation for transient resistance acquisition, which involves the targeted degradation of activated proteins and a differential turnover between pro- and anti- apoptotic proteins.

Second, we apply our approach to a synthetic spatial patterning system in yeast cells developed by collaborators. Focusing first on a sensing circuit responding to a messenger molecule, we construct a single-cell model that accurately capture the response kinetics of the circuit as observed in flow cytometry data. We then integrate this model into multi-cellular simulations and show that the response of spatially-organized micro-colonies sub-

mitted to gradients of messenger molecules is correctly predicted. Finally, we incorporate a model of a killing circuit and compare the predicted patterns of dead or alive cells with experimental data, yielding insights into how the circuit parameters translate into multicellular organization phenotypes.

Our modeling approach has the potential to accelerate the obtention of more quantitative and predictive models of cell populations that encompass the molecular scale.

# Résumé

Les sources intrinsèques d'hétérogénéité cellulaire, comme l'expression stochastique des gènes, sont de plus en plus reconnues comme jouant un rôle important dans la dynamique des tissus, tumeurs, communautés microbiennes… Cependant, elles sont souvent ignorées ou représentées de manière simpliste dans les modèles théoriques de populations de cellules.

Dans cette thèse, nous proposons une approche cellule-centrée (chaque cellule est représentée de manière individuelle), multi-échelle (les décisions cellulaires sont placées sous le contrôle de voies de signalisation biochimiques simulées dans chaque cellule) pour modéliser la dynamique de populations de cellules. La nouveauté principale de cette approche réside dans la prise en compte systématique (pour toutes les protéines modélisées) des fluctuations du niveau des protéines résultant de l'expression stochastique des gènes. Cela permet d'étudier l'effet combiné des causes intrinsèques et environnementales d'hétérogénéité cellulaire sur la dynamique de la population de cellules. Un élément central de notre approche est une stratégie parsimonieuse pour attribuer les paramètres de modèles d'expression stochastique des gènes.

Nous appliquons cette approche à deux cas d'étude. Nous considérons en premier la resistance à l'agent anti-cancer TRAIL, qui peut induire l'apoptose sélectivement dans les cellules cancéreuses. Nous construisons d'abord un modèle 'cellule unique' de l'apoptose induite par TRAIL et le comparons à des données existantes quantitatives. Le modèle explique la mort fractionnelle (le fait que seul une fraction des cellules meurent à la suite d'un traitement) et prédit correctement l'héritabilité transiente du destin cellulaire ainsi que l'acquisition transiente de résistance, deux propriétés observées mais hors de portée des modèles pré-existants, qui ne capturent pas la dynamique de l'hétérogénéité cellulaire. Dans une seconde étape, nous intégrons ce modèle dans des simulations multi-cellulaires pour étudier la résistance à TRAIL dans des scénarios virtuels intermédiaires entre les études classiques *in-vitro* et la réponse de tumeurs *in-vivo*. Plus précisément, nous considérons la réponse en temps long de sphéroides multi-cellulaires à des traitements répétés de TRAIL. L'analyse de nos simulations permet de proposer une explication originale et méchanistique de l'acquisition transiente de résistance, impliquant la dégradation ciblée des protéines activées et un différentiel dans le renouvellement des protéines pro- et anti-

apoptotiques.

Nous appliquons aussi notre approche à un système synthétique de création de motifs développé dans des levures par des collaborateurs. Nous nous concentrons d'abord sur un circuit senseur d'une molécule messager pour lequel nous construisons un modèle cellule unique qui capture de manière fine la dynamique de réponse du circuit telle qu'observée par cytométrie en flux. Nous intégrons ensuite ce modèle dans des des simulations multi-cellulaires et montrons que la réponse de micro-colonies organisées spatialement et soumises à des gradients de molécule messager est correctement prédite. Finalement, nous incorporons un modèle d'un circuit de mort et comparons les motifs prédits de cellules mortes/vivantes avec des données expérimentales, nous permettons de mieux comprendre comment les paramètres du circuit se traduisent en phénotypes d'organisation multi-cellulaire.

Notre approche peut contribuer à l'obtention de modèles de populations de cellules de plus en plus quantitatifs, prédictifs et qui englobent l'échelle moléculaire.

# Table of Contents

# Acknowledgements

First I would like to thank the members of my PhD committee, Alessandra Carbone, Mark Chaplain, Paul Macklin, Emmanuel Barillot and Jérémie Roux. I am profoundly grateful that they invested precious time in reading this manuscript, traveling to Paris and evaluating my defense.

There are two persons I cannot thank enough at this place: my PhD advisors Dirk Drasdo and Gregory Batt. They gave me the opportunity to work in an exciting field and provided me guidance and assistance throughout this long journey. I learned a lot from their mentoring, and because they have different personalities and research styles, I probably learned twice as much! Thanks to their great support, I was also able to attend many top-class conferences around the world, as well as the profoundly enriching q-bio summer school.

A special thanks goes to William Weens, who taught me a lot on c++, gave me many useful programming tricks and helped me finding my way in large and complex codes (such as his, a true masterpiece). Without him I would never have done as much as I did…

Another special thank goes to Szymon Stoma, with whom I worked closely on most of my projects. I really enjoyed the brainstorming sessions, the endless but fruitful discussions on various aspects of the modeling, the numerical algorithms, the code architecture…

I am also grateful to Ron Weiss for allowing me to visit his lab in Boston on several occasions, and I would like to thank Brian Teague and Jing-Jing Sun for what was an exciting, challenging and to me very formative collaboration.

At Inria I found an exceptional scientific environment and a very pleasant atmosphere. I would like to thank all the members of Bang (Mamba now) and Contraintes (Lifeware), especially Nick Jagiella, Géraldine Cellière, Paul Van Liedekerke, Noémie Boisset, and François Fages. I also thank Jim Pioche and Habib Jreige for the organization of weekly football games at the (old) Rocquencourt site. I'm also grateful to my coffee machine for its flawless functioning at an intense rate during many years.

Finally and most importantly, those years would not have been the same without the great moral support I received from my family and friends. Justine, my love and other half, especially contributed to the maintenance of my mental health. Our cat, Freja, also helped a lot.

# Chapter 1

# Introduction

## 1.1 Motivations

Understanding the behavior of multi-cellular assemblies such as tissues, tumors or microbial communities is a central goal of biology, and improvements of this understanding often translate into medical or biotechnological advances. Mathematical modeling is an invaluable tool to help realizing this goal (Thorne et al. 2007; Schnell et al. 2007; Anderson & Quaranta 2008; Dada & Mendes 2011; Deisboeck et al. 2011; Wolkenhauer et al. 2014). By formalizing into equations our current understanding on the mechanisms at play, it allows to evaluate the quality of this understanding, to identify knowledge gaps or inconsistencies with observations. It can also point out to new, highly informative experiments to perform, and permits to test hypotheses on yet unclear mechanisms.

### Cell-based multi-scale modeling, a natural formalism to describe the dynamics of tissues

Although many models of tissues that abstract cellular identity have brought important biological insights (consider for example the seminal work on the *chemical basis of morphogenesis* (Turing 1952), in which reaction-diffusion equations shed light on the mechanisms of pattern formation), the *cell-based multi-scale modeling* paradigm (Thorne et al. 2007; Dada & Mendes 2011; Deisboeck et al. 2011; Drasdo et al. 2014; Wolkenhauer et al. 2014) provides a natural and appealing way to formalize our understanding of the dynamics of tissues. Indeed, this paradigm adopts a cell-centric perspective, reflecting the way modern biologists think about life and how biological knowledge is organized. Turing himself considered such cell-based modeling in his paper on morphogenesis:

> *In this section a mathematical model of the growing embryo will be described. […] the cell theory is recognized but the cells are idealized into geometrical points. […] one*

*proceeds as with a physical theory and defines an entity called 'the state of the system'. One then describes how that state is to be determined from the state at a moment very shortly before. […] the description of the state consists of two parts, the mechanical and the chemical. The mechanical part of the state describes the positions, masses, velocities and elastic properties of the cells, and the forces between them. […] The chemical part of the state is given […] as the chemical composition of each separate cell; the diffusibility of each substance between each two adjacent cells must also be given. […] This account of the problem omits many features, e.g. electrical properties and the internal structure of the cell. But even so it is a problem of formidable mathematical complexity. One cannot at present hope to make any progress with the understanding of such systems except in very simplified cases.*

The cell-centric perspective provides a natural decomposition of the tissue modeling problem into sub-problems. A first sub-problem concerns the modeling of the internal dynamics of cells and how it drives cellular decisions. Then, one should describe how the composition of the cell environment impacts the cell internal dynamics. By essence, *molecular systems biology* attempts to explicit those two models in terms of molecular interactions (Aebersold 2005).

To enable the description of tissue dynamics, it is then needed to also model *1)* how cells modify their environment and *2)* how contacts between neighbor cells impacts their internal dynamics. Finally, one should describe the own dynamics of the extra-cellular environment, in which, for example, processes such as molecular diffusion, transport, or mechanical relaxation could take place. Modeling the latter sub-problems (impact of cells on environment, cell-cell interactions and the own dynamics of the environment) and coupling them to the former (internal dynamics of cells and how they are influenced by the environment) is the central goal of *multi-cellular systems biology*.

When all the sub-models are coherently integrated, the resulting model is by definition a *cell-based multi-scale model*. Such model provides a natural and intuitive formalization of how a tissue is believed to change with time.

## A bottom-up strategy for the construction of predictive models of tissues

To this decomposition of the modeling problem also corresponds a natural, *bottom-up* strategy to build, extend and refine models of tissue dynamics (Galle et al. 2008).

The first step is to build predictive models of how individual cells evolve as a function of their environment. For this purpose, models of cell internal dynamics and their response to cell environment can be constructed based on existing knowledge. Their predictive power can then be assessed by precisely probing cellular processes in well defined environmental conditions and at the single-cell level (Spiller et al. 2010). Iteration between model re-

finements and experiments can be performed until a deemed sufficient predictive power is obtained.

In a second step, sub-models for how cells influence their environment and how cells in contact interact can be constructed and integrated with single-cell models. The resulting cell-based multi-scale models can also be iteratively refined by considering experimental systems of increasingly complex cell-environment and cell-cell couplings.

## Recent progress renders this strategy technically within our reach

While this strategy for formalizing, constructing and improving our knowledge of tissue dynamics is natural and appealing, its practical applicability is hindered by important technical limitations. Among those limitations, the most critical are of experimental nature.

Indeed, while our ability to understand and predict the evolution of individual cells is a cornerstone of cell-based multi-scale modeling, it was for a long time not possible to experimentally probe the behavior of single cells with precision and scope (Elowitz et al. 2002).

Thus, current knowledge on cell biology was in large part obtained from population measurements, in which a large number of seemingly similar cells are submitted to a treatment of interest, and then are lyzed such that their molecular constituents can be pooled together for analysis. In those conditions, interpretation of the results is difficult: biologists naturally and intuitively reason with a cell-level perspective, but available measurements only reflect a population average of the behavior of many cells (Batchelor et al. 2009).

In the past decade, tremendous progress in experimental techniques was realized and enabled experimentalists to probe the behavior of single cells with molecular detail, and in certain cases without requiring cell destruction, therefore enabling live-cell observation. Key elements in the emergence of *single-cell experimental biology* are genetic engineering, i.e. the ability to manipulate the genome of cells; the discovery of fluorescent proteins (Chalfie et al. 1994) and their utilization as tags for measuring the level or activity of any protein of interest (Crivat & Taraska 2012; Regot et al. 2014); flow cytometry and live-cell microscopy (Selimkhanov et al. 2012); the development of microfluidic cell culture systems (Mehling & Tay 2014), that enable very precise control on the micro-environment in which cells are grown; and more recently, 'omics' techniques that allow large-scale, close to genome-wide measurements (Wang & Bodovitz 2010): single-cell mass spectrometry (Bendall et al. 2011; Bodenmiller et al. 2012; Bendall et al. 2014) for proteomes, single-cell RNA-seq for transcriptomes (Tang et al. 2009; Shalek et al. 2014), and single-cell genome sequencing (Blainey & Quake 2014; Lasken & McLean 2014).

The utilization of those techniques enabled the investigation of cellular processes at the single-cell level and led to major insights into the underlying mechanisms. An illustrative example is the study of extrinsic apoptosis (i.e. programmed cell death induced by external

factors) in mammalian cells. Using fluorescent reporters to measure the activity of caspases (proteins that control the initiation and execution of apoptosis via their protease activity), it was found that the initiation phase was slow and highly variable between cells while the second phase, leading to irreversible commitment to death, happened very suddenly and showed little cell-to-cell variability (Albeck, Burke, Aldridge, et al. 2008; Hellwig et al. 2008). Such behavior, which is thought to be fundamental to allow a clear, unambiguous cellular decision between life and death, is masked in population-level measurements of caspases activity kinetics.

Aside of experimental limitations, our ability to manipulate cell-based multi-scale models is restricted by the computational cost of numerical simulations. Indeed, because those models are in general very difficult to study analytically, computational costs directly limit the complexity and the size of the models we can investigate. However, while this prevent for example the simulation of many cells with full molecular details (the first whole-cell model (Karr et al. 2012) only simulated an isolated *M. genitalium* cell, although hundreds of times), it is already possible to simulate very large models and the available computational power is still growing at a high rate.

In summary, the main technical limitation to the implementation of the *bottom-up, cell-centric* strategy to understand the dynamics of tissues, that is the difficulty to probe the behavior of single cells in controlled environments, was markedly overcome in the past decade. However, as will be now discussed, the concommitently emerging picture of how single cells behave also seriously challenges our ability to *model* this behavior. In other words, out of sight behind important technical barriers, were standing at least equally important methodological barriers.

## The methodological challenge posed by cell-to-cell variability

The relevance of the cell-based multi-scale modeling paradigm depends on our ability to predict the evolution of single cells in well controlled environmental conditions. In cases of failure, the modeling of tissue dynamics could still be attempted and achieved with approaches that abstract cellular identity. However, it would likely prevent the much needed gap bridging between the molecular and tissue scales (Wolkenhauer et al. 2014).

Paradoxically, as we are more and more able to observe and measure single cell behaviors, they seem less and less predictable. Indeed, there is a growing awareness that *cell-to-cell variability* is widespread and constitutes a fundamental property of cells. It refers to the fact that *a priori* similar cells cultivated in the same environmental conditions could develop important phenotypic heterogeneity and take very different cellular decisions in response to signals (Balázsi et al. 2011).

This phenomenon challenges our ability to build predictive models of single cell behavior.

To facilitate the description of those challenges, we will consider an illustrative example. A population of cancer cells is submitted to a cytotoxic drug. We assume that all cells always sense the same environment, i.e. that cell-extrinsic sources of cell-to-cell variability can be excluded. Although at the time of treatment there are no observable differences between cells, 90% die rapidly after treatment while others survive.

A first attitude a modeler could adopt in this situation is to consider that because 'identical' cells (from the viewpoint of the observer) respond differently to the treatment, the model should include some *intrinsic noise* (or stochasticity) to account for cell fate variability. Concretely, if the model for cell dynamics and its response to drug is phenomenologic, one will simply attribute a probability of responding to the drug and apply it to each cell. If on the contrary the model is mechanistic (i.e. attempts to explicit chemical reactions at play), one can incorporate stochasticity into an originally deterministic model so as to account for molecular noise (Wilkinson 2009).

This approach was in many cases successful to explain cell-to-cell variability observations. Historical examples are the variability in the number of viral particules released by infected bacteria, that can be mainly attributed to stochastic production of viral particules in the infected cell (Delbrück 1945); or the variability in how much and how fast cells produce a protein in response to promoter activation, to which instrinsic stochasticity of transcription and translation largely contributes (McAdams & Arkin 1997). Other examples include the contribution of a triple positive feedback loop in $NF\text{-}\kappa B$ signaling to digital pathway activation (Tay et al. 2010) after exposure to very low levels of signal ligands.

Intrinsic noise is not the only possible explanation to cell-to-cell variability. Because live monitoring of the complete molecular state of single cells is way beyond our scope, it is possible that the observed apparition of cell-to-cell variability (in our example, variability in the decision between life and death) simply reflects pre-existing, hidden differences between cells (Ricicova et al. 2013).

What could indicate the presence of hidden heterogeneity? First, if a mechanistic model of how the drug acts to induce cell death is available, it could fail to reproduce sufficient cell fate variability under realistic assumptions for its parameters. In this case, one will search for plausible cellular quantities that could vary between cells and are implicated in the drug response. Those quantities might be already in the model but assumed not to vary between cells or they could lie at the boundaries of model scope, hence requiring a model extension.

Indication of hidden heterogeneity could also exist in the experimental data and can be detected without mechanistic knowledge. For example, if sibling cells are more likely to respond the same way to the drug than non-related cells, it shows that some level of hidden heterogeneity contributes to cell fate variability and that the related cellular quantities are, to some extent, conserved at division.

To model the impact of pre-existing heterogeneity (often referred to as *extrinsic noise*

in such context) on the observed variability, the standard approach (Gaudet et al. 2012; Kallenberger et al. 2014; Shokhirev et al. 2015) is to sample the value of the initially variable cellular quantities from adequate distributions and to use it as initial condition for a model of cell response (this model can be deterministic or include some stochasticity representing *intrinsic noise*).

Situations in which pre-existing heterogeneity was demonstrated to be a major cause of cell-to-cell variability in response to a stimulus are numerous. For example, cancer cells treated with the apoptosis inducing ligand TRAIL commit to death after a highly variable delay that could be linked to pre-existing differences in the level of extrinsic apoptosis signaling proteins (Spencer et al. 2009). The predominant role of pre-existing heterogeneity was demonstrated by the observation that sister cells commit to death almost at the same time. In addition, when a deterministic model of TRAIL signaling was combined with randomly distributed initial protein levels, predicted variability in death commitment delays were similar to observations. Such non-genetic variability in response to TRAIL could have implications for cancer therapy, because many TRAIL-like drugs (i.e., that target the same receptors) were or are still being developed and tested in clinical trials (Dimberg et al. 2013). Another therapeutically important example is the identification of bacterial *persistence* (or growth bistability, i.e. the pre-existence of a small fraction of non-growing cells) as a mechanism of resistance to antibiotics (Balaban et al. 2004).

In summary, modeling cell-to-cell variability in response to a stimulus is challenging because both intrinsic noise during the response and pre-existing, potentially un-observed heterogeneity could contribute. To address this, models of cell response (that can include stochasticity, see for example (Tay et al. 2010)) were combined with distributions of cell state to represent the pre-existing cell-to-cell differences. In many cases, this modeling approach enabled to successfully explain observations of cell-to-cell variability. However, as we will see, it suffers from a major shortcoming that prevent to use such models as *single-cell models* to incorporate into *cell-based, multi-scale* models of tissue dynamics.

## Towards a dynamic and integrated view on cell-to-cell variability

The main limitation of the previous approach is that pre-existing heterogeneity is seen as a *static parameter* of the cell population. The origin of this heterogeneity is ignored, as well as the possibility that this heterogeneity might emerge from dynamic fluctuations in single cells.

However, with the notable exception of genetic information, many cellular quantities exhibit fluctuation timescales that can be of the order of the cell cycle or lower. For example, in mammalian cells, the *mixing time* (i.e. the typical time for single cells to reach high or low level of the population distribution) of protein levels was observed for a dozen of genes to

range between one and three days (Sigal et al. 2006). In general, sources of *diversification* are inevitably at play within cells, because many constituents are in low copy number and thus subjected to molecular noise. While regulation mechanisms can be implemented to buffer the effect of molecular noise, completely supressing noise would have a tremendous energetic cost and would not be evolutionnary favored (Fraser et al. 2004).

When a cell population is submitted to a stimulation, such as a TRAIL treatment, cells will respond differently depending on their initial state. If the response has an impact on the proliferative behavior of the cell or on its potential to undergo cell death, *selection* will operate and the distribution of cell states in the population will be affected. In the case of TRAIL-induced apoptosis, if the amount of a given protein has a strong and positive impact on the execution of apoptosis, its level in the population is expected to decrease after treatment (independently of the modifications that could affect this protein during signaling).

Because it does not account for *diversification* of cell state in absence of signaling, the approach that models pre-existing heterogeneity via distributions will grossly over-estimate *selection* effects: if a second treatment is applied on the same population, it will be predicted to have the same, diminished effect independently on the time interval between treatments, while in reality one would expect the recovery of the initial population distribution (and hence of the treatment impact) as time goes on. For example, *transient resistance acquisition* after treatment with anti-cancer drugs was observed in several occasions (for TRAIL (Flusberg et al. 2013), for inhibitors of EGFR signaling (Sun et al. 2014) or chemotherapeutic agents (Sharma et al. 2010)), and in each case pre-existing differences were shown to contribute to cell fate variability. Models that use the distribution approach to account for those differences would not be able to reproduce transient resistance acquisition.

This approach is therefore not suited for long-term simulation of cell populations submitted to arbitrary stimulations, while this capability is required by the *cell-based multi-scale* paradigm for describing tissue dynamics. Instead, single-cell models should account for the *diversification* of the factors that in the presence of drug promote cell fate variability. In short, variability that was modeled as being *extrinsic* should be recasted as intrinsic noise, allowing to fully represent the dynamic nature of cell-to-cell variability in cell populations. For example, if there is a set of proteins for which their levels are known to impact cell fate decision in response to a cytotoxic drug, and if the natural cell-to-cell variability in those levels is such that cell fate variability is expected; the single-cell fluctuations of those levels from which population variability originates should be explicitely modeled.

This would enable to study the interplay between diversification, selection and drug-induced cellular responses. Such model could be used to predict the long-term outcome of repeated drug treatments on cancer cell populations cultivated in homogeneous and controlled conditions. Those predictions can permit the identification of efficient strategies for

drug choice, combination, dosing and scheduling that can be tested and validated experimentally.

When such models are successfully obtained, they can be integrated in cell-based multi-scale models in order to describe increasingly complex population contexts (i.e. with increased cell-environment and cell-cell couplings). For example, the response to repeated treatments of *in-vivo* tumors is likely to differ from the response of the same population cultivated *in-vitro* in homegeneous conditions, because of heterogeneous drug penetration or nutrients supply, secretion of factors by the surrounding tissues, and potentially many other factors.

By comparing model predictions with observations from adequate experimental systems (such as *in-vitro* grown multi-cellular spheroids, which provides an interesting intermediate between monolayer cultures and *in-vivo* systems (Hirschhaeuser et al. 2010)), an increasingly comprehensive, multi-scale picture of the cell population dynamics could be achieved.

## Integrating top-down and bottom-up description of tissues: expected insights from synthetic biology

From now on, we considered the problem of how to improve and formalize our understanding of the dynamics of *existing, natural* tissues and discussed the applicability of a natural, cell-centric, bottom-up strategy to do so. However, *synthetic biology*, a recent but fast growing field of biological research, aims to craft *new, artificial* living systems with desired properties and behavior (Benner & Sismour 2005).

For this purpose, synthetic biologists rely on the genetic modification of existing cells. Therefore, the first successes of synthetic biology concerned the implementation of cell-level desired behavior, such as the production of high value molecules or the sensory detection of certain compounds. But nowadays synthetic biologists aim to extend their scope by implementing tissue-level behaviors (Purnick & Weiss 2009).

This goal is highly challenging because one needs to find the molecular-level, genetic modifications to perform in single cells that would enable a tissue-level desired behavior. In other words, *multi-cellular synthetic biology* has to solve the "inverse" problem that *multi-cellular systems biology* aim to address. Thus, two distinct tasks should be performed: first, one needs to identify candidate genetic implementations that are potentially able to generate the desired tissue-level behavior; second, those implementations should be constructed and the emerging behavior experimentally tested. Because the second task is tedious, costly and time-consuming, success greatly depends on the quality with which the first task (the *design* problem) is achieved. The set of genetic implementations to test should be minimized while their likelihood to be valid should be maximized.

Without a rational, model-based approach, meeting those constraints might be impos-

sible in practice as soon as the desired behavior has a certain degree of complexity. *Cell-based multi-scale modeling* is well suited for such model-based design because it allows to decompose the design problem into two sub-problems. First, using abstract, phenomenological models of cell behavior, cell-cell and cell-environment couplings, one can identify high-level cell behaviors that would lead to the desired tissue-level behavior. Then, one would search for genetic implementations that can program those high-level cell behaviors. This second step corresponds to the simpler (but still very challenging), historical design problem addressed by synthetic biologists.

Because this top-down design strategy rely on abstract, high-level phenomenological models, deviations between the real and desired cell and tissue behavior are inevitably expected. Thus, it should be complemented by a bottom-up modeling approach to characterize the real behavior of a given genetic implementation of the system. This would allow to debug and tune the system until the desired behavior is achieved. In addition, by comparing and linking the top-down and bottom-up model descriptions of the system, one could gain insights on the functioning of natural tissues. For example, links between high-level tissue behavior and molecular networks into cells could shed light on the evolution of multi-cellular organisms from uni-cellular organisms.

## 1.2 Contributions

The principal contribution of this thesis is a cell-based modeling approach to simulate the dynamics of multi-cellular populations with molecular details. The approach can be conceptually divided in two main steps. First, in each cell, cellular decisions are driven by quantitative models of biochemical reaction pathways. Importantly, stochastic fluctuations in protein levels arising from transcriptional noise are modeled for all the proteins involved in those pathways. Second, those *single-cell* models are integrated into multi-cellular simulations in order to investigate the emergent population behavior when cell-cell and cell-environment couplings are added.

Cell-based, multi-scale modeling approach encompassing molecular details have already been proposed (Ramis-Conde et al. 2008; Walker et al. 2008; Mirams et al. 2013; Starruß et al. 2014). However, a main novelty of our approach is the fact that protein fluctuations are accounted for all modeled proteins. This is important because those fluctuations *intrinsically* generate cell-to-cell variability, even when cells sense homogenous and constant environments. If the environment is changing, this variability will be dynamically reshaped by the interplay between environment sensing and protein fluctuations. Moreover, if the environment is also heterogenous among cells, both cell-intrinsic and environmental sources of cell-to-cell variability will impact the behavior of the whole population. Systematically modeling those fluctuations is challenging, because the parameters governing the stochas-

tic expression of proteins are gene-specific, largely unknown and difficult to infer (Sherman & Cohen 2014). Building upon analytical results on a popular stochastic gene expression model, we have developped a parsimonious and principled method to enable this systematic parameterization.

We applied our cell-based multi-scale modeling approach to two biological problems. The first problem concerns the resistance of cancer cell populations to the therapeutical agent TRAIL, which induces apoptosis selectively in cancer cells. The second problem is the modeling of a synthetic yeast patterning system currently developped in the Weiss Laboratory at MIT.

In both cases, accounting for protein fluctuations appeared highly needed. For TRAIL-induced apoptosis, evidence for an important role of cell-to-cell variability in protein levels was accumulating (Spencer et al. 2009; Rehm et al. 2009); yet, despite the existence of a detailed kinetic model of TRAIL apoptosis signaling (Albeck, Burke, Spencer, et al. 2008), no model was able to explain the observed fractional killing, or observations on the transient correlation between sister cell fate (Spencer et al. 2009). For the synthetic yeast patterning system, high-throughput flow cytometry data revealed highly heterogenous response kinetics for a sensing circuit central to system functioning. Such heterogeneity is likely to impact the overall system behavior.

In both cases, integrating single-cell models into multi-cellular simulations also appeared critical. For TRAIL-induced apoptosis, it is required to investigate the long-term response of cancer cell populations repeatedly treated by TRAIL, and to study the contribution of cell-intrinsic and environmental effects in TRAIL resistance in realistic scenarios (i.e., that ressemble the treatment of *in-vivo* tumors). For the synthetic yeast patterning system, the need to simulate spatially organized cell population is obvious as the long-term objective of the project is to obtain controlled, dynamic spatial patterns within and between micro-colonies.

In both cases, our method to account for protein fluctuations was successful. For TRAIL-induced apoptosis, comparison with available single-cell data (Spencer et al. 2009) shows that it permits to greatly increase the prediction scope of a previous model of TRAIL apoptosis signaling (Albeck, Burke, Spencer, et al. 2008). Moreover, the obtained model also led to the finding that fractional killing and transient resistance acquisition after TRAIL treatment does not necessarily involve the parallel activation of others, *pro-survival* pathways as was predominantly proposed (Falschlehner et al. 2007) but could simply emerge from an interplay between protein fluctuations, selection and targeted degradation of pro-apoptotic forms. Together, those contributions resulted in a publication into a reference journal in the computational biology field (Bertaux et al. 2014). For the synthetic yeast patterning system, it enabled the construction of a single-cell model of the sensing circuit that accurately capture the heterogenous response kinetics observed by flow cytometry.

In both cases, integrating the corresponding single-cell models into multi-cellular simulations also yielded important insights. For TRAIL-induced apoptosis, it predicted that resistance acquisition by non-mutational mechanisms can be very important, in agreement with observations (Flusberg et al. 2013), and shed light on the molecular mechanisms involved, such as the targeted degradation of activated pro-apoptotic form and an imbalance between the turnover of pro- and anti- apoptotic proteins. In addition, simulations of the treatment of multi-cellular spheroids suggested that limited TRAIL penetration is unlikely to be a driving cause of resistance, but that it can exacerbate the impact of cell-intrinsic resistance acquisition. For the synthetic yeast patterning system, the multi-cellular simulations correctly predicted the spatio-temporal response of micro-colonies submitted to gradients of messenger molecule, demonstrating that biological models can be quantitative from single cells to spatially organized populations.

Another contribution of this thesis is the developpment of software tools (for example, a tool to infer parameters of a protein fluctuation model from the protein variability and mixing time, and tools that allow the automated generation of simulation code from a programmatic description of a model). Those tools could be useful to other researchers interested in the simulation of cell-based multi-scale models and are available on GitHub http://github.com/fbertaux.

## 1.3 Outline

This document is organized as follows. Chapter 2 describes the construction and validation of a single-cell model of TRAIL-induced apoptosis. The model is constructed from a previous ODE model using a generic and principled approach to systematically account for protein level fluctuations. Model predictions are compared to existing data on cell fate variability, cell fate correlation between sister cells, and transient resistance gain after treatment.

In Chapter 3, the obtained single-cell model of TRAIL-induced apoptosis is integrated in a cell-based multi-scale model to predict the long-term response of cancer cell populations to repeated TRAIL treatments in two virtual experimental *in-vitro* settings: 'monolayer' and spheroid cultures. In the spheroid setting, the fate of cells are coupled by reduced diffusion and increased degradation of TRAIL within the spheroid. In the 'monolayer' setting, no direct or indirect cell-cell couplings are assumed to play a role. We use first the 'monolayer' setting to investigate the role of time in TRAIL resistance. We vary the treatment parameters (dose and period) as well as parameters relating to the molecular implementation of TRAIL-induced apoptosis (mimicking cell lines with different resistance properties). The results provide insights on the mechanisms driving long-term resistance that we formalize in a simple phenomenological model. Finally, we use the spheroid setting to investigate the

additional impact of limited TRAIL penetration into the long-term resistance to repeated treatments.

Chapter 4 presents the results obtained by applying our cell-based multi-scale modeling approach to guide the development of an artificial yeast patterning system. We start by describing the context of the collaboration, the genetic implementation of the system and the experimental data that will be confronted to our model. Focusing on a sensing circuit, which is a central component of the full system, we construct a single-cell model that quantitatively capture the behavior of the system as quantified by flow cytometry data. Then, we integrate into multi-cellular simulations and ask wether we can correctly predict the behavior of spatially organized micro-colonies responding to a gradient of messenger molecules. Finally, we extend the model to account for the combined behavior of the sensing circuit with a cell death circuit and compare model predictions with experimental data.

In Chapter 5, we discuss methodological problems related to the description and simulation of cell-based multi-scale models such as the ones used in this thesis. We present algorithms for the simulation combined signal transduction and stochastic gene expression and discuss their computational efficiency. We then discuss a few important points regarding the simulation of cell populations. Finally, we consider the problem of minimizing the development time and code errors asociated to the simulation of cell-based multi-scale models and present a tool allowing the automated generation of simulation code from a programmatic model description.

Chapter 6 concludes this thesis. We summarize the results and try to relate them with the current efforts of the quantitative biology and system biology communities. We then propose future directions of research.

# Chapter 2

# Modeling the dynamics of cell-to-cell variability in TRAIL-induced apoptosis

## 2.1 Background: the importance of protein fluctuations in TRAIL-induced apoptosis

### TRAIL: an inducer of extrinsic apoptosis selective towards cancer cells

Apoptosis is a form of controlled cell death playing a crucial role in the development and homeostasis of multi-cellular organisms. Many human diseases are linked to anomalies in the regulation of apoptosis. Apoptosis can be triggered either from inside the cell (intrinsic apoptosis) or by external *death ligands*, i.e. proteins that are able to bind certain type of membrane-bound proteins called *death receptors* (extrinsic apoptosis).

Several types of death ligands and death receptors are known. Among them, TNF (Tumor Necrosis Factor) and its receptors (TNFRs) were the first to be characterized. TNF plays a central role in the regulation of inflammation and many successful drugs target its receptors. Other death ligands are CD95L (and its receptor CD95) and TRAIL (and its receptors DRs) (Ashkenazi & Dixit 1998).

When TRAIL was discovered in the mid-90's (Wiley et al. 1995; Pitti et al. 1996), it raised high interest because it was observed to induce apoptosis preferentially in cancer cells while being harmless to healthy cells (Walczak et al. 1999; Ashkenazi et al. 1999). This selectivity, together with the independence of extrinsic apoptosis with respect to the mutational status of p53 (p53 mutations are a common cause of resistance to chemotherapy), suggests that using TRAIL-induced apoptosis could be an efficient strategy in cancer therapy (Johnstone et al. 2008).

This led to dozens clinical trials involving drugs targeting TRAIL death receptors (such as TRAIL itself or antibodies) either in monotherapy or in combination with other anti-

cancer agents. Most trials were disappointing: although the drugs were well tolerated, anti-tumor efficacy was usually limited (Johnstone et al. 2008; Dimberg et al. 2013). Because understanding the mechanisms behind TRAIL resistance could enable to design better therapeutical strategies, a tremendous research effort was and is still undertaken to characterize TRAIL-induced apoptosis.

## How does TRAIL work: molecular events from receptor binding to death commitment

Most of the molecular events leading from TRAIL exposure to cell death are known (Johnstone et al. 2008). Figure 2.1 provides a stylized, simplified representation of those events. First, the binding of TRAIL to the death receptors DR4 or DR5 promotes their cross-linking. On the intra-cellular domain, this leads to the formation of the death-inducing signaling complex (DISC), which recruit and activate initiator caspases like caspase-8 (C8) (Falschlehner et al. 2007).

Once activated, initiator caspases cleave and activate effector caspases such as caspase-3 (C3). Effector caspases cleave essential structural proteins, inhibitors of DNase, and DNA repair proteins (PARP). Thus, as soon as the activity of effector caspases becomes important, the cell is irreversibly committed to apoptosis and the associated morphological changes usually occur within 10 to 20 minutes (Albeck, Burke, Aldridge, et al. 2008). The cellular effect of effector caspase activation is regulated by factors such as XIAP, which blocks the proteolytic activity of caspase-3 by binding tightly to its active site and promotes its degradation via ubiquitination (Deveraux et al. 1998). In addition to the direct activation of effector caspases, initiator caspases also activate Bid that then activates Bax (Luo et al. 1998).

If not kept in check by inhibitors, most notably Bcl2, activated Bax directly contributes to the formation of pores in the mitochondria outer membrane, leading to mitochondrial outer membrane permeabilization (MOMP) (Kim et al. 2006). In most cells, MOMP is required to reach lethal activity of effector caspases. Following MOMP, critical apoptosis regulators, such as Smac and cytochrome c (CyC), translocate into the cytoplasm. Smac binds to and inactivates XIAP, thus relieving the inhibition of effector caspases by XIAP (Du et al. 2000). Cytochrome c combines with Apaf-1 to form the apoptosome that in turn activates the initiator caspase-9 (C9) that activates effector caspases.

In summary, death signals are transmitted to effector caspases by initiator caspases either directly or indirectly via a mitochondrial pathway.

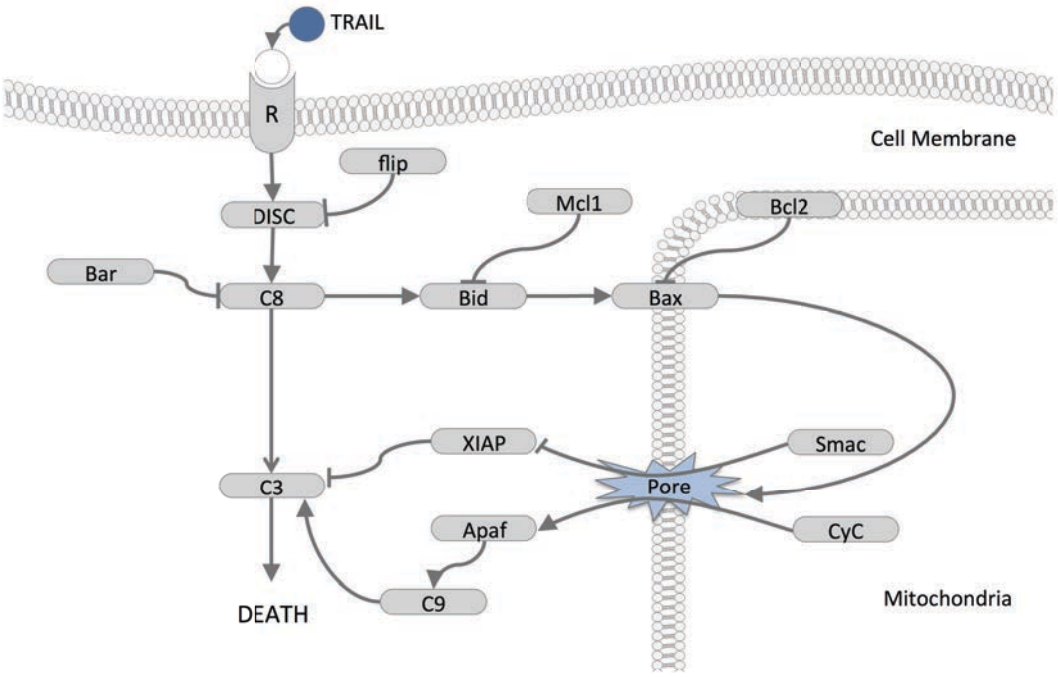



Figure 2.1: **TRAIL-induced apoptosis.** A simplified view of the TRAIL-induced apoptosis pathway is shown. The displayed interactions are abtracted from an established kinetic model (Albeck, Burke, Spencer, et al. 2008).

## Non-genetic cell fate variability in TRAIL-induced apoptosis

Resistance of cancer cell populations to TRAIL-induced apoptosis is widely observed in both *in-vitro* and *in-vivo* settings (Lemke, Karstedt, Zinngrebe, et al. 2014). Obviously, partial killing after TRAIL exposure has important therapeutical implications. Characterizing the mechanisms by which cells die in response to TRAIL is not sufficient: it is even more critical to understand why some cells can escape TRAIL-induced apoptosis.

The fate of a cell in response to TRAIL is not fully determined by its genotype. Indeed, even monoclonal populations cultivated *in-vitro* as monolayers usually exhibit partial responses to TRAIL treatment: only a fraction of the cells undergo apoptosis, even at saturating TRAIL doses (Spencer et al. 2009). Selection of subclones harboring new mutations is unlikely because the amount of such *fractional killing* seems to be a stable property of a given cell line: when two treatments are separated by at least a week, the same fraction of cells is killed in both cases (Flusberg & Sorger 2013).

However, if TRAIL treatment is repeated within a few days after initial exposure, less killing is obtained (Song et al. 2006 ; Flusberg et al. 2013). This *reversible resistance* (or *transient resistance acquisition* as we prefer to call it) could also have important implications for therapy, notably for the scheduling of drug applications. Together, those observations

demonstrate the existence of non-genetic factors in TRAIL-induced apoptosis that cause *1)* cell fate variability among naïve (i.e. previously unexposed to TRAIL) cell populations and *2)* transient resistance acquisition for the surviving population after TRAIL treatment.

## How can a cell survive TRAIL exposure: molecular events counteracting apoptotic signaling

What could be the molecular mechanisms that promote cell survival in absence of genetic defects in the TRAIL apoptosis machinery? This question was at the center of extensive research efforts. Many mechanisms of various nature were reported. Some are well understood but many are poorly characterized (Falschlehner et al. 2007).

Those mechanisms can be classified into three categories. First, *competitive binding* of anti-apoptotic proteins can prevent activated death signaling proteins to activate downstream targets. This is the case for example for decoy receptors towards TRAIL (Sheridan et al. 1997), for FLIP towards DISC (Shirley & Micheau 2013), for XIAP towards caspase-3, for Mcl-1 towards Bid (Youle & Strasser 2008) and for Bcl2 towards Bax (Oltvai et al. 1993).

A second type of defense against death signaling is the *active degradation* of activated pro-apoptotic proteins. Such degradation is achieved by the ubiquitin-proteasome pathway and mediated by proteins that bind the pro-apoptotic target and prime them for degradation (Vucic et al. 2011). For example, XIAP promotes the degradation of caspase-3 (Suzuki et al. 2001) and TRAF2 promotes the degradation of caspase-8 (Gonzalvez et al. 2012). Note that because this mechanism requires binding to the pro-apoptotic protein, it could have a dual action against death signaling: competitive binding and degradation (as for XIAP). Most activated pro-apoptotic proteins have been reported to be actively degraded (Breitschopf et al. 2000 ; Ferraro et al. 2008 ; Li & Dou 2000 ; Tawa et al. 2004 ; Thorpe et al. 2008).

Finally, an other type of defense against death signaling is the parallel activation of *pro-survival signaling*. Such signaling could induce the synthesis of, or post-translationally activate, anti-apoptotic proteins (Chaudhary et al. 1997; Son et al. 2010; Sun et al. 2011). Thus, such anti-apoptotic response ultimately rely on the two other types of defense (competitive binding and degradation). In fact, parallel activation of survival pathways is the most commonly advanced hypothesis for fractional killing in TRAIL-induced apoptosis (Falschlehner et al. 2007), although in many cases a clear mechanistic understanding of such parallel signaling is lacking.

Because of the variety and complexity of the molecular mechanisms that seem involved in determining a cell response to TRAIL exposure, there is no hope to gain a comprehensive, mechanistic picture of the cell decision between life and death without adopting a quantitative, system-level approach. In addition, because non-genetic cell-to-cell variability seems to impact this decision, both experimental and theoretical investigations probably cannot rely

solely on population level measurements and should acknowledge the single-cell level.

## Co-treatment with a protein synthesis inhibitor: a valuable tool to investigate core TRAIL apoptosis signaling

A common practice in the experimental investigation of extrinsic apoptosis is to block protein synthesis when cells are stimulated with death signals. For example, the first single-cell studies of TRAIL-induced caspases activation, in which fluorescent reporters (FRET) allowed to monitor caspase activity dynamics, used cycloheximide (CHX, a very potent inhibitor of protein synthesis) in co-treatment with TRAIL (Albeck, Burke, Aldridge, et al. 2008; Hellwig et al. 2008).

The principal motivation for this is to remove the influence of TRAIL-induced protein synthesis by pro-survival signaling. Therefore, changes in protein levels and activity are necessary mediated by protein-protein interactions and protein degradation. This facilitates the interpretation of experimental results because the *core* apoptosis signaling is functioning in isolation with respect to transcription- or translation- mediated regulations. In addition, because gene expression is generally noisy (gene and mRNAs are in low copy number in cells), blocking protein synthesis removes a source of noise in cell response to TRAIL.

Those single-cell experiments led to important insights into the timing of initiator and activator caspases activitation in response to TRAIL and CHX (Albeck, Burke, Aldridge, et al. 2008; Hellwig et al. 2008) in HeLa cells. First, the pre-MOMP delay is highly variable between cells. During that period, initiator caspases (caspase-8) activity steadily rises while effector caspases (caspase-3) are not active. After MOMP, a brutal, snap-action activation of effector caspases occur in all cells and with highly similar kinetics. Further experiments confirmed the role of XIAP and targeted degradation of caspase-3 in preventing caspase-3 activity during the pre-MOMP period.

Those precise single-cell observations were confronted to an ODE, mass-action kinetic model of TRAIL-induced apoptosis (Albeck, Burke, Spencer, et al. 2008). The model is exhaustive in the sense that it represents all events from TRAIL binding to receptors until effector caspases activation, including MOMP-related events. It describes 70 reactions between 58 protein species, including 17 native proteins. Despite this large size, it is not completely mechanistic, as some events such as receptor oligomerization upon TRAIL binding and DISC assembly are simplified into a few reactions. The model has important capabilities: it can reproduce the snap-action control over effector caspases activation and its regulation by Bcl2 and smac. Because it was constructed to reproduce experiments involving TRAIL and CHX co-treatments, it does not include protein synthesis.

In summary, single-cell experiments on TRAIL-induced apoptosis conducted in absence of protein synthesis, when complemented with kinetic modeling, enabled to gain a compre-

hensive and relatively detailed picture of the molecular events that connect receptor binding to cell death commitment. However, because in those conditions all cells eventually die, direct insights on the mechanisms behind fractional killing were out of reach.

## The role of protein level variability in death time variability

Fractional killing is the manifestation of cell-to-cell variability in the decision between life and death. As we already mentionned, cell-to-cell variability is also observed among cells that die: they commit to death after a highly variable delay from one another, including conditions in which protein synthesis is blocked (Albeck, Burke, Spencer, et al. 2008; Spencer et al. 2009; Rehm et al. 2009). Therefore, such variability cannot be explained by differences in translation-dependent TRAIL-induced survival responses.

Rather, it was proposed to originate mostly from pre-existing differences in the levels of proteins composing the apoptosis pathway. This explanation was supported by modeling: when taken as initial conditions of the previously described model from Albeck and colleagues (which is deterministic), differences in protein levels are sufficient to explain observed variability in death times (Spencer et al. 2009). Additional support to this explanation is provided by the observation that recently divided sister cells die almost synchronously (Rehm et al. 2009; Spencer et al. 2009; Bhola & Simon 2009), as expected if protein content is equally shared between daughters and if noise in protein-protein signaling reactions plays a marginal role.

Because cell-to-cell variability in protein levels is not frozen but results from a dynamical equilibrium driven by fluctuations in individual cells (Sigal et al. 2006), death synchrony between sister cells is expected to get weaker as the duration between division and treatment increased (Rehm et al. 2009; Spencer et al. 2009; Bhola & Simon 2009). This is indeed what was observed, providing further support for the predominant role of protein level differences in determining death time variability.

In summary, converging experimental and theoretical results indicate that natural fluctuations in cells generate cell-to-cell variability in protein levels that is transformed by apoptosis signaling into cell-to-cell variability in death times.

## Adopting a dynamic view on protein level variability by modeling fluctuations rather than distributions

To quantitatively assess the loss of death time synchrony between sister cells as their age at treatment time increases, the previously mentioned modeling approach, where only the consequences of protein fluctuations (cell-to-cell variability at a fixed time) are accounted for, is inadequate. Instead, protein fluctuations themselves should be modeled.

The need to account for protein fluctuations is even more stringent when considering observations after treatment with TRAIL alone. In that case, protein synthesis is not blocked and thus can impact the decision between life and death. Even if TRAIL does not change protein synthesis (via induction of survival pathways), significant differences with TRAIL and CHX treatments are expected, as the constitutively noisy protein synthesis will interact with signaling reactions.

Thus, before complexifying the model to account for eventual regulations via survival pathways induction, it is critical to assess how much can be explained when protein synthesis is not altered by TRAIL signaling. In this chapter, we investigate this question by enabling a kinetic model of TRAIL-induced apoptosis with stochastic protein turnover for all proteins, following a generic and principled approach. It enriches the model with a fundamental property as the dynamics of cell-to-cell variability is represented, allowing disentangling the effects of constitutive protein fluctuations, signaling protein-protein reactions and potentially induced changes in protein synthesis.

## 2.2 Model construction using a generic and principled approach to account for protein fluctuations in signal transduction pathways

### Extended vision of signal transduction pathways

Protein synthesis and degradation are subjected to noise, resulting in fluctuations of protein concentrations in individual cells and in cell-to-cell variability at the population level (Raser & O'Shea 2005). Such variability could have consequences on signal transduction: aside of conventional epigenetic differences (Rando & Verstrepen 2007), unequal access to ligand molecules or simply noise in signaling reactions, it often contributes importantly to heterogeneous behavior within an isogenic population (Spencer et al. 2009 ; Colman-Lerner et al. 2005 ; Geva-Zatorsky et al. 2006).

One approach to account for those differences is to incorporate protein level variability as random initial conditions of an ODE model describing the signaling reactions ("extrinsic noise approach") (Gaudet et al. 2012). However, variability is imposed at time zero and then behavior is deterministic: thus, as we argued in this thesis introduction, it is not appropriate to study transduction on long time scales, during which protein levels dynamically fluctuate (Sigal et al. 2006).

A more natural manner to account for protein level variability is to represent their stochastic synthesis and degradation ("intrinsic noise approach"). Although several studies did account for cell-to-cell differences in protein levels in an extrinsic, static manner

via random initial conditions (Spencer et al. 2009 ; Gaudet et al. 2012 ; Stoma et al. 2013 ; Kallenberger et al. 2014), and many models of signal transduction considered the effect of noise in protein-protein reactions (Shibata & Fujimoto 2005 ; Lapidus et al. 2008) or in the expression of signal transduction target genes (Tay et al. 2010 ; Lipniacki et al. 2006 ; Lipniacki et al. 2007), no kinetic model of signal transduction pathways considering systematically (i.e., for all the proteins acting in the pathway) noise in protein synthesis and degradation has been developed so far.

Here we propose a modeling approach to account for gene expression noise within kinetic models of signal transduction pathways. In brief, we use different instances of a popular stochastic model of gene expression to describe the fluctuations of all proteins in the pathways. Each of those individual *protein fluctuation models* are integrated with the ODE model describing protein-protein signal reactions, resulting in hybrid stochastic/deterministic model. Importantly, the parameter inference problem for the protein fluctuation models is addressed by adopting a principled, parsimonious strategy that rely on approximating the fluctuations of long-lived proteins.

## Modeling the fluctuations of individual proteins: a popular model of stochastic gene expression

We model protein fluctuations arising from gene expression noise with a stochastic process describing mRNA level fluctuations (promoter activity switches, mRNA production and degradation are stochastic events), and deterministic processes for protein translation and degradation (Figure 2.2). The stochastic reactions are triggered at a constant rate (or propensity), therefore, the time between two stochastic events is exponentially distributed. For example, the time needed for an inactive gene to become active is exponentially distributed with average $T_{off} = \frac{1}{k_{on}}$ and the time needed to produce one mRNA when the gene is active is also exponentially distributed with average equal to $\frac{1}{k_{sm}}$.

This popular model (Raj et al. 2006; Paszek 2007; Shahrezaei & Swain 2008; Suter et al. 2011; Dar et al. 2012; Singh et al. 2012; Viñuelas et al. 2013; Sherman & Cohen 2014) has many names, such as the *central dogma model of gene expression* or the *two-state transcriptional bursting model*. Here we will call it *protein fluctuation model* or alternatively *stochastic protein turnover model* to highlight the importance of the protein half-life in shaping fluctuations (as will be shown later).

Of note, the variant we consider does not account for noise in translation and protein degradation. This choice could be justified by the high copy number in which proteins are generally found in human cells (Schwanhäusser et al. 2011). It facilitates the numerical simulation of the model without complicating much its analytical treatment. The validity of this approximation is discussed in more details elsewhere (Paszek 2007).

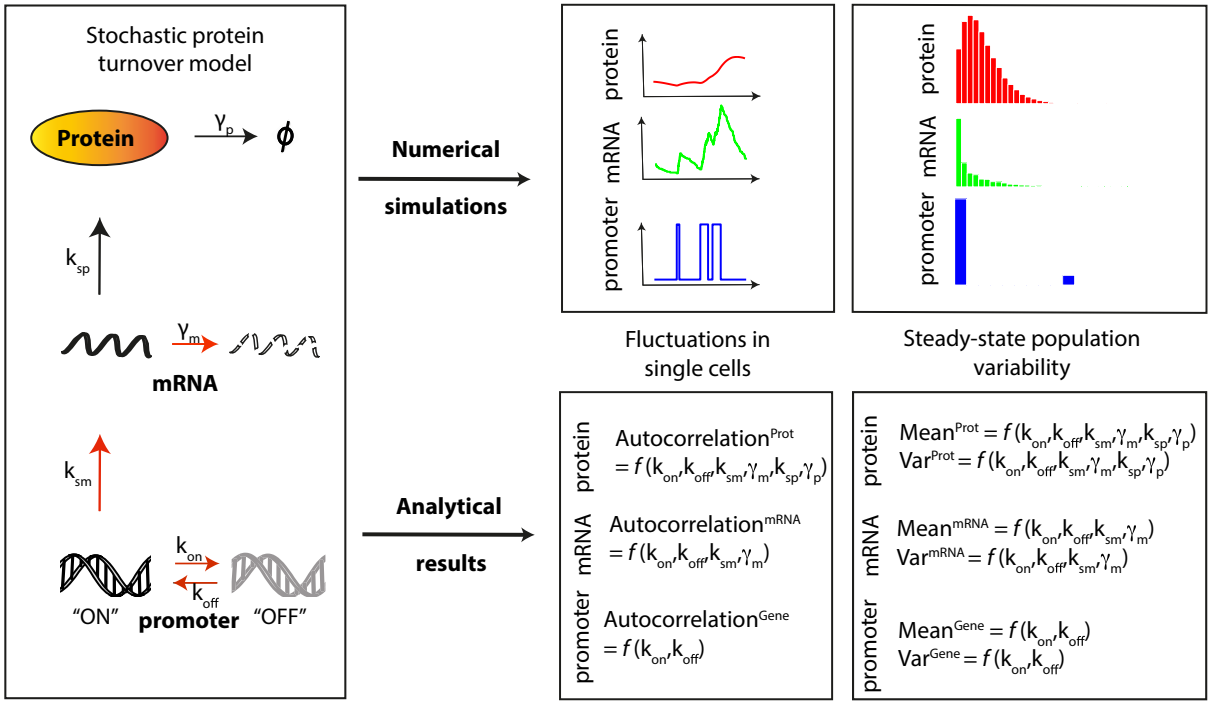



Figure 2.2: **Description, characterization and simulation of the stochastic protein turnover model.** It is defined by six rates governing model reactions. Gene activity switches, mRNA production and degradation (red arrows) are stochastic reactions. Protein synthesis and degradation reactions (black arrows) are deterministic. Numerical simulations can be used to generate the corresponding temporal fluctuations in single cells. When a population of cells is simulated, the cell-to-cell variability can be studied. After some time, cell-to-cell variability reaches a steady state. Analytical calculations provide expressions characterizing the steady-state variability (moments of the steady-state distributions), but also fluctuations (autocorrelation functions).

## Integration of individual fluctuation models within kinetic models of signal transduction pathways

Starting from an ODE, kinetic model describing the protein-protein reactions involved in signal transduction, each species corresponding to a protein in its native form is equipped with its own instance of the stochastic protein turnover model.

Those stochastic models are then integrated within the ODE model of signal transduction (Figure 2.3). The resulting model is therefore *hybrid* and features a *gene expression layer* where reactions are stochastic and a *signal transduction layer* where reactions are deterministic.

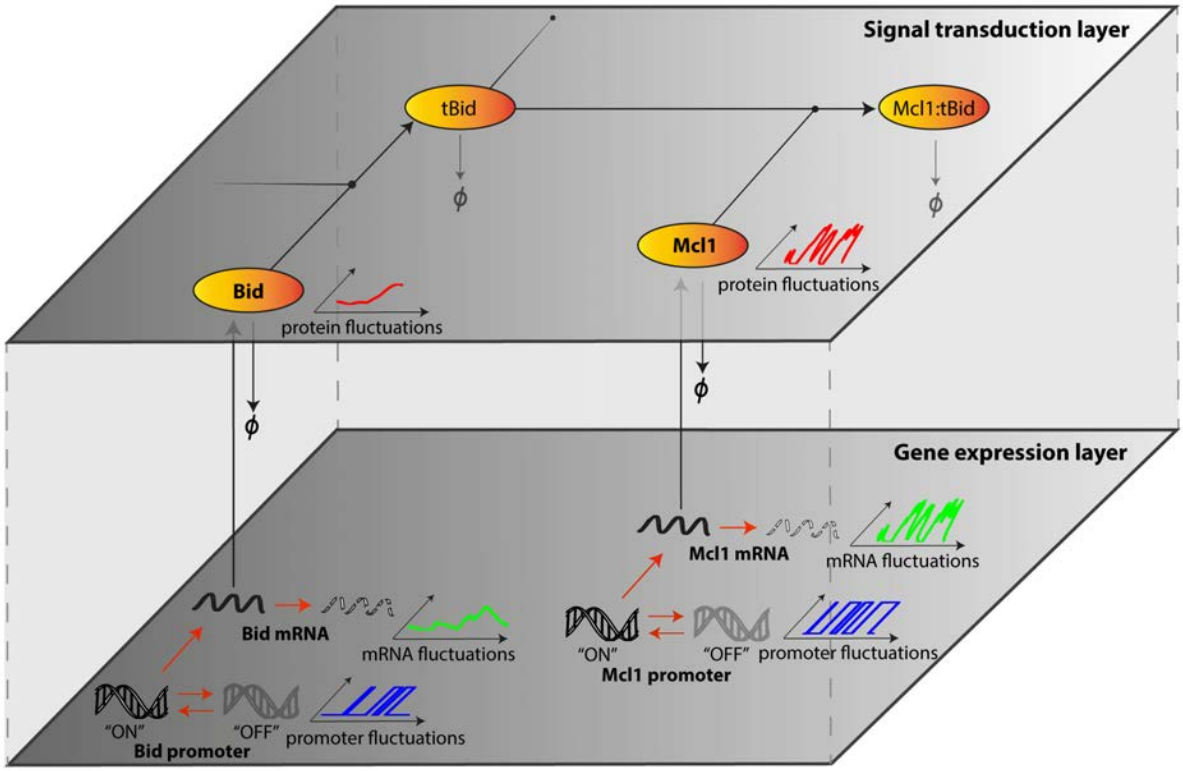



Figure 2.3: **Accounting for stochastic protein turnover in signal transduction pathways.** Scheme of the modeling approach. Protein-protein interactions mediating signal transduction (signal transduction layer) are modeled by ordinary differential equations. In parallel, promoter activity changes, mRNA production and degradation (gene expression layer) are seen as stochastic events and generate fluctuations in mRNA levels. This impacts the synthesis rates of the corresponding proteins. Together with protein degradation, it generates fluctuations in protein levels (here shown in absence of transduction). Only a fragment of the extrinsic apoptosis pathway is shown. Deterministic/stochastic interpretation of chemical reactions is represented with black/red arrows respectively.

## Inferring the rates of a stochastic protein turnover model is difficult

This approach introduces new parameters (6 rate parameters per protein) that need to be inferred. While the rates of such stochastic protein turnover models are rarely directly measurable, their value can be constrained by using experimentally measurable data and analyt-

ical results (Figure 2.2). Recently, significant progress has been made on both experimental and theoretical sides to enable this inference approach (Singh et al. 2012 ; Raj et al. 2006 ; Suter et al. 2011 ; Dar et al. 2012 ; Viñuelas et al. 2013 ; Paszek 2007 ; Munsky et al. 2009 ; Sherman & Cohen 2014 ).

First, expressions of the moments of the steady-state distribution can be obtained by performing Laplace transforms on the steady-state formulation of the chemical master equation associated with the model. It has notably been done by (Paszek 2007) and yields the following moments for the steady-state distribution:

Means of the gene status, mRNA level and protein level

$$\mathbb{E}[G] = \frac{k_{on}}{k_{on} + k_{off}} = \frac{k_{on}}{\gamma_g} \quad (\gamma_g = k_{on} + k_{off}) \tag{2.1}$$

$$\mathbb{E}[m] = \mathbb{E}[G]\frac{k_{sm}}{\gamma_m} \tag{2.2}$$

$$\mathbb{E}[P] = \mathbb{E}[m]\frac{k_{sp}}{\gamma_p} \tag{2.3}$$

Coefficients of variation (standard deviation divided by mean, abbreviated *CV* from now on) of the gene status, mRNA level and protein level

$$CV[G]^2 = \frac{1 - \mathbb{E}[G]}{\mathbb{E}[G]} \tag{2.4}$$

$$CV[m]^2 = \frac{\gamma_m}{\gamma_m + \gamma_g} CV[G]^2 + \frac{1}{\mathbb{E}[m]} \tag{2.5}$$

$$CV[P]^2 = \frac{\gamma_m \gamma_p (\gamma_g + \gamma_m + \gamma_p)}{(\gamma_g + \gamma_m)(\gamma_g + \gamma_p)(\gamma_m + \gamma_p)} CV[G]^2 + \frac{\gamma_p}{\gamma_p + \gamma_m} \frac{1}{\mathbb{E}[m]} \tag{2.6}$$

Usually, only the steady-state distribution of protein levels is experimentally available, for example using specific antibodies and flow cytometry. It has been shown that the moments of this distribution, and even its shape, is usually not sufficient to identify a single set of parameters of the stochastic protein turnover model that generates this distribution at steady-state (Sherman & Cohen 2014). This *unidentifiability* cannot be fully resolved by the additional knowledge of the protein degradation rate $\gamma_p$.

To infer the gene activity bursting kinetics with precision, another approach requiring genetic engineering and live-cell microscopy has been proposed (Suter et al. 2011). It relies on modifying the sequence of the gene to obtain both a short-lived mRNA and protein, the latter having luciferase activity enabling its quantification in live cells. Using maximum likelihood methods on obtained single-cell trajectories, estimates for $k_{on}$ and $k_{off}$ can be recovered. The amount of experimental work needed by this method, together with the fact

that additional estimations of the endogeneous mRNA and protein stabilities are required to describe fluctuations of the endogeneous protein, renders it in practice intractable for the systematic inference of protein fluctuation models to integrate in large size signal transduction pathways.

## Two statistics to characterize protein fluctuations

The difficulty to infer all the rates of a stochastic protein turnover model instance could mean that their full knowledge is not required to describe the fluctuations of protein level with a satisfying accuracy. To investigate this idea, it is needed to define a small number of *measures* or *statistics* that provide a characterization of the fluctuations generated by a stochastic protein turnover model instance.

A first natural choice is the relative wideness (*CV*) of the protein level state distribution. Its expression was given in (2.3). It informs about the typical amplitude of the protein fluctuations and depends on the 5 rate parameters $k_{on}$, $k_{off}$, $k_{sm}$, $\gamma_m$ and $\gamma_p$. Note that it does not depend on $k_{sp}$: because protein translation and degradation are deterministic, $k_{sp}$ simply acts as a scaling factor that sets the mean protein level.

The protein level steady-state distribution does not contain information about the speed at which it is visited in individual cells through fluctuations. A simple argument to this claim is that scaling all rate parameters by the same factor would leave the steady-state distribution unchanged as it amounts to rescale time.

A natural way to characterize this speed of fluctuations is to rely on the auto-correlation function of the protein level. This function measures how much information (quantified by the correlation) about the initial protein level remains after a given time. In other words, it gives the time evolution of the *memory* of single cells with respect to their protein level. Here, we provide an exact derivation of the auto-correlation function for the stochastic protein turnover model. This derivation and the resulting expression are presented in Appendix 1.

Three different autocorrelation curves corresponding to different model parameterizations are shown in Figure 2.4. To characterize this function with a single number, one can extract the time needed to lose half of the correlation with initial protein level. This value is called the *mixing time* and has already been measured experimentally for endogeneous proteins in human cells (Sigal et al. 2006).

In summary, we propose here to use the two real-valued statistics $CV$ (of the protein level distribution) and mixing time $\tau$ (of the protein level) to characterize the *amplitude* and *speed* of protein fluctuations generated by a given stochastic protein turnover model. A more detailed discussion on the parameterization of this model from constraints on $CV$ and $\tau$ is given in Chapter 5 (see 5.1, section *Parameterization of stochastic protein turnover models from protein variability and mixing time*).

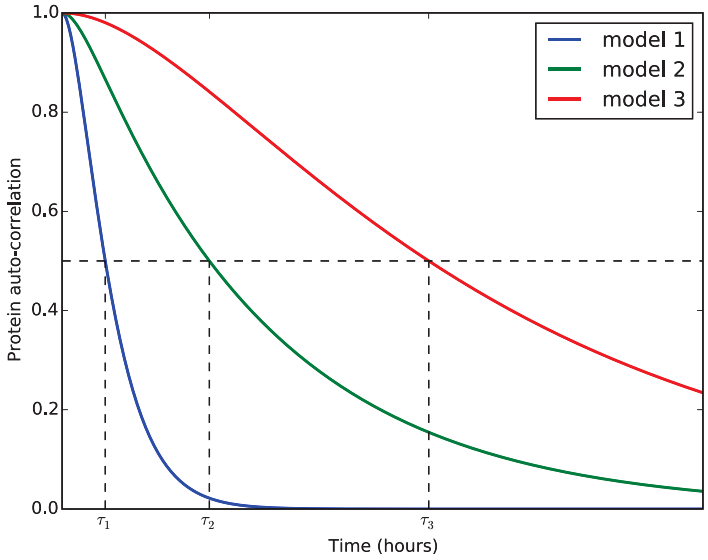



Figure 2.4: **Protein level auto-correlation function shape depends on model parameters.** Examples of auto-correlation curves for three different parameterizations of the stochastic protein turnover model. In each case, the corresponding *mixing time* $\tau$, i.e. the time at which auto-correlation equals 0.5, is indicated.

## Approximating the fluctuations of long-lived proteins by standard models

Intuitively, fluctuations of short-lived proteins are expected to be more sensitive to the precise kinetics of bursting than long-lived proteins, because the protein level at a given time reflects previous protein synthesis along a time window related to the protein half-life. Therefore, the level of long-lived proteins only informs about a cumulated amount of protein synthesis events, but not about when exactly those events occurred. On the other hand, the level of short-lived proteins is more closely related to the instantaneous rate of protein synthesis. It is this argument that led (Suter et al. 2011) to use short-lived mRNAs and proteins to infer the transcriptional bursting kinetics associated to a promoter sequence.

To investigate the validity of this hypothesis, we explored the relationship between the two statistics $CV$ and $\tau$ and the bursting parameters $T_{on} = \frac{1}{k_{off}}$ and $T_{off} = \frac{1}{k_{on}}$ in two different settings of protein and mRNA stability (Figure 2.5). In the first setting, the protein is stable (its half-life is 27 hours and is set by dilution), and the mRNA half-life is 9 hours. It corresponds to typical stabilities found in mammalian cells (Schwanhäusser et al. 2011). In the second setting, both the protein and mRNA are short-lived (half-lives of 2 and 1 hours respectively). Imposing $k_{on}$, $k_{off}$, $\gamma_m$ and $\gamma_p$ is not sufficient to determine the fluctuations of a stochastic protein turnover model: $k_{sm}$ remains unconstrained and can impact $CV$ and $\tau$. It was chosen such that the mean mRNA level $\mathbb{E}[m]$ equals 17, corresponding to the median value for the mRNAs quantified in a genome-wide study in mammalian cells (Schwanhäusser et al. 2011).

The range of both $CV$ and $\tau$ are narrower for the first setting compared to the second, pro-

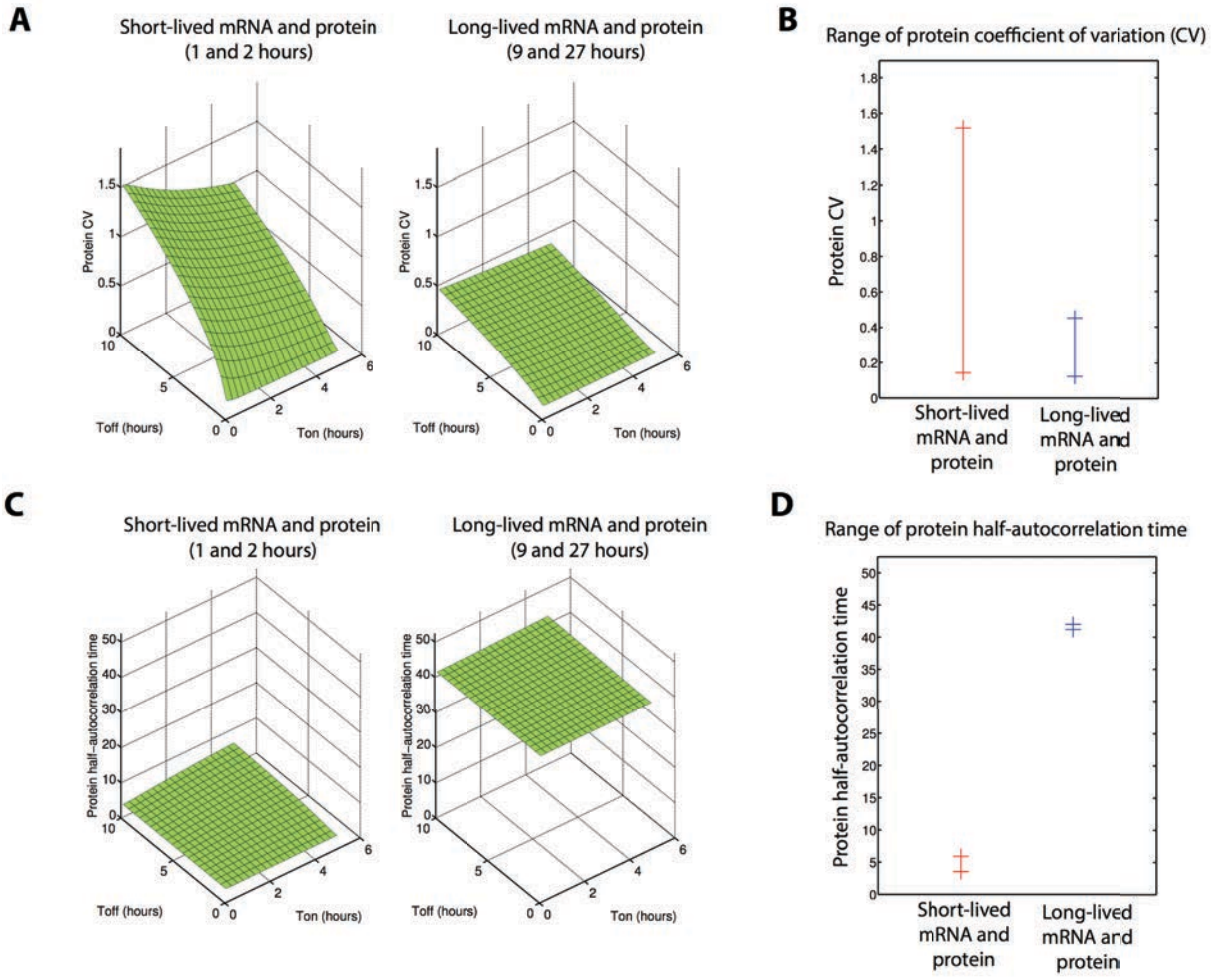



Figure 2.5: **Fluctuations of long-lived proteins are less dependent on transcriptional bursting kinetics than short-lived proteins.** Comparison of protein level coefficient of variation (A) and mixing time (C) as a function of transcriptional bursting rates for two settings: a short-lived protein and mRNA (half-lives of 2 and 1 hours, resp.) and a long-lived protein and mRNA (27 and 9 hours, resp.). Combinations of Ton and Toff values ranging from 0.1 to 5 hours and 0.1 to 10 hours respectively were tested ($T_{on} = \frac{1}{k_{off}}$ and $T_{off} = \frac{1}{k_{on}}$ are mean ON and OFF time of the gene). In each case, the remaining rate parameter $k_{sm}$ was chosen such that $\mathbb{E}[m] = 17$. (B and D) Representation of the range of values obtained for all models tested in (A) and (C).

viding support to the hypothesis that protein fluctuations are less sensitive to the bursting rates when the protein is long-lived. Otherwise stated, for typical protein and mRNA half-lives, a large set of promoter rate combinations leads to fluctuations of similar amplitude and speed as characterized by $CV$ and $\tau$.

Interestingly, in that case the obtained mixing time is around 40 hours, in the middle of the range of experimentally estimated values for twenty endogenous proteins in human cells (Sigal et al. 2006). This non-trivial consistency result adds to the relevance of using the stochastic protein turnover model to describe protein fluctuations in human cells.

Therefore, using a *standard* parameterization for the stochastic protein turnover model can provide a good approximation of protein fluctuations for most proteins. By *standard* parameterization we mean a parameterization that respect typical constraints on protein and mRNA stabilities, mean mRNA level and promoter switching times, as will be illustrated concretely on the example of TRAIL-induced apoptosis. Fluctuations and steady-state distribution of such a model are presented in Figure 2.6. This result is a cornerstone of our approach: it enables a parsimonious parameterization strategy in which only short-lived proteins necessitate particular attention, while others are equipped with standard stochastic protein turnover models.

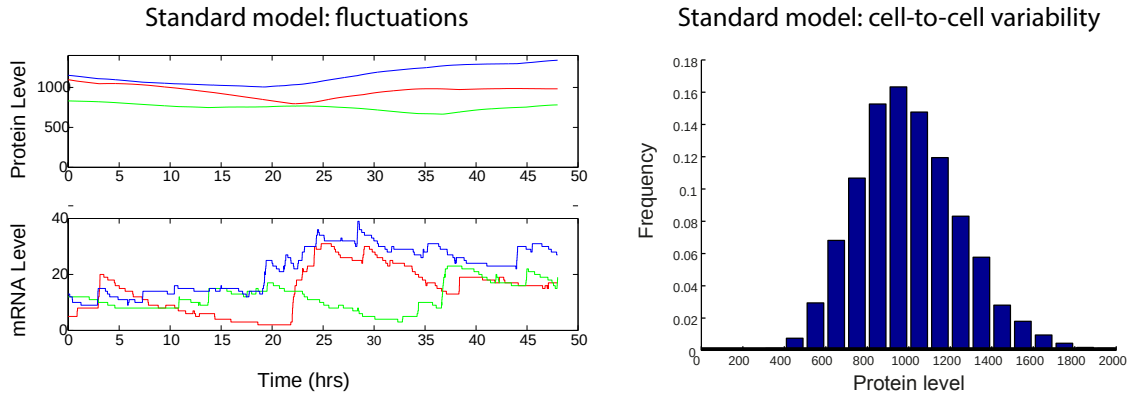



Figure 2.6: **Behavior of a standard stochastic protein turnover model.** Promoter switching rates respect typical ranges observed in (Suter et al. 2011) and lead to a protein level $CV$ of 0.25. Left plots show three representative single-cell time courses of protein and mRNA levels. Histogram at the right displays the corresponding distribution of protein level obtained when simulating a large number of cells for a long duration, corresponding to a snapshot of the cell-to-cell variability expected in a population.

## Application to TRAIL-induced apoptosis

We applied our approach to TRAIL-induced apoptosis. We used the EARM kinetic model mentionned earlier (Spencer et al. 2009) to describe protein-protein reactions taking place between TRAIL death exposure and cell death commitment. It comprises 17 native proteins and 41 other species involved in 71 reactions. The corresponding given equations are given

in Appendix 2, and simulation code is available at http://github.com/fbertaux/hEARM. The parameterization strategy for building the corresponding stochastic protein turnover models is summarized in Figure 2.7.

Unless required, we equipped all native proteins with the same default model of stochastic protein turnover. We used median values for mRNA levels, protein and mRNA half-lives from measured distributions in mammalian cells (Schwanhäusser et al. 2011). At the promoter level, switching rates were estimated for a dozen of genes (Suter et al. 2011), and we used measured values to constrain model reaction rates. Remaining rates were deduced from the mean and variance of protein level present as initial condition in (Spencer et al. 2009) using analytical expressions derived from the stochastic protein turnover model (Figure 2.7).

As mentioned earlier, short-lived proteins should be given particular attention. FLIP and Mcl-1 are known to exhibit very fast turnover (Nijhawan et al. 2003 ; Poukkula et al. 2005 ; Laussmann et al. 2012 ). Measurements in mouse ES cells (Sharova et al. 2009) also suggest that FLIP and Mcl-1 transcripts are particularly short-lived. Therefore, we considered a specific stochastic protein turnover model for those two proteins, exploring realistic ranges for promoter switching rates, mRNA half-life and protein half-life (Figure 2.7).

Importantly, all parameters have been constrained based on experimental data and analytical results, with the exception of four parameters ("ON" and "OFF" promoter switching rates, mRNA and protein half-lives for FLIP and Mcl-1). Note that because of their similar protein and mRNA half-life, we used the same couple of promoter switching rates for both proteins in our exploration of parameter space. Thus, our parsimonious parameterization strategy for the stochastic protein turnover models drastically limited the number of introduced degrees of freedom. This enabled us to systematically explore realistic ranges for remaining parameters.

Aside of the parameterization of the introduced protein fluctuation models, we made several modifications to the parameters of the model used in (Spencer et al. 2009). First, the half-life of the death ligand TRAIL, that was originally set to ~66 hours, was set to 9 hours, following an available experimental measurement (Youn et al. 2007).

Another discrepancy between the original parameters and biological knowledge was the degradation rates of the non-native forms. Almost all of them were assumed to also have a half-life of ~66 hours, while as we already mentionned, most activated apoptotic proteins are known to be targeted for proteasomal degradation and hence exhibit much shorter half-lives (Vucic et al. 2011). To account for this fact, we used a unique half-life (5 hours) for all non-native species, except for the mitochondrial pores (1.9 hours, same value as in the original model) and the complexes involving the short-lived protein FLIP and Mcl-1 (the half-life of the native form was used, based on the assumption that FLIP and Mcl-1 are targeted for degradation in a similar fashion alone or with their binding partner).

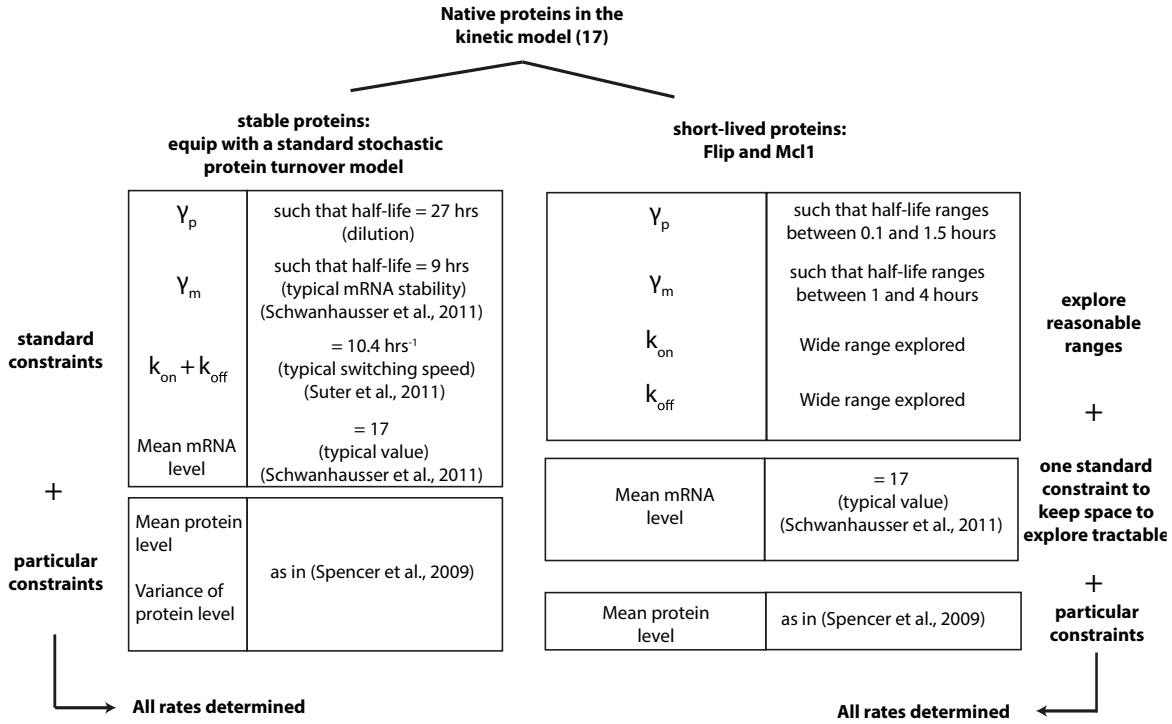



Figure 2.7: **Building stochastic protein turnover models for TRAIL-induced apoptosis.** Routine followed to choose rates of all 17 native proteins in the EARM kinetic model of TRAIL-induced apoptosis. Typical values from multi-genes studies in mammalian cells are used to constrain rate values. Specific attention is given to FLIP and Mcl-1 because they are known to be short-lived, and thus more prone to exhibit large variations.

Finally, we did not change the kinetic rates of the protein-protein reactions, except the rate of pC6 cleavage by C3, set to zero to represent the assumption that the caspase feedback loop (the ability of caspase 3 to cleave pro-caspase 6, which in turn can cleave pro-caspase 8) is absent (see authors comment to (Albeck, Burke, Spencer, et al. 2008), on editor's website). A complete list of parameter values that we used is given in Appendix 2.

The impact of those changes will be discussed in the discussion closing this Chapter. Importantly, most results presented in this Chapter are robust with respect to those changes. However, as we will see in the next Chapter, the long-term response of cell populations treated with TRAIL are strongly impacted by the rates at which activated apoptotic proteins are degraded.

To study the influence of stochastic protein turnover on fractional killing and reversible resistance, we sought to confront our model with existing quantitative data about TRAIL-induced apoptosis in HeLa cells. Those experiments, described in detail later, can be classified into two groups based on the type of information they contain: 1) quantification of the variability in cell fate, 2) characterization of the transient memory in cell state. While previous approaches using ODE models with distributions for initial protein levels (capturing a static description of cell-to-cell variability) (Spencer et al. 2009 ; Gaudet et al. 2012) are potentially able to reproduce the first type of data, a dynamic view on cell-to-cell variability as proposed in our model is needed to account for both types of data. We adopted the following strategy: first, search for models able to reproduce observations on cell fate variability; and second test whether valid models can robustly predict observed behaviors where transient memory matters.

## 2.3 Model predicts transient memory in cell sensitivity to TRAIL and CHX

Using live-cell microscopy, Spencer et al. investigated the fate of hundreds of cells after exposure to TRAIL and CHX (10 ng/mL and 2.5 µg/mL, Figure 2.8-A). All cells undergo MOMP with a highly variable delay (from 2 to 8 hours, Figure 2.8-B). To study cell fate inheritance, the authors also recorded 20 hours before treatment to identify sister cells (Figure 2.9-A). They were found to have highly correlated MOMP times (correlation coefficient close to 1 for recently divided cells, about 0.5 for older sisters - Figure 2.9-B, black curve).

Here, the MOMP time distribution provides a quantification of the cell fate variability, while MOMP time correlations between sister cells also give information on the transient memory in cell state. Within our framework, in-silico reproduction of those experiments is straightforward (Figures 2.8-D and 2.9-C), enabling us to investigate possible origins of transient cell fate inheritance.

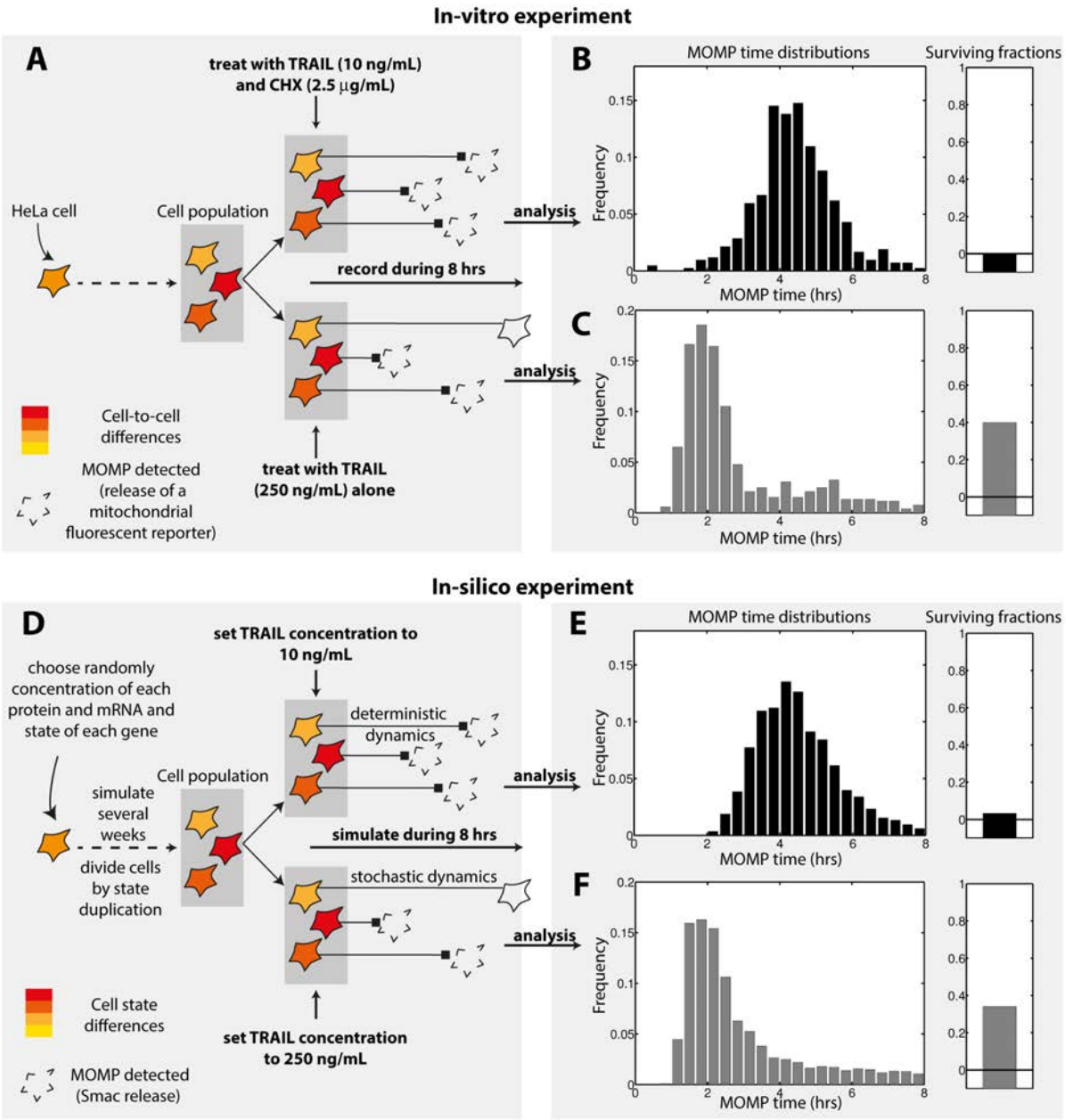



Figure 2.8: **Cell fate variability in TRAIL-induced apoptosis.** (A-C) Cell fate variability experiments performed in (Spencer et al. 2009). (A) HeLa cell populations were treated with either 10 ng/mL of TRAIL and 2.5 $\mu g/mL$ of cycloheximide (CHX) or 250 ng/mL of TRAIL alone. Cells were tracked during 8 hours by live-cell microscopy and MOMP time was detected via mitochondrial release of a fluorescent reporter. (B-C) Histograms of MOMP times and surviving fractions observed for treatment with (B) TRAIL and CHX or (C) TRAIL alone. (D-F) In-silico reproduction of those experiments with our "fitted" model (i.e. the parameterization in the explored parameter space region giving the best agreement for cell fate variability data, see detailed description in Appendix 2). (D) Simulations (see Appendix 3 for details). (E-F) Results for the (E) TRAIL and CHX or (F) TRAIL alone treatments. For the latter case, representative model trajectories are given in Figure 2.14.

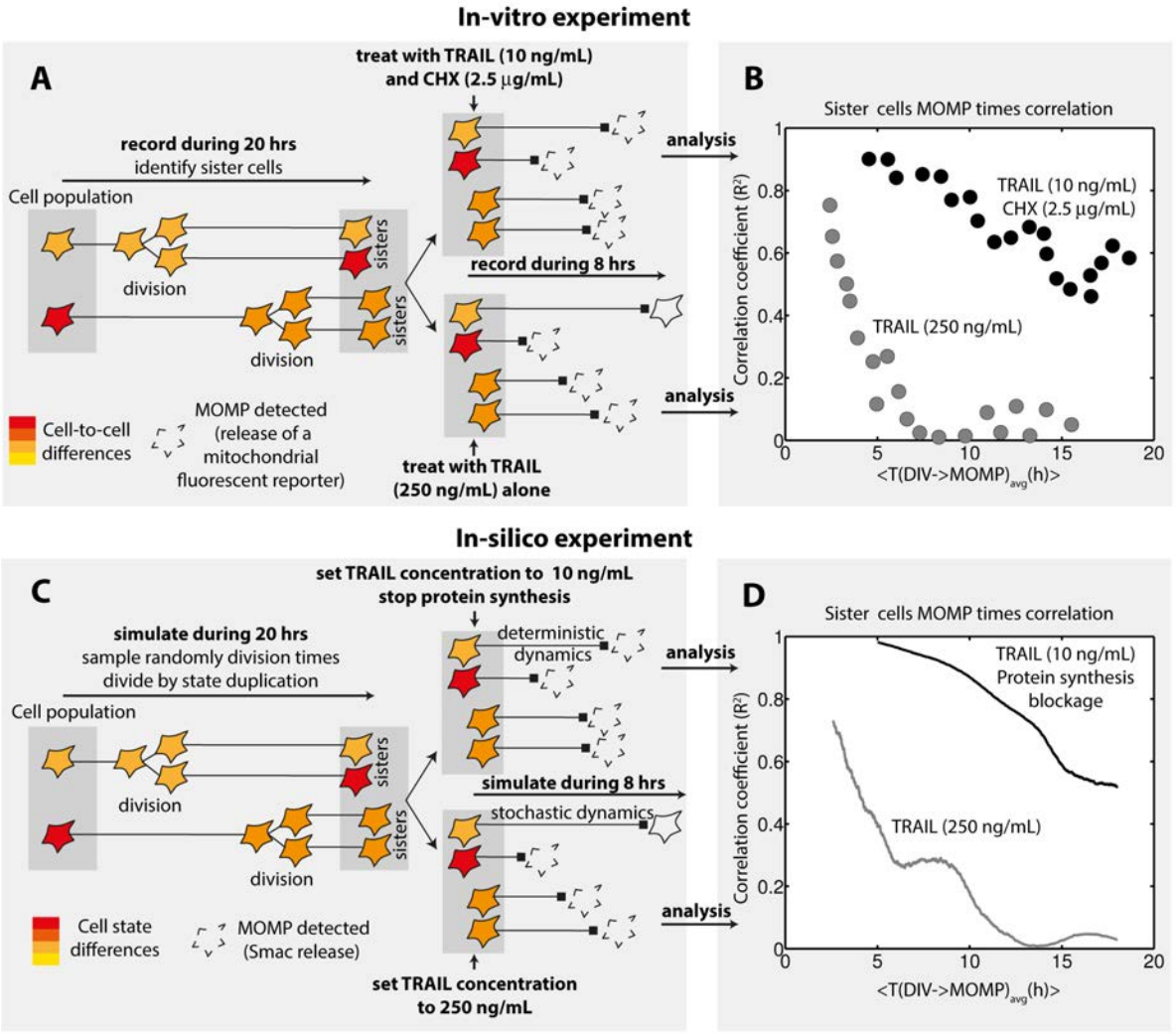



Figure 2.9: **Transient cell fate inheritance in TRAIL-induced apoptosis.** (A-B) Experiments measuring correlation of MOMP times between sister cells performed in (Spencer et al. 2009). (A) HeLa cells were recorded from 20 hours before treatment as in Figure 2.8-A. Sister cells were identified to permit comparison of their fate. (B) Quantification of cell fate inheritance was realized by computing the correlation between sister cells MOMP time as a function of the duration between division and MOMP (averaged between sisters). (C-D) In-silico reproduction of those experiments with the model of Figure 2.8. (C) Description. (D) Quantification of cell fate inheritance was applied to simulation results as in (B). See Appendix 3 for details.

We first asked if the observed cell fate variability could be reproduced. In the model, it is only determined by protein levels at treatment time (behavior is deterministic as synthesis is assumed to be fully blocked by CHX and noise in signaling reactions is neglected), and differences between sister cells are only caused by protein synthesis noise occurring between division and treatment (in agreement with the fact that recently divided sisters died almost synchronously, we assumed an equal repartition of protein content at division).

We found that excellent agreement with observed MOMP time variability can be obtained (Figure 2.8-E). Further analysis revealed that such agreement requires FLIP and Mcl-1 protein half-life to be short and to fall within a narrow range (between 0.3 and 0.6 hours, Figure 2.10-C). This model prediction is consistent with previous measurements in HeLa cells (30 and 40 minutes for FLIP short isoform and Mcl-1 respectively, (Nijhawan et al. 2003 ; Poukkula et al. 2005)). In contrast, FLIP and Mcl-1 mRNA half-life and promoter switching rates are not strongly constrained, probably because their influence on cell fate is limited by the rapid protein level decrease caused by synthesis blockade.

We then asked whether our extended model also capture transient cell fate inheritance (Figure 2.9-B). It is the case: fitted models accurately predict the MOMP time correlation between sister cells (Figure 2.9,D,black curve and Figure 2.10-C). Of note, assuming standard promoter switching rates for FLIP and Mcl-1 (but accounting for their short mRNA and protein half-life - this parameterization will later be referred as the "non-fitted" model) already provides a good agreement for both MOMP time distribution and MOMP time correlation between sister cells (Figure 2.11, upper plots, +CHX data). This non-trivial result shows that the speed at which the sensitivity to TRAIL and CHX fluctuates in single cells is well captured and thus suggests that our generic approach permits to describe fluctuations of protein levels with sufficient accuracy.

## 2.4 Model explains fractional killing and predicts transient cell fate inheritance

Spencer and colleagues repeated this experiment but treated cells with TRAIL alone (250 ng/mL). In this condition, an important fraction of cells died fast (MOMP in ~ 2 hours) but 40% were still alive after 8 hours (Figure 2.8-C), illustrating the fractional killing property. Also, cell fate inheritance between sister cells was markedly changed: only young sister cells that underwent MOMP rapidly were importantly correlated (Figure 2.9-B, grey curve).

We asked whether the observed cell fate variability, including fractional killing, could be reproduced in-silico. Within our modeling assumptions, absence of co-treatment with CHX makes a fundamental difference: as synthesis continues, the effect of gene expression noise during TRAIL-induced apoptosis could be investigated, and comparison with

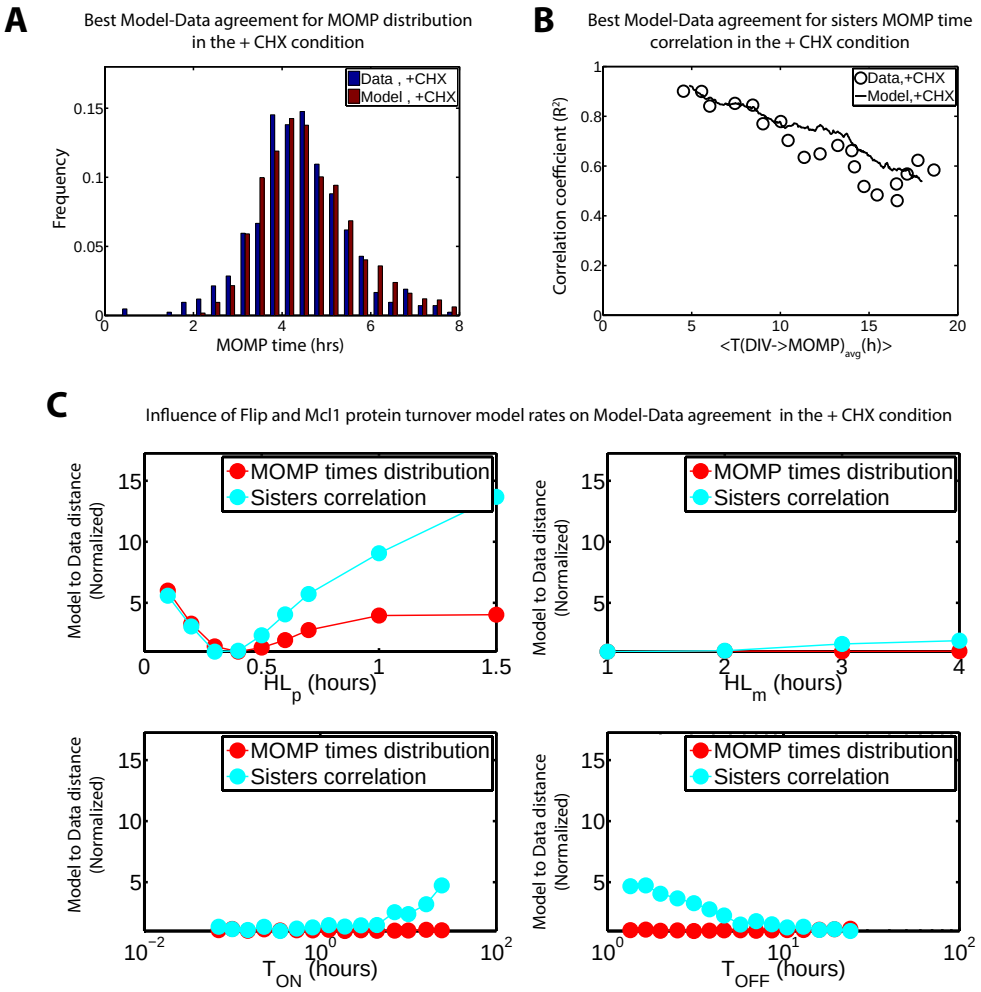



Figure 2.10: **Stochastic protein turnover models capture fluctuations of cell sensitivity to TRAIL and CHX.** (A) Best agreement found between model and data for MOMP times distribution in the +CHX condition. Obtained for FLIP and Mcl-1 model rates such that protein/mRNA half-life and mean ON/OFF promoter activity duration equaled 0.4/1 and 1.9/3.1 hours respectively. See Appendix 3 for the description of how model data agreement was quantified. (B) Best agreement found between model and data for MOMP time correlation between sisters in the +CHX condition. Obtained for FLIP and Mcl-1 model rates such that protein/mRNA half-life and mean ON/OFF promoter activity duration equaled 0.3/1 and 0.35/24 hours respectively. (C) Influence of FLIP and Mcl-1 model rates on Model-Data agreement in the +CHX condition. For each parameter, we plot the model to data distance corresponding to the best model when all other three parameters are varied.

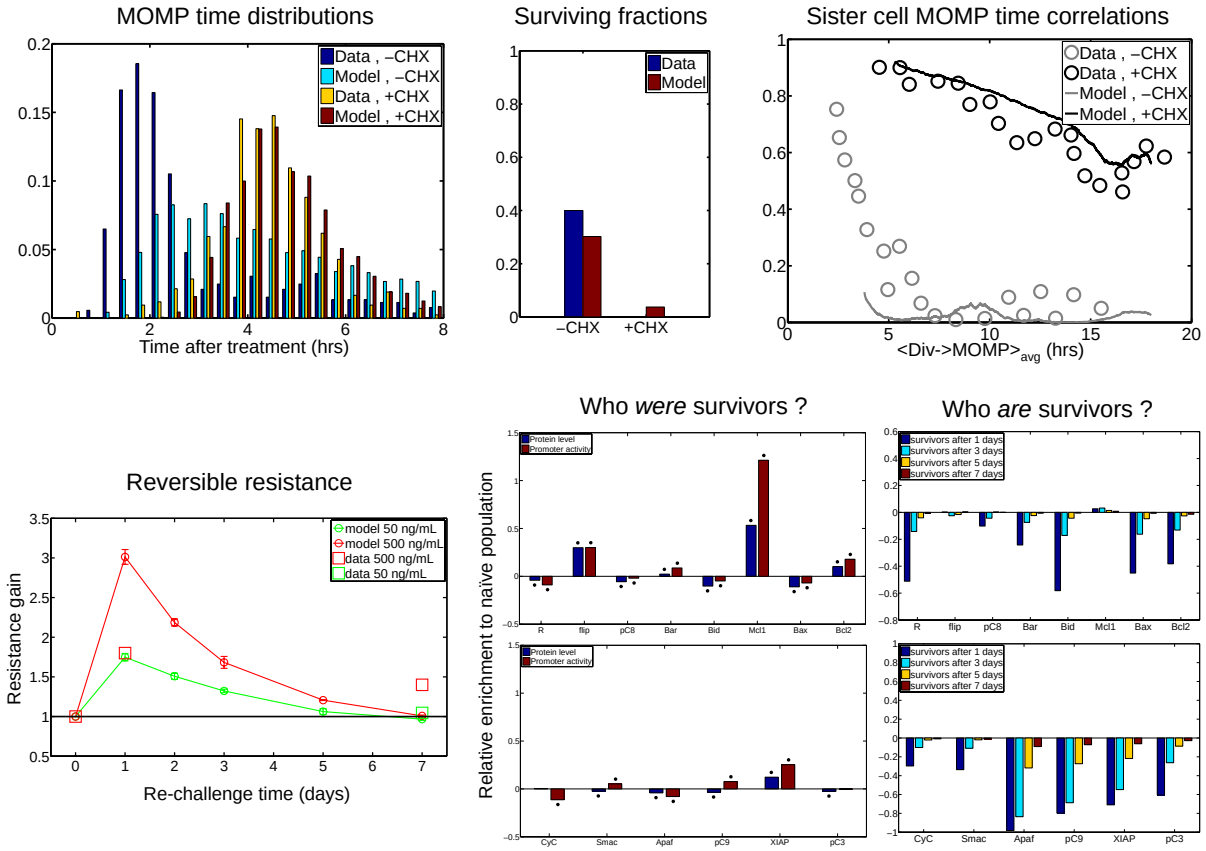



Figure 2.11: **The "non-fitted" model quantitatively predicts TRAIL+CHX single-cell data and lead to fractional killing and reversible resistance for TRAIL alone treatments.** In the non-fitted model, FLIP and Mcl-1 promoter switching rates are standard (Ton=0.1 hours and Toff=2.6 hours) but the short half-life of their mRNA and protein is accounted for (2 hours and 0.5 hours respectively). We reproduce here for this model all the results presented in Figures 2.8, 2.9, 2.16 and 2.17 for the "fitted" model. Quantitative agreement is obtained for TRAIL+CHX single cell data from (Spencer et al. 2009) (MOMP time distribution and sister cell MOMP time correlations). No quantitative agreement is obtained in the case of TRAIL alone treatments, but the existence of fractional killing and reversible resistance is nevertheless predicted. Note that because fewer cells were simulated compared to main text figures ($5.10^4$ instead of $10^5$ for sister cell experiments), sister correlation curves appears slightly noisier.

the TRAIL and CHX condition is insightful. Examples of single-cell trajectories in this setting are shown in Figure 2.14. Strikingly, we found that quantitative agreement for both MOMP time distribution and surviving fraction could be obtained (Figure 2.8-F).

Robustness analysis showed that rates of the FLIP and Mcl-1 stochastic protein turnover model, and particularly promoter switching rates, are in this case strongly constrained. Interestingly, MOMP time distribution and surviving fraction constrain those values differently (Figure 2.12-A,B), resulting in an narrow ranges for their values: agreement for both observations together is obtained only when promoter switching rates are both low (Figure 2.12-C). Such low switching rates lead to large, rare fluctuations of protein levels (Figure 2.12-D). Those atypical fluctuations phenotypes are expected to leave a signature at the population level: the shape of the protein level distribution would be bimodal rather than resembling a lognormal distribution (Figure 2.13). This property is thus a model prediction.

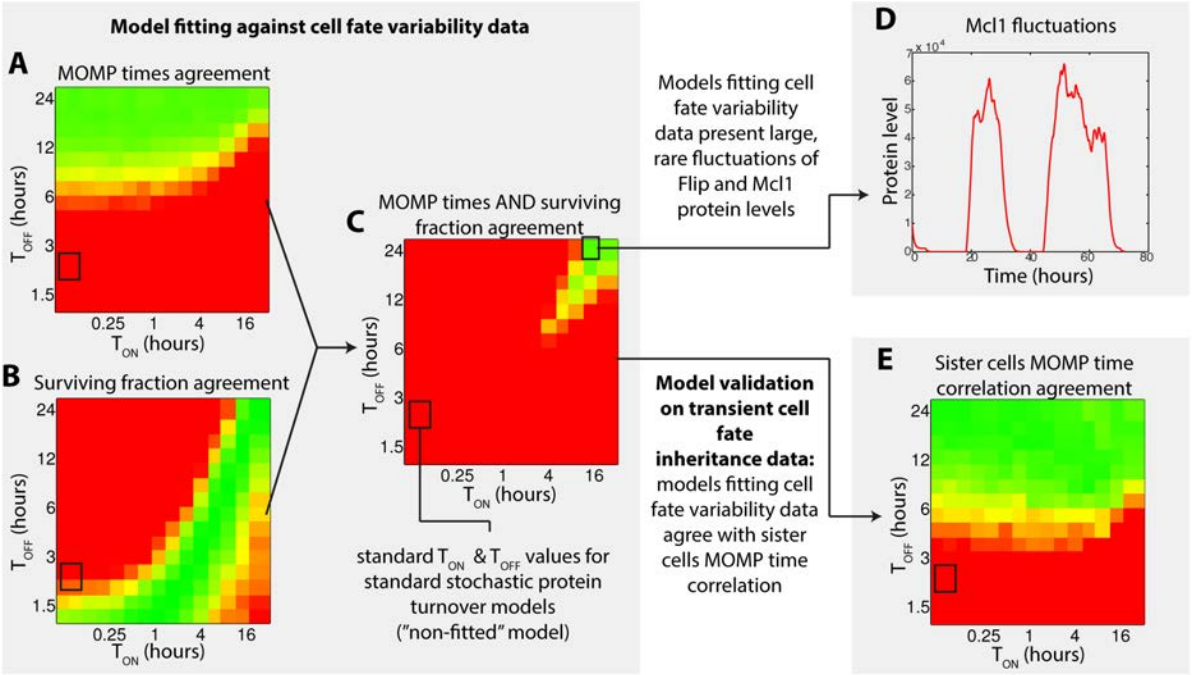


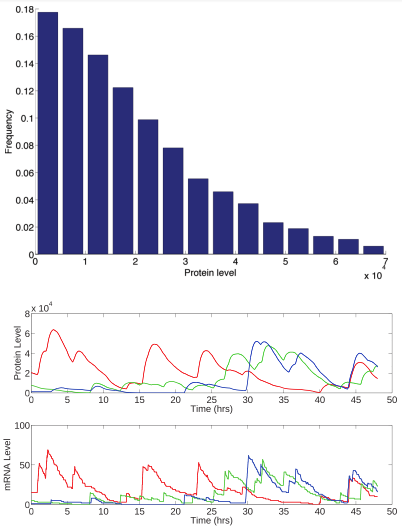

Figure 2.12: **Model fitting to cell fate variability data predicts large, rare fluctuations of FLIP / Mcl-1 and transient cell fate inheritance.** (A-C) Agreement between model prediction and experimental data for (A) death (i.e. MOMP) time distribution, (B) surviving fraction after 8 hours, and (C) both together, for treatment by TRAIL alone (250 ng/mL), as a function of FLIP / Mcl-1 promoter switching times (other parameters as in Table S1). (D) Representative protein level fluctuations of Mcl-1 described by a stochastic protein turnover model allowing good agreement for both MOMP time distribution and surviving fraction. This model has been used for Figures 2.8, 2.9, 2.16 and 2.17. (E) Model-data agreement for MOMP time correlation between sister cells. For (A), (B), (C) and (E), agreement quality increases from red to green. The quantification algorithm is detailed in Appendix 3.

Those fluctuations are likely to impact how the fate of sister cells diverge with time. Thus, we asked whether the model could also account for the observed fast loss of cell fate inheritance. Remarkably, the fitted models accurately and robustly predict MOMP time correlations between sister cells (Figures 2.9-D and ref{model_fitting_no_chx}-E). As mentioned earlier, the same couple of promoter switching rates was used for FLIP and Mcl-1 during

**Mcl1 (and Flip) fluctuations for the "non-fitted" model**

Standard promoter switching rates

| | Mcl1 | Flip |
|---|---|---|
| protein half-life (hrs) | 0.5 | |
| mRNA half-life (hrs) | 2 | |
| Ton (hrs) | 0.1 | |
| Toff (hrs) | 2.6 | |
| protein distribution CV | 0.85 | |
| Protein distribution modality | unimodal | |




**Mcl1 (and Flip) fluctuations for the "fitted" model**

Best agreement found with Spencer et al. cell fate variability data

| | Mcl1 | Flip |
|---|---|---|
| protein half-life (hrs) | 0.4 | |
| mRNA half-life (hrs) | 1 | |
| Ton (hrs) | 16 | |
| Toff (hrs) | 24 | |
| protein distribution CV | 1.15 | |
| Protein distribution modality | bimodal | |

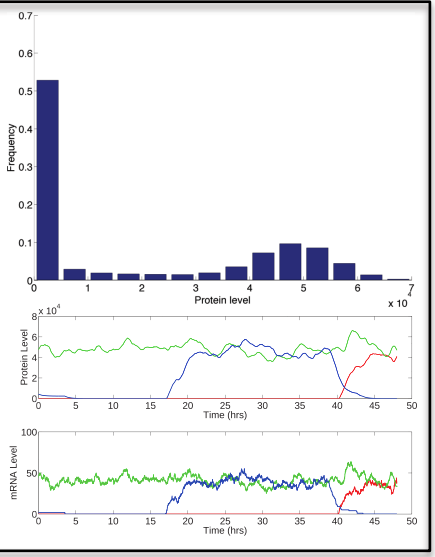



Figure 2.13: **Mcl-1 and FLIP fluctuations for standard or "fitted" promoter switching rates.** For the top frame, promoter switching rates are standard (as in Fig. 2). Because mRNA and protein half-lives are short, protein level fluctuates more rapidly and the steady-state distribution is changed (it is wider and the mode is in 0) compared to the standard stochastic protein turnover model (Fig. 2). On the bottom frame, the steady-state distribution becomes bimodal because the promoter switching rates are low compared to mRNA and protein degradation. In both cases, fluctuations and distribution are shown for Mcl-1; they are similar for FLIP as only the protein synthesis rate changes to account for a different mean protein level.

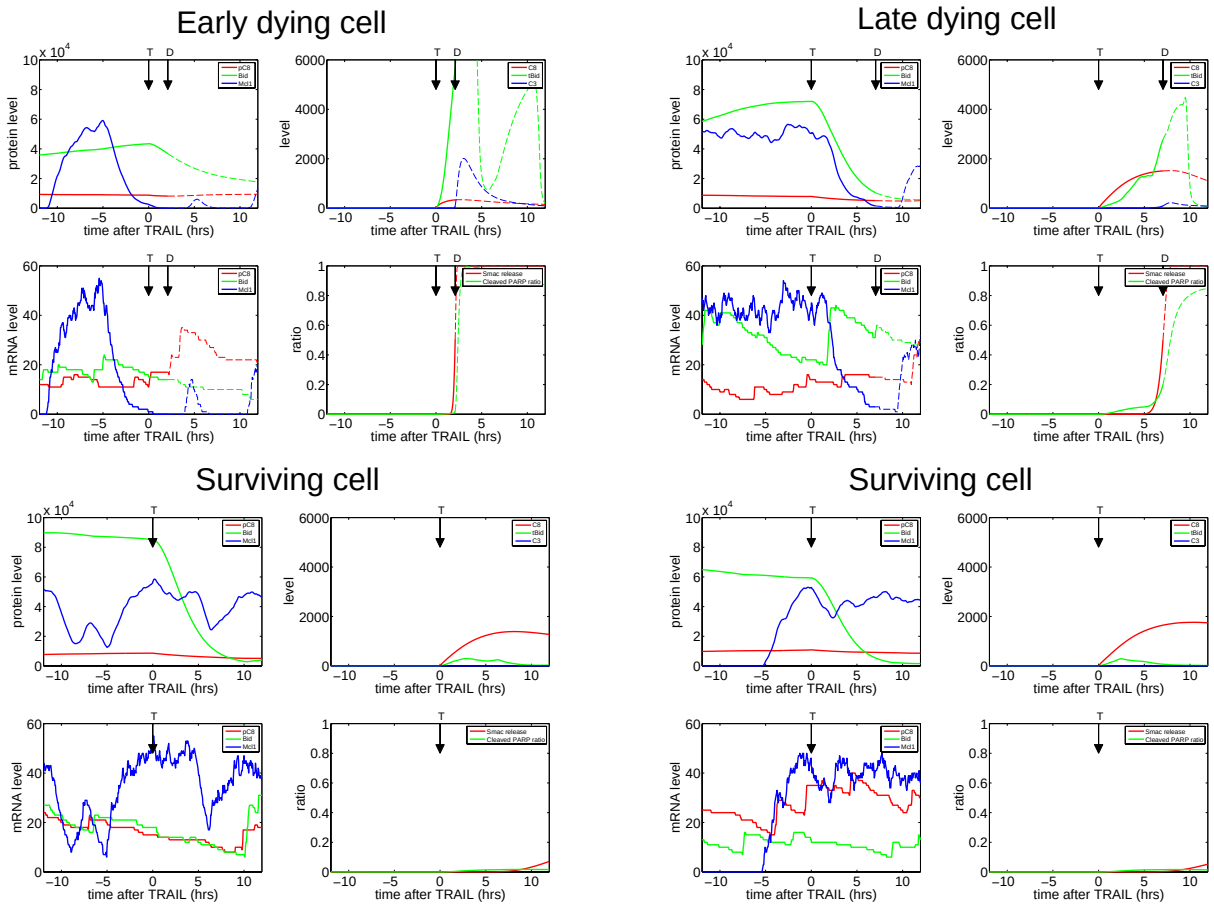



Figure 2.14: **Representative single-cell trajectories before and after TRAIL treatment for the "fitted" model.** Trajectories for two dying and two surviving cells (after 12 hours of TRAIL treatment) are shown. T-marked arrows denote the time of TRAIL addition (250 ng/mL), D-marked arrows denote the time of death commitment (MOMP). mRNA (lower left of each panel) and native form protein levels (upper left of each panel) are shown for pro-caspase 8, Bid and Mcl-1. Levels of activated caspase 8, truncated Bid and activated caspase 3 are also shown (upper right of each panel), as well as the ratio of released Smac and of cleaved PARP (lower right of each panel).

exploration, but further analysis showed that assuming low promoter switching rates for Mcl-1 alone was sufficient to obtain quantitative agreement for MOMP distributions surviving fractions, and that sister cells MOMP time correlations were still correctly predicted (Figure 2.15). Thus, comparison with transient cell fate inheritance data supports that large, rare fluctuations of Mcl-1 could be responsible for the observed cell fate variability.

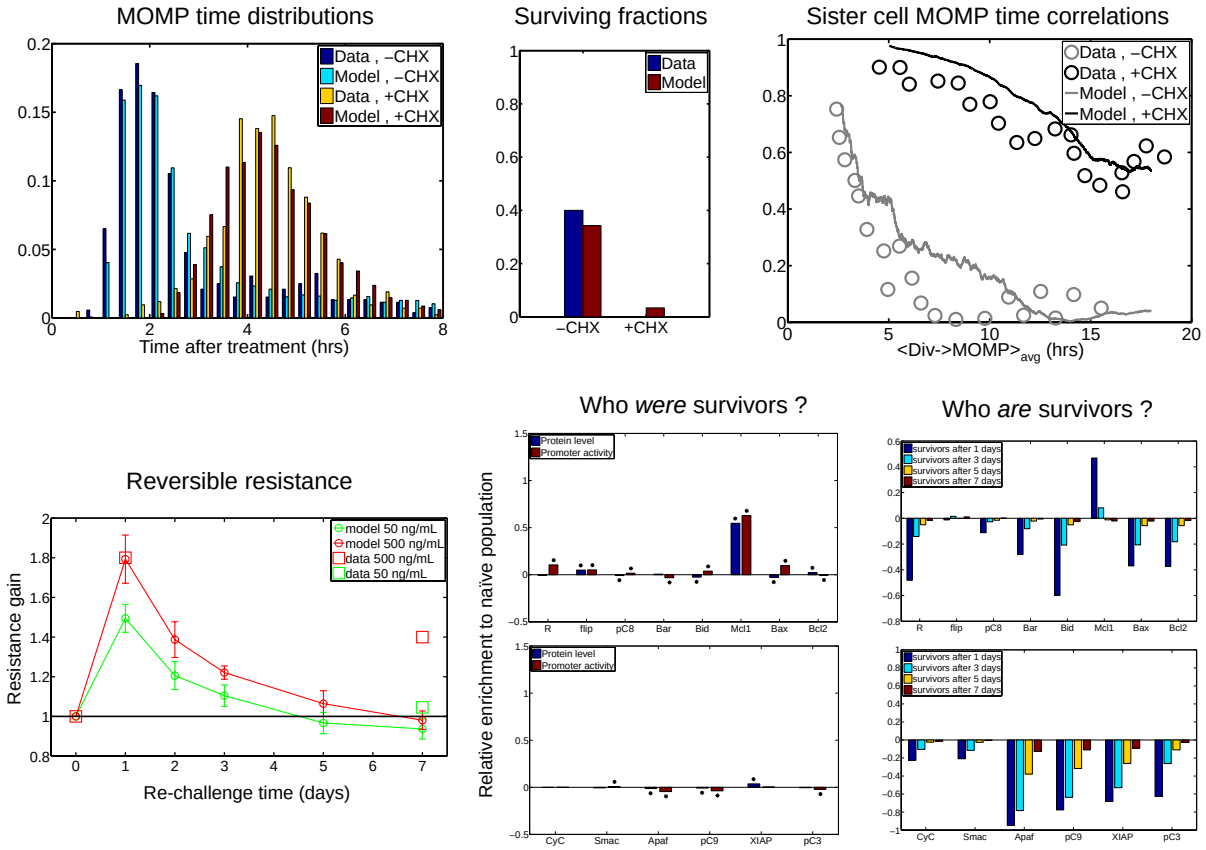



Figure 2.15: **Large, rare fluctuations of Mcl-1 alone are sufficient to explain cell fate variability and transient inheritance in both conditions.** While FLIP and Mcl-1 protein and mRNA half-lives were the same as for the "fitted" model (0.4 and 1.0 hours respectively), only the Mcl-1 promoter was assumed to have low switching rates (Ton and Toff are 16 and 24 hours resp.). The switching rates of the FLIP promoter were assumed to be standard (Ton=0.1 hours and Toff=2.6 hours). All the results presented in the main text for the "fitted" model are reproduced here. Note that because fewer cells were simulated compared to main text figures ($5.10^4$ instead of $10^5$ for sister cell experiments), sister correlation curves appears slightly noisier.

## 2.5 Model predicts reversible resistance

Recently, reversible resistance was observed among various cell lines (Flusberg et al. 2013). Cell populations were submitted to two consecutive TRAIL treatments. The duration between treatments was varied from 1 day to 1 week (Figure 2.16-A). One-day survivors were significantly more resistant than the initial population, but such resistance was significantly decreased or even lost in one-week survivors. Thus, cells surviving a first TRAIL treatment are transiently resistant.

Remarkably, in-silico reproduction of those (Figure 2.16-B) showed that our model predicts the presence of reversible resistance (Figure 2.16-C): one-day survivors exhibit a dose-dependent increase of resistance to a second TRAIL treatment, which disappears after 3 to 5 days. This is surprising since our model does not include induced regulation mediated by survival pathways. Moreover, the presence of reversible resistance is a robust property of the model as it is also obtained when assuming standard promoter switching rates for Mcl-1 and FLIP ("non-fitted" model, Figure 2.11).

However, agreement with related experimental data (Flusberg et al. 2013 ; Flusberg & Sorger 2013) is only qualitative. While we cannot exclude that model parameterizations allowing a quantitative agreement exist, it might be needed to include additional mechanisms such as survival pathways induction to explain the observed sustained resistance gain after one week when treating cells with a high TRAIL dose (Flusberg et al. 2013). It should also be noted that the two experimental measurements of resistance gain in one week survivors were obtained with different methods. It would thus be interesting to confirm the sustained resistance observation for high TRAIL dose in the video-microscopy experimental setting, which allows a more direct and precise measurement of changes in the overall resistance of populations (provided that a sufficient number of cells are analyzed).

## 2.6 Molecular determinants of fractional killing and reversible resistance

What are the mechanisms behind cell escape to TRAIL-induced apoptosis, either on the short-term (fractional killing) or the long-term (reversible resistance)? Using the fact that in-silico, all protein, mRNA levels and gene activity states can be monitored in single cells, we investigated those questions at the molecular level.

To study the influence of pre-existing differences on cell fate, we compared at the time of stimulation the sub-population of 'future survivors' with the whole population (Figure 2.17-A,B). Future survivors strongly stood out by their Mcl-1 protein level and gene activity state (Figure 2.17-B). FLIP also appeared to play an important role in determining cell decision, and smaller but significant effect was also seen for Bid, Bax, Bcl2 and XIAP. Although it is a good predictor of cell fate, initial Mcl-1 gene activity status does not completely determine survival: neither all Mcl-1 "ON" cells survived nor all Mcl-1 "OFF" cells died. Thus, pre-existing differences in protein levels and promoter activities are major determinant of cell fate but stochastic events in gene expression occurring during signal transduction also play a role.

While timing of death for cells treated with TRAIL and CHX appeared to be multi-factorial (Spencer et al. 2009), our results suggest that cell survival is predominantly

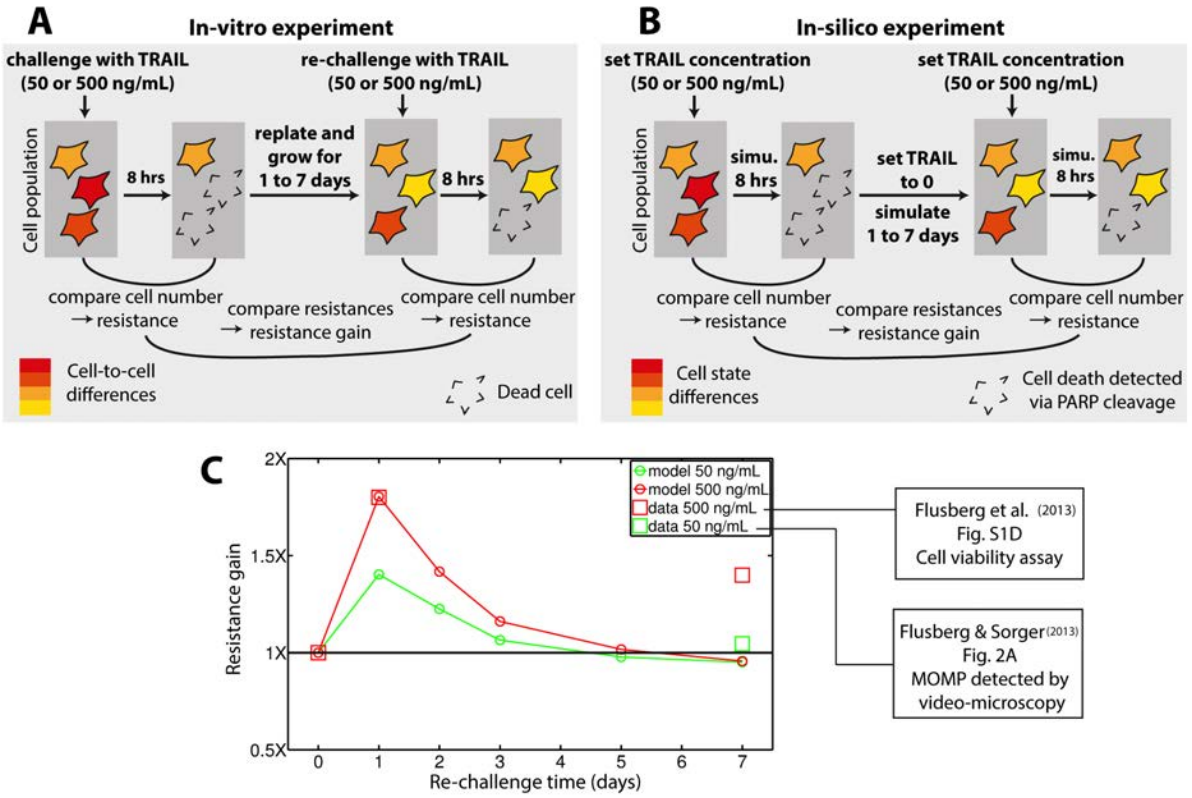



Figure 2.16: **Reversible resistance in repeated TRAIL treatments.** (A) Schematic description of the 'repeated TRAIL' experiments performed in [14,15] to characterize reversible resistance in HeLa cells. (B) In-silico reproduction of these experiments with our model (details in Appendix 3). (C) Resistance gains in surviving cells relative to naïve cells as a function of time between the two TRAIL treatments. Data are shown for experimental observations (Flusberg et al. 2013; Flusberg & Sorger 2013) and model predictions (our study). Comparison of resistance gain experimental measurements between the two TRAIL doses (500 and 50 ng/mL for (Flusberg et al. 2013) and (Flusberg & Sorger 2013) respectively) should be done with care, as the measurement and quantification method differed.

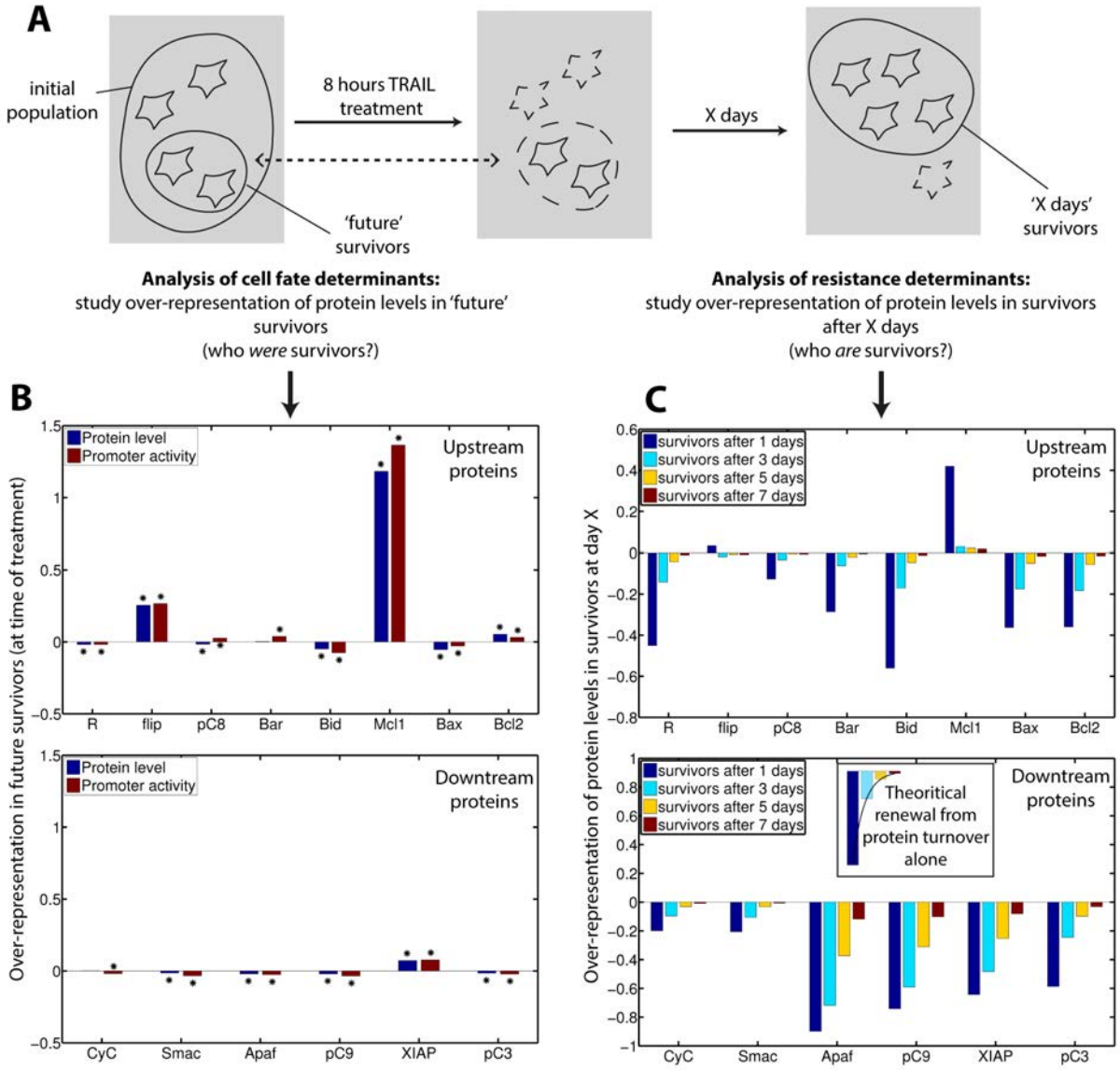



Figure 2.17: **Molecular determinants of cell fate and resistance to repeated TRAIL treatments.** (A) Cartoon illustrating that the determinants of cell fate and resistance can be studied by analyzing the over-representation of protein levels in 'future survivors' (cells that will still be alive after treatment) at the time of treatment, and in surviving cells at day X, respectively. (B) Cell fate determinants analysis: over-representation (compared to initial population) of protein level (blue) and promoter activity (red) at the time of treatment in 'future survivors'. Asterisks mark differences that passed a 5% significance test. (C) Resistance determinants analysis: over-representation of protein levels in surviving cells at day X. Inset illustrates the recovery kinetics expected from protein turnover only (i.e. in absence of significant selection effect or residual signaling activity). Therefore, deviation from such kinetics indicates the presence of a selection effect or residual signaling activity.

determined by Mcl-1 (Figure 2.17-B). This important role of Mcl-1 is robustly predicted. Indeed, it also holds for the "non-fitted" model, which assume standard promoter switching rates for all proteins, including Mcl-1 and FLIP (Figure 2.11).

To investigate the determinants of reversible resistance, we tracked the temporal evolution of protein levels in surviving cells (Figure 2.17-A,C). The protein level composition of one day survivors contrasts with the protein content observed in future survivors: almost all protein levels differ importantly from the naïve population composition, while that was the case only for Mcl-1 and FLIP in future survivors. This is expected as all proteins are partly activated during signal transduction, leading to a higher degradation (active forms have a shorter half-life).

Therefore, the distinction between the causes of cell survival and the consequences of partial apoptosis induction cannot be easily resolved by the sole observation of protein levels in survivors. When signaling stops, recovery of protein levels is expected to follow exponential kinetics governed by the turnover rate (Figure 2.17-C, inset – see the death receptor (R), pro-caspase 8, Bar and Bid). Deviation from such kinetics indicates either the persistence of signaling reactions that continue to consume proteins (it is the case for Apaf, pro-caspase 9 and XIAP, as further analysis confirmed) or is a consequence of important selection.

Indeed, while Mcl-1 and FLIP should recover normal levels in a few hours in absence of selection because of their high turnover rate, Mcl-1 levels (but not FLIP levels) are still strongly higher than in naive cells one day after TRAIL treatment, consistently with previous observations on the relative selection strength that operated on them. Together, those results indicate that recovery phenotypes in surviving cells result from a complex interplay of three distinct effects: selection during apoptosis, transcriptional noise and protein turnover as a driving force tending to reset protein levels to their initial, pre-stimulus distribution, and long-term residual signaling activity. This explains why it is difficult to understand the recovery process and justifies the use of modeling to disentangle the various contributions.

## 2.7 Discussion

Rehm et al. and Spencer et al. (Rehm et al. 2009 ; Spencer et al. 2009) made two insightful observations about TRAIL-induced apoptosis. First, recently born sister cells died almost synchronously when treated with TRAIL and a protein synthesis inhibitor, while in contrast, unrelated cells died after highly variable durations. This demonstrated that TRAIL signaling is mostly deterministic when protein synthesis is blocked and that the timing of death is determined by the cell internal state at the time of treatment.

Second, they observed that such synchrony in sister cells death is gradually lost as the time between division and treatment increases. This showed that the cell 'TRAIL sensitivity state' (the part of cell internal state involved in death timing determination) naturally

fluctuates over a dozen of hours. In addition, the modeling results in (Spencer et al. 2009) highly suggested that such state is mainly composed by the various levels of the proteins acting in the extrinsic apoptosis pathway.

In parallel, important progress on the characterization of the stochasticity in gene expression has been made: the two-state transcriptional bursting model was shown to permit high accuracy and several approaches to infer its parameters were proposed, enabling the quantitative modeling of protein fluctuations in single cells (Singh et al. 2012; Raj et al. 2006; Dar et al. 2012; Viñuelas et al. 2013; Paszek 2007; Munsky et al. 2009).

## Modeling protein fluctuations in TRAIL-induced apoptosis

In this Chapter, we presented how we merged those two approaches by integrating such stochastic models of gene expression within an existing kinetic model of TRAIL-induced apoptosis (Albeck, Burke, Spencer, et al. 2008) in a systematic and principled manner. Doing so provides advantages compared to previous approaches to account for cell-to-cell variability in protein levels (Spencer et al. 2009; Gaudet et al. 2012; Kallenberger et al. 2014). First, variability is not considered as an "input" parameter but arises naturally from stochastic fluctuations. The dynamics of this variability is thus intrinsically represented within the system, allowing investigating the effects of transient memory in protein levels. Second, the influence of protein synthesis noise during TRAIL-induced apoptosis could also be investigated. Importantly, we followed a parsimonious parameterization strategy, motivated by the fact that fluctuations of long-lived proteins are rather insensitive to the precise kinetics of transcriptional bursting, enabling us to equip most proteins (long-lived proteins) with reasonably accurate fluctuation models even in absence of gene expression data for each and every promoter.

The sister cells experiment for which cells were treated with TRAIL and CHX provided ideal data to validate our modeling approach: in that case, behavior is mostly deterministic as soon as treatment starts and only fluctuations occurring before treatment are responsible for death time variability and de-correlation between sister cells. Moreover, gene regulation via survival pathways induction is ineffective as protein synthesis is blocked.

Because our model was able to quantitatively reproduce the MOMP time distribution and then accurately predicted sister cells correlation, our modeling approach appears as a promising tool to investigate the effect of protein fluctuations on signal transduction, despite the limitations inherent to its simplicity (for example, stochastic gene expression events were assumed to be independent between proteins, neglecting the fact that levels of different proteins can be partially correlated (Gaudet et al. 2012; Sigal et al. 2006), possibly because of common transcription factors or coordinated chromatin-state transitions). Of note, good agreement was readily obtained when assuming for FLIP and Mcl-1 standard pro-

moter switching rates (but short protein and mRNA half-lives, in agreement with available knowledge - "non-fitted" model, Figure 2.11).

Finally, the transposition of fluctuation timescales from individual proteins into 'TRAIL sensitivity states' is not trivial: while in our model, stable proteins levels are mixed in about 40 hours, cells were switching between 'fast dying' and 'slow dying' phenotypes more rapidly (about 10-15 hours). As combinatorial and non-linear effects are at play, mechanistic models of protein-proteins reactions are needed to link protein-level timescales with more high-level phenotypic transitions (Gupta et al. 2011).

## Questioning the role of survival pathways

Several studies reported that TRAIL can induce survival pathways (Chaudhary et al. 1997; Son et al. 2010; Sun et al. 2011). How such induced changes affect signal transduction and eventually stop apoptotic signaling remains unclear. On the other hand, the contribution of constitutive protein synthesis noise, which is responsible for pre-existing differences between cells, has not been evaluated.

Although it does not exclude the existence of other mechanisms, an important result of our study is that fractional killing can be obtained without assuming any TRAIL-induced regulation. Alternatively, we find that because of its fast turnover, constitutive expression of the Mcl-1 protein has the potential to rescue cells from TRAIL apoptotic signaling. In this context, solely accounting for protein fluctuations within the TRAIL apoptosis pathway predicts the fractional killing property (Figures 2.8 and 2.11).

While our results challenge current opinion on the role of survival pathways in TRAIL-induced apoptosis, they are consistent with observations made on wild type HeLa cells that neither blocking NF-κB response nor inhibiting the Akt pathway do significantly change the surviving cell fraction after TRAIL treatment (Braeuer et al. 2006; Lalaoui et al. 2011). The pivotal role for Mcl-1 in TRAIL-induced apoptosis predicted by our model is consistent with the recent finding that Mcl-1 silencing by shRNA in HeLa cells completely sensitize cells to TRAIL (Lemke, Karstedt, Abd El Hay, et al. 2014).

While moderate fluctuations of Mcl-1 levels were sufficient to obtain fractional killing, a quantitative agreement with the Spencer et al. (2009) single-cell data (MOMP time distribution and surviving fraction) required large and rare Mcl-1 fluctuations, caused by rare switches between long periods of gene activity or inactivity. Interestingly, in that case, the observed rapid loss of MOMP time correlation between sister cells quantitatively emerged from model simulations. However, this result was obtained for a single TRAIL dose. Given the predicted importance of Mcl-1 in determining cell fate, it would be valuable to investigate model ability to reproduce surviving fraction, MOMP time distribution and sister cell fate correlation for lower and higher TRAIL doses.

FLIP is often mentioned as a key factor in cell resistance to TRAIL (Shirley & Micheau 2013), but in our model FLIP has less impact on cell survival than Mcl-1. Consistently, Lemke et al. (2014) silencing experiments demonstrated a dominant role for Mcl-1 and a synergy with FLIP. However, our model might under-estimate the role of FLIP: the representation of DISC-related events in EARM is simple and thus does not account for recent biological findings, including the stoichiometry between its components (Dickens et al. 2012; Schleich et al. 2012). Improving how DISC assembly is modeled might thus be needed to elucidate the precise role of FLIP in fractional killing and reversible resistance, especially for cell lines that express higher FLIP amounts than HeLa.

## Origins of reversible resistance: joint effect of selection and stochastic protein turnover

A second significant result reported here is that our model predicts the phenomenon of reversible resistance, showing that constitutively noisy protein synthesis, protein-protein interactions and protein degradation are by themselves sufficient to explain a dose-dependent, significant increase of resistance in recent survivors and its gradual loss within 3-5 days.

This result is consistent with the observation that NF-κB blockade does not change resistance acquisition after TRAIL treatment (Flusberg et al. 2013) (in MCF10A cells; HeLa cells have not been tested). In-silico analysis at the molecular level revealed that reversible resistance as predicted by the model was shaped by a complex interplay between 1) selection based on protein levels and transcriptional activity, 2) protein turnover and 3) residual signaling activity.

As opposed to the death process, which involves a sharp and complete activation of effector caspases, our results suggest that recovery in cells that did not commit to death is a slow and complex process. While one should not conclude from our results that parallel activation of survival pathways by TRAIL plays no role in reversible resistance, our results show that the sole contribution of protein level fluctuations occurring within the extrinsic apoptosis pathway can partly lead to reversible resistance. Thus, protein fluctuations should be accounted for to gain quantitative insights into reversible resistance.

## Impact of caspase feedback loop and non-native forms degradation on model behavior

As previously justified in the section presenting model construction, we assumed in all our simulations 1) absence of the caspase feedback loop (i.e., the ability for caspase 3 to cleave pro-caspase 6, which can then activate pro-caspase 8) and 2) non-native forms for which a degradation rate could not be specifically attributed (all non-native forms except TRAIL,

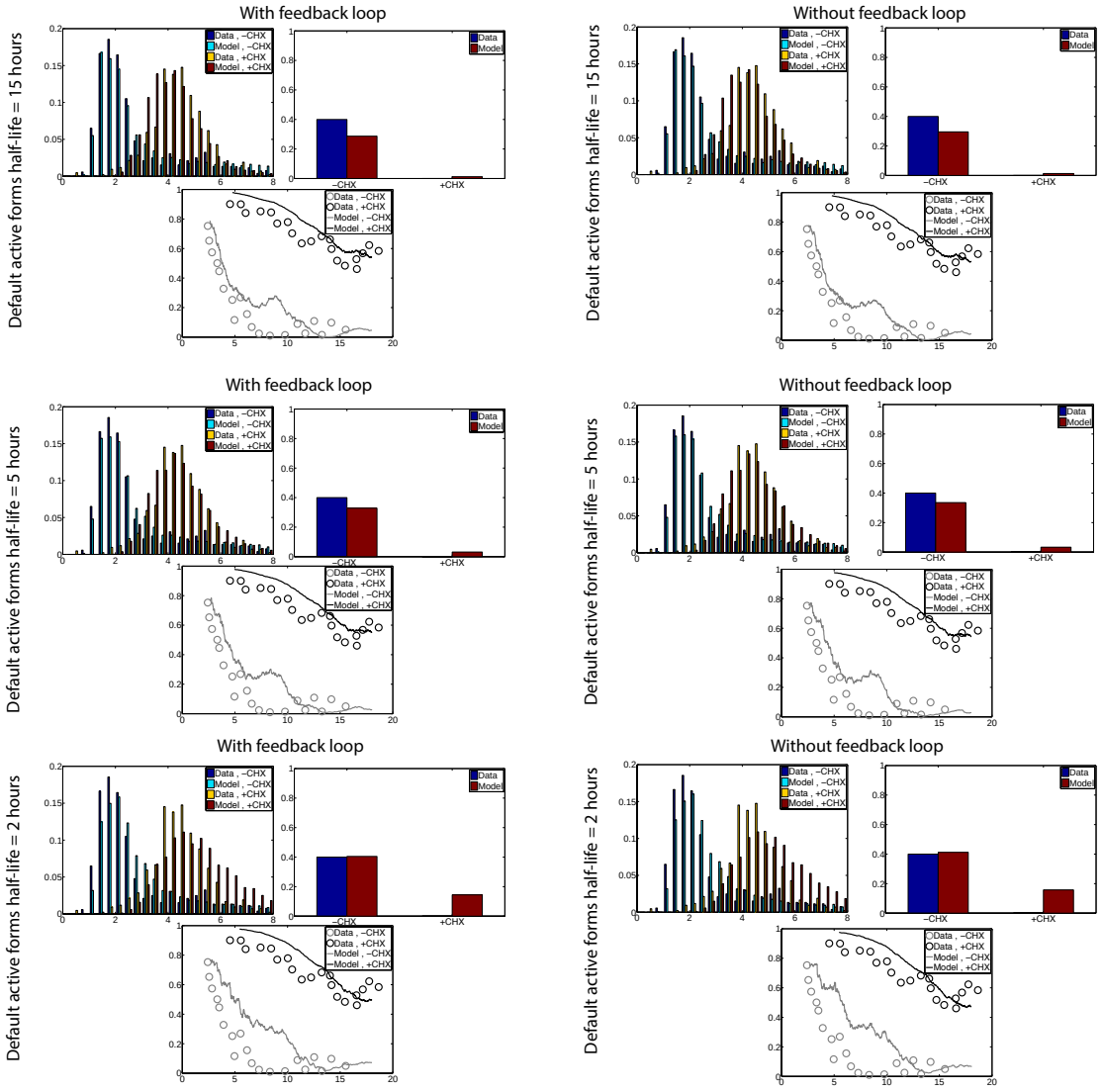



Figure 2.18: **Impact on short-term model behavior of the presence/absence of caspase feedback loop and the degradation of non-native ('active') forms.** Model-Data agreement is shown for MOMP time distributions, surviving fractions and sisters correlation of MOMP time in both treatment conditions for model variants when the caspase feedback loop is either present/absent and the default active forms half-life is 15, 5 or 2 hours.

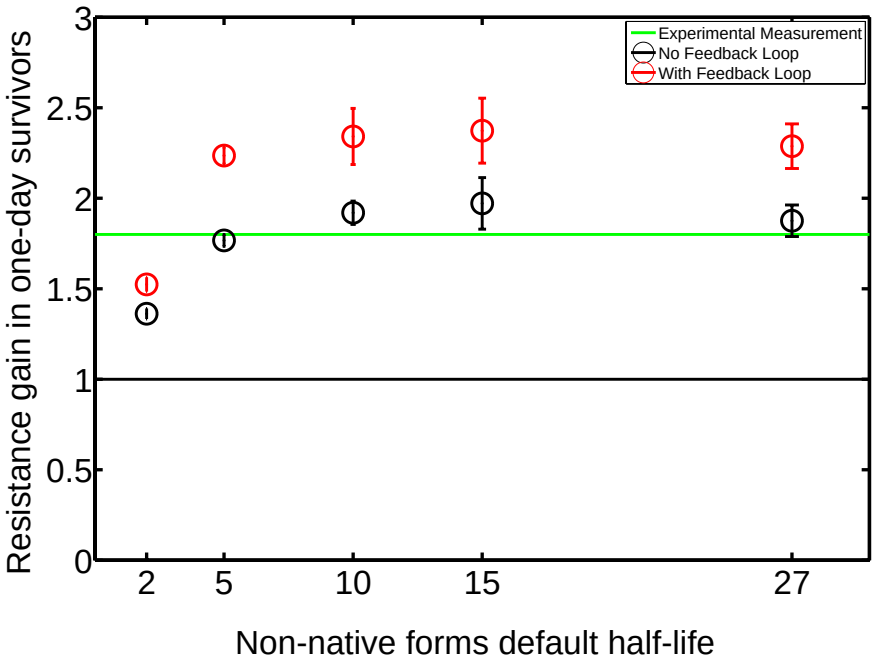



Figure 2.19: **Impact on resistance gain in one-day survivors of the presence/absence of caspase feedback loop and the degradation of non-native ('active') forms.** In-silico repeated TRAIL experiment (as in Figure 2.16) was repeated for variants of the "fitted" model regarding presence/absence of the caspase feedback loop and the default active forms half-life. Resistance gain in one-day survivors is shown. Simulations were repeated 4 times with $10^4$ cells, error bars indicate standard deviation of estimated resistance gain between replicates.

the species representing mitochondrial pores and the complexes involving FLIP and Mcl-1) are degraded with a 5 hours half-life.

We investigated the impact of those choices on our main results by repeating the corresponding in-silico experiments for different values for the non-native forms half-life and status of the caspase feedback loop. The results of the short-term, single-cell experiments are almost un-affected when the non-native forms are longer-lived (15 hours) and similar when they are shorter-lived (2 hours) (Figure 2.18). Here, the presence or absence of the caspase feedback loop had no effects.

Regarding model predictions for transient resistance acquisition, the quantified resistance gain in one-day survivors is not strongly affected. Still, the presence of the caspase feedback loop systematically increases it, while it decreases when non-native forms half-life is below 5 hours (Figure 2.19). Note that here resistance acquisition was quantified by comparing cell numbers at treatment time and 8 hours after (to follow Flusberg and colleagues) and hence does not necessarily reflect the total killing efficacy of a TRAIL treatment. In fact, we found the long-term evolution of population size for cells treated with TRAIL to strongly depend on non-native forms degradation. The source of this dependence and its potential consequences for the outcome of repeated TRAIL treatments are investigated in details in Chapter 3.

## Accounting for gene expression noise appears necessary to investigate signal transduction

While here we focused on TRAIL-induced apoptosis, our modeling approach is generic and can be applied to other signal transduction pathways. Our results showed that even in absence of induced gene regulation, gene expression noise interacts with signaling dynamics on a non-trivial manner. Thus, even in contexts where the influence of induced gene-regulation is indisputable, its sound quantification probably requires to investigate first the role of constitutive gene expression noise.

Only then models could be enriched parsimoniously with well-characterized regulatory links until all observations are successfully explained. Significant advances to allow such detailed characterization of gene regulation occurred recently (Tay et al. 2010; Molina et al. 2013; Neuert et al. 2013). Following such approaches could significantly extend the reach of models of signal transduction towards accurate, single-cell level description of populations submitted to varying signaling contexts. On the methodological side, we illustrate in Chapter 4 how to extend our modeling approach to account for regulated stochastic gene expression and demonstrate that it allows to capture induction dynamics of an artificial yeast signal transduction pathway with single-cell resolution.

# Chapter 3

# A multi-scale model for investigating TRAIL resistance in multi-cellular tumor spheroids

## 3.1 Motivations: towards a multi-scale understanding of TRAIL resistance

In the previous Chapter, we developped a single-cell model of TRAIL-induced apoptosis that accounts for 1) protein-protein signaling reactions linking TRAIL exposure to commitment to apoptosis, 2) stochastic gene expression for the proteins involved in this signaling and 3) protein degradation. An important result that was obtained is that under parsimonious and realistic assumptions for parameter values, *fractional killing* and *transient resistance acquisition* readily emerged from model simulations. Those two properties relating to TRAIL resistance are observed *in-vitro* for many different cancer cell lines.

A priori, such resistance can have implications for the therapeutical usage of TRAIL-like drugs (i.e. death receptor agonists, abbreviated DRAs), notably regarding the choice of the DRA (potency and pharmaco-kinetic properties) and the treatment strategy (dosing, scheduling and co-drugging for decreasing resistance). Also, in addition to the mechanisms of TRAIL resistance implicated in *in-vitro* experiments, other factors could impact the response of *in-vivo* tumors to TRAIL treatments, such as limited and heterogenous drug exposure, context-dependent changes in cell state affecting TRAIL sensitivity, influence of other cells such as endothelial or immune cells, etc. (Pampaloni et al. 2007; Yamada & Cukierman 2007).

Gaining a quantitative understanding of how those different factors shape TRAIL resistance *in-vivo* would permit to rationally design treatment strategies to maximize their efficacy. However, despite numerous *in-vivo* studies on DRA treatments (Walczak et al. 1999;

Ashkenazi et al. 1999; Chinnaiyan et al. 2000; Fulda et al. 2002; Merchant et al. 2004; DeRosier et al. 2007; Lee et al. 2010; Huet et al. 2014), our understanding of TRAIL resistance in such settings and how it relates to the *in-vitro* resistance of the same cell lines remains limited. Indeed, those studies mostly focused on demonstrating qualitatively the anti-tumor activity of a new DRA or a new combination of a DRA with another agent.

Mathematical modeling and numerical simulations are potentially very helpful to build such quantitative understanding because it allows to translate previous knowledge and hypotheses on the mechanisms at play into predictions that can be compared with experimental observations. However, no mathematical models enabling the investigation of TRAIL resistance over long time scales and for cell populations exhibiting spatially heterogenous micro-environments have yet been developped. A potential explanation for this is that so far no model able to quantitatively account for the observed dynamics of cell-to-cell variability was available. Moreover, studying spatial and long-term effects requires the integration of quantitative intra-cellular models into quantitative cell-based models, which represents a challenge.

In this Chapter, we integrate our single-cell model of TRAIL-induced apoptosis into multi-cellular simulations to investigate TRAIL resistance in two virtual experimental settings that could bridge the gap between standard *in-vitro* and *in-vivo* settings. The first setting corresponds to the long-term *in-vitro*, monolayer culture (several weeks) of cell populations repeatedly treated by TRAIL. The second setting extends the first but instead cells are grown as 3D multi-cellular spheroids, allowing us to investigate the additional role of spatial heterogeneity in TRAIL exposure. Although spheroids are smaller than *in-vitro* tumors, their size is higher or comparable to the typical distance between two vessels within *in-vivo* tumors (around 100-200 micrometers, (Primeau et al. 2005; Baker et al. 2008)), making it relevant to study heterogeneity in drug spatial distribution. Those two settings are intermediate between standard assays of TRAIL cytotoxicity and pre-clinical *in-vivo* assays in mice. Importantly, they are experimentally relatively easy to perform, and much more amenable than *in-vivo* experiments. Therefore, our modeling predictions are readily testable experimentally.

## 3.2 The role of time: could resistance acquisition protect tumor monolayers from repeated TRAIL treatments?

### Modeling populations of single cells submitted to TRAIL treatments

In this section, we investigate the long-term consequences of fractional killing and transient resistance acquisition as predicted by our model. To allow the simulation of the long-term evolution of cell populations submitted to arbitrary TRAIL stimulations, the single-cell

TRAIL model is coupled with a simple model of cell proliferation (Figure 3.1). This model accounts for cell-to-cell variability in the cell cycle distribution in order to reproduce the loss of synchrony in cell divisions that is observed in cancer cell micro-colonies cultured *in-vitro*.

Cells of the population are assumed to experience the same concentration of TRAIL in their environment. This unique environmental TRAIL concentration can be time-varying and is assumed not to be influenced by cells (Figure 3.1). This situation is meant to represent the long-term culture of cancer cell populations as monolayers.

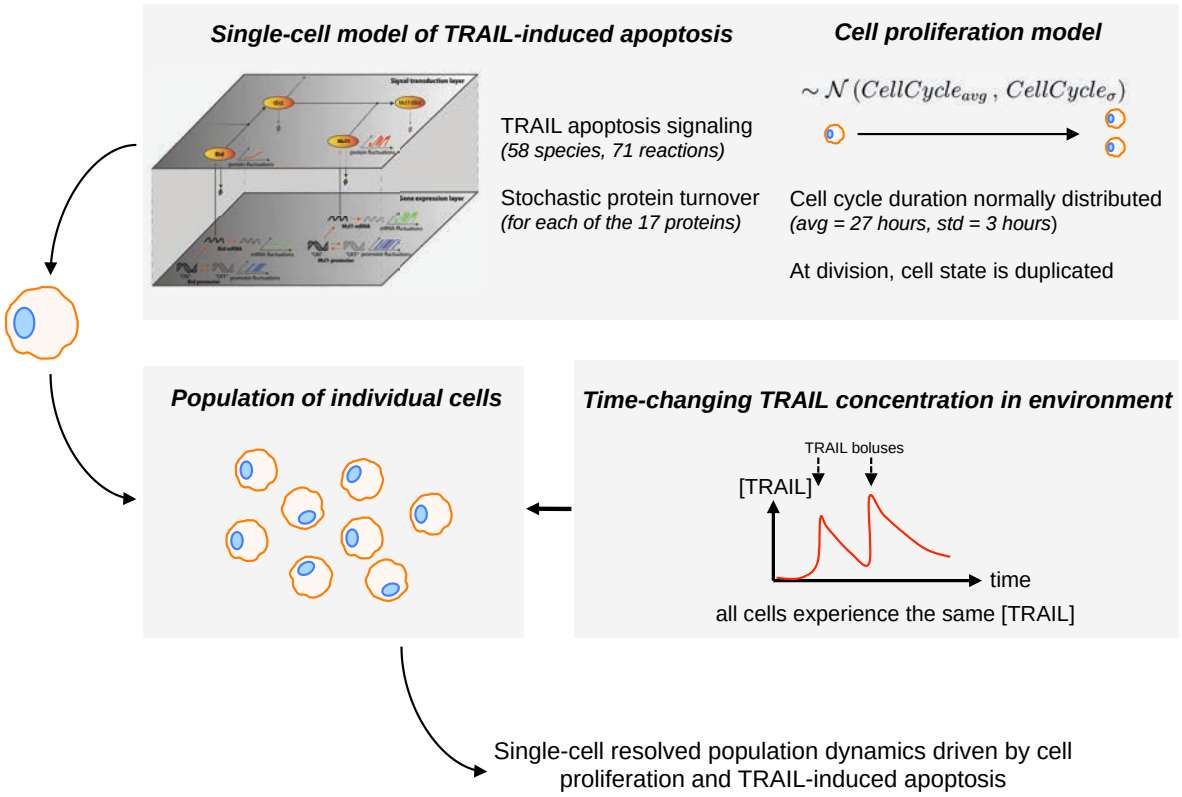



Figure 3.1: **Modeling populations of single cells submitted to TRAIL treatments.** Our single-cell model of TRAIL-induced apoptosis is simulated in each cell of a population. In addition, cell proliferation is modeled by dividing cells after a random, normally distributed cell cycle duration. At division cell state is duplicated in the two daughter cells. Such population can be submitted to arbitrary time-varying TRAIL concentrations (for example generated by bolus administration at discrete times), leading to single-cell resolved population dynamics driven by cell proliferation and TRAIL-induced apoptosis.

## A reference model for HeLa cells

As a basis for our analysis, we consider one of the most parsimonious parameterization of our single-cell model of TRAIL-induced apoptosis, because this parameterization readily predicts the phenomena of fractional killing and transient resistance acquisition. This allow us to investigate the long-term consequences of those phenomena without relying on too specific, yet not validated assumptions (such as the predicted Mcl-1 bimodal distribution and slow transcriptional bursting). Hence, we assumed that Mcl-1 and Flip transcriptional

bursting rates have standard values ($T_{on} = 0.1$ hours, $T_{off} = 2.6$ hours, see Figure 2.13). The non-native forms are assumed to be degraded with a half-life of 3 hours when no specific values can be set. The caspase feedback loop is assumed to be absent. Finally, we brought a minor modification to the previously tested parameterizations: cPARP (which value provides a readout for the integrated amount of caspase-3 activity and is used for determining irreversible commitment to apoptosis) is assumed to be stable (hence 'degraded' only via dilution).

In the previous Chapter, model-data agreement in terms of cell killing was tested for a single TRAIL dose (in absence of CHX co-treatment) and for short-term data (up to 8 hours after treatment). Because we are now interested in the long-term cytotoxic effect of arbitrary TRAIL dose temporal profiles, we first tested the ability of this model parameterization to reproduce existing data that spans a wide range of TRAIL doses and in which killing efficiency is measured after 24 hours of treatment (Lemke, Karstedt, Abd El Hay, et al. 2014). We obtained a reasonable agreement for both wild-type and Mcl-1 shRNA expressing HeLa cells dose-response curves (Figure 3.2). Thus, the predicted killing efficiencies for HeLa cell populations treated with various TRAIL doses for 24 hours resembles experimental observation. Of note, the dose-response curves for Flip shRNA and Flip + Mcl-1 shRNAs qualitatively differed from observations, as expected because the representation of Flip in the model is simplistic (not shown).

## In-silico quantification of initial resistance: reproducing a standard dose-response curve

The standard way to evaluate the potential of drugs as anti-cancer cytotoxic agents is to perform dose-response experiments on cancer cell lines cultured *in-vitro*. We performed such experiment *in-silico* by proceeding as follows. First, we started with small populations of single cells that were let grown without TRAIL until they reached a size of 5000 cells. Then, we applied different doses of TRAIL to those populations and simulated their evolution for 3 days.

Figure 3.3 shows the resulting curves for the temporal evolution of population size. For each dose, two independent replicates were simulated, to provide indication on the variability inherent to the finite size of our population. Here, this variability exists but is rather limited (for the initial population size we used). As expected, low doses only moderately affect cell growth, while higher doses are able to decrease population size. Both the extent and duration of cell killing increases with TRAIL. From those curves, one can extract a standard dose-response curve, which is the common output of standard cytotoxicity assays (Fallahi-Sichani et al. 2013). A typical end-point duration for anti-cancer drug cytotoxicity assays is 3 days. The simulated dose-response curve corresponding to that end-point is also shown

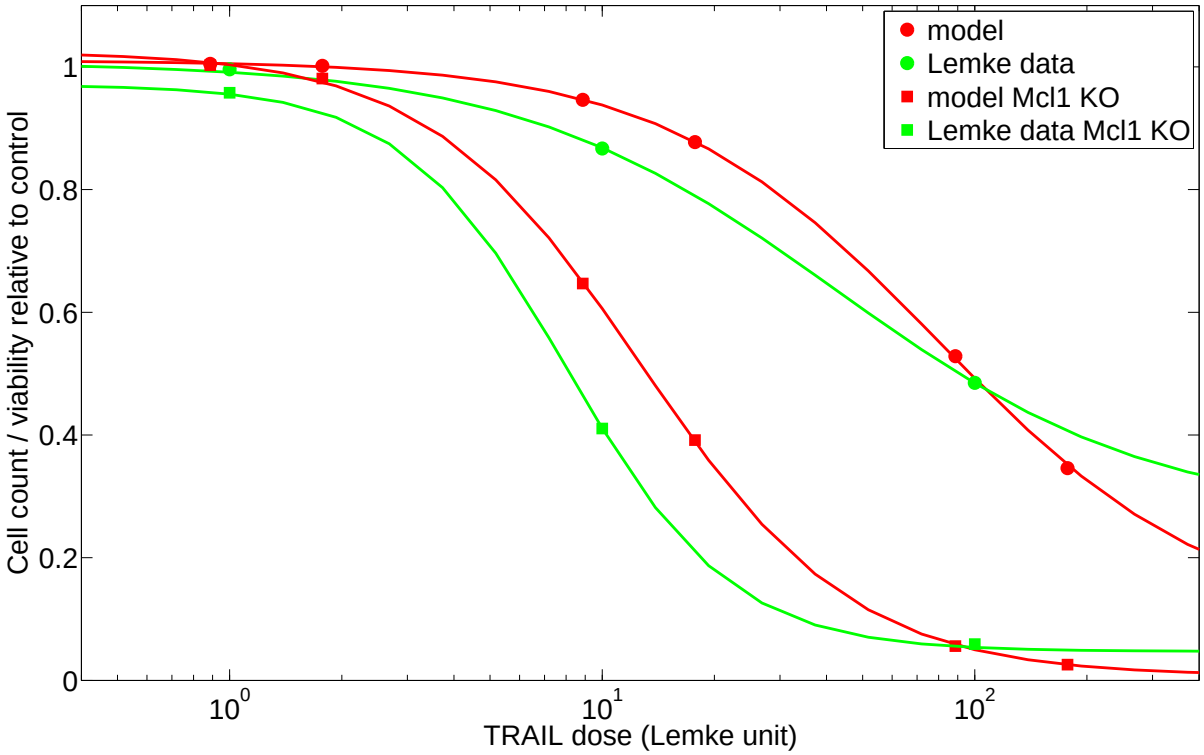



Figure 3.2: **Predicted killing at 24 hours as a function of TRAIL dose resembles experimental observations on HeLa cells.** Simulated dose-response curve (the principle of such simulations is detailed in Figure 3.3) for a 24 hours end-point to reproduce experiments from (Lemke, Karstedt, Abd El Hay, et al. 2014). In-addition, to allow comparison with the Mcl-1 shRNA silencing experiment performed by Lemke et al, we also performed simulations in which Mcl-1 synthesis rate was reduced by a factor 10. Markers (circles for WT, squares for Mcl-1 shRNA knock-out) indicate simulated (red) or experimental (green) cell count/viability 24 hours after treatment. Solid lines indicate corresponding sigmoid fits. Because the TRAIL formulation (different strategies are used to enhance the cross-linking capabilities of the original TRAIL protein) used by (Lemke, Karstedt, Abd El Hay, et al. 2014, isoleucine zipper-tagger TRAIL) differs from the one used in the Sorger group ('SuperKiller' TRAIL), we introduced a scaling factor that was chosen such that sigmoid fits of the simulated and observed dose-response curve (in absence of silencing) agree the most (here, one unit of 'Sorger group' TRAIL corresponds to 17.7 units of 'Lemke' TRAIL).

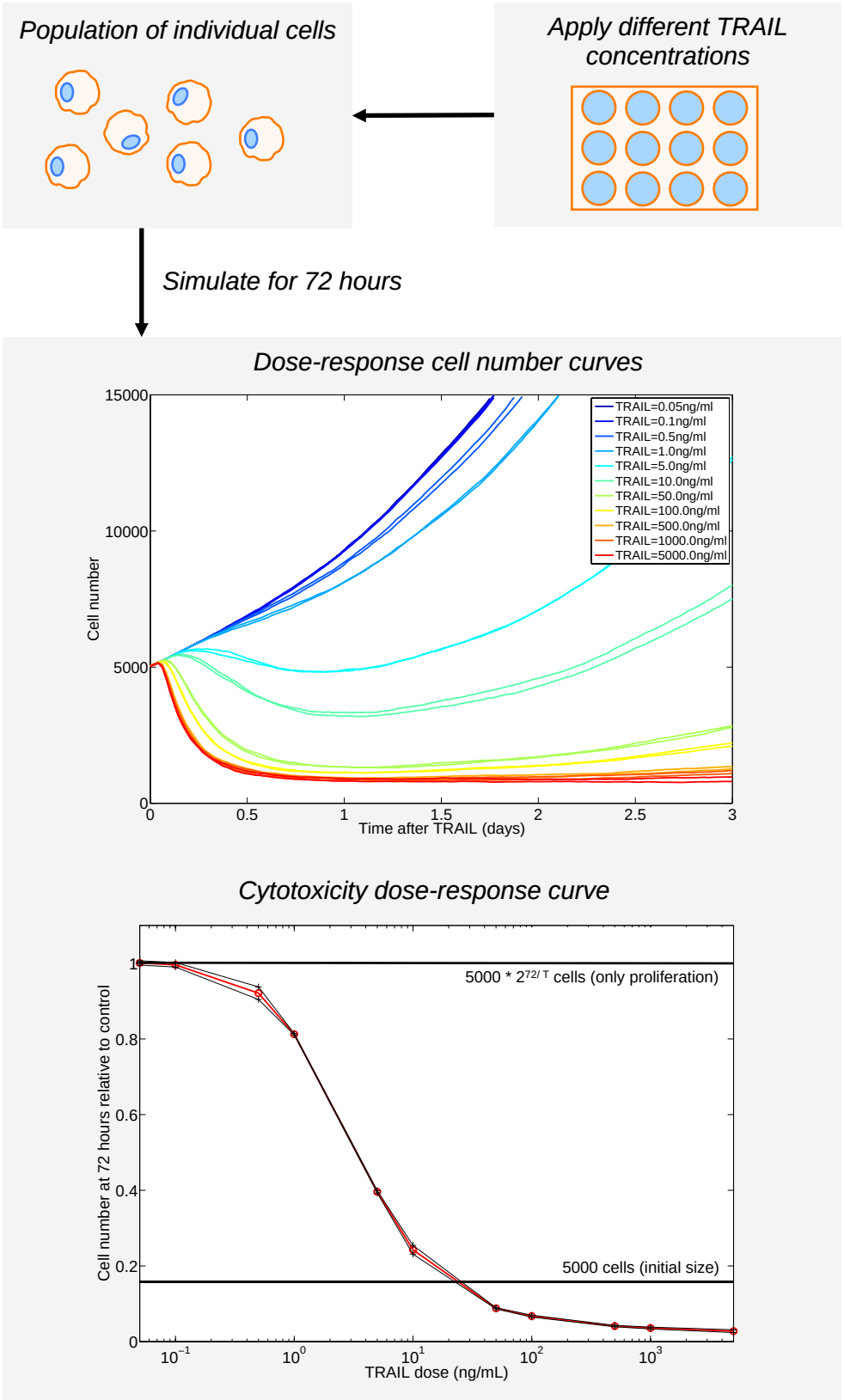



Figure 3.3: **Simulating the TRAIL dose-response of populations of single cells.** Populations of growing cells were submitted to different TRAIL concentrations when they reached a size of 5000 cells. The resulting cell number curves are shown (two independent simulations per dose were performed). Note that the TRAIL concentration sensed by each population decreases exponentially with time because TRAIL is unstable (9 hours half-life). From such curves, one can extract a standard cytotoxicity dose-response curve by dviding the number of cells after 72 hours by the cell number expected in absence of TRAIL. Black thin lines indicate standard deviations for the two replicates. The horizontal lines indicate either the no killing level or the killing level required to balance growth.

in Figure 3.3.

What informations on the ability of TRAIL to eradicate or at least control a population of growing HeLa cells is provided by such dose-response curves? Here, it suggests that there exists a dose (between 10 and 50 ng/mL) such that treating every 72 hours would be just sufficient to maintain the population size below its initial size (as could also be seen from the corresponding cell number curves). Moreover, higher doses would apparently be sufficient to rapidly eradicate the population, as a 5-fold reduction of cell number could be achieved in 3 days.

## Long-term resistance to repeated TRAIL treatments cannot be predicted from the response to single treatments

The previous considerations ignored the possibility that the population of cells surviving a first TRAIL treatment can be transiently more resistant than the initial population. But such non-genetic transient resistance acquisition was observed experimentally (Flusberg et al. 2013) and is an emerging property of our model of TRAIL-induced apoptosis (see Chapter 2). By performing longer *in-silico* experiments, in which TRAIL treatments are repeated every 72 hours, we can investigate the long-term consequences of non-genetic resistance acquisition as predicted by our model for timescales that are close to *in-vivo* timescales for TRAIL-based anti-tumor assays. From such simulations, the temporal evolution of cell number curves informs about the long-term cytotoxic efficiency of the repeated TRAIL treatments (Figure 3.4).

Resistance acquisition between the first treatment and the subsequent treatments is clearly visible. For example, while a dose of 100 ng/mL resulted in more than a 2-fold reduction in cell number 3 days after the first treatment, all the following treatments killed cells but not enough to balance growth, causing a net growth in population size on the long-term. Resistance does not accumulate at each treatment: rather, it seems to stabilize after the first one or two treatments. Thus, it is possible to define a "long-term" dose response curve, by computing for each dose the net growth rate between the second and last pulse. This net growth rate can be translated into a ratio of cell number difference with respect to a control for a 3 days time period, allowing comparison with the "standard" dose-response curve (Figure 3.5).

This comparison highlights the potential therapeutical consequences of non-genetic resistance acquisition as predicted by our model: the dose that is needed to "control" the population size is 10 times higher that what is predicted from the standard cytotoxicity dose-response curve. The maximum killing efficiency ($E_{inf}$ with the notations of (Fallahi-Sichani et al. 2013)) is also markedly changed. Therefore, our results suggest that when choosing the dosage of an anti-cancer therapeutic strategy, relying solely on standard, single-treatment

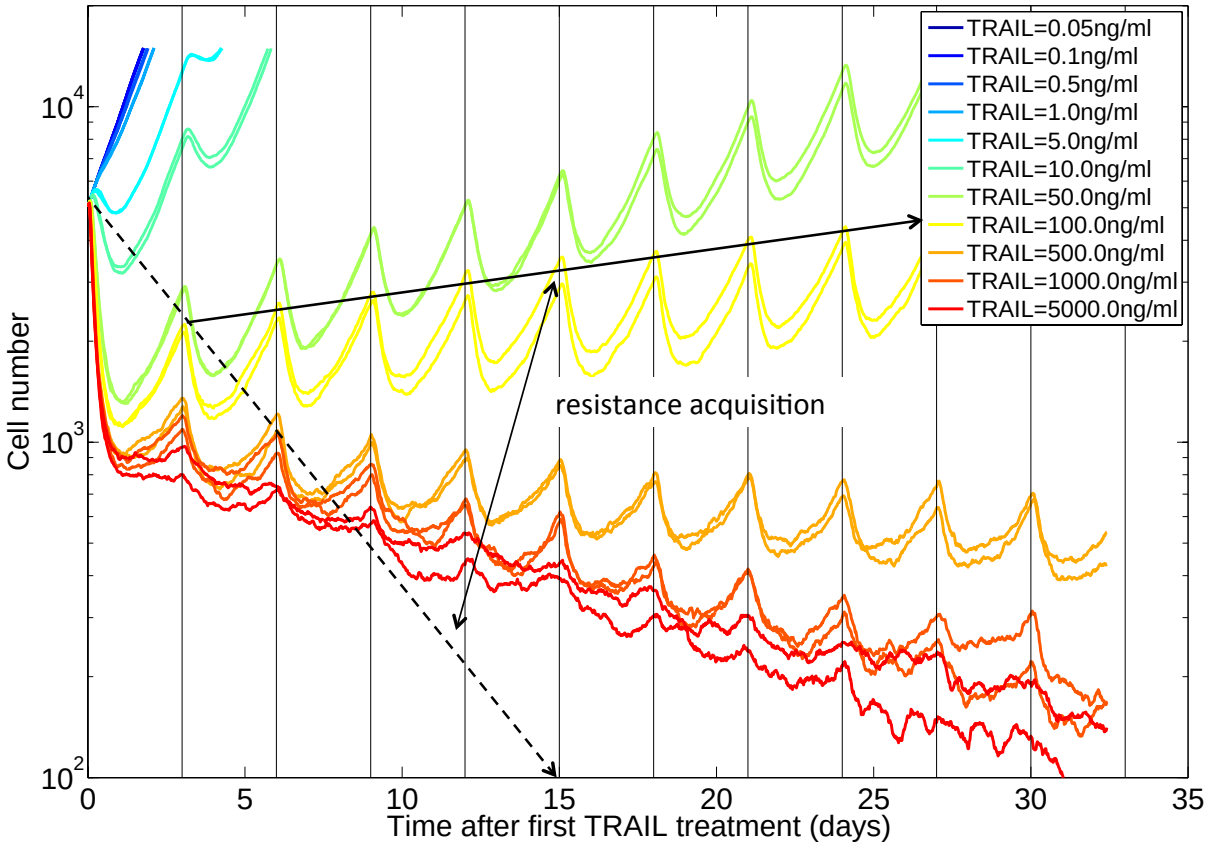



Figure 3.4: **Cell populations repeatedly treated with TRAIL acquire resistance.** The long-term evolution of population sizes when TRAIL treatment is repeated every 72 hours is shown. Simulations in which population size reached 15000 cells were stopped. Although the highest doses allows for strong reduction of the population size, resistance acquisition is clearly visible: later treatments are always less efficient than the initial treatment. Treatment times are indicated with vertical black lines. A log-scale is used such that constant killing efficiency among each consecutive treatment translates into linear progression between the peaks. Note that as expected, variability and noise in those curves increases as cell number decreases.

cytotoxicity assays could lead to failure of the therapy even in absence of mutational effects if non-genetic resistance acquisition takes place.

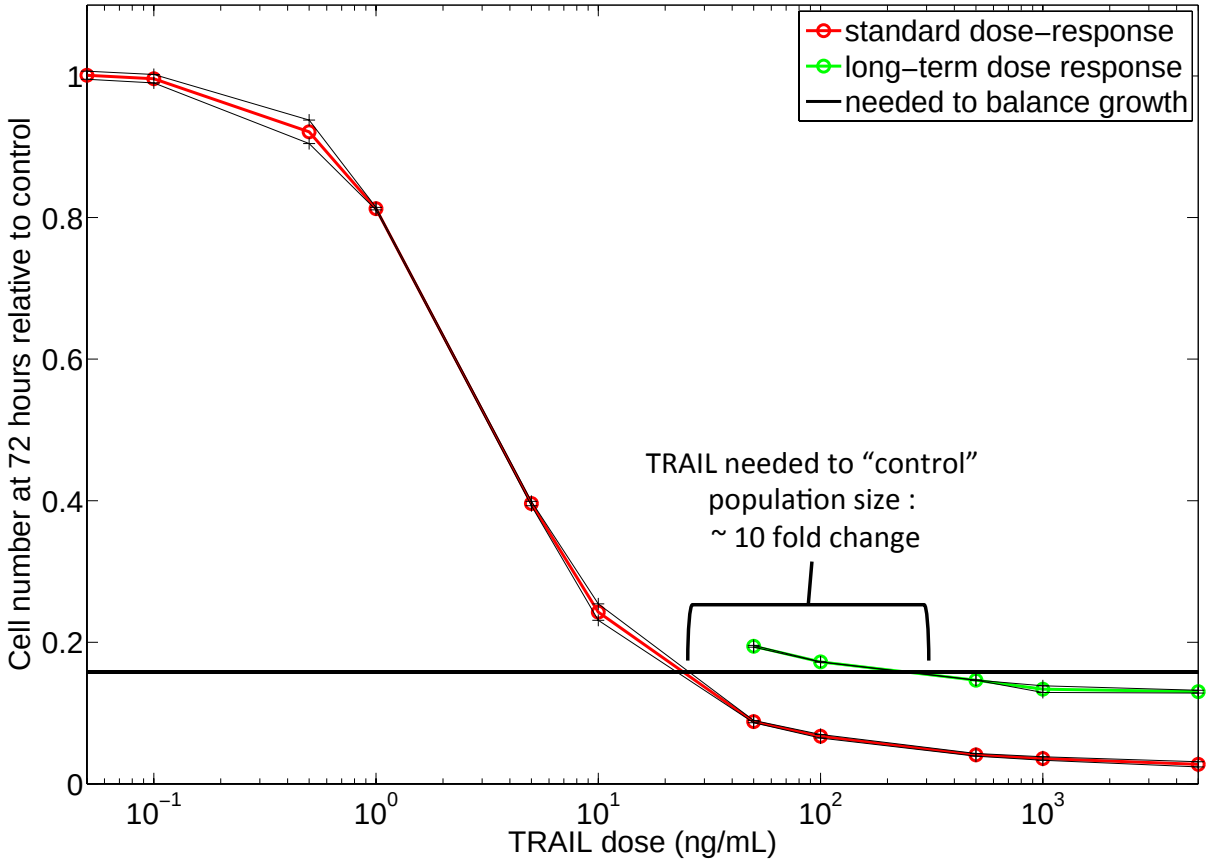



Figure 3.5: **A much higher TRAIL dose than predicted from single treatment killing efficiency is needed to outbalance cell proliferation on the long-term.** Extraction of a "long-term" dose-response curve from the long-term repeated TRAIL treatments simulations (green). The "standard" dose-response is shown in red for comparison. The long-term dose response curve does not extend to low TRAIL values because in those cases population sizes grew too fast to allow an accurate computation of the long-term net growth rate with reasonable computational times. The TRAIL dose that would allow to balance proliferation on the long-term is highlighted and compared to its equivalent as predicted by the standard dose response curve.

## Influence of scheduling on the outcome of long-term repeated TRAIL treatments: despite transient resistance acquisition, waiting between treatments is not favorable

In the previous analysis, we considered a fixed period of 3 days between treatments. It is a priori unclear whether increasing or decreasing this period could improve the long-term outcome of the repeated treatments. On one hand, if the period is too long, after a treatment, killing might stop to be important before the next treatment (hence cell proliferation would dominate in the interval). On the other hand, as suggested recently (Flusberg et al. 2013), because of transient resistance acquisition, increasing this period could enable to recover more sensitivity and thus more killing efficiency for the next treatment.

To study the influence of scheduling, we took the treatment of 100 ng/mL every 3 days as a reference (yellow curves in Figure 3.4) and tested different periods between treatments while keeping total TRAIL dose delivered constant (Figure 3.6). First, while for unfrequent treatments, the net growth rate of the population displays large variations in between treatments (strong killing followed by dominant proliferation), the amplitude of those variations decreases with the treatment period. Second, the efficiency of the repeated TRAIL treatments to limit population growth increases with treatment frequency until reaching a plateau: the 48 hours or 24 hours period treatment strategies were equally good.

Hence, our model predicts that the trade-off between waiting for the initial sensitivity to recover and the cell proliferation that occurs during this waiting time is such that waiting is not favored but that maintaining sustained (hence lower) killing is preferable. This important model prediction can be tested experimentally.

## Distinguishing population-level or cell-level resistance acquisition

We found that non-genetic resistance acquisition as predicted by our model could have a strong impact on the long-term outcome of repeated TRAIL treatments. But what are the mechanisms driving this resistance acquisition? Two distinct types of contribution to resistance acquisition could be conceptually distinguished.

First, the resistance to TRAIL of individual cells could increase because of molecular changes induced by previous TRAIL treatments. Note that the notion of "TRAIL resistance" for individual cells is probabilistic: for example, a cell having a probability of 65% to survive the next treatment (because of its current molecular state) could indeed survive; and the probability to survive another treatment could have rised to 80% because of TRAIL-induced changes. We term such 'adaptation' of individual cells to subsequent treatments *cell-level resistance acquisition*.

The second type of mechanism involved in resistance acquisition operates at the population level. Even if the individual cells "TRAIL resistance" is not changed in surviving cells, cells with low surviving probability are *by definition* eliminated in higher proportions than cells with high surviving probabilities. Hence, the amount of cells with high surviving probabilities is expected to increase by a pure selection effect. In that case, we speak of *population-level resistance acquisition* as it does not involve changes of resistance of individual cells.

In our model, transcriptional noise is assumed not to be affected by TRAIL. Therefore, changes in the average mRNA levels in a population necessarily imply a selection effect. In Figure 3.7 we show the temporal evolution of population averages of mRNA levels for cells repeatedly treated with different doses of TRAIL. Those curves clearly indicate selection effects, with, as expected, transcripts coding for anti-apoptotic proteins (Mcl-1 and XIAP

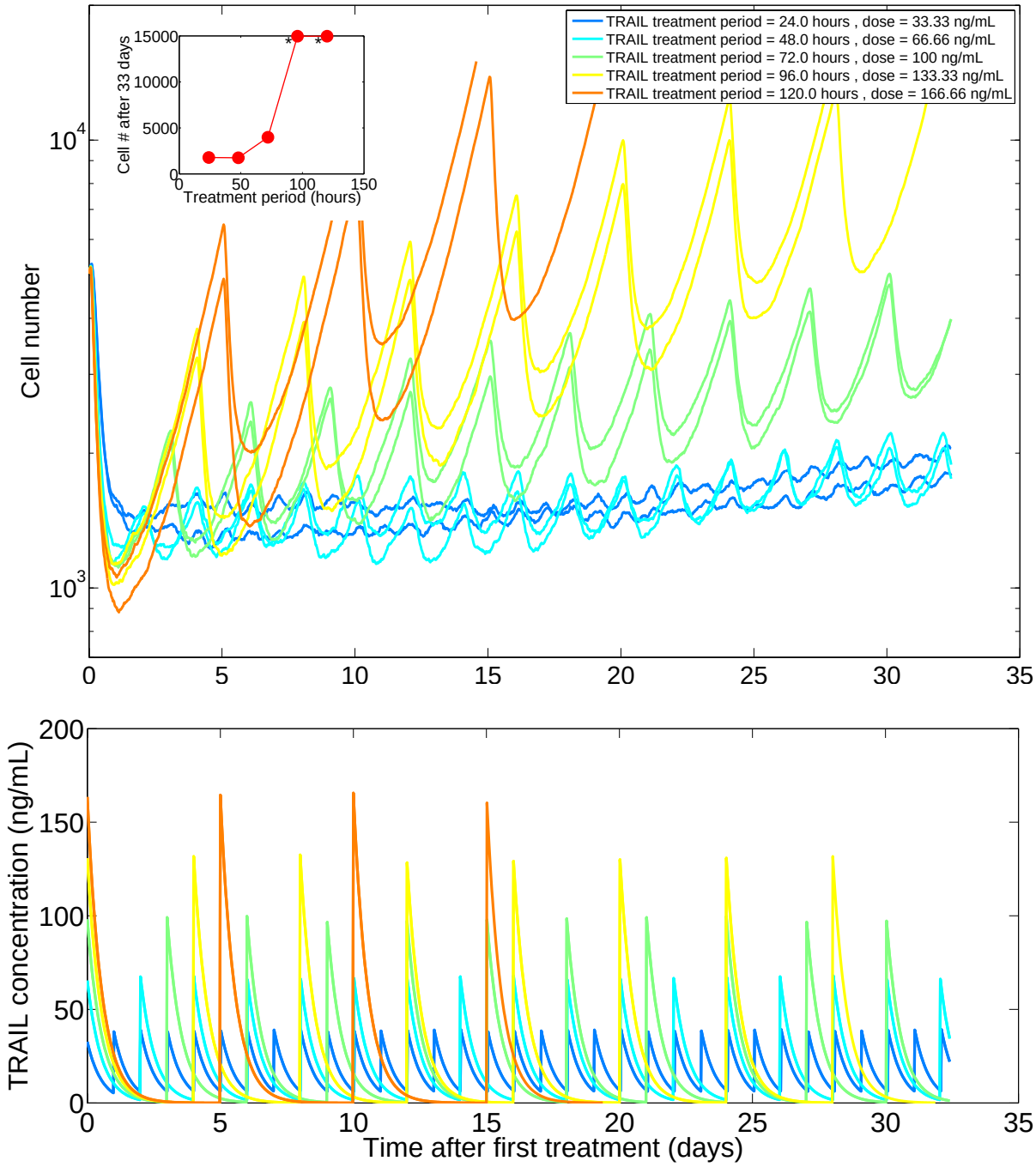



Figure 3.6: **At constant overall TRAIL usage, frequently administrating low TRAIL doses is more efficient.** Influence of scheduling on the long-term outcome of repeated TRAIL treatments. Different doses are applied at various frequencies such that in average, a dose of 100 ng/mL is delivered every 3 days. In other words, the same total amount of TRAIL is delivered either at strong doses un-frequently or at low doses frequently. Top: temporal evolution of alive cell number. Bottom: temporal evolution of TRAIL concentration. In each case, two independent replicates are shown. The inset in the top plot shows the endpoint cell number as a function of treatment period (marks indicate simulations that were stopped earlier because cell number was too high). Note that for treatment periods below 48 hours, the long-term killing efficiency is approximately constant.

are shown) being transiently enriched after TRAIL treatments while transcripts coding for pro-apoptotic proteins (pC8 and Bid are shown) are transiently depleted after treatments. However, the quantitative impact on average mRNA levels is small (15% at maximum) and is in general almost completely erased 72 hours after each treatment. Still, it cannot yet be excluded that the combined effect of those small differences can have an impact on the population resistance, as they all go in the same direction (enrichment/depletion of pro-apoptotic/anti-apoptotic mRNAs).

As opposed to mRNAs, the levels of native proteins can change because of cell-level effects caused by TRAIL signaling. For example, pro-apoptotic proteins can be cleaved and then degraded, causing a decrease in the level of the native, uncleaved protein. Thus, the temporal evolution of native protein level population averages reflects both selection and cell-level effects. Those evolutions are shown in Figure 3.8. Almost all protein levels decrease after treatment before recovering, even for anti-apoptotic proteins such as XIAP. The example of XIAP is interesting because mRNA levels are enriched (Figure 3.7, bottom-right) while protein levels are depleted, demonstrating that cell-level decrease via degradation is dominant over population-level selection, at least regarding XIAP evolution.

Overall, those results suggest that protein degradation in surviving cells is a main driver of TRAIL resistance acquisition as predicted by our model. A heuristic explanation based on the model structure for our observations is that in order to survive, cells need to limit the transformation of death signals into effector caspase activation; and this is achieved by binding of anti-apoptotic proteins to activated pro-apoptotic proteins followed by targeted degradation, resulting in decreased amounts of both pro- and anti- apoptotic proteins. However, because some anti-apoptotic proteins exhibit fast turnover, the balance between pro- vs anti- apoptotic proteins is transiently in favor of anti-apoptotic proteins, resulting in an increased resistance to subsequent treatments.

To rigorously demonstrate the dominant role of TRAIL signaling related degradation of proteins in resistance acquisition, we ran a long-term, repeated TRAIL treatments simulation in which its influence was artificially removed (Figure 3.9, green curve). More precisely, before applying the second (or third, etc…) TRAIL treatment, we reset the state of each (alive) cell to what it could have been if no TRAIL had been added at the first (or second, etc…) TRAIL treatment. By state we mean the complete state of the intra-cellular model, i.e. the levels of protein species (native or non-native) and the level of mRNAs and promoter activity status for the 17 native proteins. Hence, only selection effects are kept. As expected, although it is still visible, long-term resistance acquisition is importantly decreased compared to the normal simulation (red curve).

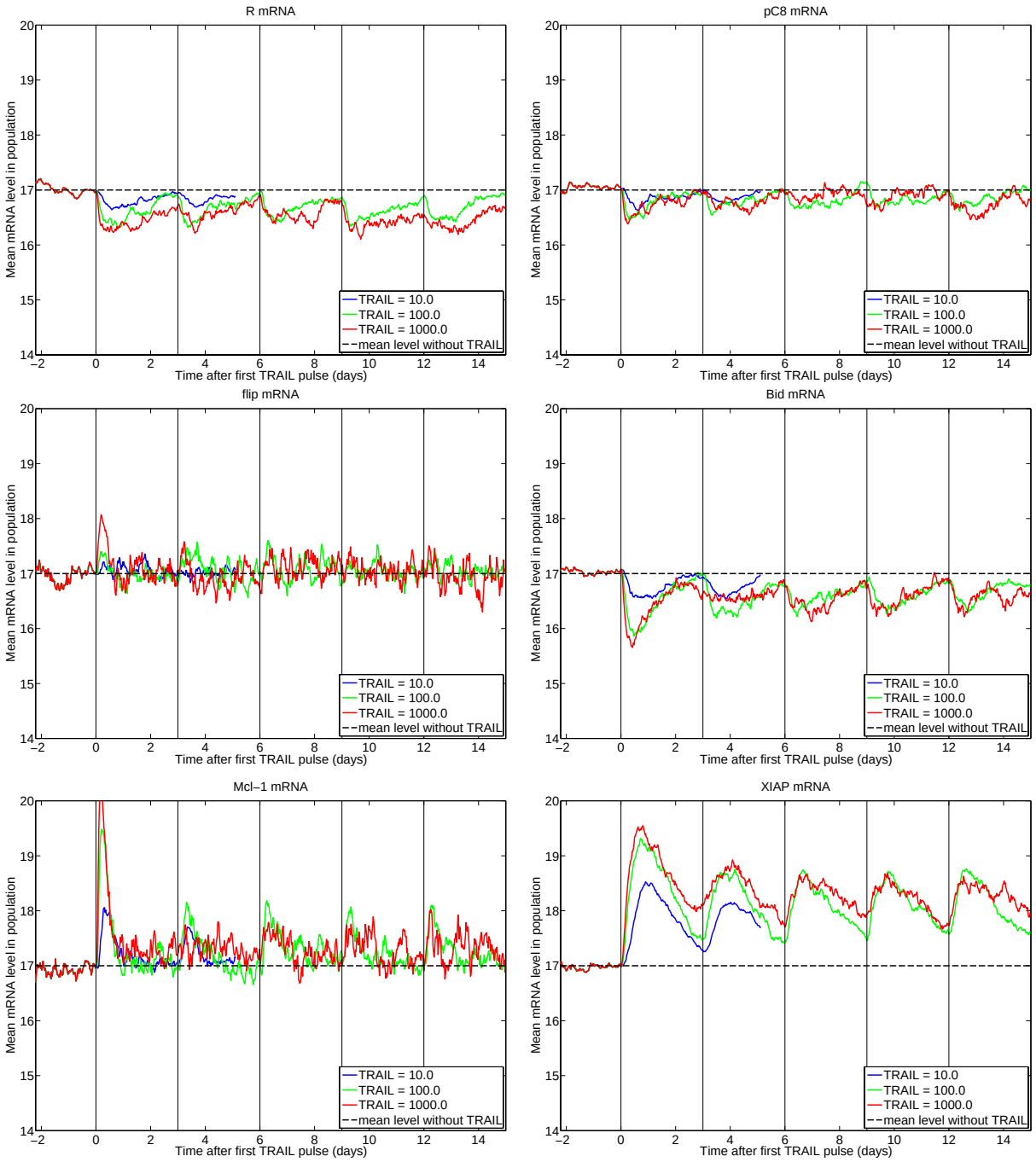



Figure 3.7: **Moderate selection effects revealed by changes in average mRNA levels in the population**. Evolution of average mRNA levels in the population for long-term repeated (every 3 days) TRAIL treatments, for 3 different TRAIL doses (in ng/mL). Only R, pC8, Flip, Bid, Mcl-1 and XIAP are shown. vertical thin black lines indicate the times of TRAIL treatments. Note that the y-axis scale only range from 14 to 20 mRNA copies. In absence of TRAIL, an average of 17 is observed as expected (horizontal dashed line). Deviations above/below 17 indicate positive/negative selection of cells with respect to that mRNA (and hence the protein it encodes).

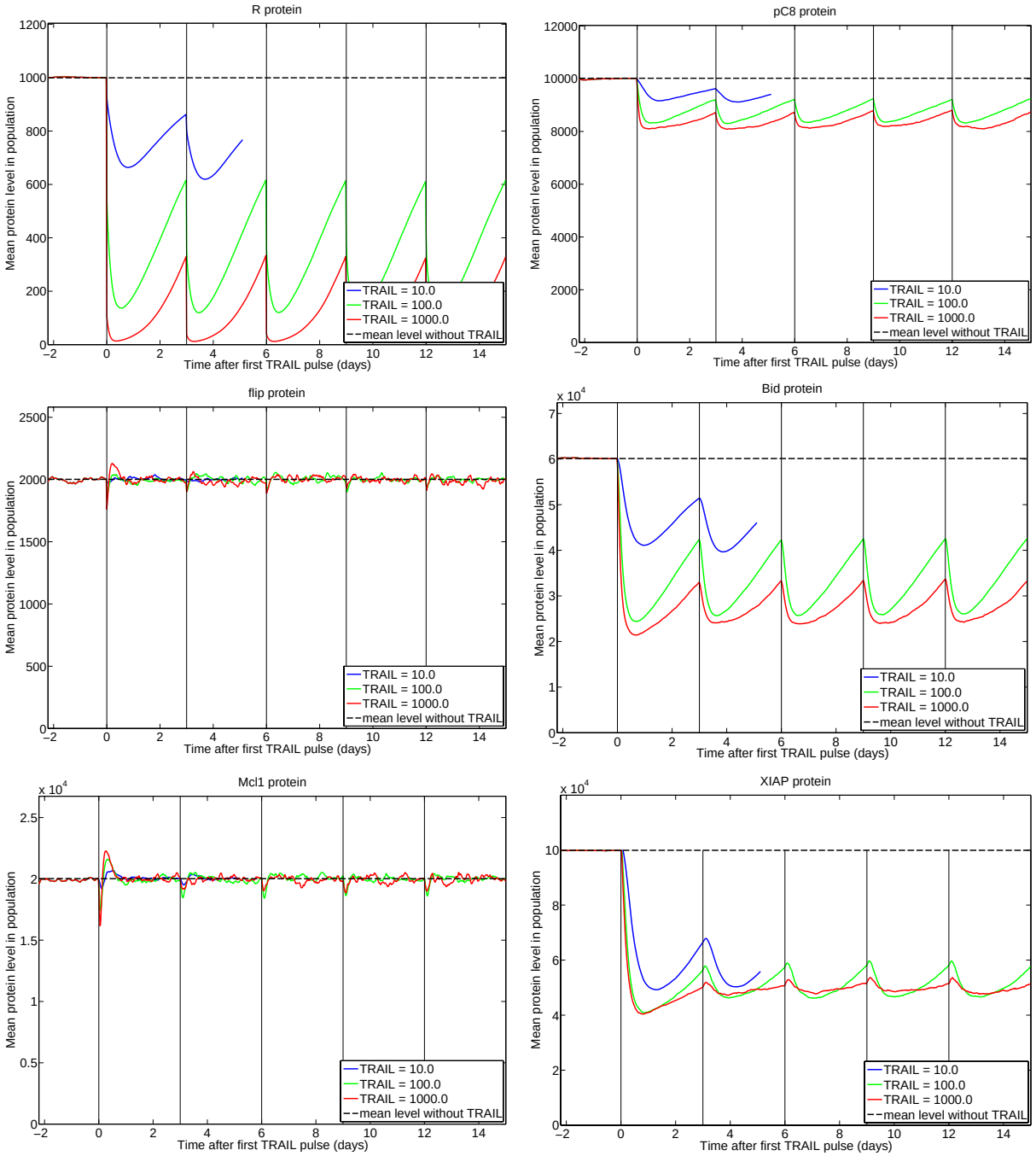



Figure 3.8: **Evolution of average protein levels in the population suggests that cell-level, TRAIL signaling related degradation plays an important role.** Evolution of average native protein levels in the population for the same *in-silico* experiments as Figure 3.7. Only R, pC8, Flip, Bid, Mcl-1 and XIAP are shown. Vertical thin black lines indicate the times of TRAIL treatments. Expected levels in absence of TRAIL are indicated by the horizontal dashed line. Deviations above/below those levels is caused either by selection effects or cell-level TRAIL signaling related degradation.

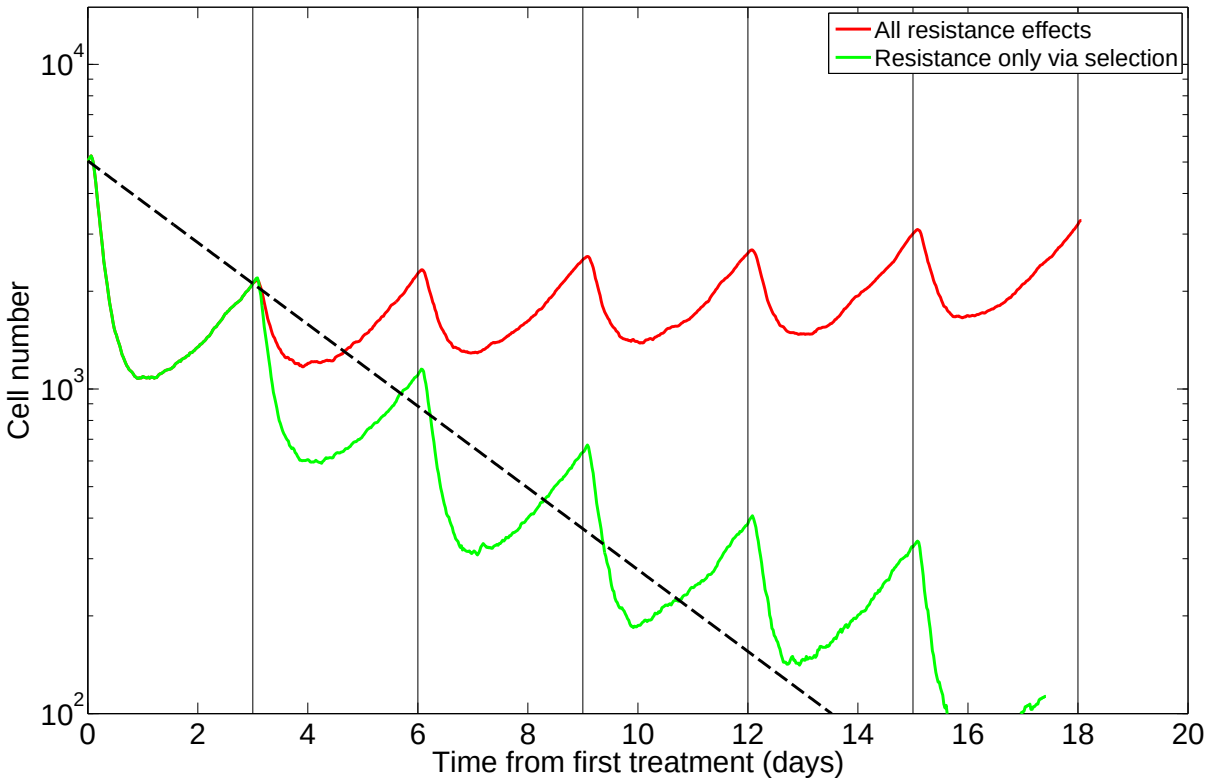



Figure 3.9: **Dominant role of TRAIL signaling related protein degradation in resistance acquisition.** The same simulation as in Figure 3.4 was performed for the dose 100 ng/mL (red curve), except that states of cells alive at the i+1$^{th}$ treatment were reset to what they could have been if no TRAIL had been added at the i$^{th}$ treatment (green curve). Horizontal thin black lines indicate times of treatments. The dashed black line highlights the evolution of cell number at treatment times that is expected in absence of resistance acquisition.

## Key role of targeted protein degradation in determining both initial and long-term resistance

The previous results were obtained with a single model parameterization. To investigate further the role of targeted degradation in TRAIL resistance, we varied the aggregate parameter in our model that used for the half-life of most non-native forms (all except Flip:R* & Mcl-1:tBid, for which we use the degradation of the non-native form; the abstract species representing mitochondrial pores; and cPARP, which is our readout for effector caspases activity).

In each case, we simulated the long-term dose-response as done previously for our reference model parameterization. The resulting cell number curves for the highest dose (1000 ng/mL) are shown in Figure 3.10 (upper plot). The rate of non-native forms degradation has a strong impact on both the initial and long-term resistance to TRAIL: compare for examples the curves corresponding to a half-life of 2.5 and 3.5 hours. The resistance to TRAIL increases with the degradation of non-native forms.

From those curves, it seems that this sensitization effect applies on both the initial and long-term resistance. To analyze further the relationship between the initial and long-term resistance, we computed in each case the initial killing efficiency (i.e., the effective killing rate obtained between the first and the second treatment) and the long-term killing efficiency (i.e., the effective killing rate obtained between the second and last treatment). Suprisingly, we found that the ratio between those two quantities was independent of the non-native forms degradation rate and hence of the initial resistance (Figure 3.10, bottom plots). In summary, our model predicts that altering the rate at which non-native forms are degraded can strongly decrease TRAIL resistance, and that initial and long-term resistance are equally affected.

The most direct way to alter those rates is to treat cells with proteasome inhibitors such as bortezomib (Velcade as commercial name, this drug being already approved as an anti-cancer drug). Indeed, it is now recognized that TRAIL resistant cell lines can be strongly sensitized to TRAIL by co-treatment with bortezomib (Menke et al. 2011; Fulda 2011), a combination which has been also successful *in-vivo* (Wilt et al. 2013; Christian et al. 2009; Shanker et al. 2008). While the molecular mechanisms behind the sensitizing effect of proteasome inhibition to TRAIL remain unclear, our model provides a simple, mechanistic explanation that do not involve genetic regulation or cross-talks with survival pathways.

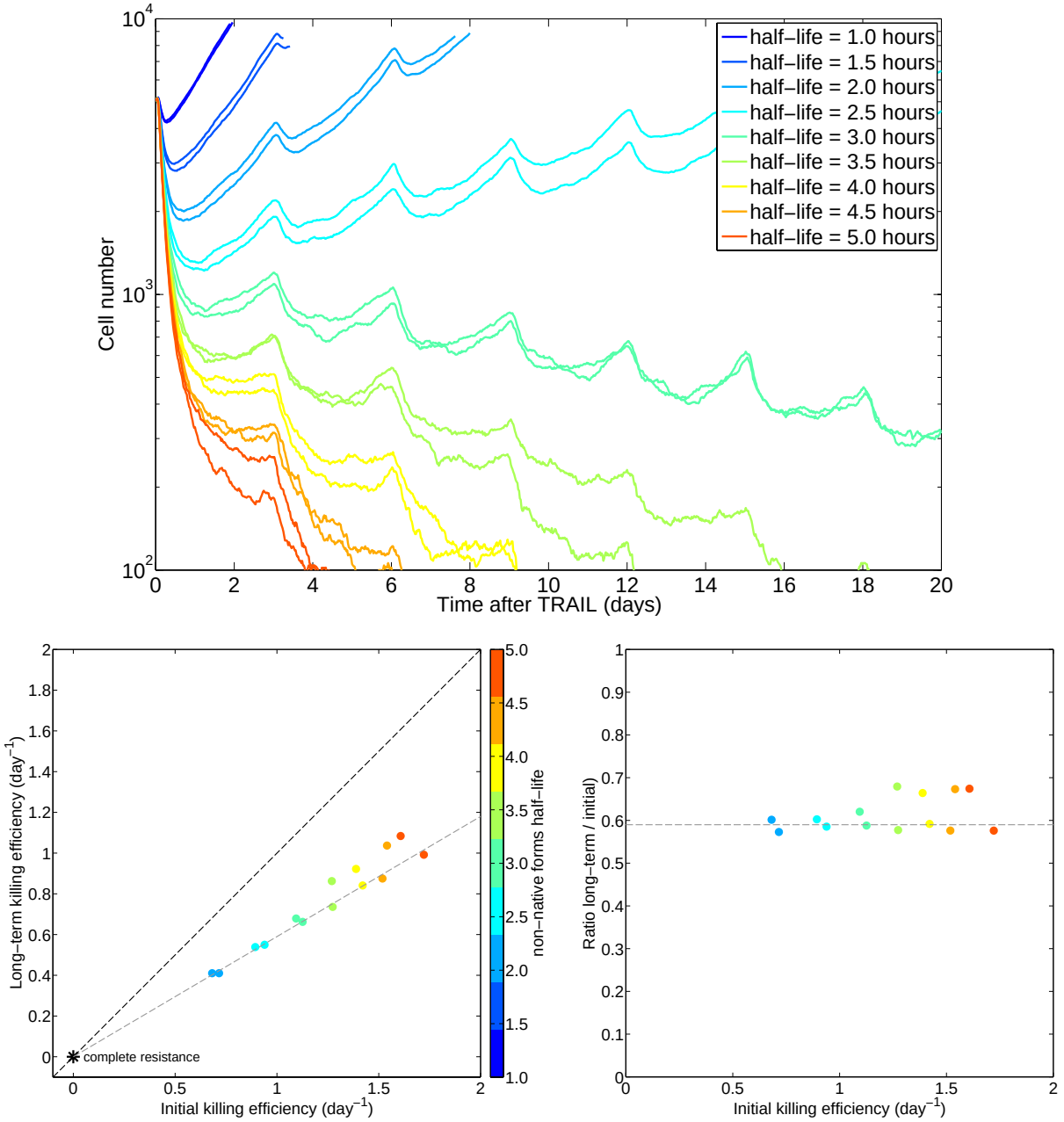



Figure 3.10: **Increasing non-native forms degradation equally decreases both initial and long-term TRAIL resistance.** In all plots, different colors correspond to different values for the aggregate parameter describing the degradation of non-native forms. Top: cell number curves for 1000 ng/mL TRAIL treatments repeated every 3 days. In each case, two independent replicates are shown. From those curves, the initial and long-term killing efficiencies can be computed. Bottom left: Long-term versus initial killing efficiency. Bottom right: ratio between the initial and long-term killing efficiencies ("resistance factor") versus the initial killing efficiency. The grey dashed lines highlight the apparent linear relationship between initial and long-term killing efficiencies.

## A 3-parameters phenomenological model captures the observed dynamics of resistance under TRAIL stimulation

The previous results show that the predicted resistance acquisition when TRAIL treatments are repeated is not driven by targeted degradation, as altering it equally impacts initial and long-term resistance. But then, what parameters are implicated in resistance acquisition? We asked whether altering the balance between anti- and pro- apoptotic protein expression levels could impact resistance acquisition. To answer this question, we repeated the previous analysis except that we either 1) changed the mean expression levels of all anti-apoptotic proteins while keeping the mean expression levels of pro-apoptotic proteins unchanged or 2) the opposite. Again, in both cases, those alterations were able to impact initial and long-term resistance, but not their ratio (not shown).

Next, we tested the impact of the 'dilution' rate, i.e. the rates at which the stable native proteins (all but Flip and Mcl-1) are degraded. Interestingly, we observed the opposite behavior compared to previous variations of non-native forms degradation or protein expression levels: resistance acquisition is strongly affected (it increases as dilution rate decreases), while initial resistance is almost un-affected (Figure 3.11).

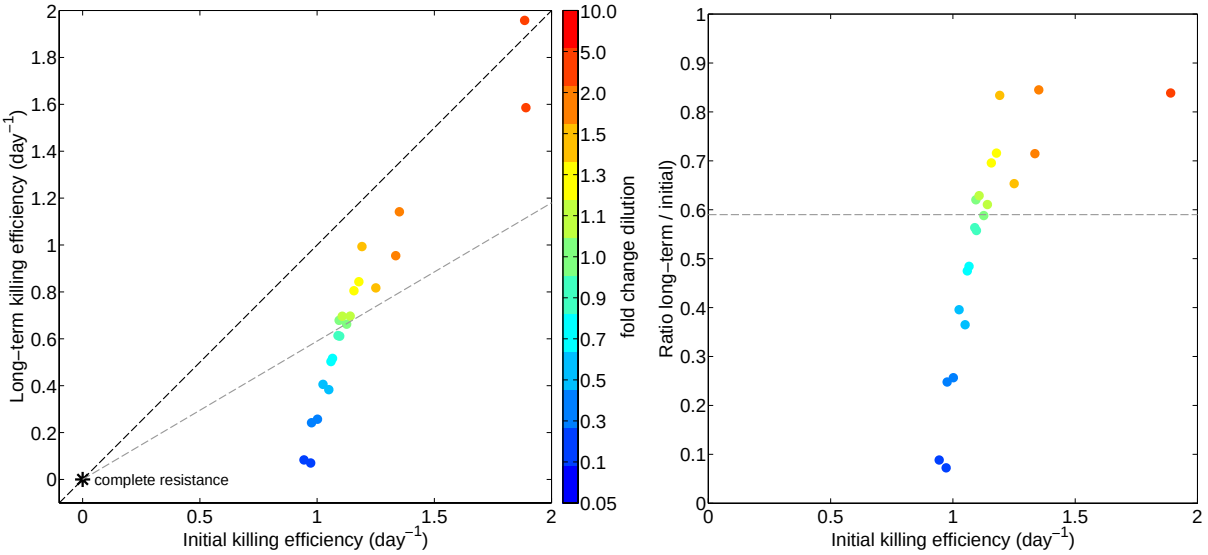



Figure 3.11: **Turnover rate of stable proteins governs resistance acquisition.** Same analysis as in Figure 3.10 except that the turnover rate of stable proteins is varied (i.e., dilution rate). The fold change corresponding to the reference model is indicated in the colorbar (1 corresponds to a half-life of 27 hours, 10 to a half-life of 2.7 hours, etc.). In each case, the corresponding protein synthesis rates are scaled by the same factor such that the mean expression level is unchanged. Grey dashed lines indicate the linear relationship between initial and long-term killing efficiency that was valid for variation in non-native forms degradation (see Figure 3.10). It does not hold here, instead, resistance acquisition is strongly affected while initial resistance does not change much.

The fact that some types of parameter modifications can change the overall TRAIL resistance without changing resistance acquisition itself while others can change resistance acquisition without affecting much initial resistance suggests that the model behavior can

be described with a "factorized", low-parametric phenomenological model in which the separate contributions of initial (or 'baseline') resistance and resistance acquisition could be easily identified. To investigate this hypothesis, we formalized our heuristic explanation of the model behavior into a simple, 3-parameters model (depicted in Figure 3.12).

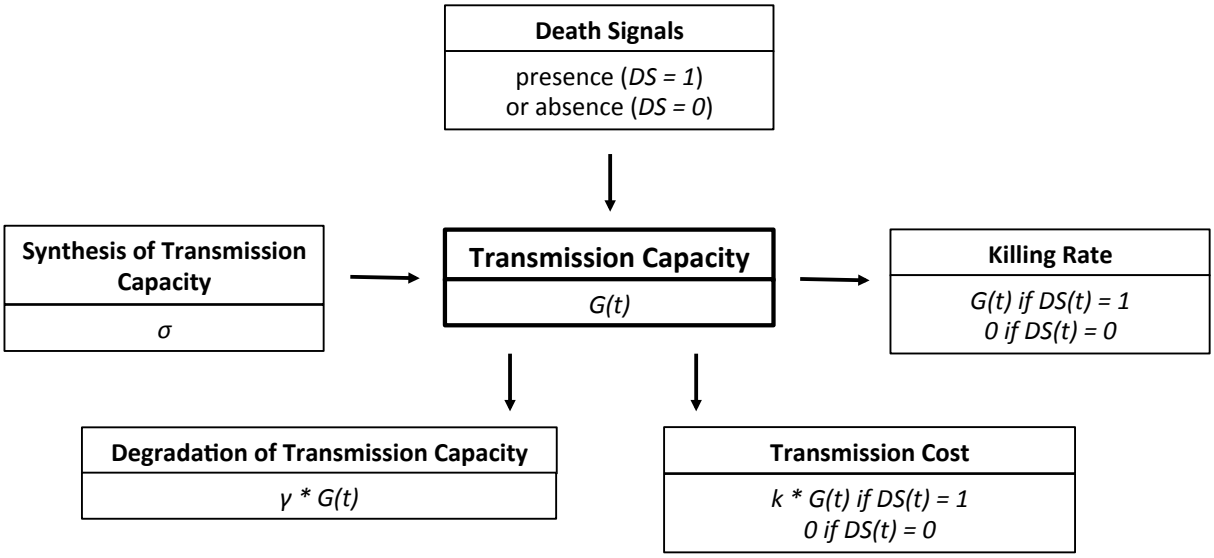



Figure 3.12: **A 3-parameters phenomenological model to describe the population-level killing dynamics in presence or absence of death signals.** A single variable represents the instantaneous ability of the population to transform death signals into cell death (transmission capacity, noted $G$). $G$ is simply the death rate that is obtained in presence of death signals. Resistance acquisition is modeled via the parameter $k$, which represents the cost of transmission in terms of transmission capacity: in presence of death signals, killing occurs but is consumed. If $k = 0$, there is no resistance acquisition. The 'synthesis' of $G$ and its normal (i.e. in absence of death signals) 'degradation' are modeled with the two parameters $\sigma$ and $\gamma$ respectively.

This model features a single, population-level variable $G$ describing the instantaneous ability of the population to transform death signals into actual killing (*transmission capacity*). For simplicity, we use a binary representation of the death signal ($DS$): it is either present ($DS = 1$) or absent ($DS = 0$). The death rate of the population is then given by $\lambda(t) = G(t)\,DS(t)$: $G(t)$ is simply the 'potential' death rate that is 'realized' if death signals are present at $t$. The equation describing the dynamics of $G$ is the following:

$$\frac{dG}{dt} = \sigma - \gamma\,G(t) - k\,DS(t)\,G(t)$$

($G$ = death rate obtained in presence of death signals)

In this equation, $\sigma$ and $\gamma$ represent the normal synthesis and degradation rates of the transmission capacity $G$. The additional term $k\,DS(t)\,G(t)$ accounts for the fact that transforming death signals into killing has a cost in terms of transmission capacity (for example because cells with low transmission capacity will be positively selected or because degradation will decrease the balance of pro-apoptotic proteins over anti-apoptotic proteins). In summary, the killing dynamics in cell populations facing TRAIL stimulation is described

with only three parameters: $\sigma$, $\gamma$ and $k$.

This equation is easy to solve analytically for any temporal pattern of presence/absence of death signals (i.e. the treatment strategy). In absence of death signals, $G$ converges exponentially to $G_{off} = \frac{\sigma}{\gamma}$, and hence the population would have the maximum response (i.e., death rate $\lambda_{off} = G_{off}$) when death signals are added. On the contrary, in presence of death signals the transmission capacity converges exponentially towards $G_{on} = \frac{\sigma}{\gamma+k} \leq G_{off}$. Thus, the long-term death rate $\lambda_{off} = G_{off}$ in presence of death signals is lower than the initial death rate $\lambda_{off}$ obtained when a naïve population is treated.

A prediction of this phenomenological model is that the "resistance factor" (i.e. the relative change of killing efficiency of subsequent treatments compared to the initial treatment, as shown in Figure 3.10, bottom right) would converge exponentially to 1 as the period between treatments is increased. Interestingly, this behavior is also observed for the original, mechanistic model (Figure 3.13).

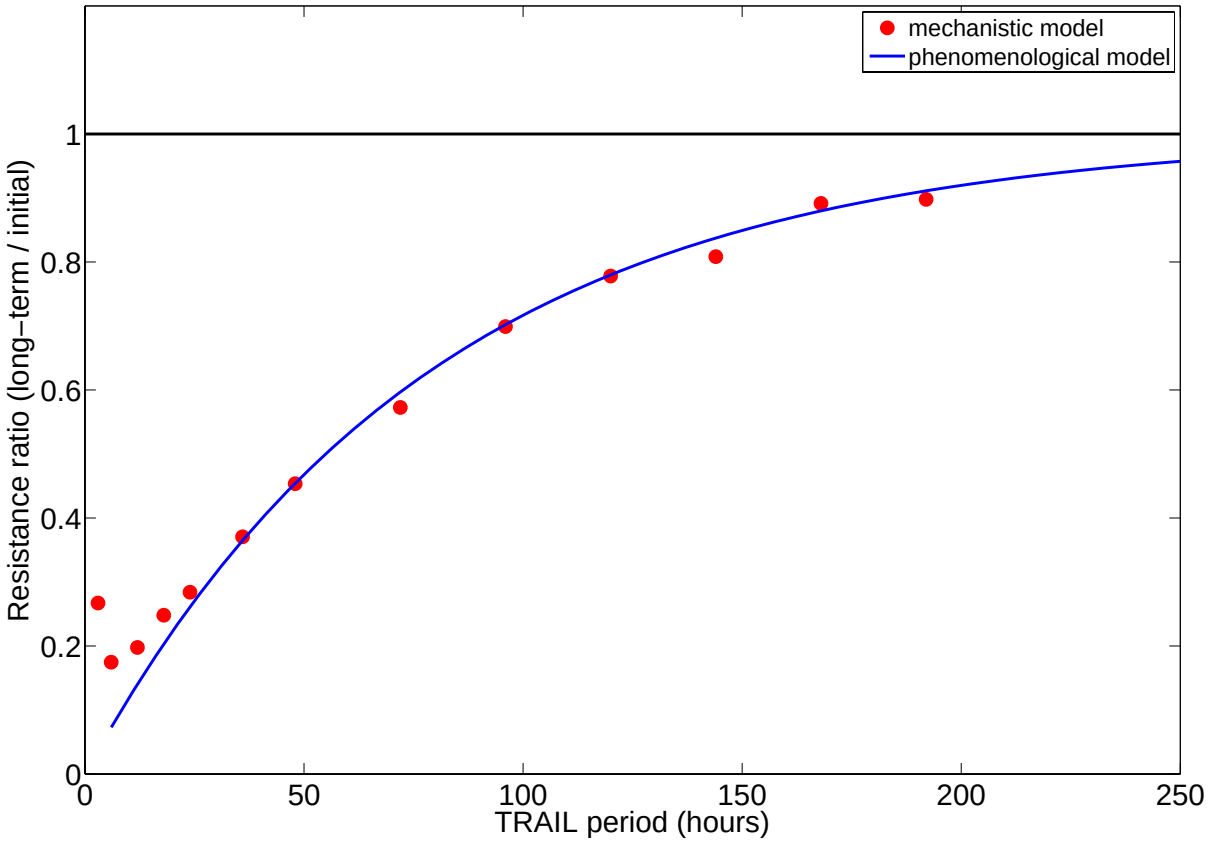



Figure 3.13: **Resistance factor dependency on treatment period of the mechanistic model can be reproduced with the 3-parameters phenomenological model.** For the mechanistic model, repeated treatments with 1000 ng/mL TRAIL were performed for different treatment periods. In each case, the resistance ratio (killing efficiency of late treatments divided by killing efficiency of the first treatment, where killing efficiency is the ratio between cell number at the time of next treatment and cell number at treatment time) was computed. Parameters of the phenomenological model were adjusted to reproduce the same dependency, assuming that each treatment leads to 3 hours of death signal presence. Those parameters are $\sigma = 1$, $k = 2$ and $\gamma = \frac{ln\,2}{55}\ hrs^{-1}$. For the mechanistic model, increase of the resistance factor at low TRAIL periods is expected, as delays between TRAIL exposure and cell death result in death caused by the first treatment being attributed to the second treatment, etc.

Another prediction of the phenomenological model is that on the long-term, sustaining death signals is the best strategy to limit cell proliferation (this can be demonstrated analytically). Again, this is consistent with the original model predictions (see Figure 3.6). Hence, our simple model can capture essential features of the original model behavior regarding resistance to long-term repeated treatments.

In principle, it should be possible to map each parameter in the original, mechanistic model into its contribution to the phenomenological parameters $\sigma$, $\gamma$ and $k$. It seems intuitive that altering protein synthesis rates would impact $\sigma$ only, while changing the turnover rate of native proteins should impact $\gamma$ mostly, and those intuitions are consistent with our observations. Interpreting the impact of non-native forms degradation in terms of the phenomenological parameters $\sigma$, $\gamma$ and $k$ seems less immediate and would require further investigation.

## 3.3 The role of space: could geometrical effects protect tumor spheroids from repeated TRAIL treatments?

In the previous section, we investigated the role of time in the resistance of cancer cell populations treated with TRAIL by performing long-term single-cell resolved population dynamics simulations. In this section, we extend the model to investigate the role of cell-extrinsic heterogeneity that arises from the 3-dimensional organization of cells growing as spheroids. We focus on the effect of limited TRAIL penetration within the spheroid that is expected if its diffusion is slow and its degradation is high.

### Modeling cell growth and apoptosis for multi-cellular spheroids

For this purpose, it is needed to explicit the spatial localization of cells as well as the spatial distribution of TRAIL. This leads to a cell-based, 3D multi-scale model composed of several components highlighted in Figure 3.14.

First, to describe the motion of individual cells, we model the physical forces exerted on them and solve the corresponding equations of motion. Three types of forces are considered: adhesive/repulsive forces for cells in contact, micro-motility forces to mimic the ability of cells to explore their environment by active migration, and friction forces to account for the viscosity of the extra-cellular matrix.

To model adhesion/repulsion between cells in contact, we use the modified Hertz-model as in (Galle et al. 2005). In this model, the adhesion/repulsion force for a given contact

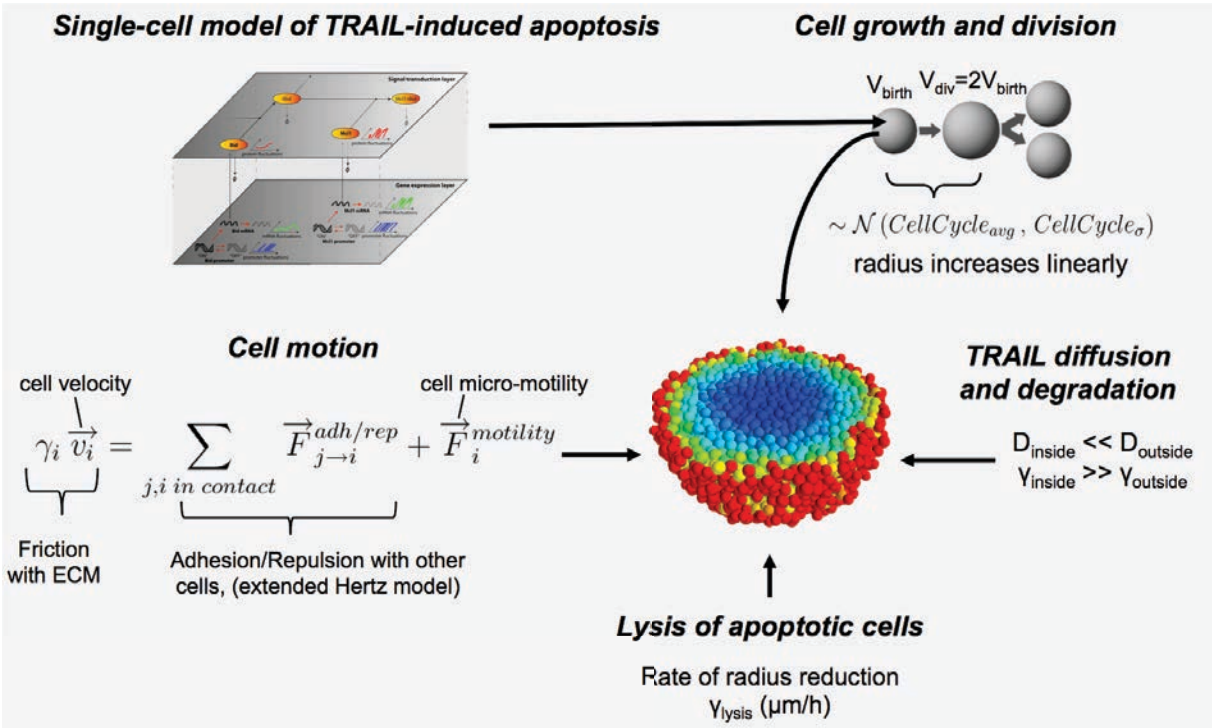



Figure 3.14: **Multi-scale model of multi-cellular spheroids treated with TRAIL.** Cells are physically represented as growing spheres. Cell motion is governed by adhesive/repulsive forces for cell-cell contacts, micro-motility and a friction term. Inertia is neglected because motion is friction-dominated. The rate at which cell radius grow is chosen to match a normally distributed cell-cycle duration as for previous section. The diffusion and degradation of TRAIL is also solved. Diffusion is assumed to be small when cell density is high, degradation is assumed to be high when cell density is high. In each cell, an instance of the single-cell model of TRAIL-induced apoptosis is simulated and governs cell decision to enter apoptotis. Apoptotic cells have their radius decreasing at a constant rate.

between two cells $i$ and $j$ derives from the following potential energy:

$$V_{i,j} = \frac{2}{5}\tilde{E_{ij}}\left(R_i + R_j - d_{ij}\right)^{5/2}\sqrt{\frac{R_i R_j}{R_i + R_j}} - \sigma A_{ij} \tag{3.1}$$

$$\text{where } \tilde{E_{ij}} = \left[\frac{3}{4}\left(\frac{1-\nu_i^2}{E_i} + \frac{1-\nu_j^2}{E_j}\right)\right]^{-1}$$

$$\text{and } A_{ij} = \pi\frac{R_i R_j}{R_i + R_j}\left(R_i + R_j - d_{ij}\right)$$

In this equation, the first term corresponds to elastic repulsion. $R_i$ denotes the radius of cell $i$, $d_{ij}$ the distance between the two cell centers, $E_i$ the Young modulus of cell $i$ and $\nu_i$ its Poisson ratio. $\tilde{E_{ij}}$ is an effective Young modulus for this particular contact. The Young modulus has the dimension of a pressure and measures cell elasticity, while the Poisson ratio is a dimensionless parameter smaller than $0.5$ measuring volume conservation upon deformation (i.e. compressibility). The second term corresponds to adhesion between the two cells. $\sigma$ is the adhesivity (dimension of an energy per unit surface) and $A_{ij}$ the contact area between the two spheres (not the sphere overlap surface but an approximation of the true contact surface valid for small deformations (Landau et al. 2012)). The corresponding force is directed along the axis joining the two cell centers and has the following magnitude:

$$F_{ij}^{adh/rep} = \tilde{E_{ij}}\left(R_i + R_j - d_{ij}\right)^{3/2}\sqrt{\frac{R_i R_j}{R_i + R_j}} - \pi\sigma\frac{R_i R_j}{R_i + R_j} \tag{3.2}$$

We model cell micro-motility by a force which direction and magnitude stochastically switches at a rate $\frac{1}{\tau_{MM}}$ such that the mean duration between micro-motility force changes for a given cell is $\tau_{MM}$ (the persistence time). The direction of this force chosen randomly (isotropic) and its magnitude is sampled from a gaussian distribution. Thus, 3 parameters describe the micro-motility model: the persistence time, the average micro-motility force, and its standard deviation representing cell-to-cell variability.

Finally, each cell motion is slowed down by a friction force that follows the expression:

$$\overrightarrow{F}_i^{friction} = -4\pi R_i^2\,\mu\,\overrightarrow{v}_i \tag{3.3}$$

Here for simplicity we have aggregated cell-cell/cell-ECM, normal/parallel friction coef-

ficients into a single friction coefficient $c_i$ and used the functional form $c = \mu\, S_C$ (Galle et al. 2005) where $\mu$ is the friction constant and $S_C$ the surface of contact.

It is generally accepted that inertia terms are negligible in comparison to friction terms (Galle et al. 2005). Indeed, given that cell mass $m \sim 10^{-12}\, kg$, the time constant with which velocity adapts to changes in forces other than friction is $\frac{m}{4\pi R_i^2\, \mu} \sim 0.001\, \mu^{-1} s$ (for a cell radius of $10\, \mu m$), that is less than $10^{-11}\, s$ for the value of $\mu$ used by (Galle et al. 2005). Therefore, the equation of motion for cell $i$ can be written as:

$$4\pi R_i^2 \mu\, \overrightarrow{v_i} = \sum_{j,i\ in\ contact} \overrightarrow{F}^{adh/rep}_{j \to i} + \overrightarrow{F}^{motility}_{i} \tag{3.4}$$

For a given configuration of cells, this equation is straight-forward to solve as soon as cell-cell contacts have been detected. It provides a velocity vector for each cell that is used to compute its displacement for a given time step. In our simulations, the time step is adaptively chosen based on computed velocities such that no cell moves more than $0.1\, \mu m$, ensuring a certain accuracy and that no contacts are missed. All the parameters needed to simulate the motion of cells are given in Table 1.

Table 3.1: List of parameter values for computing the motion of cells.

| Parameter | Symbol | Value | Unit |
|---|---|---|---|
| Young modulus | $E$ | 400 | $Pa$ |
| Poisson ratio | $\nu$ | 0.4 | none |
| Adhesivity | $\sigma$ | $10^{-4}$ | $J.m^{-2}$ |
| Friction coefficient | $\mu$ | $10^{7}$ | $Pa.s.m^{-1}$ |
| Micro-motility persistence time | $\tau_{MM}$ | 2 | $min$ |
| Micro-motility average force | $F^{MM}_{avg}$ | $10^{-10}$ | $N$ |
| Micro-motility force variability | $CV_{FMM}$ | 0.5 | none |
| Maximum displacement per motion step | $\Delta l_{max}$ | 0.1 | $\mu m$ |

To model cell growth and division, as in the previous section each cell is attributed a cell cycle duration sampled from a normal distribution. During the cell cycle, cell radius is assumed to increase linearly with a rate ensuring that volume doubles. At division, the mother cell is replaced by two half-volume daughter cells that are just in contact. The axis of division is assumed to be randomly isotropic. Parameters needed to simulate cell growth

and division are summarized in Table 2.

Table 3.2: List of parameter values for simulating cell growth and division.

| Parameter | Symbol | Value | Unit |
|---|---|---|---|
| Average cell cycle duration | $CC_{avg}$ | 27 | $hours$ |
| Cell cycle duration variability | $CV_{CC}$ | 0.11 | none |
| Cell radius at birth | $R_{birth}$ | 7.5 | $\mu m$ |
| Cell radius at division | $R_{div}$ | $2^{1/3}\,R_{birth}$ | $\mu m$ |
| Daughter cells distance at division | $d_{div}$ | $2\,R_{birth}$ | $\mu m$ |

To model TRAIL penetration, we discretize a cuboïdal domain surrounding the spheroid with a regular grid, in which TRAIL diffusion and degradation is computed using an explicit first order scheme. To account for the fact that TRAIL diffusion (and degradation) is expected to be slower (higher) in regions of high cell density, we attribute at each step a different diffusion coefficient ( and TRAIL half-life) dependending on the presence ($D_{cells}$ and $HL^{cells}_{TRAIL}$) or absence ($D_{medium}$ and $HL^{medium}_{TRAIL}$) of cells in that voxel.

Because our aim is to study the potential impact of insufficient TRAIL penetration on the outcome of repeated TRAIL treatments, we made conservative assumptions for $D_{cells}$ and $HL^{cells}_{TRAIL}$ values to ensure the existence of gradients within compact spheroids of more than 5000 cells. For comparison, the diffusion coefficient of water in water is $2000\,\mu m^2.s^{-1}$ and a typical diffusion coefficient for an antibody in a tissue is $10\,\mu m^2.s^{-1}$ (Thurber et al. 2008).

TRAIL treatments are modeled by using time-dependent Dirichlet boundary conditions that follows the same evolution as for 'monolayer' simulations (exponential decrease with a 9 hours half-life in-between administrations). The parameters needed to simulate TRAIL penetration are summarized in Table 3.

Table 3.3: List of parameter values for simulating TRAIL penetration in the spheroid.

| Parameter | Symbol | Value | Unit |
|---|---|---|---|
| TRAIL diffusion in absence of cells | $D_{medium}$ | 50 | $\mu m^2.s^{-1}$ |
| TRAIL diffusion in presence of cells | $D_{cells}$ | 0.2 | $\mu m^2.s^{-1}$ |
| TRAIL half-life in absence of cells | $HL^{medium}_{TRAIL}$ | 9 | $hours$ |

| Parameter | Symbol | Value | Unit |
|---|---|---|---|
| TRAIL half-life in presence of cells | $HL_{TRAIL}^{cells}$ | 0.5 | $hours$ |
| Spatial discretization for solving diffusion | $l_{TRAIL}$ | $2\,R_{div}$ | $\mu m$ |
| Boundaries dimension | $L_{medium}$ | 600 | $\mu m$ |

Finally, our single-cell model of TRAIL-induced apoptosis is simulated in each cell as previously except that the TRAIL concentration depends on the cell position. When a cell commits to apoptosis (i.e., its internal amount of cleaved PARP reaches a threshold), we assume that its radius is reduced at a given rate representing the speed of the cell lysis process (when a dying cell radius goes below a threshold, it is simply removed from the simulation). This parameter has potentially an impact on the spatial re-organization of the spheroid when cell death is high, which in turn could affect TRAIL penetration and hence subsequent killing.

Table 3.4: List of parameter values for simulating TRAIL-induced apoptosis and lysis of dying cells.

| Parameter | Symbol | Value | Unit |
|---|---|---|---|
| Radius reduction rate of dying cells | $\gamma_{lysis}$ | varied (1 or 10) | $\mu m.hr^{-1}$ |
| Radius under which cells are removed | $R_{removal}$ | $0.2\,R_{division}$ | $\mu m$ |
| Parameters of the TRAIL-induced apoptosis model | - | as reference model in previous section | - |

We have fully described the mathematical equations describing the model but did not completely describe how those equations are numerically solved. Briefly, the three processes (cell growth and motion, extracellular TRAIL diffusion and degradation, intra-cellular TRAIL signaling) are simulated asynchronously with a global timestep of 5 minutes, and each process has its own internal timestep (the timestep for cell growth and motion is computed adaptively based on maximum displacement, the timestep for extracellular TRAIL diffusion and degradation is computed from the CFL criterion, the timestep for intra-cellular TRAIL signaling is controlled by the ODE solver based on absolute and relative error tolerance parameters). The numerical aspects related to the simulation of such models are discussed in more details in Chapter 5 (section 5.2). In particular, tests of numerical convergence justifying the choice of the global timestep and the diffusion and

degradation internal timestep are given (Figure 5.5).

## Simulation of repeated treatments on spheroids

In the previous section, we found that because of resistance acquisition as predicted by our reference model of TRAIL-induced apoptosis in HeLa cells the dose required to control the size of a cell population (when repeated every 3 days) growing as a monolayer was 200 ng/mL, while 20 ng/mL would have appeared sufficient from standard killing assay.

We asked whether insufficient TRAIL penetration could also impact the long-term outcome of repeated treatments. For this purpose, we ran several simulations of the multi-scale model described above as follows. We start by seeding a small population (500 cells) within a small volume until it reaches a size of 8000 cells. Of note, this allows to obtain a realistic correlation between internal cell state (variable because of protein fluctuations) and spatial position: neighbors cells are more likely to be close siblings and hence to have correlated cell state.

We then apply TRAIL every 3 days for more than 30 days. A example visualization of the resulting spheroid evolution is shown in Figure 3.15. It displays 6 different time points and 3 cellular properties per time point: TRAIL local concentration, active caspase-8 levels and normal/apoptotic cell status. It is interesting to see how the 'non-noisy' cell-extrinsic information (TRAIL concentration) is transformed in a more noisy pattern of caspase-8 activation and an even more noisy pattern of cell fate decision because of cell-intrinsic heterogeneity and stochasticity.

Different doses between 125 and 225 ng/mL were tested. In addition, two values for the radius reduction rate for apoptotic cells were tested (1 and 10 $\mu m$ per hour). First, for 200 ng/mL the long-term evolution of population size of the spheroids was undistinguishable from the corresponding 'monolayer' simulation (Figure 3.16, top). Thus, treatments that were sufficient to control the proliferation of monolayers are predicted to also be sufficient for spheroids. However, for slightly lower doses significant differences of the long-term evolution are seen (Figure 3.16, bottom). Importantly, in that case the speed at which apoptotic cells are lysed has an influence. Those results suggest that if only geometrical effects are at play, the key factor to consider for the long-term treatment of spheroids is resistance acquisition, which can be assessed on monolayer cultures. Thus, if experimentally strong differences between monolayers and spheroids are observed, it would point out to other effects, for example ECM-mediated changes in gene expression. Our modeling framework can easily be used to study the impact of such potential effects.

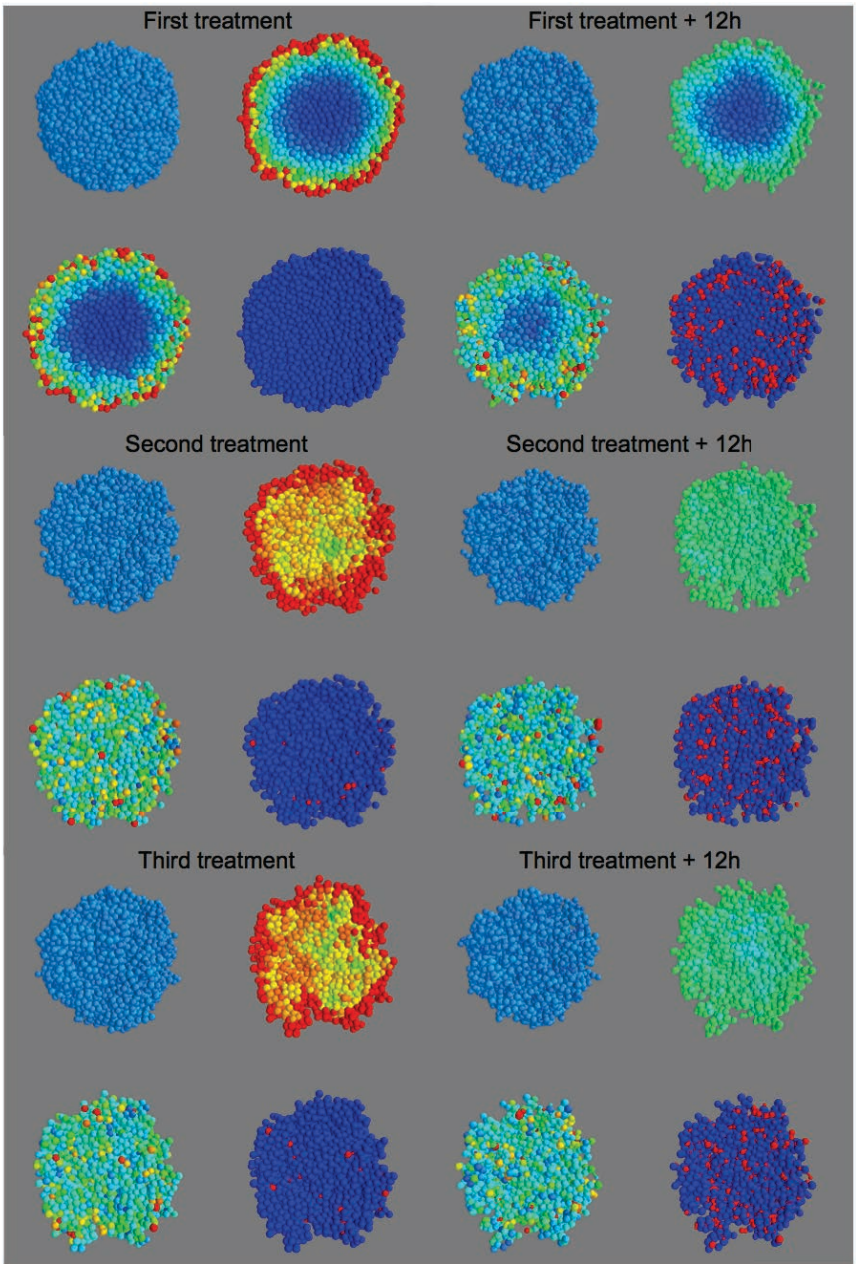



Figure 3.15: **Repeated TRAIL treatments on multi-cellular spheroids.** Visualization of spheroid evolution under repeated TRAIL treatments (125 ng/mL every 3 days, starts when spheroid size reaches 8000 cells). Six time points are shown. In each case, full spheroid (upper-left), TRAIL penetration (upper-right, dark blue: 0 ng/mL, red: 100 ng/mL), caspase-8 activation (bottom-left, dark blue: 0 copies/cell, red: 400 copies/cell) and apoptotic cells (bottom-right, in red) are displayed. Radius reduction rate for apoptotic cells is set to $1.0\,\mu m$ per hour.

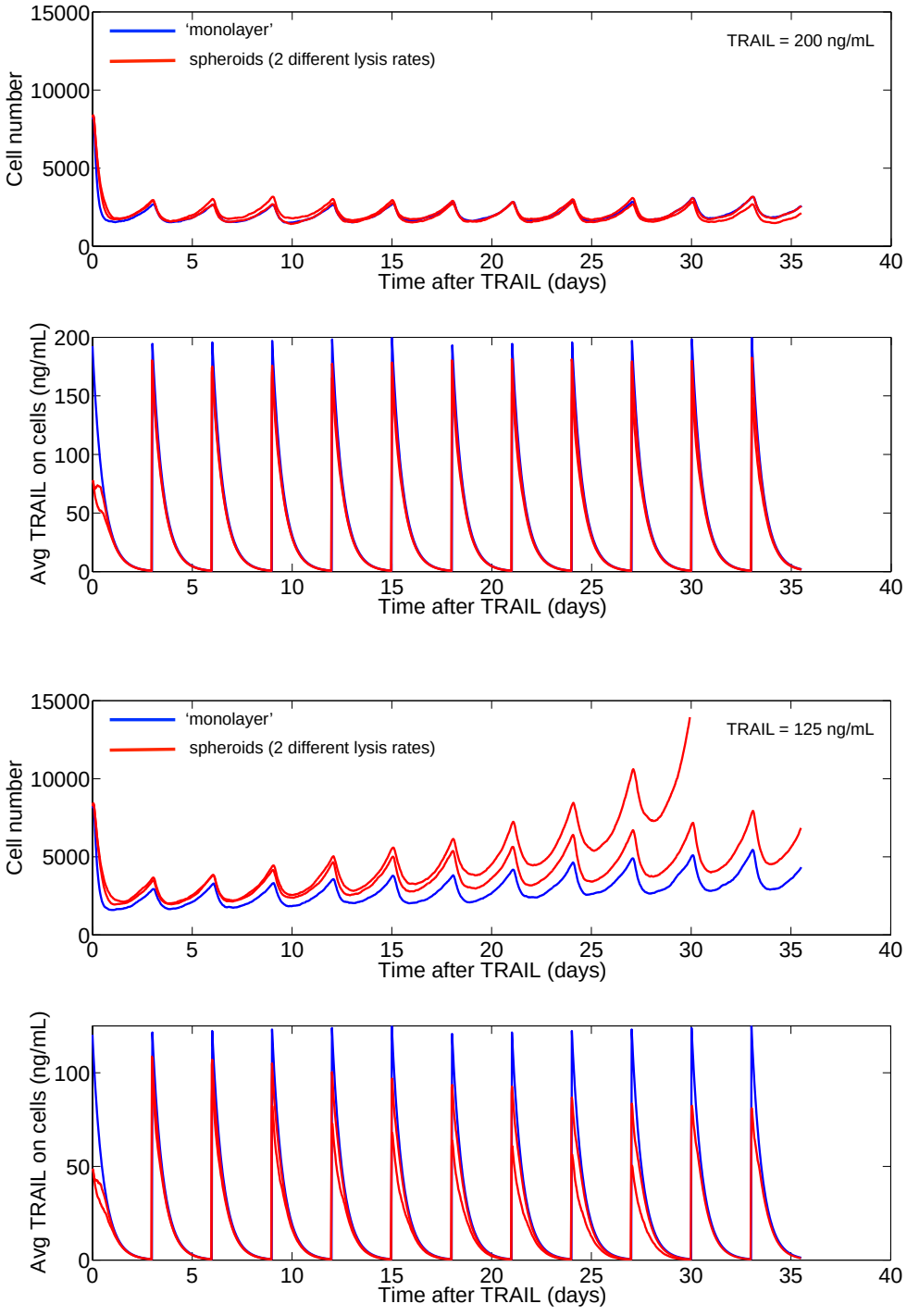



Figure 3.16: **Limited TRAIL penetration in spheroids can exacerbate resistance as compared to monolayers but does not by itself confer long-term protection.** Comparison of long-term outcome of repeated TRAIL treatments between spheroids and 'monolayers'. Two doses are tested: 200 ng/mL and 125 ng/mL. For spheroids, two different radius reduction rates for apoptotic cells are tested (0.1 and 1.0 $\mu m$ per hour). In each case, the upper plot shows the evolution of alive cell number and the lower plot shows the evolution of the (average for spheroids) TRAIL concentration that cells are experiencing.

## 3.4 Discussion

Our current knowledge of the mechanisms that shape the response of *in-vivo* tumors to DRAs is still largely incomplete and imprecise. Yet, the following facts are established: 1) protein synthesis is noisy and causes protein level fluctuations in individual cells; 2) DRA-induced apoptosis is mediated by a signaling cascade involving caspases and a mitochondrial pathway; 3) anti-apoptotic proteins can decrease death signaling by competitive binding; 4) most activated pro-apoptotic proteins as well as two anti-apoptotic proteins are degraded rapidly via the proteasome-ubiquitin pathway.

A quantitative model was constructed to represent how those molecular mechanisms could interact to dictate the response of HeLa cells to TRAIL (see Chapter 2). A parsimonious, realistic parameterization of this model can predict the fractional killing and transient resistance acquisition properties (Figure 2.11), and when combined with a simple model of cell proliferation, it leads to a semi-quantitative agreement with the observed dose-response cytotoxicity assayed 24 hours after TRAIL exposure and can account for the impact of Mcl-1 silencing on this dose-response (Figure 3.2). Thus, despite its limitations (discussed in Chapter 2), this model is uniquely positionned to investigate the population dynamics of HeLa cell populations under TRAIL stimulation.

In addition, because the model is detailed at the molecular level, the insights it can provide are not necessary limited to the HeLa cell line. Indeed, studying the relationship between its parameters and the predicted behavior could improve our understanding of the molecular mechanisms that shape TRAIL resistance for other cell lines, provided those mechanisms rely on species and interactions that are represented in the model.

In this chapter, we have developped a multi-scale model to investigate TRAIL resistance in virtual experimental settings that could bridge the gap between standard *in-vitro* settings and *in-vivo* settings. The first setting corresponds to the long-term (several weeks) response of cancer cell populations grown *in-vitro* in standard conditions (e.g. monolayers) but repeatedly submitted to TRAIL treatments. It is meant to bridge the *temporal* dimension of the gap that exists between standard cytotoxicity assays and *in-vivo* experiments; otherwise stated, it allows to study the contribution of time (and treatment repetitions) to *in-vivo* TRAIL resistance, independently of the resistance that is related to differences in the cells *context* between *in-vitro* and *in-vivo* settings. The second virtual experimental setting we have considered is the treatment of cells grown *in-vitro* as 3-dimensional spheroids (Figure 3.14). This setting is complementary to the first as it allows to investigate the *contextual* dimension (i.e., differences in the micro-environment felt by cells) of the gap between standard *in-vitro* and *in-vivo* experiments.

## Testable predictions about the long-term response of HeLa cells to repeated TRAIL treatments

A first important model prediction that was obtained by simulating repeated TRAIL treatments on HeLa cells is that killing efficiency is strongly reduced between the first and later treatments, this effect being present at all doses that enable killing (Figure 3.4). This prediction has implications for dosing strategies, as here the dose needed to balance cell proliferation on the long-term is 10 times higher that what one would have predicted based on 'standard' (single treatment) cytotoxicity assay.

Testing this prediction experimentally seems technically feasible, and yet it would be highly valuable to quantify the amount of resistance acquisition and the timescales at which it takes place. More generally, we did not find such quantitative experiments (cell number evolution as a function a time for repeated treatments) in the literature, despite the high number of studies investigating TRAIL resistance (rather, researchers often conducted escalating treatments over several weeks to obtain resistant cell populations used as a material to study the biochemical basis of TRAIL resistance (Cheng et al. 2006; Lane et al. 2006; Wang et al. 2008; Yoshida et al. 2009)). Still, based on a recent experimental study in which the efficiency of two consecutive TRAIL treatments were compared (Flusberg et al. 2013), it is likely that the strong increase of long-term killing efficiency (with respect to the initial killing efficiency) that we predict for HeLa cells also holds for other cancer cell lines.

Despite the resistance acquisition between initial and consecutive treatments, simulations also predicted that the best strategy to eliminate or control the size of a cell population using TRAIL is to maintain a constant stimulation. The obtained killing rate would be lower than the maximum instantaneous killing rate obtained when applying a single treatment on a 'naïve' population, yet it would be higher than the average killing rate obtained on the long-term when repeating 'pulse' treatments. This prediction can also be tested experimentally.

Testing this could be relevant for the therapeutical use of DRA's, since the frequency of administration and the stability of the DRA (that spans more than an order of magnitude between rhTRAIL and DR4/DR5 antibodies) dictate the shape of the temporal drug profile sensed by the tumor. This would require a precise control on the TRAIL concentration in the cell micro-environment. This can be achieved with microfluidic devices, an approach that was successfully applied to investigate NF-κB signaling (Kellogg & Tay 2015). However, such technology has not yet been applied to investigate TRAIL resistance (but one study used such a device to study the efficacy of combinatorial treatment of TRAIL and Doxorubicin (Kim et al. 2012)).

## Potential impact of non-canonical signaling

Although we did not check whether this predicted property (i.e. that constant TRAIL stimulation the best strategy to limit cell proliferation on the long-term) is valid across a wide range of parameters, we believe it is likely the case and that this property is related to the model structure rather than specific parameter values. In fact, different but comparable model structures might also lead to the same behavior, given that it was also predicted by our 3-parameters phenomenological model (Figure 3.12).

If it is experimentally invalidated, it would indicate that effects not captured by the model have an important role. A likely candidate would be a pleiotropic effect (Hart et al. 2014) of TRAIL, which might have a positive influence on cell proliferation (Ehrhardt et al. 2003) and/or the level or activity of anti-apoptotic proteins (through transcriptional, post-transcriptional or post-translational regulation) (Falschlehner et al. 2007) that is masked by cell death at high doses but dominant at low doses.

Recently, TRAIL resistance acquisition has been experimentally characterized in detail in another cell line (MCF10A, (Flusberg et al. 2013)). One important finding of this study is that a cell population can gain resistance to TRAIL in absence of cell killing: pre-treatment with TRAIL in presence of a caspase-8 inhibitor (hence preventing cell death even if caspase-8 is activated) led to reduced killing (as compared to naïve cells) when a normal TRAIL treatment was subsequently applied. This shows that at least for this cell line, resistance acquisition cannot be attributed solely to selection effects, and that non-lethal TRAIL treatments can increase the resistance of individual cells. This is consistent with our simulation results (Figure 3.9), which also showed (for our reference model parameterization for HeLa cells) that the contribution of selection effects to resistance acquisition is not dominant. Their data also shows that this cell-level resistance increase occurs at the level of or upstream of caspase-8, as downstream signaling was blocked. Again, in order to get deeper insights into the mechanisms at play, it will be interesting to investigate this in HeLa cells and quantitatively compare the results with model predictions.

## Dissecting distinct types of contributions to TRAIL resistance

In agreement with experimental evidence (Flusberg et al. 2013), our simulation results show that even in absence of genetic mutations, the resistance to TRAIL of a cell population is a not a fixed, intrinsic trait but is dynamically shaped by the TRAIL stimulation profile it is submitted to. Thus, to predict the outcome of repeated TRAIL treatments, knowing the killing efficiency of single treatments is not sufficient. In addition, the 'capacity' of the cell line to gain resistance when treatments are repeated should also be quantified. This capacity can be defined as the ratio between the 'long-term' killing efficiency (i.e. for a population that has been already treated several times) and the 'initial' killing efficiency (for a naïve popu-

lation). By studying the relationship between the model parameters and the corresponding values of those two quantities, we found that different types of parameter changes affect them differently.

First, modifying the mean expression levels of pro- or anti-apoptotic proteins has an impact on both the initial and long-term killing efficiency, however, the ratio between the two quantities remain unchanged. Thus, the extent to which cell population can gain resistance is not controlled by the expression levels of those proteins.

Second, modifying the rates of targeted degradation for non-native species (such as activated caspase 8, etc…) had a very strong impact on both initial and long-term killing efficiency, with half-lives varying from 1 to 5 hours predicting a large spectrum of highly resistant to highly sensitive cell lines (Figure 3.10). Those results are consistent with the well established observation that co-treatment with proteasome inhibitors such as bortezomib strongly sensitize cancer cells to TRAIL-induced apoptosis (Menke et al. 2011; Leverkus et al. 2003; Naumann et al. 2011). While initial systems biology studies on TRAIL-induced apoptosis focused on linking difference in expression levels with the sensitivity of different cell lines (Rehm et al. 2006; Spencer et al. 2009; Aldridge et al. 2011; Albeck, Burke, Aldridge, et al. 2008), our modeling results point towards a crucial role of the targeted degradation of activated pro-apoptotic proteins. Recently, experimental analysis of caspase-8 activity dynamics in single HeLa cells revealed that proteasomal degradation is responsible for stopping and reverting the increase in activated caspase-8 proteins a few hours after TRAIL induction (Roux et al. 2015). More generally, it is becoming more and more appreciated that protein degradation often plays a key role in the quantitative functioning of signal transduction pathways (Loriaux & Hoffmann 2013).

It should be noted that in our model, most of the targeted degradation reactions are modeled with constant rate first order reactions. However in reality, many biochemical steps are probably involved, with the rate-limiting step being controlled by another protein (in fact, this is already the case for XIAP-mediated degradation of activated caspase-3, as in the original model (Albeck, Burke, Aldridge, et al. 2008); also, as we assume that Mcl-1 and Flip complexes with tBid and R* are degraded at the same rate as Mcl-1 and Flip, those two proteins are effectively targeting tBid and R* for degradation). Explicitely modeling all the proteins involved in the targeted degradation of the core apoptosis signaling might be needed to gain a more accurate and complete picture of the regulation of TRAIL-induced apoptosis, and hence of TRAIL resistance acquisition. Indeed, a study has found that the half-life of Flip and Mcl-1 was markedly changed in lung cancer cell lines that acquired TRAIL resistance through weeks-long treatments with escalating TRAIL doses (Wang et al. 2008).

While our model predicts a strong impact of non-native species degradation on TRAIL resistance, both initial and long-term killing are predicted to be affected equally, as their ra-

tio remained constant. Thus, such degradation is not responsible for the gain of resistance arising when treatments are repeated. Searching for parameters that are implicated in resistance acquisition *per se*, we found that the 'dilution' rate, i.e. the rate at which stable proteins (all except Flip and Mcl-1 in our model) are 'degraded' because of cell growth, strongly impacts the ratio between initial and long-term killing efficiency without affecting much initial killing efficiency. Thus, our model predicts that resistance acquisition is highly dependent on the turnover of 'stable' proteins.

Those results can be interpreted as follows: upon the first TRAIL stimulation, targeted degradation of activated pro-apoptotic proteins, allows some cells to survive at the price of decreased levels of native pro-apoptotic proteins (and hence increased resistance). Recovery of those levels occurs slowly, at the rate of cell proliferation, because those proteins are mostly stable. This view can be contrasted by noting that the decrease in native pro-apoptotic proteins could be accompanied by a similar decrease of anti-apoptotic proteins that participated in the targeted degradation, such as for R/Flip, Bid/Mcl-1 and Bax/Bcl2 in our model. However, because some anti-apoptotic proteins exhibit a fast turnover, they rapidly recover their steady-state levels and hence there is a transient imbalance towards relatively more abundant anti-apoptotic proteins.

To formalize this heuristic explanation at a level that abstracts specific biochemical reactions, we have proposed a 3-parameters phenomenological model (Figure 3.12). In that model, both selection effects and cell-level increase of resistance in surviving cells are abstracted by a single parameter: the 'cost' that translating death signals into cell killing has in terms of the capacity to translate subsequent death signals. Interestingly, this model captures essential features of the mechanistic TRAIL-induced apoptosis model, such as the prediction that constant stimulation is the best strategy to limit long-term proliferation. This generic model provides a conceptual link between molecular mechanisms and the high-level, population-scale dynamics of cell killing for any kind of cytotoxic drug and hence might be useful in the context of high-dimensional drug screening studies where the response of many cell lines to many drugs need to be interpreted. It would be also interesting to develop a variant of this model in which a pleiotropic effect of the drug is accounted for.

## Relevance of non-genetic resistance to DRAs *in-vivo*

It is clear that for isogenic cancer cell lines cultivated *in-vitro* for timescales of several weeks or below, the contribution of mutational mechanisms to resistance acquisition is limited. Resistance acquisition rather arises from transcriptional and post-transcriptional mechanisms, such as in our model. Still, when longer time-scales are considered, it is possible that the contribution of mutational mechanisms to resistance acquisition starts to be dominant.

However, when month-long *in-vitro* selection of TRAIL resistant 'clones' is performed

by cultivating cells in media with escalating TRAIL concentrations, the resistance of those clones is often found to be reversed following culture in TRAIL-free medium (for example, in a study using leukemia cells it was the case for one out of two of tested 'clones' (Cheng et al. 2006)). Thus, non-mutational resistance acquisition is also relevant for month-timescales.

In addition to the role of timescales, the difference in cellular context between *in-vitro* and *in-vivo* cancer cell populations might also influence to relative contribution of mutational versus non-mutational mechanisms to TRAIL resistance. Although in most *in-vivo* experimental studies of TRAIL resistance, the data and protocol is such that it is difficult to conclude about the presence or absence of resistance acquisition when treatment is repeated over time, we found strong indications for non-mutational resistance acquisition to DRA treatment in a recent study (Huet et al. 2014): in Figure 5 of the corresponding paper, the first treatment allows a reduction of the tumor volume, while its not the case for the second treatment applied one week later (for both types of DRAs tested). Thus, non-mutational resistance acquisition to DRAs is probably relevant *in-vivo*. This means that in order to quantitatively understand the dynamics of TRAIL resistance for *in-vivo* tumors, it might be needed to combine intra-cellular models of extrinsic apoptosis such as ours (capable of predicting cell-intrinsic, non-mutational resistance acquisition) with agent-based models of tumor growth (in which the effect of cell-extrinsic factors can be modeled).

## The role of cell-extrinsic factors in resistance acquisition

As a step towards quantitative, cell-based multi-scale models of TRAIL resistance for *in-vivo* tumors, we have simulated the response of multi-cellular spheroids to repeated TRAIL treatments. Our results suggest that limited TRAIL penetration within a tumor is probably not a major source of resistance (Figure 3.16), as even when assuming slow diffusion and fast degradation of TRAIL within the spheroid, no strong differences of long-term outcome of repeated treatments were predicted for treatment strategies that were efficient on 'monolayer' (in the sense that all cells face the same environment) populations.

Still, it is yet well accepted that the drug resistance of 2D vs 3D cultures of the same cells can be strongly different (Pampaloni et al. 2007; Yamada & Cukierman 2007). If those differences do not arise from limited drug penetration, it means that cell-level differences in drug resistance exist. Those differences will be caused by differences in gene expression profiles that would themselves be caused by differences in cell micro-environment.

In the specific case of TRAIL, the differential response of monolayers versus spheroids has almost not been investigated. However, a study of the response of mesothelioma cell lines to bortezomib (which by itself can induce apoptosis in some cancer cell lines) has been conducted (Barbone et al. 2011). They found that the same cells were much more resistant when grown as spheroids, and found important differences in the expression levels of many

pro- and anti- apoptotic proteins, but could not make specific links between those changes and the change in resistance. Our modeling approach provides a valuable framework to quantitatively investigate this question, and it will be interesting to quantify experimentally the differential response of HeLa spheroids and monolayers to repeated TRAIL treatments.

Even if we manage to successfully map changes in gene expression between the two culture conditions with the corresponding response to TRAIL resistance, the cause of those gene expression changes will still be unknown. Understanding the underlying mechanisms would potentially reveal molecular targets to prevent pro-resistance gene expression changes to take place. Striking evidence for such changes was obtained for Ewing's sarcoma cell lines (Merchant et al. 2004). In this study, tumor cells implanted and grown on mice (in absence of treatment) were then explanted and grown *in-vitro*. Their resistance was much higher than the parental cell line, showing that simply growing in an *in-vivo* context led to the development of TRAIL resistance in those cells, and that this resistance is not lost immediately after change of the environment (but it was lost after several passages *in-vitro*, showing it is does not involve mutational mechanisms). The underlying causes for this non-TRAIL mediated resistance acquisition are yet unknown. Again, our modeling framework could be useful to investigate them.

# Chapter 4

# Modeling an artificial yeast sensing device from single cells to spatially organized micro-colonies

## 4.1 Motivations: towards spatial self-organization of cell populations

Multi-cellular synthetic biology aims to extend the scope of synthetic biology by implementing tissue-level functions in living systems. This task is highly challenging because of the multi-scale gap between what should be modified (genes in cells) and what should be obtained (a desired behavior of the whole cell population).

An early success of multi-cellular synthetic biology is the implementation of *population control* in liquid cultures of *E. coli* bacteria (You et al. 2004). This system relies on a simple design: cells are implemented with the ability to sense the size of the population and to undergo cell death for too high population sizes. This was realized by using the AHL cell-cell communication system from *Vibrio fischeri* (AHL is a small, diffusible molecule that can be synthized or sensed by cells) in combination with a gene coding for the toxic protein CcdB (i.e. causing cell death when present in sufficient amount in the cell), which production was placed under the positive control of AHL sensing. This system worked well in the sense that different target densities could be chosen by tuning AHL half-life and that those target densities were maintened for durations as long as 60 hours.

### Engineering spatial multi-cellular systems

In this example, cells of the population were cultured in liquid medium. Several others multi-cellular synthetic systems were demonstrated for populations growing in liquid cul-

tures, for example, a prey-predator system (Balagaddé et al. 2008) and a system involving altruistic cell death (Tanouchi et al. 2012). However, in such cases the cell population do not exhibit any spatial organization, hence it does not really form a *tissue*. For many potential applications of multi-cellular synthetic biology (such as tissue engineering, synthetic organogenesis, engineered biomaterials…), cells of the population are expected to be spatially organized.

"Quorum sensing" cell-cell communication (when cells can both synthetize and sense a messenger molecule) is in principle useful for implementing spatial organization, as it provides cells information about their social context. An early demonstration of such potential was brought by Basu & Weiss. They engineered *E. coli* to be either AHL "senders" (synthetizing and releasing AHL) or "receivers" (expressing fluorescent proteins in response to AHL) and obtained with this system geometric patterns of fluorescence, which shapes depend on the spatial configuration of sender cell colonies. In that example as well as several others (Chen & Weiss 2005; Tabor et al. 2009; Payne et al. 2013), control over the population spatial organization is only exerted on the *phenotypes* of the cells (expression/secretion of certain molecules, etc.), but the spatial configuration of cells is itself not controlled, and is mainly dictated by the endogenous growth program and the environment. In other terms, the size or the shape of the tissue are not subjected to self-organization (other than natural).

Engineered spatial systems with such self-organization capabilities still remain very limited. One notable exception is the obtention of ring-shaped patterns of the cell density in bacteria by combining cell-cell communication with a genetic circuit controlling cell micromotility (Liu et al. 2011). A possible explanation for this limited success is the involvement of multiple spatial and temporal scales (cell growth, proliferation and death; diffusion; signal transduction; gene expression; etc.) that make it difficult to reason about the system. In addition, the impact of the (unavoidable) cell-to-cell variability on system functioning is hard to predict and might be fundamentally different between non-spatial (e.g., populations grown in liquid culture) and spatial systems.

Cell-based multi-scale modeling is a priori well suited to address those challenges, as it allows the explicit representation of individual cells in space and renders it possible to account for both cell-intrinsic (such as stochastic gene expression) and cell-extrinsic (such as gradients of messenger molecules) sources of variability in cell state and fate. In this Chapter, I present the application of cell-based multi-scale modeling for guiding the development of an artificial yeast patterning system.

## Case study: implementing spatial patterning in yeast

In principle, spatial patterns in the configuration of cells in a population can be generated from simple *local rules* driving cell fate. A famous example is the mathematical object called

the 'Game of Life' (Conway 1970). In the 'Game of Life', a 2D square lattice composed of cells that are either 'occupied' (i.e. alive) or 'un-occupied' evolves with time in a discrete and deterministic fashion: alive cells can die either from loneliness or over-crowding, and un-occupied cells can become alive if they have an adequate number of alive neighbors (to represent cell division). The fact that such simple rules can generate rich and complex dynamic patterns depending on the initial configuration fascinated many scientists.

Drawing inspiration from the 'Game of Life' rules, the Weiss lab at MIT proposed to construct a patterning system in yeast. The high-level design for this system is illustrated in Figure 4.1. It relies on combining density-sensing (the ability for cells to sense the local cell density) with cell death programs such that individual cells would die if they are either isolated or overcrowded while growing normally for intermediate local cell densities.

A natural choice to implement density-sensing in yeast cells is the previously developped artificial cell-cell communication system based on the plant cytokine IP (Chen & Weiss 2005). Therefore, they proposed a modular implementation of the patterning system (Figure 4.2). In addition to two modules for IP production and IP sensing, it contains two cell death modules connected to IP sensing meant to function either for too high IP sensing (high-threshold killing) or for too low IP sensing (low-threshold killing).

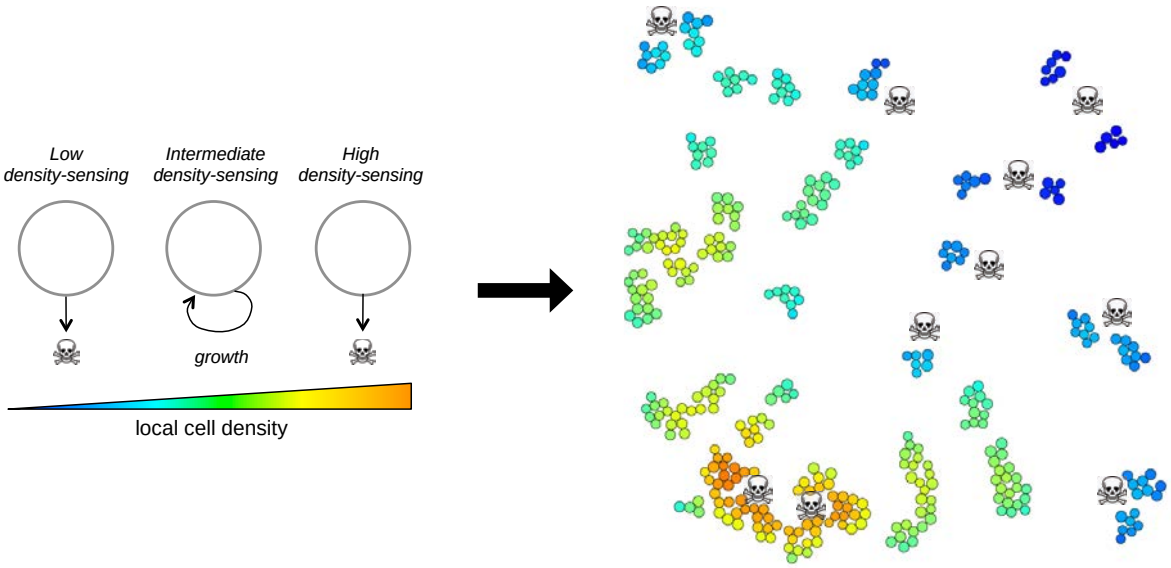



Figure 4.1: **High-level design of a patterning system inspired from Conway 'Game of Life'.** Engineered cells would be able to sense their local cell density and to die if it is either too low or too high (left). Only intermediate densities would allow the cell to grow. The functioning of this design for a given 2D spatial configuration of cells is illustrated (right). The color of each cell represents the local density it is sensing. Group of cells that would die because of 'overcrowding' or 'isolation' are indicated. Other cells would grow normally.

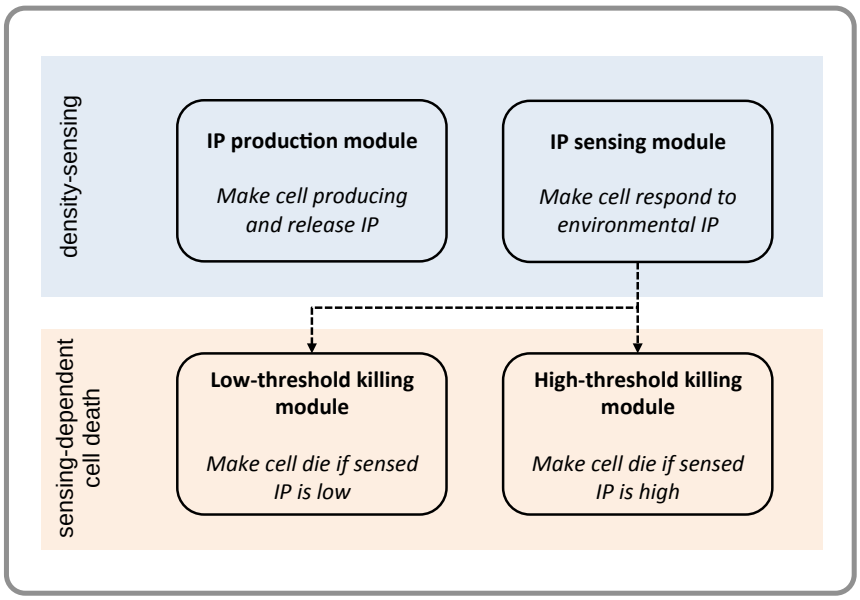



Figure 4.2: **Proposed modular implementation of the yeast patterning system.** Density-sensing is achieved using the IP-based cell-cell communication system (IP is a small, diffusible cytokine from the plant *A. thaliana*) demonstrated earlier (Chen & Weiss 2005). It relies on two separate modules for IP production on one side and IP sensing on the other. In addition, two distinct killing modules implement IP sensing -dependent cell death. The "low-threshold" (LT) killing module is active when IP sensing is low, while the "high-threshold" (HT) killing module is active when IP sensing is high. For intermediate IP sensing, both killing modules are inactive.

## Context and objective of the collaboration

When we started collaborating with the Weiss lab on this project, a genetic implementation of this system was already constructed by Jing-Jing Sun (Sun 2014), building upon work from Ming-Tang Chen (Chen 2008). A description of the corresponding genetic circuits is shown in Figure 4.3. Briefly, high-threshold killing is achieved via excision of the vital gene pkc1 (involved in cell wall integrity pathway) while low-threshold killing is achieved via the apoptotis-inducing protein Bax.

However, testing, tuning and debugging of this system is difficult. First, a brute-force exploration of the system behavior is not feasible given the high dimensionality of the space of possible experiments (three different knobs for tuning the IP production and the killing modules; also, the initial spatial configuration of cells is in principle critical in determining its evolution; and finally the dimension and physical properties of the solid medium in which cells are cultured a priori also impacts system functioning by modulating the dynamics of IP gradients). Moreover, as already mentioned, intuitive reasoning about the system behavior is also difficult because it involves multiple spatial and temporal scales.

Hence, our objective in this collaboration was to use bottom-up cell-based multi-scale modeling to help building a quantitative understanding of the system functioning. Such understanding is expected in turn to permit more rational and efficient model testing, tun-

ing and debugging. Importantly, we aimed to exploit as much as possible the system modularity.

More precisely, we used liquid culture data on the IP-induced response of receiver cells (i.e. bearing only IP sensing module driving GFP expression, see Figure 4.3) to construct a quantitative, single-cell model of IP sensing. Then, we integrated this model into a spatial cell-based multi-scale model reproducing solid culture experiments characterizing the response of receiver cell micro-colonies to gradients of IP. We used this model to 1) infer the diffusion coefficient of IP in the medium and 2) assess the ability of the knowledge gained from liquid experiments to predict the behavior of the same cells cultured as spatially organized populations (thereby assessing transposability from liquid to solid). Finally, we parsimoniously modified the model of IP sensing to model the behavior of the high threshold killing in response to IP. We used it to predict the evolution of patterning cell micro-colonies for which high threshold killing was activated in response to IP, and compared the results with corresponding experiments.

## 4.2 Existing experimental data characterizing system behavior

Here we describe the experimental data that we used either to estimate parameters of our models or to verify their prediction capabilities. First, in order to quantitatively characterize IP sensing module behavior for receiver cells grown in liquid culture, a high-throughput flow cytometry experiment was performed by our collaborators to measure IP-induced GFP expression dynamics with high dose and temporal resolution and at the single-cell level (Figure 4.4).

However, the final system is meant to function in solid cultures and in the presence of IP gradients. In addition, such gradients should exist at a sufficiently small scale, for example between close micro-colonies. To characterize the behavior of IP sensing for receiver cell micro-colonies in presence of IP gradients, an experimental setup making use of a *cell printer* (a technology that enables to deposit very few cells at defined locations on a surface with micrometric precision) was developped by our collaborators (Figure 4.5). As we will see, a differential response in space and time was obtained between micro-colonies, indicating that sufficient gradients of IP to allow system functioning can exist at this spatial scale.

Finally, the functioning of the high-threshold killing module was tested by printing patterning cells on a agar gel containing Dox (needed for its functioning) and IP (for activating it) (Figure 4.6). This data showed significant killing in the induced micro-colonies, albeit many hours after seeding cells. Interestingly, not all cells died, and the localization of dead or alive cells in the micro-colonies does not seem random: large clusters of dead cells are apparent. Because ultimately, the desired behavior of the system is the obtention of clear, dynamic, predictable spatial patterns of dead or alive cells, it would be highly valuable to

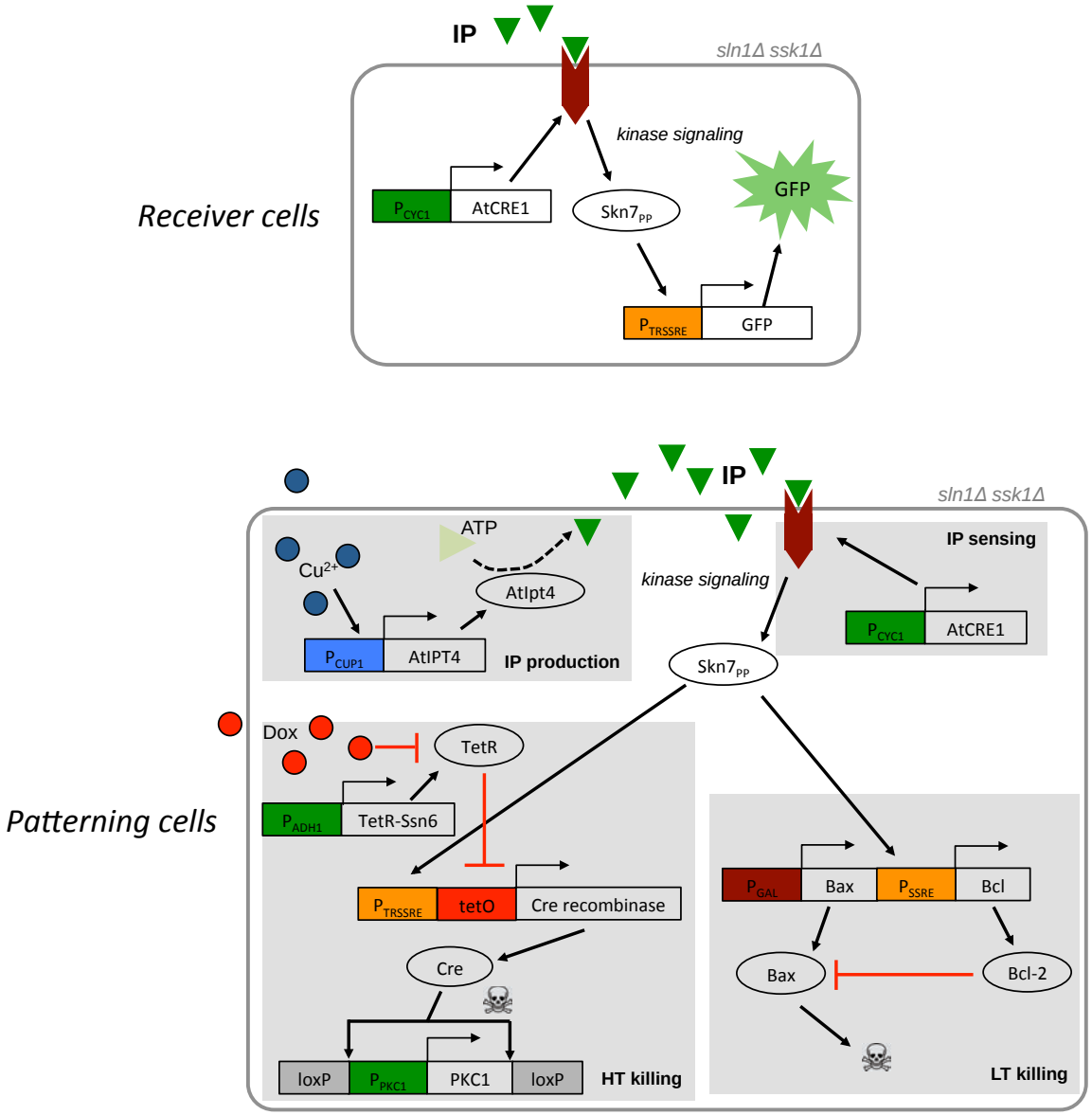



Figure 4.3: **Genetic implementation of the yeast patterning system.** Receiver cells, that only bear the IP sensing module driving GFP expression, are also shown. For patterning cells, the distinct modules (IP production, IP sensing, high- and low- threshold killing) are highlighted. In both cases, constitutive promoters are indicated in green. Other promoters are either inducible by copper (blue), galactose (brown) or by IP via the IP sensing module (orange). For the high-threshold killing module, strict repression of Cre synthesis is achieved by TetR and is relieved only when Dox is added to the medium. Therefore, in principle, in patterning cells IP production, high-threshold killing and low-threshold killing are active only if respectively copper, Dox or galactose are present in the medium. High-threshold killing is achieved by Cre-mediated excision of the essential gene pkc1 (which has been flanked with loxP sequences). Low-threshold killing is achieved by Bax-mediated apoptosis induction (in presence of IP, this induction is prevented via the production of the Bax antagonist Bcl-2). In both strains, endogenous genes sln1 and ssk1 have been deleted to minimize non IP-induced activation of Skn7. Those circuits were constructed by Jing-Jing Sun (Sun 2014) and used parts developped earlier by Ming-Tang Chen (Chen 2008).

gain a quantitative understanding of the mechanisms driving the apparition of such death clusters.

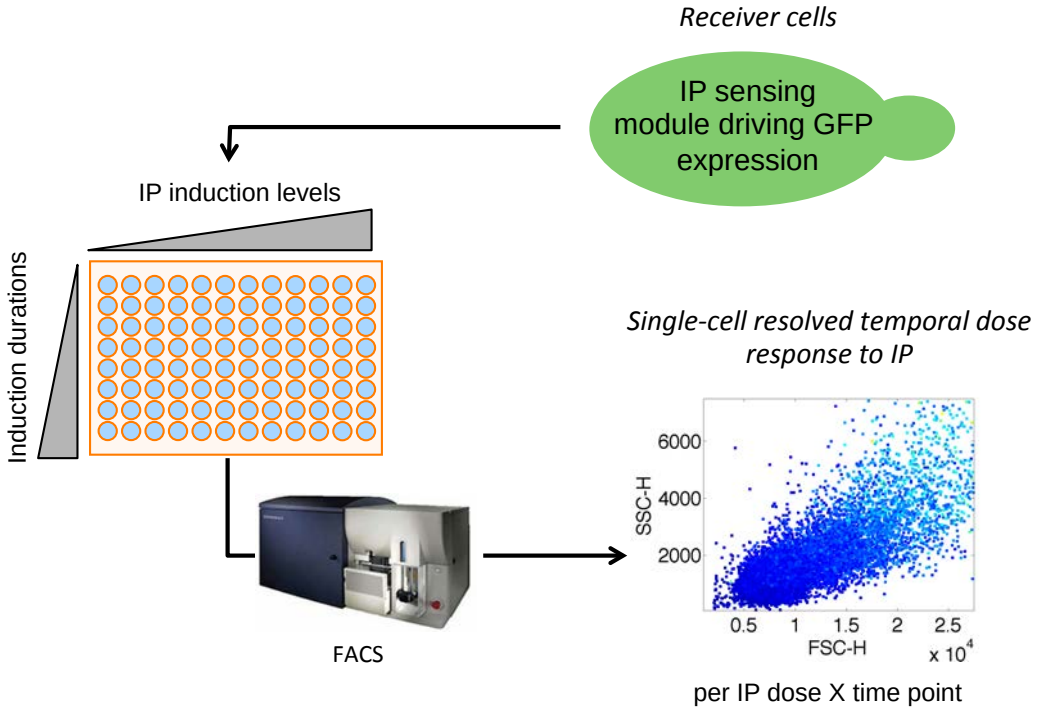



Figure 4.4: **Experimental characterization of the IP sensing module behavior in liquid culture.** Performed by Brian Teague at MIT. *Receiver cells*, a yeast strain in which an IP sensing module (see Figure 4.3) driving GFP expression has been integrated in the genome, were grown in liquid cultures containing various IP concentrations and their fluorescence was analyzed by flow cytometry at various time points after induction. This experiment was repeated for two different initial cell densities ($OD = 0.05$ and $OD = 0.2$).

## 4.3 Modeling system behavior

In this section we describe the different computational models of the system that we developped. Their calibration and the assessment of their prediction capabilities will be presented afterwards.

### A semi-mechanistic ODE model of IP sensing

IP sensing is central to the functioning of patterning cells because both killing circuits rely on it. Hence, we focused first on obtaining a quantitative model of IP sensing that will be calibrated from the data characterizing receiver cells response to IP in liquid culture (Figure 4.4).

Because the precise mechanisms by which yeast cells expressing the (exogenous) AtCre1 receptor can activate the nuclear factor Skn7 in response to IP are yet not fully understood,

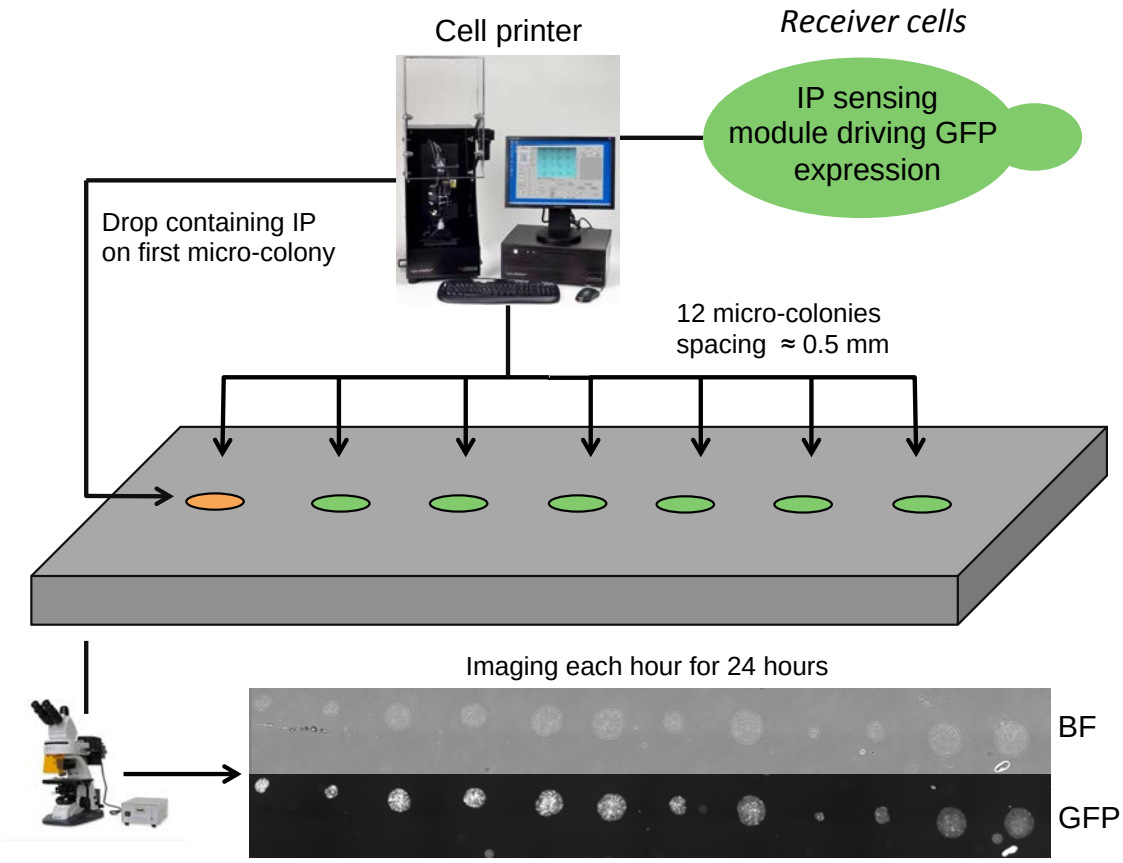



Figure 4.5: **Experimental characterization of the IP sensing module behavior in solid medium and in presence of IP gradients.** Performed by Brian Teague at MIT. A line of receiver cell micro-colonies was seeded using a cell printer (~5-20 cells per colony initially) on a polyacrylamide gel and a drop containing IP was deposited on the first colony, hence initiating IP gradients in the medium. The colonies were grown for 24 hours in a microscope chamber and brightfield and fluorescence images were acquired every hour.

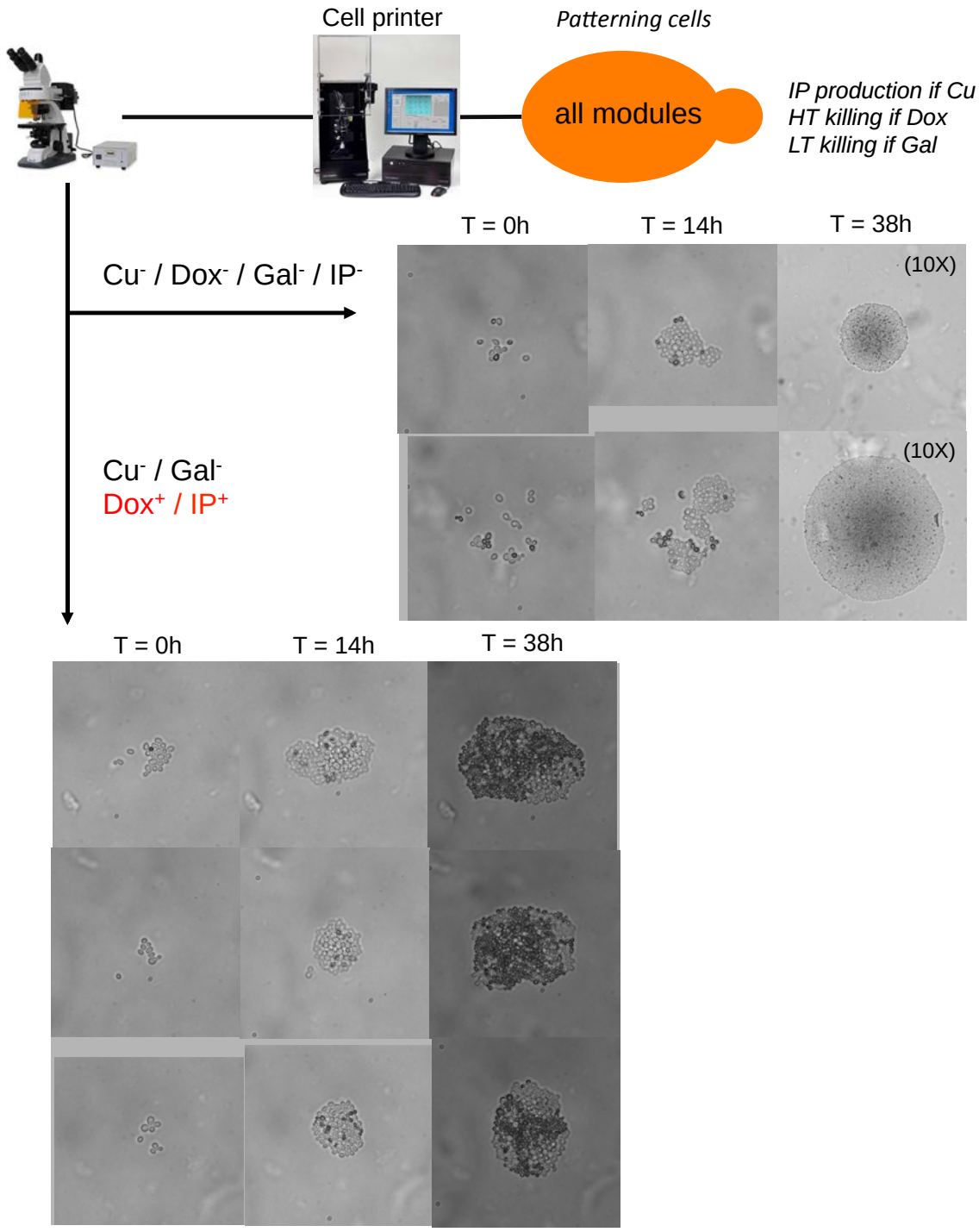



Figure 4.6: **Experimental characterization of the high-threshold killing module behavior in solid medium.** Performed by Brian Teague at MIT. Patterning cell micro-colonies were seeded using a cell printer on a agar gel. All gels also contained Phloxine B (a dye that accumulates in dead cells) and imaging was performed such that dead cells appear black. In a control experiment, the gel did not contain Dox, Cu, Gal or IP and the cells eventually formed large micro-colonies (right). To test the functioning of the high-treshold killing module, a gel containing Dox and IP was used. Strong killing occured between 14h and 38h and resulted in clusters of dead and alive cells.

we follow the approach adopted in (Palani & Sarkar 2011) by considering a semi-mechanistic ODE model of IP signaling, in which a single intermediate between receptor binding and transcription activation is represented (Figure 4.7, top). The corresponding reactions are given in Table 4.1. In addition to this structural choice, we also decided to normalize the expression levels of R, TF and (basal) GFP in order to minimize identifiability issues.

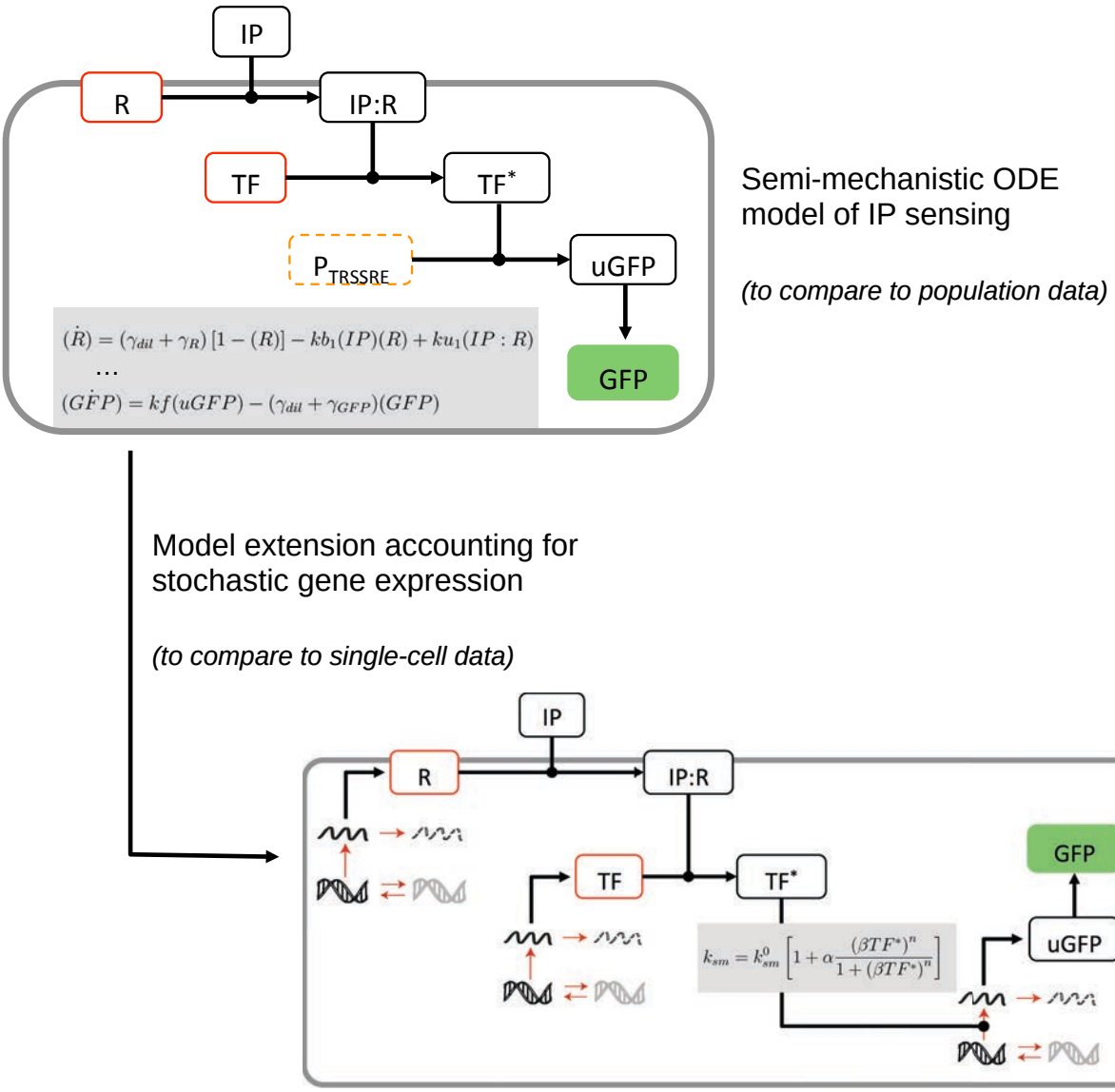



Figure 4.7: **Modeling the IP sensing module behavior in cells.** Top: diagram of a semi-mechanistic kinetic model to describe the behavior of the IP sensing module in cells. It abstracts IP signaling with 17 reactions (15 rate parameters) involving 8 species: IP, its receptor and the complex with IP, an inactive transcription factor, its complex with the IP-bound receptor, the activated transcription factor and finally GFP (for which a maturation step is considered). Corresponding reactions are given in Table 4.1. They are assumed to follow mass-action kinetics. Bottom: extension of this model accounting for stochastic gene expression. Stochastic transcriptional bursting models are driving the synthesis of R, TF and uGFP as in Chapter 2 for the TRAIL-induced apoptosis model. In addition, to account for TF* -mediated synthesis of uGFP, the transcription rate of uGFP mRNA is made dependent on TF* level.

Table 4.1: Reactions of the IP sensing semi-mechanistic ODE model. Note that because IP is supposed to be in large excess in medium, IP binding does not reduce extra-cellular IP concentration. (*): such that steady-state level (in absence of IP) is 1. (**): such that GFP steady-state level (in absence of IP) is 1.

| Reaction | Description | Rate |
|---|---|---|
| $IP + R \rightarrow IP:R + IP$ | IP binding | $kb_1$ |
| $IP:R \rightarrow R$ | IP unbinding | $ku_1$ |
| $IP:R + TF \rightarrow IP:R:TF$ | TF binding | $kb_2$ |
| $IP:R:TF \rightarrow IP:R + TF$ | TF unbinding | $ku_2$ |
| $IP:R:TF \rightarrow IP:R + TF^*$ | TF activation | $kc_1$ |
| $TF^* \rightarrow TF^* + uGFP$ | TF-dependent uGFP synthesis | $\sigma$ |
| $uGFP \rightarrow GFP$ | GFP maturation | $kf$ |
| $R \rightarrow \emptyset$ | R degradation | $\gamma_{dil} + \gamma_R$ |
| $\rightarrow R$ | R synthesis | $\gamma_{dil} + \gamma_R$ (*) |
| $TF \rightarrow \emptyset$ | TF degradation | $\gamma_{dil} + \gamma_{TF}$ |
| $\rightarrow TF$ | TF synthesis | $\gamma_{dil} + \gamma_{TF}$ (*) |
| $uGFP \rightarrow \emptyset$ | uGFP degradation | $\gamma_{dil} + \gamma_{uGFP}$ |
| $\rightarrow uGFP$ | uGFP synthesis | $(\gamma_{dil} + \gamma_{uGFP} + kf)\frac{\gamma_{dil}+\gamma_{GFP}}{kf}$ (**) |
| $GFP \rightarrow \emptyset$ | GFP degradation | $\gamma_{dil} + \gamma_{GFP}$ |
| $IP:R \rightarrow \emptyset$ | IP:R degradation | $\gamma_{dil} + \gamma_{IP:R}$ |
| $IP:R:TF \rightarrow \emptyset$ | IP:R:TF degradation | $\gamma_{dil} + \gamma_{IP:R:TF}$ |
| $TF^* \rightarrow \emptyset$ | TF* degradation | $\gamma_{dil} + \gamma_{TF^*}$ |
| $IP \rightarrow \emptyset$ | IP degradation in medium | $\gamma_{IP}$ |

## Model extension accounting for stochastic gene expression

This model is deterministic and is meant to be compared to population averaged data. However, under this form it cannot capture cell-to-cell variability in the response to IP, while such variability might be important for the functioning of the full patterning system. As we have shown in Chapter 2, combining this model with randomly distributed initial conditions to mimic cell-to-cell differences in protein level is not appropriate for timescales of the order of the cell cycle or longer. Therefore, we followed the same approach as for TRAIL-induced

apoptosis by extending the model with stochastic gene expression ( see Figure 4.7, bottom).

Importantly, as opposed to our model of TRAIL-induced apoptosis, here one the protein species regulates the transcription of another gene. We propose to model this transcriptional regulation by allowing the transcription rate of GFP (i.e., the stochastic rate at which mRNA are produced when the gene is active) to depend on the time-varying concentration of the transcription factor TF*. We assume that this dependence can be modeled with a Hill function. Note that under those assumptions, cells in which the promoter driving GFP expression is inactive would not directly produce mRNA when the transcription factor becomes activated, and that the transcription factor do not modulate the rates of promoter activity switches.

## Cell-based multi-scale model of receiver cell micro-colonies

To model the behavior of receiver cell micro-colonies in presence of IP gradients, we proceeded essentially as for modeling TRAIL treatments on multi-cellular spheroids (Chapter 3). More precisely, we integrated an instance of the single-cell IP sensing model into each cell of spatial, agent-based multi-cellular simulations in which cell growth, division and motion as well as diffusion of IP are represented (Figure 4.8). Note that here the cell-based paradigm is advantageous in the sense that the setup of the solid culture experiment can be precisely reproduced (seeding of individual cells at determined locations, geometry of the gel) and that a single-cell model of IP sensing calibrated from liquid culture experiments can be integrated in a straightforward manner.

## Towards modeling patterning cells behavior

Here we describe an extension of the above models in order to model the behavior of patterning cells instead of receiver cells. We decided to focus first on the high-threshold killing module, hence mimicking conditions in which IP production and low-threshold killing are not activated, as in the experiments described in Figure 4.6.

First, the single-cell model IP sensing is modified according to the molecular implementation of the high-threshold killing module (Figure 4.3): GFP is replaced by the recombinase Cre, a protein fluctuation model is added to describe the vital Pkc1 protein, a stochastic reaction describing Cre-mediated deletion of the pkc1 gene is added, and finally cell death is triggered when Pkc1 protein level goes below a lethal threshold.

Then, a corresponding spatial cell-based multi-scale model is constructed as for modeling receiver cell micro-colonies. Importantly, in order to closely mimic the experimental reality, dead cells are not removed from the simulation and can still be passively displaced.

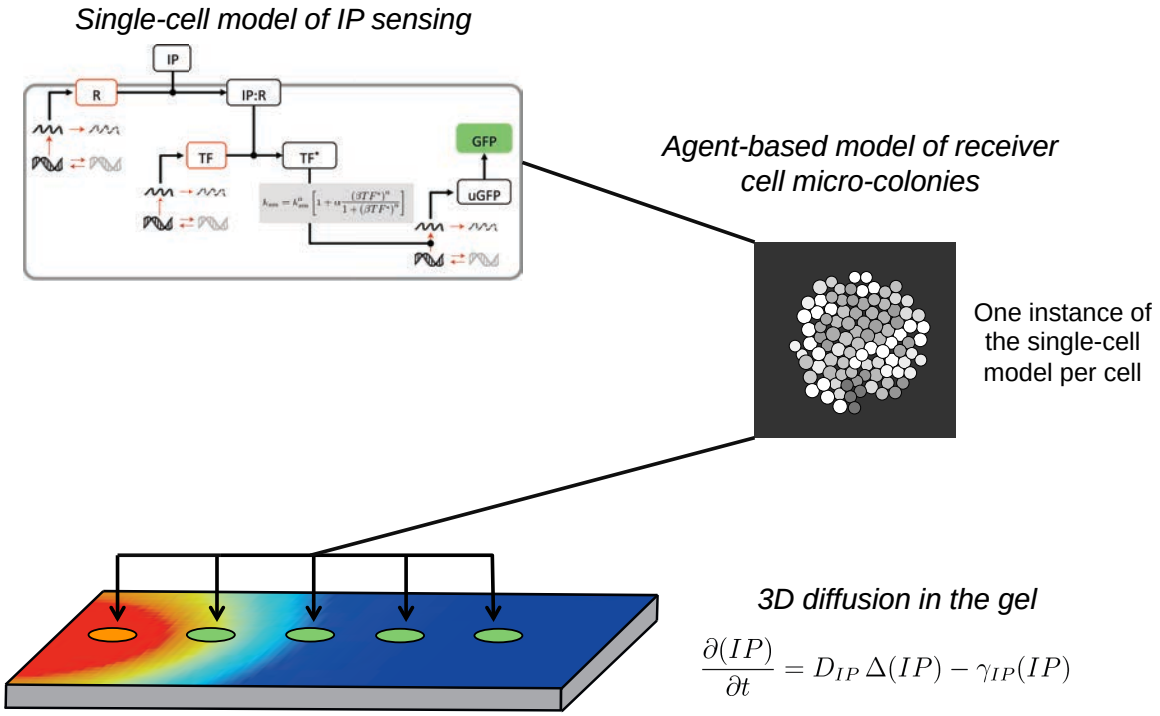



Figure 4.8: **Cell-based multi-scale model of receiver cell micro-colonies submitted to IP gradients.** Individual cells are modeled as growing and interacting spheres (similarly to the modeling of tumor cell spheroids in Chapter 3, except that parameters are adapted for yeast cells ($E = 1\,kPa$, no cell-cell adhesion, $R_{birth} = 4\,\mu m$, we did not implement budding as it did not appear necessary at this stage, and motion is limited in 2D). Each cell is equipped with an instance of the single-cell model of IP sensing described above. In addition, 3D diffusion of the IP molecules deposited in the gel is simulated and the resulting spatio-temporal profiles of IP are read by cells based on their localization.

## 4.4 Data analysis and parameter estimation

### Analysis of liquid culture data

To allow comparison of our IP sensing models with the experimental data characterizing the response of receiver cells in liquid culture, we extracted from the raw FACS data single-cell distributions of measured fluorescence corresponding to each IP induction dose and induction duration (Figure 4.9). Those distributions revealed an important cell-to-cell variability in IP response. For example, with a dose of $0.18\,\mu M$ and 4 hours after induction, about half of the cells exhibit strong fluorescence while the other half have fluorescence levels close to uninduced levels. The shape of this bimodal distribution remains largely unchanged during the following hour.

From those distributions, the population level induction dynamics can be quantified by simple averaging (Figure 4.9, bottom). Our strategy to calibrate our single-cell model of IP sensing is as follows: first, we fit the ODE model to the population level induction dynamics data to obtain a first estimate of the IP sensing 'signaling' parameters. Then, we ask whether the extended, single-cell model can predict the single-cell induction dynamics data under reasonable assumptions for the stochastic gene expression parameters.

### Estimation of signaling parameters from population-level induction dynamics

Fitting the IP sensing model on the population level induction curves identified a parameterization allowing a good agreement with the data (Figure 4.10). The corresponding parameters are given in Table 4.2. During the fitting procedure, two parameters of the model were not varied because they were known: the dilution rate was computed from the observed doubling time and the IP half-life was estimated to be 23 hours in a previous study (Chen & Weiss 2005). Of note, at very high induction ($IP = 5\ \mu M$), the experimentally observed population level response is slightly lower than the one observed for $IP = 3.3\ \mu M$. Because this non-monotonic behavior very likely corresponds to an effect that cannot be captured by the structure of our model, it was excluded from the fitting procedure.

Interestingly, the fitted maturation rate for GFP is realistic, with $\tau_m = ln(2)/kf = 12.5$ minutes: for comparison, a variant that folds faster (sfGFP) was estimated to mature in 5.6 minutes (Khmelinskii et al. 2012). This consistency suggests that overfitting was avoided. Of note, the IP receptor is estimated to be short-lived (half-life of 6.5 minutes). We could not find previous estimates of this parameter (remember that AtCre1 is not endogenously expressed by yeast cells).

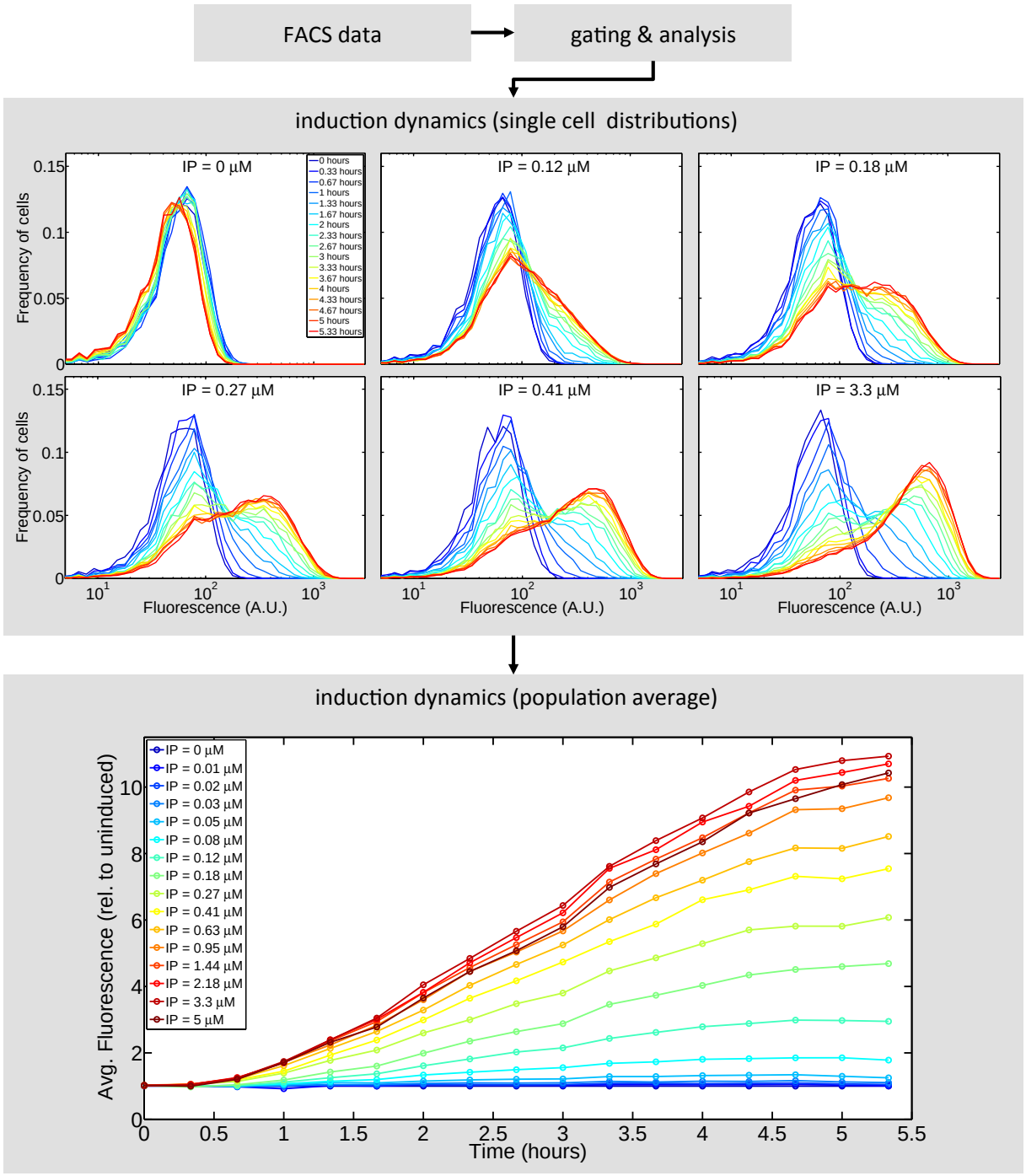



Figure 4.9: **Quantification of IP induction dynamics in liquid culture.** Analysis of the experimental data characterizing the IP sensing module behavior in liquid culture (see Figure 4.4). Gating was performed based on FSC-H and SSC-H to remove debris or cell aggregates. Resulting distributions of cell fluorescence for various IP induction levels and for different times after inductions are shown. Corresponding population average time-courses (normalized to the uninduced time-course) is also shown.

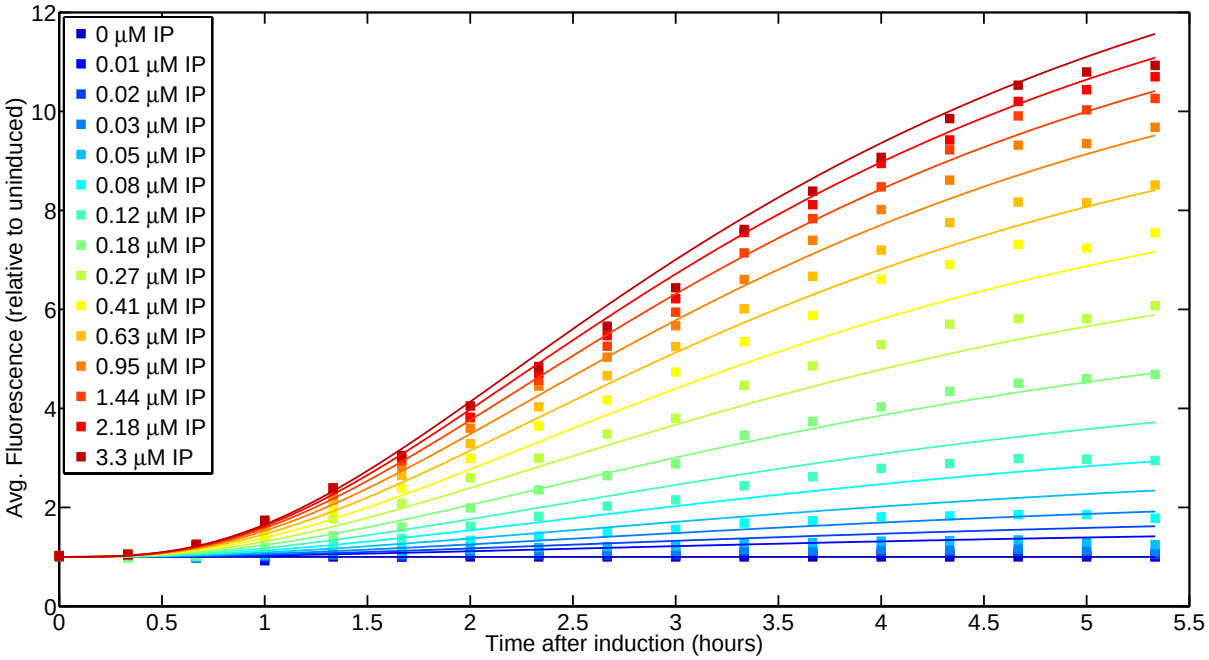



Figure 4.10: **Estimating IP sensing signaling parameters.** The semi-mechanistic ODE model of IP sensing (see Figure 4.7) was used to reproduce the population-level induction dynamics curves (data normalized to uninduced time-course, the $5\,\mu M$ IP dose was excluded as explained in text). The Downhill simplex method was used (implementation from Numerical Recipes). Dilution rate and IP degradation rate were fixed based on previous estimations. Fitted parameter values are given in Table 4.2. In the plot, data points are squares and continuous lines correspond to simulated induction curves with the fitted parameters.

Table 4.2: List of parameter values for the semi-mechanistic model of receiver cells behavior calibrated to liquid culture data. (*) Specific degradation rates that were estimated to be lower than $10^{-5}\ hr^{-1}$ were set to zero for simplicity, as this did not affect the model behavior (dilution is then largely predominant).

| Parameter symbol | Value | Comment |
|---|---|---|
| $\gamma_{dil}$ | $0.40\ hrs^{-1}$ | imposed (from measured doubling time) |
| $\gamma_{IP}$ | $0.03\ hrs^{-1}$ | imposed (previous estimation (Chen & Weiss 2005)) |
| $k_f$ | $3.32\ hrs^{-1}$ | fitted, value close to existing knowledge |
| $kb_1$ | $3.63\ min^{-1}$ | fitted |
| $ku_1$ | $0.54\ min^{-1}$ | fitted |
| $kb_2$ | $0.012\ min^{-1}$ | fitted |
| $ku_2$ | $1.88\ min^{-1}$ | fitted |
| $kc_2$ | $0.716\ min^{-1}$ | fitted |
| $\sigma$ | $0.128\ min{-}1$ | fitted |
| $\gamma_R$ | $6.47\ hrs^{-1}$ | fitted (imply a fast turnover of the receptor) |
| $\gamma_{TF}$ | $0.157\ hrs^{-1}$ | fitted |
| $\gamma_{GFP}$ | $0.117\ hrs^{-1}$ | fitted |
| $\gamma_{IP:R}$ | $0$ | fitted (*) |
| $\gamma_{IP:R:TF}$ | $0$ | fitted (*) |

| Parameter symbol | Value | Comment |
|---|---|---|
| $\gamma_{TF*}$ | 0 | fitted (*) |
| $\gamma_{uGFP}$ | 0 | fitted (*) |

## Estimation of stochastic gene expression parameters using single-cell level induction dynamics

We then asked whether the single-cell IP induction dynamics (Figure 4.9), which revealed an important cell-to-cell variability, can be captured with our single-cell extension of the IP sensing model (Figure 4.7). As for modeling protein fluctuations in TRAIL-induced apoptosis (Chapter 2), we followed a parsimonious parameterization approach. First, none of the parameters estimated previously from the population averaged data and that still appear in the single-cell model were changed. Moreover, we used the same transcriptional bursting parameters for the three proteins considered in the model (R, TF and GFP) and started with reference values obtained from a previous study of transcriptional bursting in budding yeast (Brown et al. 2013).

In a first step, we manually tuned those parameters in order to get a distribution of GFP in absence of IP that resemble observations. Then, we also manually adjusted the three parameters of TF* -mediated transcription of GFP by comparing the simulated induction dynamics single-cell distributions with observations. Strikingly, we very rapidly obtained a good semi-quantitative agreement (Figure 4.11). The model correctly predicts detailed features of the induction distributions present in the data, such as the straight, non-smooth angle at the mode of the distribution observed for inducing with $IP = 0.12\,\mu M$, and the bimodal distributions for doses ranging from $0.18\,\mu M$ to $0.41\,\mu M$. Thus, our single-cell model of IP sensing is able to reproduce observed single-cell IP induction dynamics with a good accuracy. Note that although at this stage we were satisfied with this result, an even better agreement could probably be obtained by using an automated fitting procedure.

Table 4.3: Parameters of the single-cell model of IP sensing.

| Parameter | Value | Comment |
|---|---|---|
| Mean mRNA level | 7.5 | taken from (Brown et al. 2013), consistent with (Gygi et al. 1999; Iyer & Struhl 1996) |
| mRNA half-life (dilution included) | 20 min | average value across genome is 11 min (Miller et al. 2011) |
| $T_{on}$ | 1.25 hrs | manually adjusted to roughly reproduce uninduced distribution |

| Parameter | Value | Comment |
|---|---|---|
| $T_{off}$ | 1.5 hrs | manually adjusted to roughly reproduce uninduced distribution |
| $n$ (Hill coefficient) | 2 | manually adjusted to roughly reproduce induction distributions |
| $\alpha$ | 20 | manually adjusted to roughly reproduce induction distributions |
| $\beta$ | 2 | manually adjusted to roughly reproduce induction distributions |

## Analysis of solid culture data

In order to assess if our model-based characterization of IP sensing obtained from liquid culture data can predict the observed behavior of receiver cell micro-colonies in presence of IP gradients (Figure 4.5), we first had to quantify this response in way that allows meaningful comparison with model predictions. To this aim, we developped an image analysis pipeline to compute a fluorescence temporal profile for each micro-colony (Figure 4.12). Of note, because piling up occured at the colony centers and inter-colony growth was variable, we segmented the outer rim of each colony for which a single layer of cells contribute to the measured fluorescence. We show the resulting profiles for three representative micro-colonies (Figure 4.12). As could be expected if IP diffusion is sufficiently slow, clear differences of fluorescence profiles between the colonies are obtained, consistently with their distance to the location at which IP was deposited: close colonies responded faster and to a higher extent than distant colonies.

## Estimation of the IP diffusion coefficient

Under the assumption that the physiology of receiver cells does not change much between liquid and solid culture conditions (at least with respect to the functioning of the IP sensing module), an *in-silico* reproduction of this experiment, using our model of IP sensing calibrated on liquid culture data and in which only the IP diffusion coefficient parameter is varied, should be able to predict this data. Indeed, the geometry of the gel is known, the stability of IP is probably not much affected, and the exact amount of IP deposited is known.

We tested a few values of the IP diffusion coefficient and found a good agreement between model predictions and data (Figure 4.13) for $D_{IP} = 10^{-4}\, mm^2.s^{-1}$. This value is in the range of what could be expected from previous knowledge (see for example (Brown & Johnsen 1981)). Thus, our model-based characterization of IP sensing derived from liquid culture data is quantitatively consistent with the behavior of receiver cell micro-colonies in

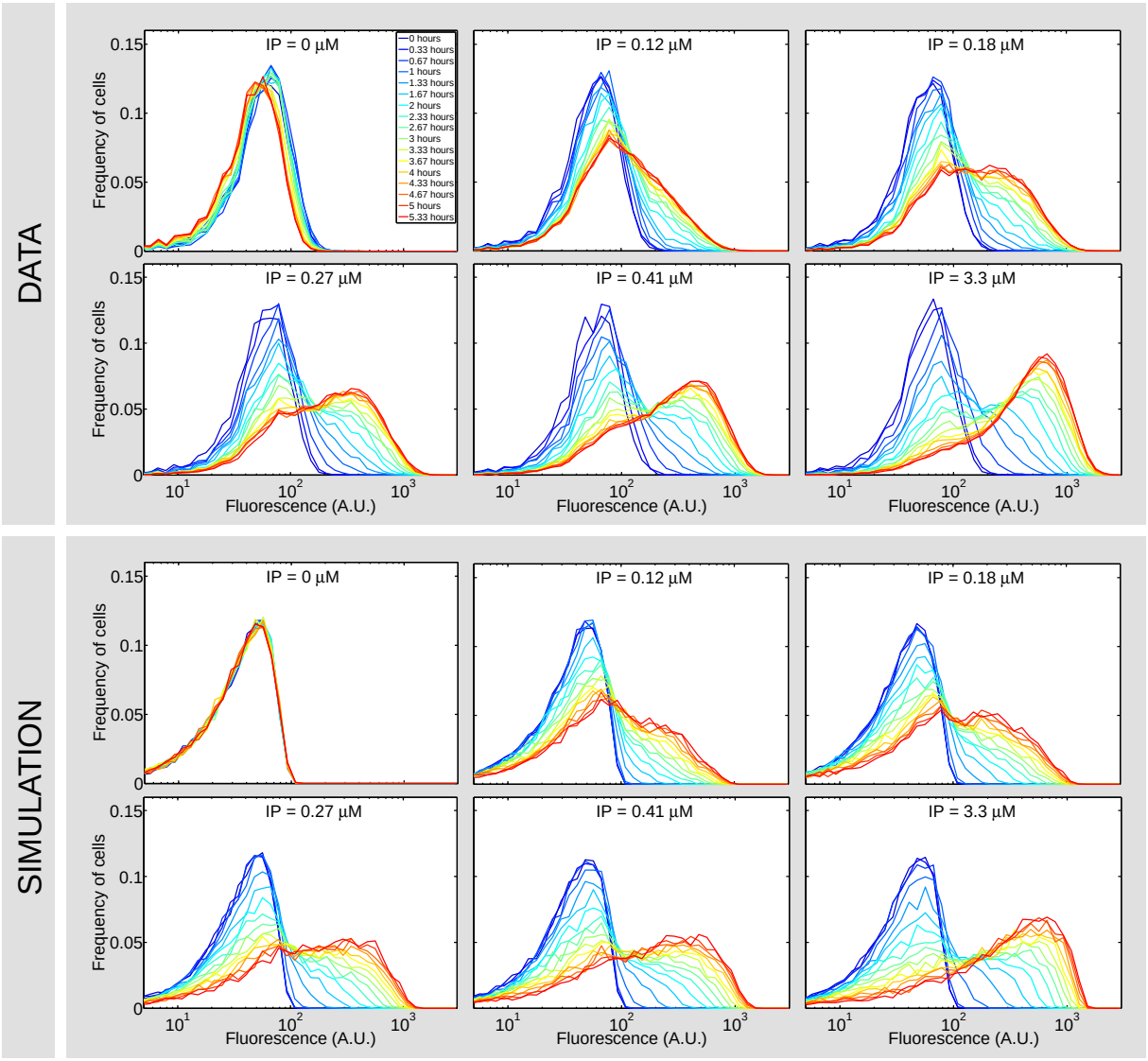



Figure 4.11: **IP sensing model accounting for stochastic gene expression captures observed single-cell induction dynamics.** *In-silico* reproduction of the single-cell induction dynamics and comparison to data. 10 thousands cells per induction dose were simulated for the same duration as in experiments and their fluorescence was stored at the same timepoints as in the experiment. The simulated data was scaled such that the mode of the uninduced distribution correspond to the experimental one, allowing direct comparison between model and data. Then the same binning was used to construct histograms. Corresponding parameters are given in Table 4.3.

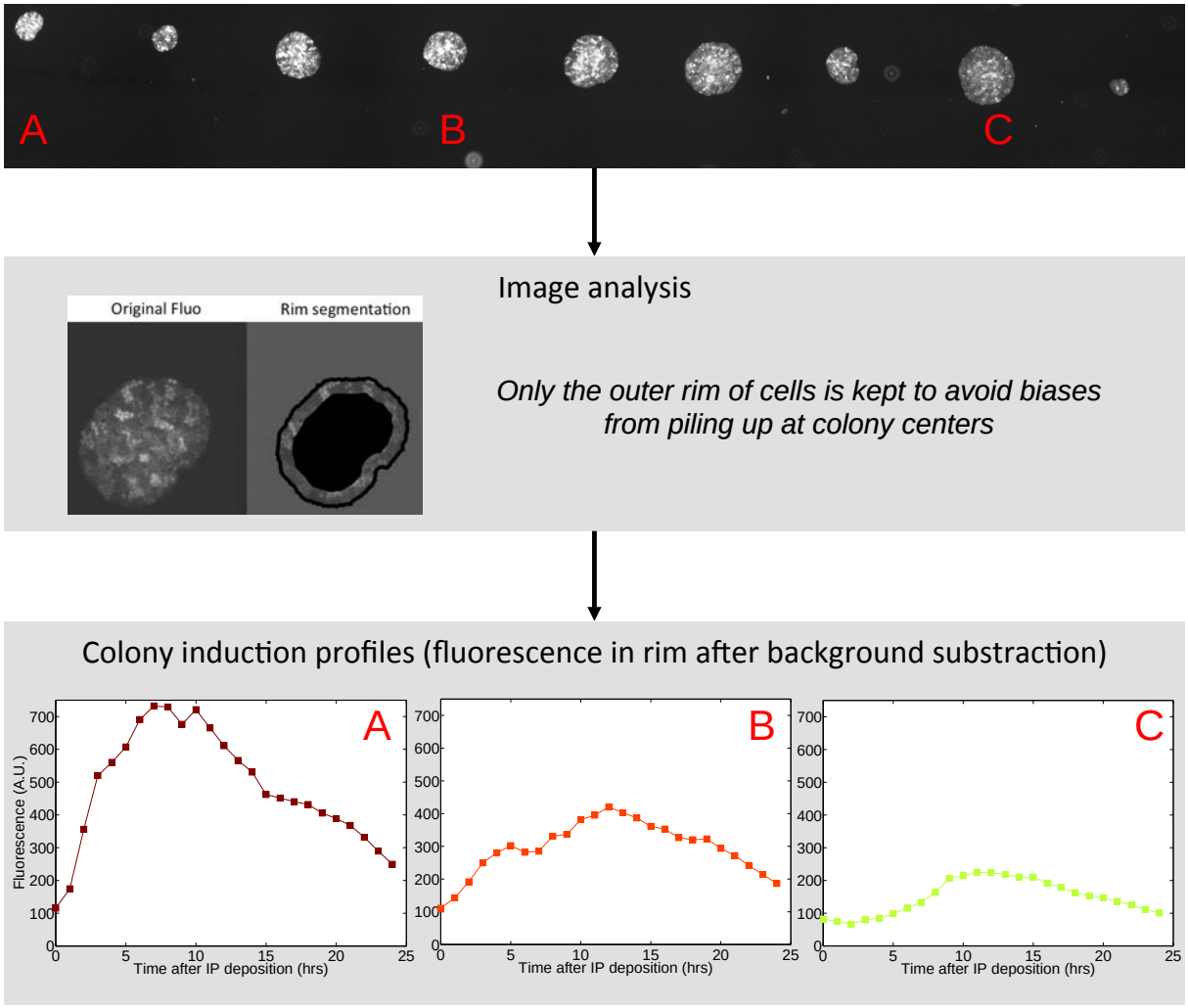



Figure 4.12: **Quantification of induction profiles for individual micro-colonies.** Analysis of the experimental data characterizing the IP sensing module behavior in solid culture (see Figure 4.5). The fluorescence images were used for segmenting colonies. For quantification of induction dynamics per colony, only the outer rim of each colony was considered to avoid biases introduced by piling up at colony centers. Background fluorescence was substracted such that the measurement reflects fold-change of cell fluorescence with respect to un-induced colonies. Three colony induction profiles are shown and their locations indicated by letters (A is the first colony on which IP was deposited).

presence of IP gradients.

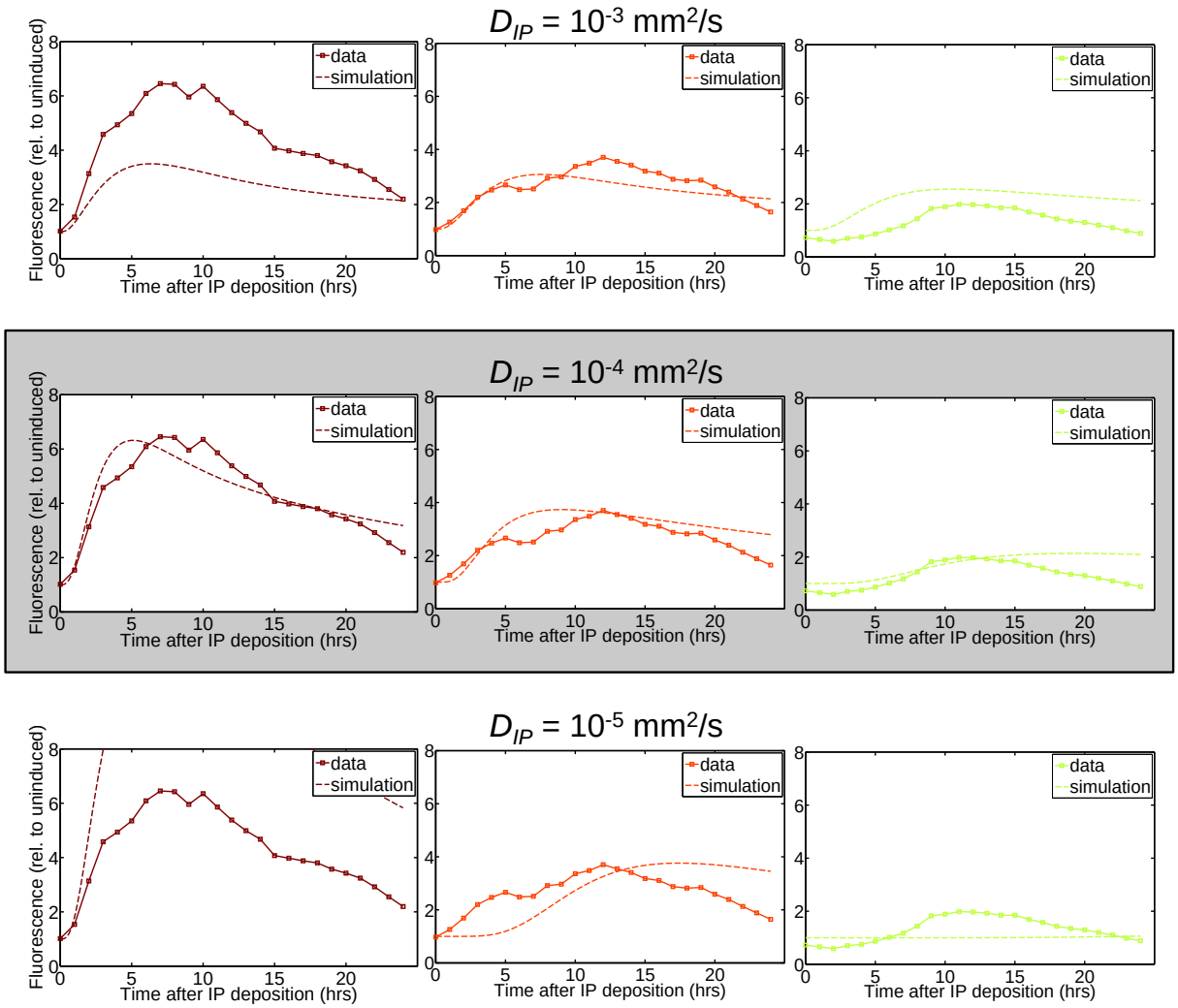



Figure 4.13: **IP sensing model calibrated from liquid culture data is consistent with solid culture data and provides an estimation of IP diffusion coefficient.** In-silico reproduction of the solid culture data. Three different values of the IP diffusion coefficient are tested. The parameter values for the ODE model describing IP sensing are the one estimated from liquid culture data (Figure 4.10 and Table 4.2). The simulated spatio-temporal profiles for the three colonies shown in Figure 4.12 are compared to the corresponding data. Good agreement is obtained for $D_{IP} = 10^{-4}\, mm^2.s^{-1}$. In simulations, for simplicity we used the IP concentration at the colony center as being representative of the whole colony response.

## 4.5 Reproducing the spatio-temporal response of receiver cell micro-colonies

In liquid culture, cell-level response to IP induction was found to be highly heterogenous, and our single-cell model of IP sensing was able to reproduce this variability (Figure 4.11). However, in FACS data no information about the population structure is available, while for micro-colonies grown on solid medium, close cells are more likely to be closely related.

Because new cells inherits molecular traits from their mother cell, such as protein concentrations, a certain spatial correlation of the fluorescence response within a micro-colony is expected (independently of IP gradients). This might have important implications for the functioning of the patterning system.

Therefore, we asked whether our cell-based multi-scale model of receiver micro-colonies could predict cell-to-cell variability in fluorescence and its spatial correlation. Hence, we reproduced *in-silico* the experiment and compared the resulting images to data (Figures 4.14, 4.15 and 4.16 for colonies A, B and C respectively). The overall agreement appears good, and clustering of 'responders' or 'non-responders' seems to be qualitatively reproduced by the model. Of note, the growth of receiver micro-colonies (with respect to cell-level cell cycle durations and colony-level 3-dimensional growth) has not yet been precisely calibrated, but we don't expect strong qualitative differences.

Finally, we confirmed the visual impression of spatial correlation in the simulations by computing the Moran's $I$ spatial correlation metric within micro-colonies (Figure 4.17, values of $I$ can range between -1 (negative spatial autocorrelation) and 1 (positive spatial autocorrelation), 0 indicates a random, non-correlated spatial pattern). Unfortunately, we could not apply the same quantification of spatial correlation on the data because single-cell segmentation at this magnification was not reliable.

## 4.6 Predicting death clusters in patterning cell micro-colonies

This carefully constructed cell-based multi-scale model of the behavior of receiver cells from the single-cell level up to spatially organized micro-colonies in presence of IP gradients establishes solid grounds on which to build a model of the full patterning system. To illustrate this, we investigated under which conditions the clearly apparent clusters of dead/alive cells that were observed experimentally (Figure 4.6) can be reproduced by the extended model accounting for high-threshold killing (as described at the end of section 4.3).

Among the few new parameters needed to describe high-threshold killing, two are a priori of central importance: the efficiency of Cre-mediated excision of the pkc1 gene, and the tolerance of cells to low levels of the Pkc1 protein. Intuitively, the former would govern the efficiency of the *commitment* to cell death, while the latter would somehow set the *delay* between commitment and death: after excision, Pkc1 protein levels would decrease because of degradation and dilution, but if cells are tolerant to low levels of Pkc1, a committed cell might survive a long time after deletion and might even divide.

Results of simulations in which those two parameters were varied are shown in Figure 4.18. First, as expected, overall killing efficiency increases with both increasing pkc1 deletion rate (higher commitment rate) and reduced tolerance for low Pkc1 protein levels. Second, clearly distinct shapes of death/alive clusters are obtained for different values of those

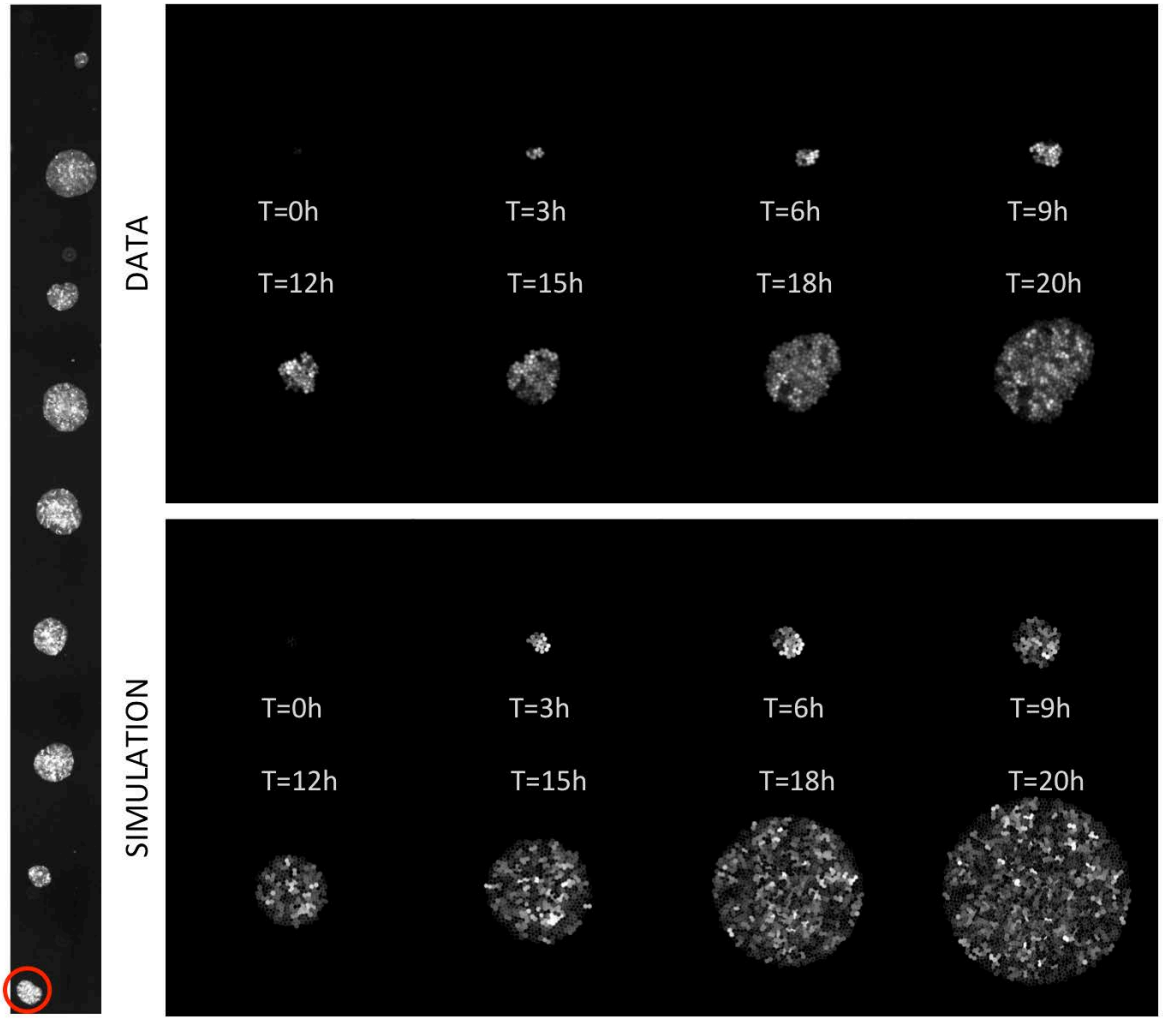



Figure 4.14: ***In-silico* reproduction of the spatio-temporal response of receiver cell micro-colonies.** The cell-based multi-scale model was used to mimick the experiment. Images of predicted single-cell fluorescence in growing micro-colonies are shown and compared to data (colony A here). The gray colorscale used to display simulation results was chosen according to the experimental one to allow a fair comparison. Clustering of 'responders' or 'non-responders' seems to be qualitatively reproduced by the model. A slightly decreased cell-cycle duration (2.2 hours in average) was used as compared to the doubling time of receiver cells measured in liquid culture (1.73 hours). Adapting the model to precisely reproduce colony growth dynamics would require single-cell tracking to quantify cell-level growth and a quantification of the 3-dimensional shape of micro-colonies driven by piling up.

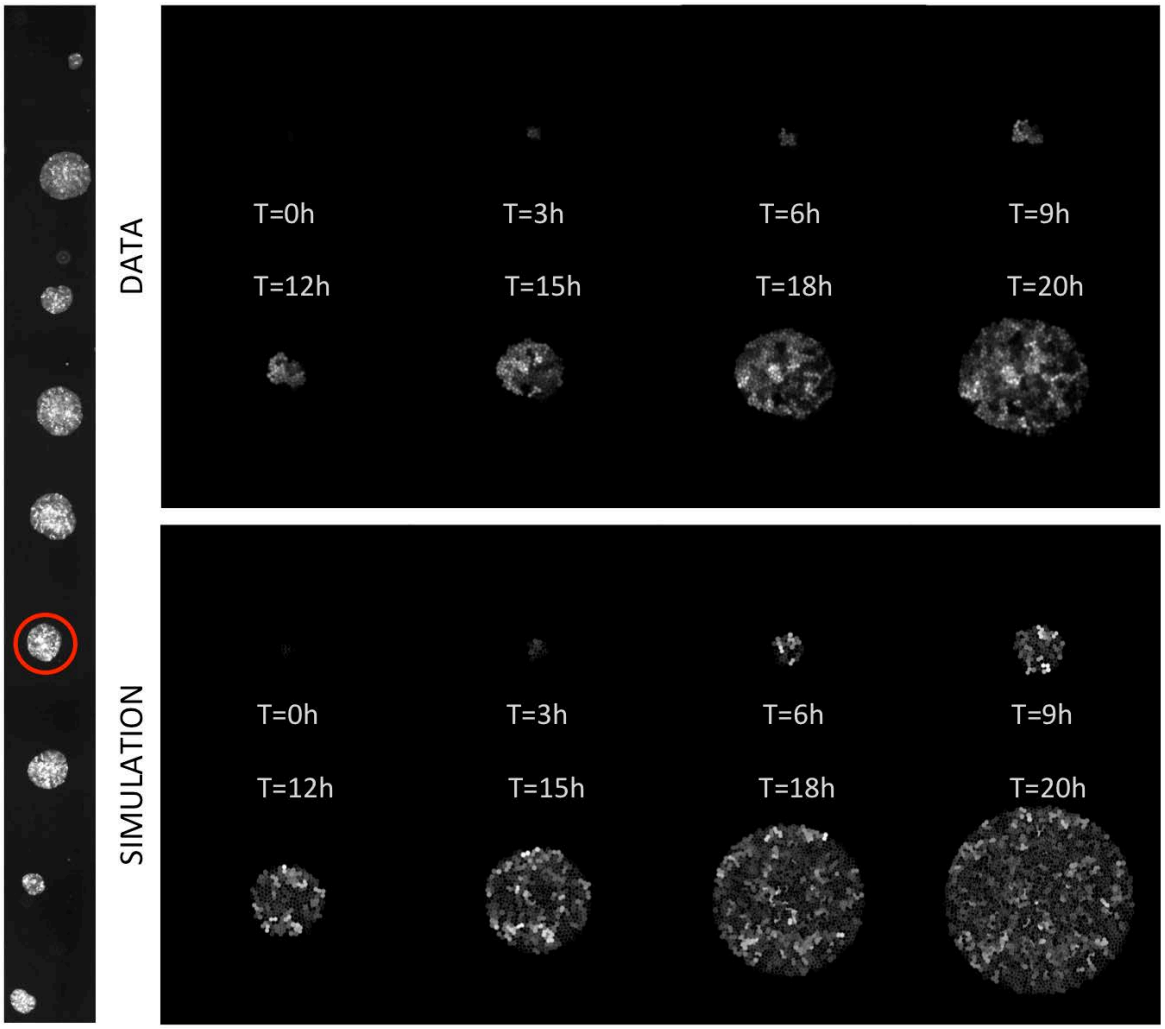



Figure 4.15: Same as Figure 4.14 but for colony B.

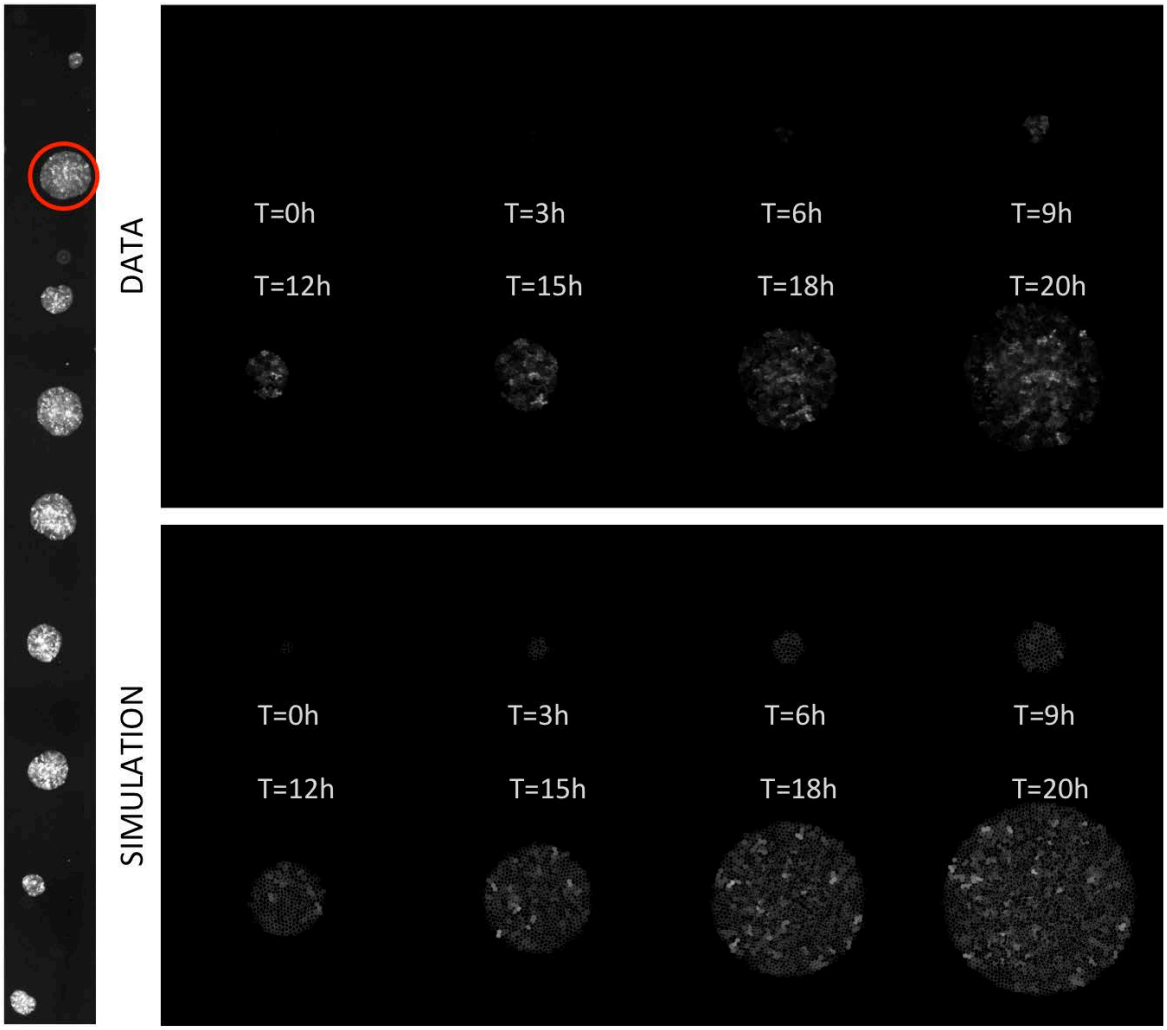



Figure 4.16: Same as Figure 4.14 but for colony C.

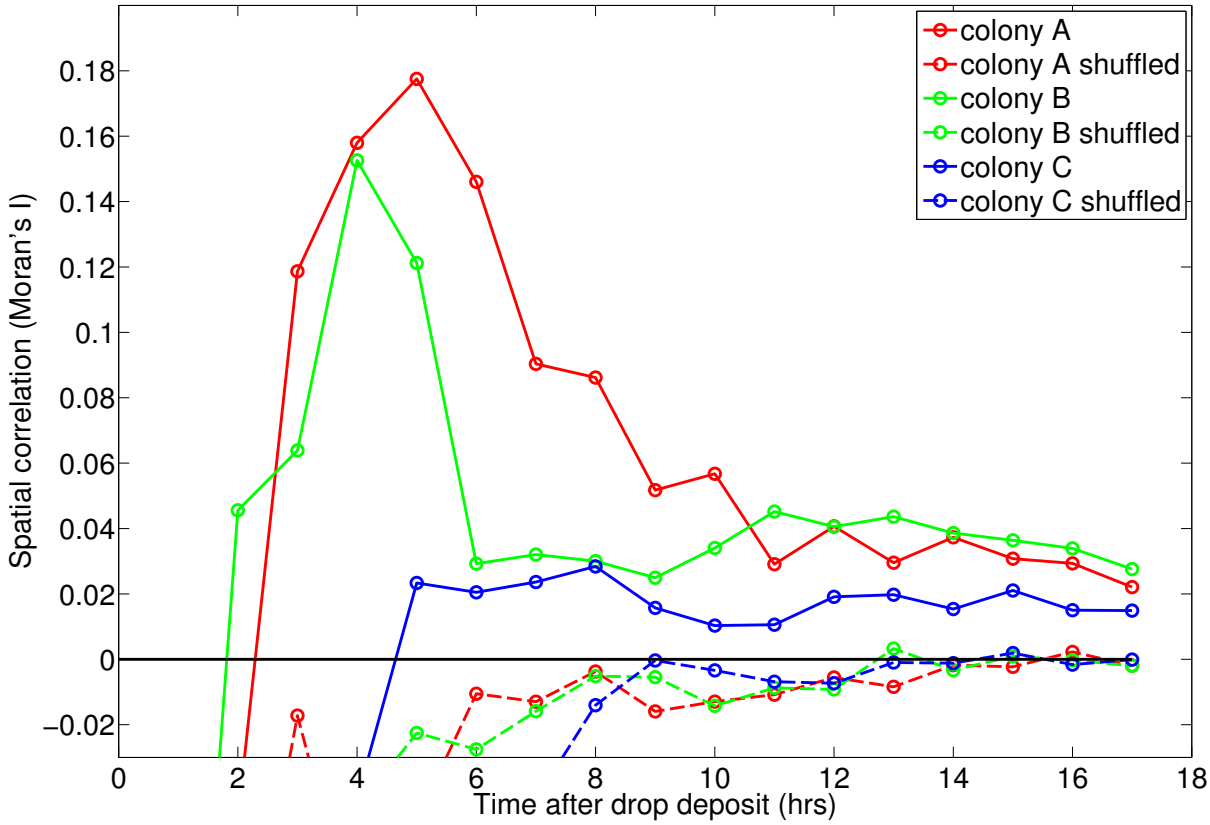



Figure 4.17: **Predicted individual cell fluorescence within micro-colonies is spatially correlated.** Temporal evolution of spatial correlation of simulated cellular GFP levels within micro-colonies. The Moran's I metric was used, with the inverse of cell-cell distances as spatial weights. In each case, a negative control is obtained by shuffling the GFP levels among cells in order to evaluate the statistical significance of the results. For times below 4 hours, the cell number per colony is too low to conclude to a spatial correlation.

parameters, even when the overall killing is similar. Larger death clusters are obtained for higher tolerance to low Pkc1 levels, corresponding to longer delays between pkc1 deletion and actual cell death. Thus, while we already showed that cellular variability and memory in IP sensing can promote a spatial correlation within micro-colonies of the protein expression in response to IP, we identified for the high-threshold killing module an additional effect, e.g. a long delay between cell commitment to death and actual death, that promotes the apparition of clusters of dead/alive cells.

In principle, quantitative comparison of such simulations with corresponding data would allow us to infer parameter values for the high-threshold killing module. Note however that patterning cells grow much slower than receiver cells. This might have quantitative consequences not only on the overall growth of micro-colonies, but also on the functioning of IP sensing and high-threshold killing because the corresponding slower dilution can impact both cellular variability and memory and the overall dynamics. Work is still ongoing to quantify and understand better those growth differences such that they can be incorporated in the model.

## 4.7 Discussion

In this Chapter, we illustrated the potential of our cell-based multi-scale modeling framework to build predictive models of engineered multi-cellular spatial systems via a rational, bottom-up approach. We have focused on an artificial sensing device implemented in yeast (IP sensing, Figure 4.3 (top), (Chen & Weiss 2005)) which is a central element of a patterning system currently developped in the Weiss Lab at MIT (Figure 4.3 (bottom)).

### Towards model-based, quantitative, single-cell level characterization of gene circuits behavior using flow cytometry

A first achievement was the construction of a single-cell model of the IP sensing and its calibration from flow cytometry data quantifying IP induction dynamics in receiver cells. Importantly, instead of comparing directly predictions of the single-cell model with observed induction distributions, an intermediate step in which an ODE version of the IP sensing model (Figure 4.7, top) was fitted to population averaged induction dynamics allowed to obtain a first estimation of IP signaling parameters (Figure 4.10 and Table 4.2). Once this ODE model was calibrated, we applied a parsimonious strategy for choosing the stochastic gene expression parameters of the single-cell model, similarly to what was done for modeling TRAIL-induced apoptosis (Chapter 2).

This approach was successful: a very small number of parameterization tests was then sufficient to obtain a good semi-quantitative agreement between the single-cell model and

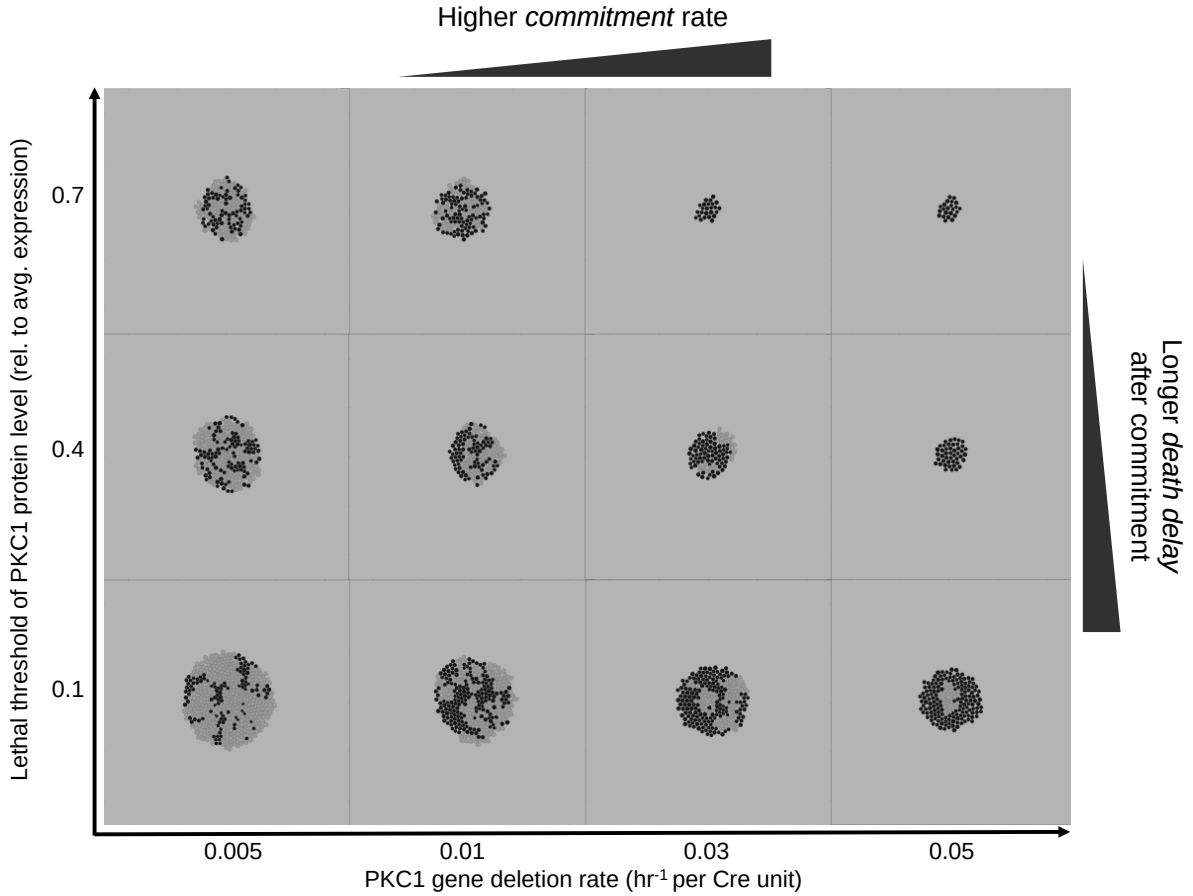



Figure 4.18: **Predicting the shape of death clusters in micro-colonies submitted to high-threshold killing.** An *in-silico* reproduction of the experiment presented in Figure 4.6 was performed with the cell-based multi-scale model for the patterning cells (in which only the high-threshold killing module is activated). Different pairs of value for the two main parameters of high-threshold killing (Cre-mediated pkc1 deletion rate and lethal threshold of Pkc1 protein levels) were tested. In each case, an image of a micro-colony 15 hours after seeding (and IP induction as it is present in the gel) is shown. Overall killing increases with both parameters, but for similar overall killing the clustering effect increases when the Pkc1 lethal threshold gets lower.

the corresponding observed distributions (Figure 4.11, Table 4.3). To our knowledge, the dynamic, quantitative functioning of synthetic sensing circuits have rarely been characterized with such accuracy at the single-cell level. Those results adds to the potential of the modeling approach proposed in Chapter 2 (which consists in extending ODE models of the protein-protein interactions involved in a given signaling pathway with parsimoniously parameterized stochastic gene expression models for all the proteins) to be broadly applicable.

In addition, the approach is here extended, as one of the protein modeled acts as a transcription factor and hence impacts the stochastic expression of another protein, while no transcription regulation was considered in Chapter 2. In the case of the IP sensing circuit in receiver cells, we have found that assuming a TF-dependent transcription rate ($k_{on}$) was sufficient to explain the data. We note that in mammalian cells, transcription factors have also been reported to be able to modulate transcriptional bursting ($k_{on}$ and $k_{off}$) (Senecal et al. 2014; Molina et al. 2013). It is likely that the effect of the transcription factor concentration on the values of $k_{on}$, $k_{off}$ and $k_{sm}$ does not follow an universal rule but depends on the specific mode of action of this transcription factor and also on the location of the gene on the chromosome. When needed, specific dependencies other than the one used here could be introduced and tested.

Early on, synthetic biologists felt that having a reliable workflow to build quantitative models of the behavior of gene circuits was needed for allowing the successful design of circuits achieving complex functions (Purnick & Weiss 2009). One major hurdle towards this goal is the ubiquitous cell-to-cell variability in circuit behavior (Miller et al. 2012; Toni & Tidor 2013), which limits the potential of purely ODE-based strategies to characterize gene circuits. Our modeling approach could enable the systematic characterization of the quantitative, dynamic behavior of synthetic circuits with single-cell resolution. Other approaches have also been proposed to infer the rates of stochastic models of induced gene expression, for example based on moment closure techniques (Zechner et al. 2012). Although our approach is presumably less powerful, it constitutes a technically simple, approximation-free alternative to such approaches, which here was proven successful on a 'real-life' example.

## Predicting circuit behavior in spatially organized cell populations from its characterization in liquid culture

The proposed approach enabled to quantitatively reproduce the response to IP of receiver cells grown in liquid culture. However, in order to be useful in guiding the design of the patterning system, this model should be able to predict the behavior of receiver cells grown as spatially organized populations. Such predictive power is not guaranteed, as cells grown in solid medium might exhibit physiological differences compared to cells grown in liquid cultures, and such differences might in turn affect the functioning of the IP sensing circuit

via context-dependency, a central issue for synthetic biology (Del Vecchio 2015). Whether or not we can successfully transpose quantitative knowledge about a system obtained from liquid culture experiments to predict its functioning in solid culture is particularly important, because the former are relatively easy and rapid to perform as compared to the latter, hence greatly increasing the number of circuits that can be characterized in a given amount of time.

Here, we have found that a model of IP sensing calibrated from liquid culture data could predict the response of receiver cells micro-colonies grown in solid medium submitted to IP gradients: indeed, simply reproducing the diffusion of IP in agar without changing IP sensing parameters led to an agreement with the data (Figure 4.12 and 4.13). Note that in this case, obtaining such agreement was not guaranteed, even if the physiology of the cells does not change between liquid and solid culture conditions. This is due to the fact that in the liquid culture data used to calibrate the model, cells were submitted to IP levels that were monotonically (and exponentially) decreasing with time. On the contrary, in the solid culture experiment cells experienced IP induction profiles of different shapes, increasing first as IP diffuses to a micro-colony, and then decreasing as it diffuses away and get degraded. In fact, an earlier version of the ODE model of IP sensing was able to explain the liquid culture data as precisely as the current one, but resulted in a signficantly bad agreement with the solid culture data (not shown). In general, performing de-induction dynamics in liquid might be useful to eliminate such incorrect model parameterizations.

While here our results show that cell physiology changes between liquid and solid culture conditions do not strongly impact the behavior of the IP sensing circuit, it is unlikely to be always the case. For example, changes in growth rate could significantly impact the quantitative behavior of a given circuit, via changes in dilution rates or global feedback on transcription (Klumpp et al. 2009). In such conditions, knowledge gained from liquid culture data could still be highly valuable in predicting the functioning of the circuit in solid culture, given that the model is sufficiently mechanistic and was cautiously calibrated to avoid overfitting. Indeed, then in principle only a subset of the model parameters, associated with relevant differences in cell physiology, would have to be adapted in order to capture the circuit behavior in solid culture (for example the dilution rate).

Another central question for synthetic biology is the *modularity* of the synthetic genetic circuits (Del Vecchio 2015): does a circuit constructed from the combination of two smaller circuits behave as one would predict by composing two models characterizing their behavior in isolation ? How precisely the concept of modularity should be applied to enable synthetic biologists to build increasingly complex systems is still an open question (Neal et al. 2014).

Here, the patterning system (Figure 4.3) rely on connecting the IP sensing module with two killing circuits, and the IP-responsive promoters ($P_{TRSSRE}$ and $P_{SSRE}$) are the genetic

elements realizing that connection. In principle, if the IP sensing model calibrated from data on receiver cells faithfully represents the reality of the involved biochemical reactions, then a model of the high-threshold killing in which the signaling parameters (from IP binding to activation of the transcription factor), and the stochastic gene expression parameters for the receptor and the transcription factor remain the same as in the IP sensing model, should be successful in predicting the high-threshold killing circuit behavior; provided that additional parameters needed to describe TF-dependent expression of Cre, pkc1 gene deletion and the lethality associated to low Pkc1 levels are calibrated.

We have here tested this idea by constructing such extension of the IP sensing model and we have found that death clusters (within micro-colonies submitted to IP) that are very similar to experimental observations can be obtained (Figures 4.6 and 4.18). We also obtained mechanistic insights about the system behavior, by pointing out the differential effect of Cre efficiency and Pkc1-associated lethality on the shape of death clusters. Such insights are valuable for the tuning and debugging of the killing circuits in order to get a desired behavior.

However, we did not yet fully characterize the high-threshold killing behavior from data because the growth of patterning cells is markedly slower than receiver cells. The causes for such difference and the expected impact on circuit functioning (which might be significant and non-trivial, although this problem has been studied mainly in bacteria (Klumpp & Hwa 2014)) are still being investigated. Also, our models currently account for cell-to-cell heterogeneity in growth in a simple manner that might overlook potential coupling with the system circuits behavior. Such potential coupling is also under investigation.

## Towards modeling the full patterning system

While we focused first on characterizing the IP sensing module (because of its central role in the design), and then started to characterize the high-threshold killing module (because in principle a single IP-dependent killing circuit could generate spatio-temporal patterns of dead/alive cells when combined with IP production and diffusion), the IP production module and the low-threshold killing module remain to be characterized in order to obtain a complete, bottom-up, model-based characterization of the patterning system (Figure 4.3).

Regarding the IP production module, estimation of the IP production rate as a function of Cu induction ( Figure 4.3) has in fact already been realized in liquid cultures, via an experimental setup that utilize receiver cells as quantitative sensors of IP concentrations. However, the quality of this characterization remain to be tested in solid culture experiments. For example, patterning cell micro-colonies, induced only for IP production, could be printed on a gel together with receiver cell micro-colonies positioned at different distances, the latter responsing to the spatio-temporal IP gradients created by the former, similarly to the exper-

iment shown in Figure 4.5.

Once such a complete model of the patterning system is constructed, it can be used to predict the outcome of well-defined experiments (gel geometry, initial positions of the microcolonies, induction levels (Cu, Dox and Gal) for the different modules) involving patterning cells. Iterative comparison between predicted and experimental results would allow refinment of the model until a deemed satisfying predictive power is reached. Then, extensive in-silico exploration of the system functioning can be performed to identify promising initial configurations and levels of induction in terms of the spatio-temporal patterns of dead/alive cells that are predicted.

At this stage, it is still not clear what kind of patterns can robustly be obtained with the current implementation of the system. It is possible that fundamental limitations apply, originating for example from the IP diffusion coefficient and stability, the delay between commitment and cell death, the overlap of those timescales with cellular growth and division, etc. Such limitations do not necessarily impact different implementations of the system equally (in other words, some implementations might function more robustly than others). While here we used a bottom-up approach to gain a detailed, quantitative understanding of the behavior of a particular implementation of the system, it might be useful to use a top-down approach to derive simpler, more generic models of the system from its high-level design (Figure 4.1). Such models can be used to broadly (but coarsely) explore the 'design space' of the system, and to provide hints on how robust system functioning could be implemented.

# Chapter 5

# Simulation of cell-based multi-scale models

The task of constructing and simulating the cell-based multi-scale models of TRAIL-induced apoptosis (Chapters 2 and 3) and of the yeast patterning system (Chapter 4) was challenging. Indeed, those models are high-dimensional and involve a large number of equations. Moreover, their dynamics is described by a mixture of different mathematical formalisms: ordinary differential equations (for example for intra-cellular signaling reactions or the motion of cells in 3D), continuous time Markov chains (for example for stochastic gene expression), agent-based/population modeling and partial differential equations (for example for the diffusion of extra-cellular molecules). This renders the efficient description and simulation of those models a difficult and challenging problem. To our knowledge, no existing tool directly enable the description of simulation of such complex models.

The goal of this chapter is to describe methodological difficulties that appeared critical during this thesis, and to discuss potential solutions. We start by describing challenges relative to the choice of adequate simulation algorithms, considering the problem of intra-cellular models first and of the whole multi-cellular simulation afterwards. Finally, we also discuss the more high-level problem of how to optimize the 'workflow' of computational biologists, notably in terms of code re-usability and safety regarding potential programming errors, and propose a tool for the automated generation of simulation code from a programmatic model description (programmatic means taking the form of a computer program, such as in PySB, (Lopez et al. 2013)).

## 5.1 Simulating combined signal transduction and stochastic gene expression

Cellular decisions are governed via biochemical pathways involving protein-protein reactions. Those pathways are traditionally modeled via ordinary differential equations. Therefore, an effort has been spent by the community of cell-based modeling towards the development of frameworks that enable the simulation of ODEs into each cell of a multi-cellular simulation, in such a way that the state variables of those ODEs can be linked to the extracellular part of the simulation, hence allowing a coupling between cells and their environment (Sütterlin et al. 2013; Starruß et al. 2014).

However, as we argued in Chapter 1, the problem with this approach is that cell-to-cell variability cannot emerge from the cell itself and can arise only from external cues, in contradiction with experimental observations. A key concept in our modeling framework is to allow for cell-instrinsic sources of cell-to-cell variability by simulating stochastic gene expression for all the proteins that take place in the cellular decisions.

This choice poses a computational challenge regarding the simulation of cell state. The Gillespie algorithm requires to interpret all reactions as being stochastic and is highly inefficient when some species are abundant. Because of the marked separation in copy numbers between genes and mRNAs on the one side and proteins on the other, it appeared reasonable to model all reactions as deterministic except gene activity switches, transcription and mRNA degradation (see Chapter 2, Figure 2.3). This type of approximation is referred to as 'partitioning' and has been reviewed elsewhere (Haseltine & Rawlings 2005).

An alternative would have been to use an approximate stochastic simulation algorithm, such as tau-leaping (Cao et al. 2007). One advantage of our choice is that the approximation is done at the level of the mathematical model rather than its simulation, which might facilitate reproducibility of the results, and allows to use existing analytical results (Paszek 2007) to judge the validity of this approximation.

## Simulation algorithms

To simulate such hybrid stochastic/deterministic models, a straightforward, 'naïve' simulation algorithm can be proposed:

**Data**: vector of cell state (gene activities and mRNA levels as integers, protein levels as floats) at time $0$, duration $T$ to simulate

**Result**: vector of cell state at time $T$

$t \leftarrow 0$;

**while** $t \leq T$ **do**

  $dt \leftarrow$ time of next reaction using Gillespie algorithm (next reaction not chosen);

  **if** *dt* < *T-t* **then**

    call ODE solver for duration $dt$ to advance deterministic reactions;

    choose the stochastic reaction that has occurred;

    apply the stochastic reaction into cell state;

    $t \leftarrow t + dt$;

  **else**

    call ODE solver for duration $T - t$;

    $t \leftarrow T$;

  **end**

**end**

**Algorithm 1:** Naïve algorithm for the simulation of combined stochastic gene expression and signal transduction.

Note that this algorithm leverages the fact that in our case, the rate of stochastic reactions depends only on the discrete variables (gene activity status and mRNA levels), allowing the use of the Gillespie algorithm for the stochastic reactions. More generally, for a given partition of a reaction scheme into stochastic and deterministic reactions (the species changed by the stochastic reactions are then the discrete variables, and the rest the continuous variables), this algorithm is valid only if the rates of stochastic reactions do not depend on continuous variables. Otherwise stated, the evolution of stochastic variables is independent from the evolution of deterministic variables (but the inverse dependence is possible).

This algorithm calls the ODE solver between each stochastic event, which might be costly if the ODE solver has some constant component in its execution time. Thus, we propose a potentially improved algorithm, that utilizes the independence of stochastic variables evolution: we simulate many steps of gene activity switches, mRNA production and mRNA degradation and we store the resulting mRNA trajectories in a look-up table before calling an ODE solver, which will read mRNA levels in that table as time-dependent parameters of the differential equations, as described in Figure 5.1.

We consider two different ODE solver methods, for which the implementation is taken from Numerical Recipes (Press 2007). The first is the Dormand-Prince '853' (DOPR853)

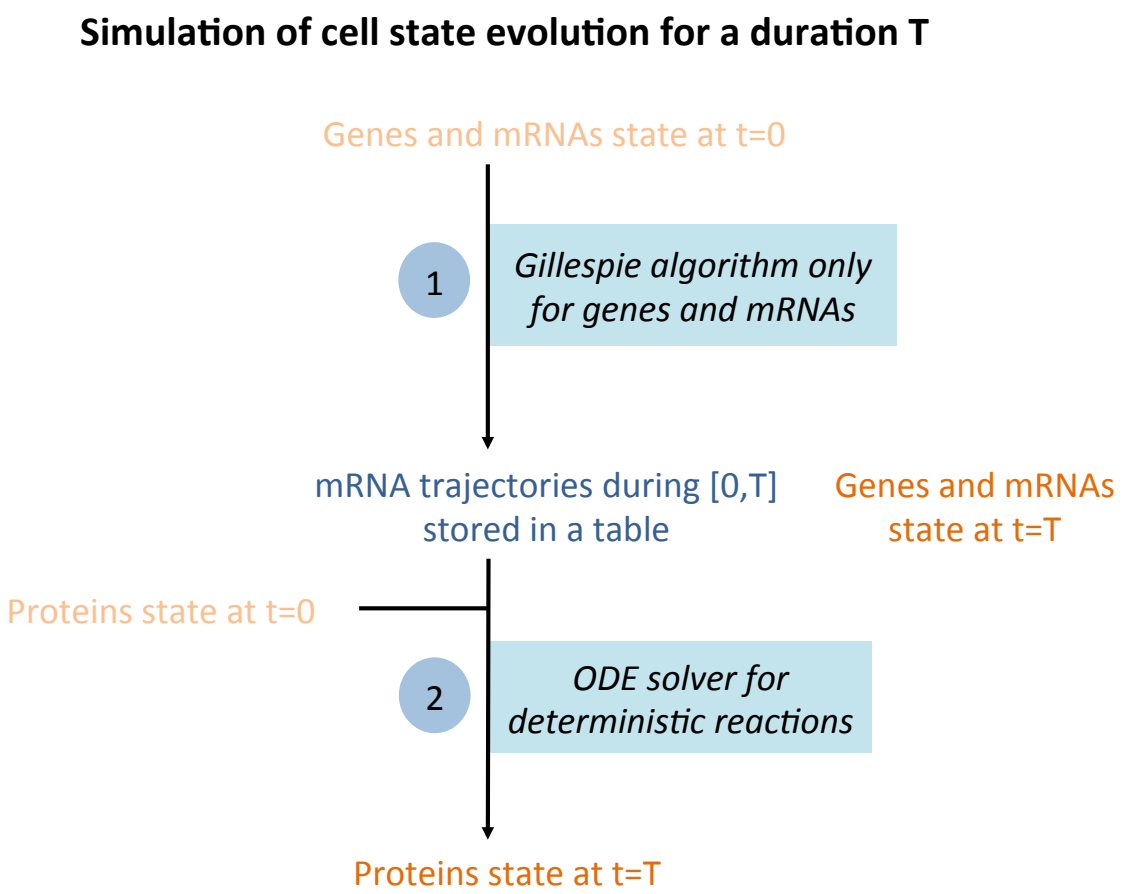



Figure 5.1: Alternative simulation algorithm for combined stochastic gene expression and signal transduction. In a first step, only the stochastic gene expression events (gene activity switches, mRNA synthesis, and mRNA degradation) are simulated using the Gillespie algorithm. The corresponding mRNA level trajectories are stored in tables. At this point, the evolution of protein levels has not been computed. In a second step, an ODE solver is used to simulate the deterministic reactions, i.e. protein synthesis and degradation and protein-protein reactions. The ODE solver can read the stored mRNA trajectories, such that the time-dependent translation rate is accounted for.

method ((Hairer et al. 2009), which is an eigth-order Runge-Kutta method. The other is an the Semi-Implicit Extrapolation (SIE) method ((Deuflhard 1985)). It has been developped for stiff problems, and requires that an expression for the jacobian is provided (although it can be an approximation).

## Computational efficiency

Using our model of TRAIL-induced apoptosis (see Chapter 2) as a real-life example, we benchmarked the computational efficiency of the mRNA storage simulation algorithm for the two ODE solvers described above (Figure 5.2). In each case, different storage durations (i.e. for how long the genes and mRNA are simulated alone and the corresponding mRNA trajectories stored) were tested.

The results strongly depended on the ODE solver method. For the Runge-Kutta based method (Dopr853), the storage duration has almost no impact on the computational efficiency. Moreover, this efficiency is very similar to the one obtained with the naïve algorithm (in which the ODE solver is called between each stochastic event).

However, the results are very different for the SIE method. First, the computational efficiency strongly depends on the chosen storage duration, up to an optimal storage duration of about 10 minutes. For those storage durations, the gain compared to the naïve algorithm is large. Notably, the total gain of using a sophisticated solver and storing mRNA trajectories is more than 10-fold and is much larger than the gain of ODE solver change when the naïve algorithm is used.

In summary, the choice of the ODE solver and the simulation algorithm is critical for the computational efficiency. Sophisticated methods such as SIE, which are already faster than more common methods with the naïve simulation algorithm, might strongly benefit from the mRNA trajectory storage strategy, probably because they usually have a strong 'fixed' cost that is paid each time the solver is called. This speed-up has been instrumental for our investigation of TRAIL-induced apoptosis presented in Chapter 2 and 3, as it allows to explore more regions of the parameter space, and to simulate more cells and longer.

A small software tool allowing the generation of c++ code for the combined simulation of signal transduction and stochastic gene expression from a programmatic description of the model reactions has been deposited on GitHub: http://github.com/fbertaux/FluctuProtST (it might be slower than the code used for hEARM in this thesis as it does not yet allow the use of the SIE solver, which requires the generation of the Jacobian function).

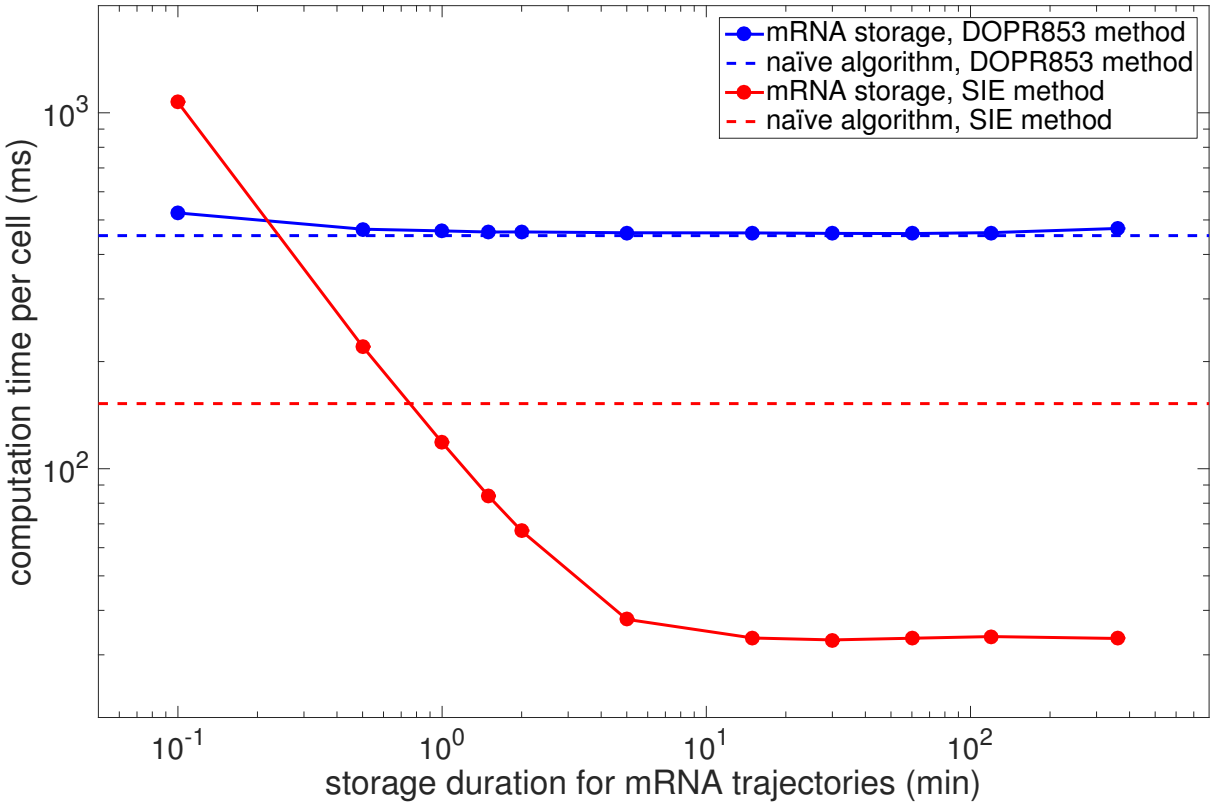



Figure 5.2: Computational efficiency of the mRNA trajectories storage algorithm as a function of the storage time for two ODE solver methods: Dopr853 and SIE. The absolute and relative error tolerances were $10^{-3}$ in both cases. For each method, comparison to the performance of the naïve algorithm is provided. In each case, 100 cells have been simulated for 8 hours after a treatment of 250 ng/mL TRAIL. A loglog scale is used.

## Parameterization of stochastic protein turnover models from protein variability and mixing time

As we have seen in Chapter 2, the transcriptional bursting parameters ($k_{on}$ and $k_{off}$) are usually unknown, not measurable directly and difficult to infer. However, knowing their exact values might not be needed if we are only interested in reproducing the protein level fluctuations, as is the case in most applications.

For this reason, we have proposed to use the protein level coefficient of variation and half-autocorrelation time as two 'statistics' that quantitatively characterize the fluctuation phenotypes of a given protein. This would mean that when the impact of the fluctuations of a given protein on the overall response has to be studied, one could explore the two-dimensional space $CV \times \tau$ instead of the larger 5-dimensional space of stochastic protein turnover models (we excluded the mean expression level as it is usually available by standard population-level measurements and does not impact $CV$ and $\tau$).

To enable this kind of parsimonious exploration, a procedure allowing to find a set of parameters that are compatible with a given $CV$ and $\tau$ is needed. Relying on the analytical expressions for $CV$ (given in Chapter 2) and $\tau$ (given in Appendix 1), we have developped such a procedure. Note that constraining $CV$ and $\tau$ is *a priori* not sufficient to uniquely determine a single parameter set, as three degrees of freedom remain. For example, if one fixes the value of the rates $\gamma_p$, $\gamma_m$ and $\gamma_g = k_{on} + k_{off}$ (an 'aggregate' switching rate for gene activity), the two parameters $EG = \frac{k_{on}}{k_{on}+k_{off}}$ (mean time fraction the gene is on) and $EM = EG\,\frac{k_{sm}}{\gamma_m}$ (mean mRNA level) remain free and might be chosen to match the two constraints on $CV$ and $\tau$.

By examining the analytical expressions for $CV$ (given in Chapter 2) and $\tau$ (given in Appendix 1), we noticed the following:

- The expression giving $CV^2$ can be decomposed in a 'gene' and 'mRNA' part: $CV^2(\gamma_g, \gamma_m, \gamma_p, EG, EM) = CV^2_{gene}(\gamma_g, \gamma_m, \gamma_p, EG, EM) + CV^2_{mRNA}(\gamma_m, \gamma_p, EM)$
- The first part $CV^2_{gene}$ of this expression varies from from $0^+$ to $+\infty$ when $EG$ is varied (between $0^+$ and $1^-$) and other parameters are fixed
- The second part $CV^2_{mRNA}$ depends only on $\gamma_m$, $\gamma_p$ and $EM$, and varies from $0^+$ to $+\infty$ when $EM$ is varied and $\gamma_m$, $\gamma_p$ fixed
- The protein autocorrelation function can be expressed solely as a function of the three rates $\gamma_p$, $\gamma_m$ and $\gamma_g$ and the 'gene contribution ratio to $CV^2$', that we call $\alpha = \frac{CV^2_{gene}}{CV^2}$
- Thus, when $\gamma_p$, $\gamma_m$ and $\gamma_g$ are fixed, all the possible values for the mixing time $\tau$ can be enumerated by varying $\alpha$ between $0^+$ and $1^-$. $\tau$ seems to vary monotically with $\alpha$

These findings led us to a procedure (described in Figure 5.3) that always find the couple $(EG, EM)$ (when it exists) compatible with given $CV$ and $\tau$ constraints, when $\gamma_p$, $\gamma_m$ and $\gamma_g$ are fixed. We implemented this procedure in *python*, and the code is available at http://github.com/fbertaux/sgeLytics.

To illustrate the type of analysis that can be done with this procedure, we investigated how different parameterizations that share the same $CV$ and $\tau$ occupy the parameter space (Figure 5.4). The slice $\gamma_m$ - $\gamma_p$ shows that both mRNA and protein half-lives cannot be large together to meet the imposed constraints, probably because of the $\tau$ value, which increases with each half-life. The slice $EG$ - $EM$ show a region (large $EG$ and large $EM$) for which no solution to the $CV$ and $\tau$ constraints exist.

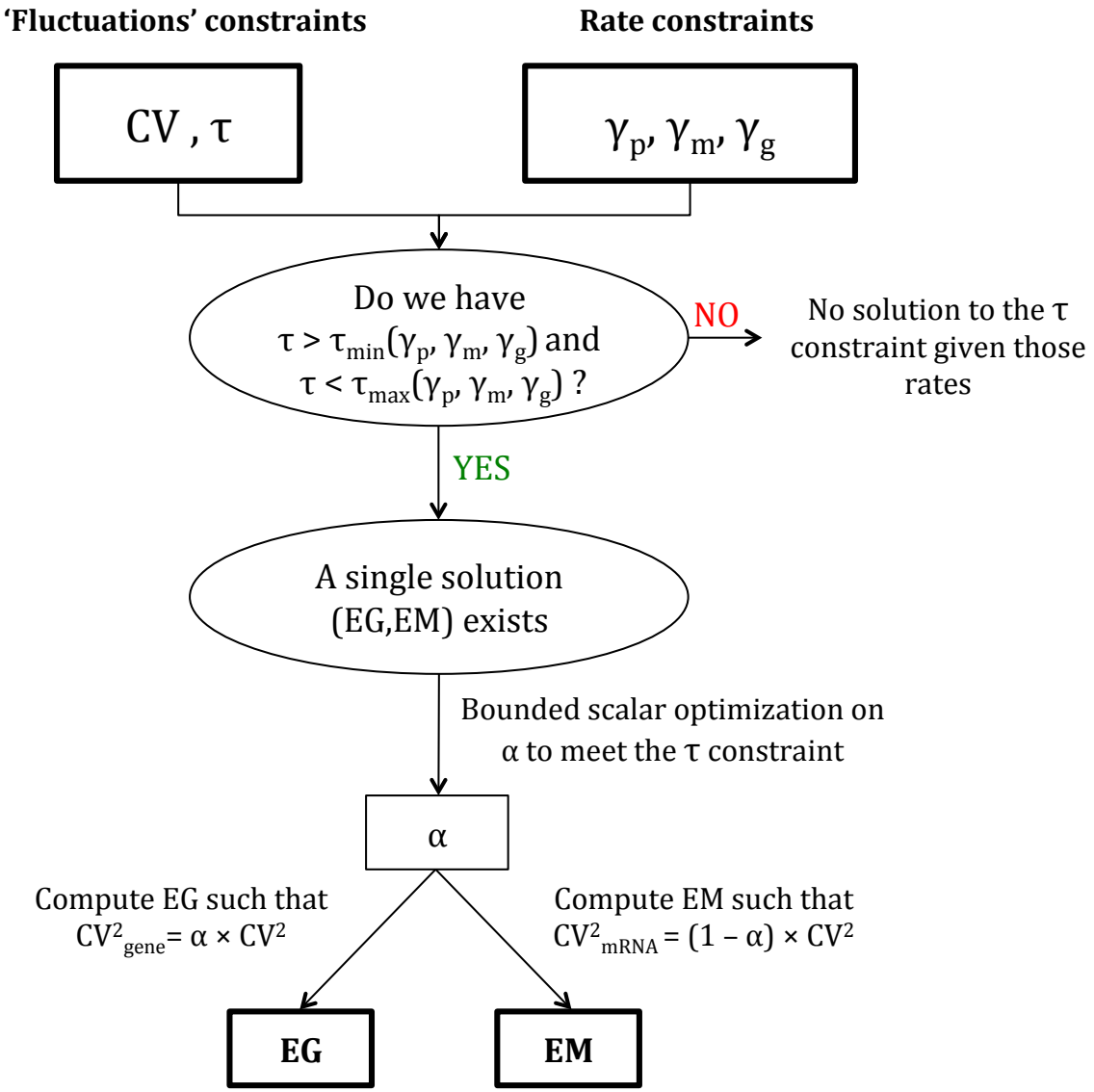



Figure 5.3: Schematic of a procedure allowing the parameterization of stochastic protein turnover models from constraints on the protein level fluctuations (i.e. the protein $CV$ and the protein mixing time $\tau$) with additional constraints on the three rates $\gamma_p$, $\gamma_m$ and $\gamma_g = k_{on} + k_{off}$. The procedure is capable of determining whether the rate constraints are compatible with the $\tau$ constraint, such that the user can change them until they are compatible.

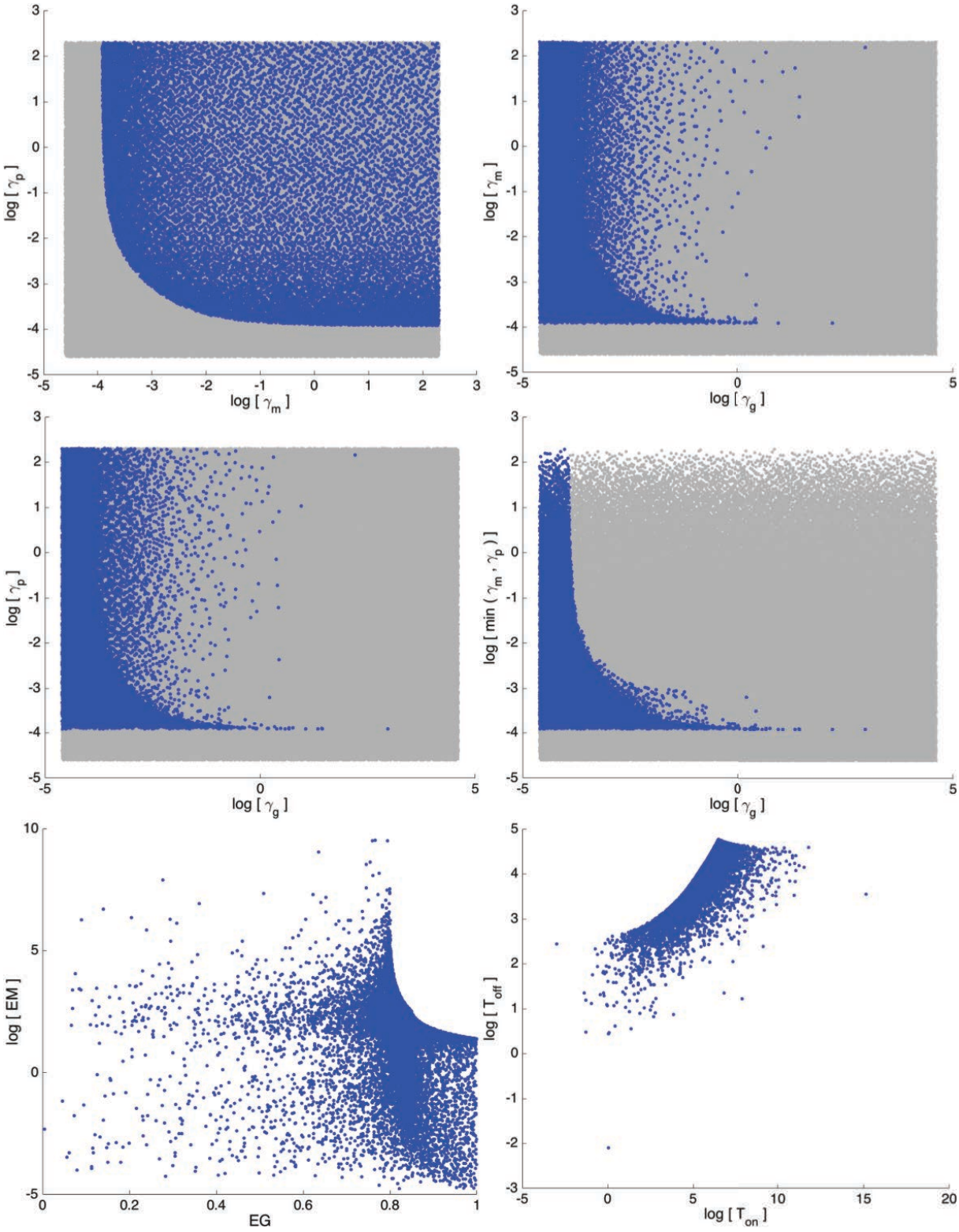



Figure 5.4: Exploration of parameter sets sharing the same protein fluctuation statistics $CV$ and $\tau$. More than $10^5$ sets of rate constraints ($\gamma_g$, $\gamma_m$, $\gamma_p$) were sampled in a log-space hybercube according to the Sobolev algorithm. In each case,the parameterization procedure described in Figure 5.3 was used together with the fluctuation constraints $CV = 0.5$ and $\tau = 35\,hrs$. The sets for which a solution exist are displayed in blue in different slices of the parameter space. In the slices involving only $\gamma_g$, $\gamma_m$ or $\gamma_p$, the sets for which no solution exist are displayed in grey.

## Simulating regulated gene expression

While this is not the case for the canonical TRAIL-induced apoptosis pathway, signal transduction pathways often involve transcriptional regulation, via signaling reactions that change the activity and/or localization of transcription factors. As a result, the transcription of target genes for those transcription factors is altered (often positively). The products of the target genes can either participate in the same pathway, hence implementing a feedback in the pathway response; or they can be involved in other pathways. For example, in $NF\text{-}\kappa B$ signaling, the cytosolic/nuclear localization of a protein complex that can bind DNA and activate the transcription of target genes is controlled by other signaling proteins that respond to external stimuli such as the cytokine $TNF\text{-}\alpha$. Among the target genes, $I\kappa B\alpha$ and $A20$ are involved in the signaling pathway responding to $TNF\text{-}\alpha$ (Tay et al. 2010).

In such context, if one is interested in modeling the expression of a gene that is targeted by a transcription factor, the rates of the two-state transcriptional bursting model cannot all be assumed to be constant: instead, at least one and maybe all of the rates $k_{on}$, $k_{off}$ and $k_{sm}$ depends on the level of the transcription factor, depending on the molecular mechanisms by which the transcription factor influences transcription. For example, in our model of the $IP$ sensing circuit implemented in yeasts (Chapter 4, Figure 4.7, bottom), we have modeled the effect of the transcription factor on its target gene by assuming that $k_{sm}$ follows a Hill function of the transcription factor concentration.

This has implications for the computational simulation of the model. Because the propensities of some stochastic reactions now evolve continuously according to the differential dynamics of the continous variables (the protein levels), the Gillespie algorithm, which rely on a constant vector of propensities in-between two reactions, cannot be used anymore. An 'exact' simulation algorithm exists: it relies on integrating the propensities of stochastic reactions together with the differential dynamics of the continuous dynamics (Salis & Kaznessis 2005; Alfonsi et al. 2005). But this algorithm is difficult to implement and requires to solve an event detection problem to stop integrating propensities when one of the stochastic reactions has fired, which might have a high computational cost.

Alternatively, a pragmatic, approximate solution is to adapt the naïve algorithm described previously (or the storage algorithm) with a maximum time step and update the rates of the stochastic reactions according to the values of the continuous variables at each time step. In this case, the algorithm approximates those rates as stepwise functions, with the size of the steps depending on 1) the chosen maximum time step and 2) the rate of changes of the continuous variables involved in the stochastic rate functions. Thus, the accuracy of the approximation can be tested by varying the maximum time step and evaluate whether the simulation results change significantly or not.

We used this approximate algorithm for the $IP$ sensing simulations (See Chapter 4) and found that for time steps below 5 minutes, the results did not depend on the time step any-

more (not shown). Thus, this appears to be a viable solution. However, it might not be computationally optimal, and further investigation is required to compare the computational efficiency with the 'exact' algorithm. Still, the computational efficiency of the approximate algorithm was sufficient for our needs, as we were able to simulate the full multi-scale model of spatially-organized receiver cell micro-colonies (Chapter 4, for example Figure 4.14) in reasonable computational times on a standard computer.

## 5.2 Simulating spatially organized cell populations

Because many of the computational challenges that arise when simulating cell populations with an agent-based framework have been described previously, notably by previous members of our group (Hoehme 2010; Jagiella 2012; Weens 2012), we will here simply highlight points that appeared critical for the present work.

### Simulating cell motion

The spatial arrangement of multi-cellular assemblies such as spheroids or micro-colonies is constantly changing because of cell growth, division, death, migration through physical interactions between cells in contact as well as between cells and their environment. To simulate this, we have used in Chapters 3 and 4 an existing lattice-free approach in which cells are assumed to behave as visco-elastic spheres (Drasdo & Höhme 2005; Drasdo et al. 2007). In this approach, an equation of motion relates at each instant velocities of individual cells with cell-cell contact forces (3.4), which values depend on the geometry of those contacts.

Thus, the positions of all cells evolve according to a set of $3N$ coupled differential equations of the form $\frac{dx_i}{dt} = F_i^x(x_1, y_1, z_1, ..., x_N, y_N, z_N)$, ..., $\frac{dz_i}{dt} = F_i^z(x_1, y_1, z_1, ..., x_N, y_N, z_N)$ where $N$ is the number of cells. However, solving this system with a standard ODE solver is challenging because 1) $N$ is not constant and can be very large compared to traditional ODE systems and 2) the right hand side is generally costly to evaluate, as it requires that all cell-cell contacts are known.

The first point was originally adressed by using a simple first-order, explicit scheme to solve cell motion, with time-step control based for example on a maximum distance that cells are allowed to travel in a single time-step (Weens 2012, used here in Chapter 3). (Of note, in a large population of cells where division takes place, this time-step choice strategy is not ideal as there will always be a division event imposing a 'worst-case' time-step. More sophisticated strategies that prevent local constraints caused for example by division to have a global impact on the computational efficiency of the simulation would be beneficial.) Alternatively, we also adapted the implementation of existing ODE solvers to solve the cell motion equation in synchrony with cell growth as well as intra-cellular dynamics

(more details are given in a following section).

The second point is critical for computational efficiency, as using a naïve contact detection algorithm (in $O(N^2)$ where $N$ is the cell number) will prevent the simulation of large cell populations (~5-20K cells on standard machines) in reasonable computational times. A efficient algorithm is to partition space into cuboïdal 'contact' voxels which size is chosen such that cells in contact have necessarily their centers in neighbor voxels (here neighbors in the sense of sharing one or more vertices). Then, contact detection is linear in $N$ (provided their is a limit to the number of cells that can belong to the same voxel): in each step, each cell is mapped to its containing voxel, and then contacts are tested only for cells in either the same voxel or neighbor voxels. More details about such algorithm can be found elsewhere (Drasdo et al. 1995).

An additional difficulty arises when cell-cell friction is accounted for. In that case, terms involving the velocities of cells in contact appears in the equation of motion (3.4). As a result, the velocity of each cell cannot be extracted directly, instead a non-diagonal linear system $AV = B$ has to be solved. Given the size of the system, the associated computational cost can be very high and adds up to the contact detection cost in the total cost of the derivatives computation for the ODE system of cell positions. An approximate method such as the conjugated gradient is likely a good choice to optimize this cost. More details on this can be found elsewhere (Weens 2012).

Another difficulty also arises when a stochastic term is used to represent micro-motility. If micro-motility is modeled as brownian motion, the resulting equations are stochastic differential equations. Depending on the numerical scheme that is used, this might cause problems, notably when a time step control procedure is used to limit the maximum distance traveled by a cell. In this thesis, we have decided to model micro-motility (see Chapter 3) as a persistent random walk instead. Briefly, the migration force (direction and magnitude) is stochastically changing at a given rate but is constant between two consecutive changes (the 'persistence' time between changes is exponentially distributed and in average is equal to the inverse of the change rate). Oppositely, for the brownian motion model, the migration force is changing within all potential time windows. When a change occurs, a new migration direction and magnitude is sampled into distributions. This facilitates the numerical resolution of the motion equations.

## Simulating the diffusion of extra-cellular molecules

To simulate the diffusion of TRAIL within spheroids (Chapter 3, see Figure 3.14) or of IP in the solid medium on which yeast micro-colonies were grown (Chapter 4, see Figure 4.8), we have taken a simple approach: we used a finite differences, Euler (first-order, explicit) numerical scheme on regular cuboïdal grid. This scheme is robust (stability is ensured when

the $CFL$ criterion is met and accuracy is usually satisfying under those conditions) and easy to implement.

However, this method can be computationally burdensome when a fine spatial resolution (relative to the domain size) is required. In our case, we usually used a spatial resolution (i.e., the length of voxels composing the discretization grid) of about a cell diameter. We believed it was not meaningful to go below that value as the surface of our cell agents is assumed to be homogenous and hence was not spatially resolved.

If needed for more complex models (for example in which cell surface receptors are not distributed homogenously on the surface, etc…), more sophisticated methods such as finite element methods with adaptative meshes can be much more powerful, but they are more difficult to implement and existing libraries are not straightforward to handle. In our group, a cell-based multi-scale modeling software currently in development (*TiSim*) will have such capabilities.

## The global synchrony problem

Most existing simulation algorithms for cell-based multi-scale models rely on the composition of sub-models representing different processes occuring at different scales (Sütterlin et al. 2013). For example, in our simulations of TRAIL treatments on multi-cellular spheroids (Chapter 3), TRAIL diffusion, cell motion and intra-cellular apoptosis signaling were not simulated in a completely simultaneous fashion.

Rather, a global time step was defined, for which those processes were simulated sequentially, and shared variables between processes (for example TRAIL concentration from diffusion and TRAIL concentration as seen by the intra-cellular model) are updated only discretely between each global steps. As a result, simulated dynamics are only approximating the true theoretical dynamics, and the global time step should be taken small enough such that this approximation remains acceptable (we used a global time step of 5 minutes for spheroid TRAIL treatment simulations).

In some circumstances, the timescales at which shared variables evolve can be such that a very small global time step should be used to ensure sufficient accuracy, which can result in excessively long computation times. For example, when a antibody with very high affinity to some cell surface receptors is released in a tissue, the spatial spread of the antibody would be governed by a 'saturate then diffuse' mechanism that will require a global time step smaller than the receptor binding timescale to be correctly reproduced (e.g. molecules that should be bound at the first cell layer will be incorrectly moved to the next cell layer during the step computing diffusion). To address such issues, one might want to construct an 'asynchrony error' measure such that it can be controlled via an adaptive global time step procedure. But rigorously choosing such measure is not straigthforward and might depend

on the specifics of the model.

A radical alternative would be to abandon the 'composition of sub-models' simulation paradigm and unify the continuous dynamics of the whole model into a single set of equations that will be fed to a single solver. After all, mathematically, the equations representing protein-protein reactions, cell motion due to cell-cell contacts, the evolution of TRAIL concentration in a voxel because of diffusion, etc… are all differential equations. In such case, the simultaneity of all continuous processes is automatically accounted for.

One difficulty with such approach is that the overall system of differential equations changes each time a cell agent is created or destroyed. We adapted the code of the DOPR5 ODE solver from Numerical Recipes (a $5^{th}$ order Runge-Kutta based scheme with adaptative time step) and integrated it into multi-agent simulation code to allow for changing the differential equations each time such event occurs. We tried this approach for simple multi-scale models of the yeast patterning system and obtained comparable results to the standard 'composition of sub-models' algorithm. We also rely on a similar approach for our tool allowing the automated simulation code generation from a programmatic model description (see next section).

Another potential problem with such approach is that the ODE solver will be fed with a system of equations of an unusually large size for which it might not be adapted. Moreover, we considered first only models with deterministic dynamics (except for agent events that can be stochastic), preventing for example to have stochastic gene expression modeled into each cell. Still, in principle this 'unified and synchronous simulation across agents' paradigm can be extended to models in which the resulting equations are hybrid stochastic/deterministic or stochastic differential equations, provided a suitable solver (for arbitrary hybrid models or arbitrary stochastic differential equations) is available.

## Numerical convergence tests

Here we present numerical tests related to the choice of timesteps to simulate the repeated application of TRAIL on multicellular spheroids (Chapter 3).

First, the global synchrony problem discussed above suggest that the global timestep should not be too large. We tested a range of global timestep values from 6 seconds to 20 minutes for a reference simulation in which a spheroid is repeatedly treated with TRAIL (Figure 5.5, top). The temporal evolution of the spheroid cell number (top left) was very similar in all cases. The TRAIL gradient within the spheroid (top right) was also similar when we compared the global timestep used throughout Chapter 3 (5 minutes) and a more stringent global timestep (0.5 minutes).

Second, we tested the influence of the internal timestep governing TRAIL diffusion and degradation and found no difference on the cell number curve when 90% or 10% of the CFL

timestep is used (Figure 5.5, bottom).

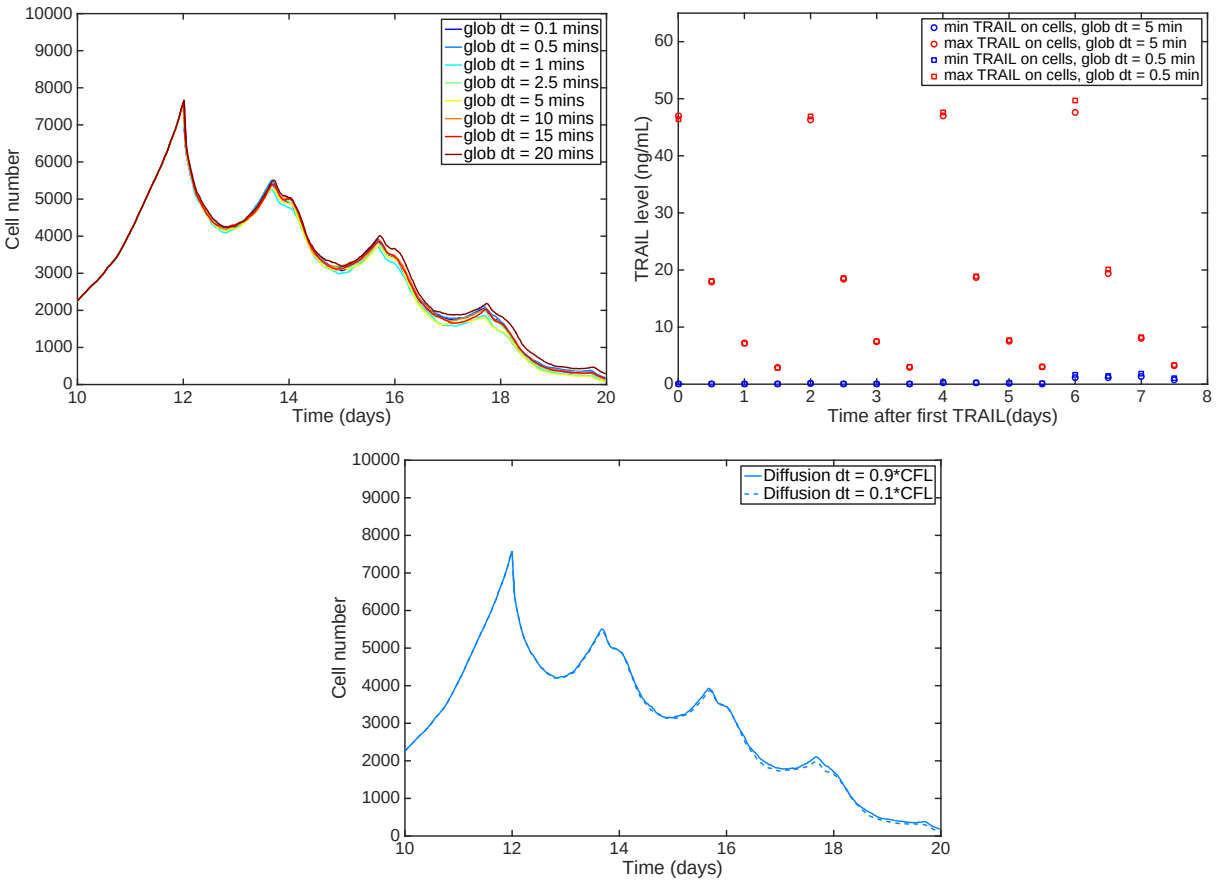



Figure 5.5: Numerical tests for choosing timesteps. Top: influence of the global timestep on cell number temporal evolution (left) and TRAIL gradient (right, characterized via the maximum and minimum TRAIL concentration felt by cells). Bottom: influence of the internal timestep for TRAIL diffusion and degradation, chosen as a fraction of the CFL timestep. To facilitate the comparison between simulations, we turned off stochastic gene expression in cells (replaced by equivalent constant protein synthesis).

## 5.3 Automated simulation code generation from a programmatic model description

To construct simulation programs for models such as the ones presented in this thesis, we initially proceeded by custom composition and integration of algorithms specialized to different sub-simulation tasks. For example, we used existing c++ implementations of ODE simulation algorithms (Press 2007) for solving intra-cellular signaling into each cell; an existing c++ implementation of the Gillespie algorithm for solving stochastic gene expression into each cell (Press 2007); a custom 3D explicit first-order finite differences algorithm for solving the diffusion of extra-cellular molecules; and an existing code previously developped in our group (Weens 2012) for solving the physics of cell-cell contacts.

With this approach, we were able to obtain functional simulation programs for the spe-

cific models we were working on. However, this required a lot of time and efforts: designing, testing and debugging such very large (>50K lines for the original code simulating cell-cell physics) and heterogenous code was extremely tedious. Moreover, the obtained simulation programs lack flexibility and re-usability: slight changes in the model assumptions would at best require error-prone manual modifications at many locations in the code, or in many cases involve significant re-writing if the structure of the simulation algorithm is impacted by the change. In addition to model analysis and simulation, the task of simply *describing* the model in an unambiguous mathematical form can be extremely tedious because of the large number of variables, parameters and equations. This is already true for single formalism models, such as large ODE models of signal transduction (for which rule-based modeling can be an effective solution, Hlavacek et al. 2006), but this problem is even more stringent in the context of cell-based multi-scale models that combine several mathematical formalisms and hierarchical levels.

To solve some of those issues, we developped a tool that enables the automated generation of simulation code from a programmatic description of the corresponding mathematical model (Figure 5.6). The tool rely on the formalization of a class of *multi-agent hybrid models* and the corresponding class of simulation algorithms. Different agents with real-valued properties can be defined. The properties of those agents can evolve according to differential dynamics or stochastic or deterministic events. Coupling between the properties of agents is possible by either defining *shared* agent (which exist in a unique instance and can be 'seen' by all agents), defining *spatially resolved agents* (which have spatial coordinates and a radius as properties, thus allowing contact detection between such agents) or defining *grid agents* (typically used for spatially resolving concentrations of molecules in the extra-cellular space).

An utilization example of the tool is illustrated by the Figures 5.7 (model description in *python*) and 5.8 (corresponding generated simulation code in *c++*). The code for this tool is available here: http://github.com/fbertaux/CellPop (no spatial capabilities) and http://github.com/fbertaux/CellPop3D (some spatial capabilities, but currently very limitedly tested).

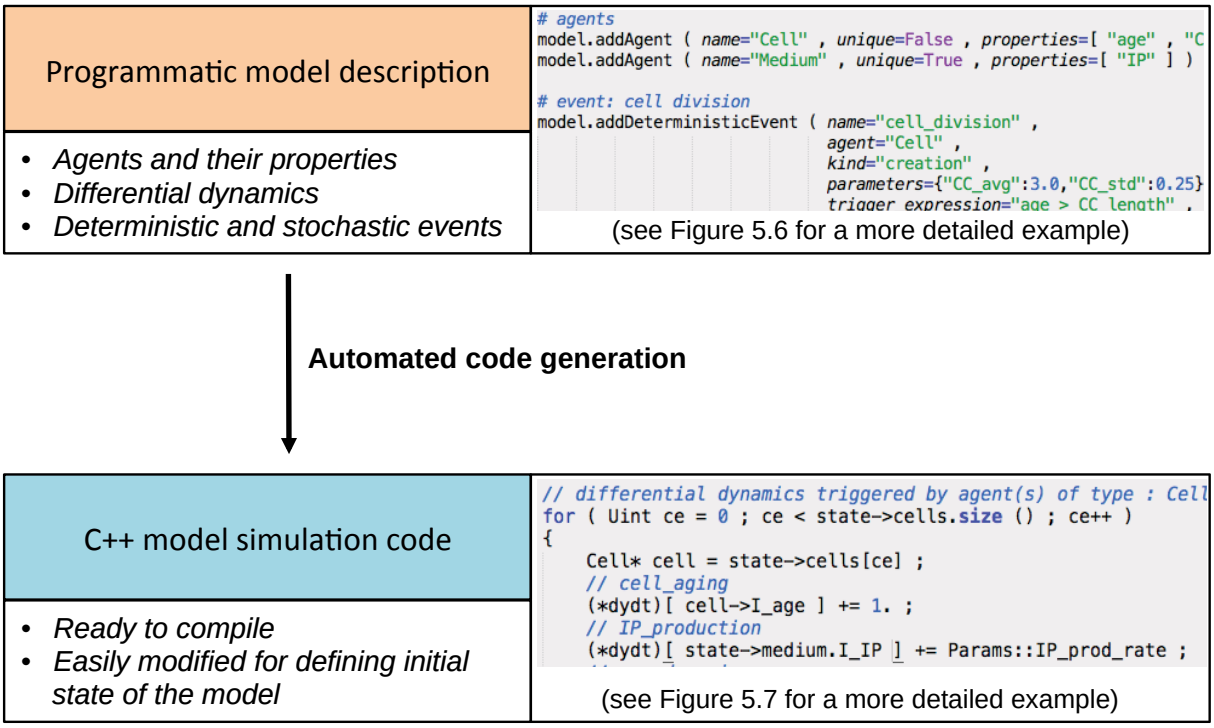



Figure 5.6: Principle of the automated code generation procedure. A model is constructed by writing a small *python* program using functions of a model description module. Those functions allow to define agents and their properties and dynamical rules (see Figure 5.7 for a more detailed example). From such *python*-encoded description of a model, c++ simulation code corresponding to that particular model can be generated using an other *python* module. The generated c++ code is ready to compile and can be easily modified to define custom state initialization and model state outputs.

```
import modeling
import simulating

### model construction
model = modeling.Model ("liquid_gol")

# agents
model.addAgent ( name="Cell" , unique=False , properties=[ "age" , "CC_length" , "Cre" , "pkc1_s
model.addAgent ( name="Medium" , unique=True , properties=[ "IP" ] )

# event: cell division
model.addDeterministicEvent ( name="cell_division" ,
                              agent="Cell" ,
                              kind="creation" ,
                              parameters={"CC_avg":3.0,"CC_std":0.25} ,
                              trigger_expression="age > CC_length" ,
                              realization_function=
                                    ["new.age = age - CC_length" , "age = new.age" ,
                                     "new.CC_length = ran.norm ( CC_avg , CC_std )" ,
                                     "CC_length = ran.norm ( CC_avg , CC_std )" ,
                                     "new.Cre = Cre" , "new.pkc1 = pkc1" , "new.pkc1_status = pk

# event: cell death
model.addDeterministicEvent ( name="cell_death" ,
                            agent="Cell" ,
                            kind="destruction" ,
                            parameters={ "pkc1_threshold":0.1 } ,
                            trigger_expression="pkc1 < pkc1_threshold" )
# event: pkc1 deletion
model.addStochasticEvent ( name="pkc1_deletion" ,
                           agent="Cell" ,
                           parameters={ "pkc1_deletion_rate":0.1 } ,
                           propensity_expression="Cre * pkc1_deletion_rate" ,
                           realization_function=["pkc1_status = 0."] )

# continuous dynamics : cell aging
model.addContinuousChange ( name="cell_aging" , agent="Cell" , property="age" , rate_expression=
# continuous dynamics : IP degradation
model.addContinuousChange ( name="IP_degradation" , agent="Medium" , property="IP" ,
                            parameters={ "IP_degradation_rate":100. } ,
                            rate_expression = "- IP * IP_degradation_rate" )
# continuous dynamics : IP production
model.addContinuousChange ( name="IP_production" , agent="Medium" , property="IP" , agent_source
                            parameters={ "IP_prod_rate":0.5 } ,
                            rate_expression="IP_prod_rate" )
# continuous dynamics : Cre production and dilution
model.addContinuousChange ( name="cre_dynamics" , agent="Cell" , property="Cre" ,
                            parameters={ "Cre_prod_rate_per_IP":1.0 } ,
                            rate_expression="IP * Cre_prod_rate_per_IP - Cre * log(2.) / CC_leng
# continuous dynamics : pkc1 production and dilution
model.addContinuousChange ( name="pkc1_dynamics" , agent="Cell" , property="pkc1" ,
                            rate_expression="pkc1_status * log(2.) / CC_length - pkc1 * log(2.)


### model simulation (i.e. generation of cpp code for simulation)
simulating.writeAllCode ( model=model , target_folder="liquid_gol_cre_pkc1" )
```

Figure 5.7: Example of programmatic model description in *python*. The model represents a population of yeast cells bearing the high-threshold killing module and the IP production module (see Figure 4.3 in Chapter 4) that grow in liquid culture. Events (deterministic or stochastic) and differential dynamics are defined using dedicated functions.

state.hpp derivatives.hpp

```
void IntegratorDopr5::computeDerivatives ( State* state , VecDoub *y , VecDoub *dydt )
{

    for ( Int i = 0 ; i < y->size () ; i++ ) (*dydt)[i] = 0. ;

    // differential dynamics triggered by agent(s) of type : Cell
    for ( Uint ce = 0 ; ce < state->cells.size () ; ce++ )
    {
        Cell* cell = state->cells[ce] ;
        // cell_aging
        (*dydt)[ cell->I_age ] += 1. ;
        // IP_production
        (*dydt)[ state->medium.I_IP ] += Params::IP_prod_rate ;
        // cre_dynamics
        (*dydt)[ cell->I_Cre ] += (*y)[ state->medium.I_IP ] * Params::Cre_prod_rate_per_IP - (
        // pkc1_dynamics
        (*dydt)[ cell->I_pkc1 ] += (*y)[ cell->I_pkc1_status ] * log(2.) / (*y)[ cell->I_CC_len
    }

    // differential dynamics triggered by agent(s) of type : Medium
    // IP_degradation
    (*dydt)[ state->medium.I_IP ] += - (*y)[ state->medium.I_IP ] * Params::IP_degradation_rate
}
```

det_events.hpp stoch_events.hpp

```
Bool cell_division_trigger ( Cell* cell )
{
    return cell->age > cell->CC_length ;
}

Bool cell_death_trigger ( Cell* cell )
{
    return cell->pkc1 < Params::pkc1_threshold ;
}

void cell_division_realization ( Cell* cell , Cell* new_cell )
{
    new_cell->age = cell->age - cell->CC_length ;
    cell->age = new_cell->age ;
    new_cell->CC_length = ran.norm ( Params::CC_avg , Params::CC_std ) ;
    cell->CC_length = ran.norm ( Params::CC_avg , Params::CC_std ) ;
    new_cell->Cre = cell->Cre ;
    new_cell->pkc1 = cell->pkc1 ;
    new_cell->pkc1_status = cell->pkc1_status ;
}

void applyDeterministicEvents ()
{
    list <Cell*> kept_cells ;
    Uint num_cells = state->cells.size () ;
    for ( Uint ce = 0 ; ce < num_cells ; ce++ )
    {
        Cell* cell = state->cells.back () ; state->cells.pop_back () ;
```

Figure 5.8: Example of c++ simulation code generation from a programmatic model description. Portions of the c++ code generated from the python model description shown earlier (Figure 5.7) are shown.

# Chapter 6

# Outlook

## 6.1 Summary and relation to other works

### Extending the scope of biochemical pathway kinetic models by accounting for stochastic protein fluctuations

Kinetic models of biochemical pathways based on ordinary differential equations are now widespread and enabled key insights into the molecular-level, quantitative functioning of those pathways (Le Novère 2015). Yet, those models have difficulty to explain cell-to-cell variability observed in single-cell data, and there is an increasing awareness that cell-to-cell variability, for which gene expression noise is a major driver, can have important consequences at the level of cell populations (Blake et al. 2006; Balázsi et al. 2011). In this thesis, we initially reasoned that extending ODE-based models of biochemical pathways with stochastic protein fluctuation models for all its proteins could greatly extend the prediction scope of those models (Chapter 2). More precisely, the temporal scope should be extended compared to approaches that account for cell-to-cell variability in protein levels via static distribution in initial conditions (Spencer et al. 2009; Gaudet et al. 2012; Kallenberger et al. 2014).

Importantly, we have hypothesized that a precise knowledge of all the biochemical parameters involved in the expression of a given protein is not needed, provided that the amplitude and speed of the resulting protein fluctuations are sufficiently accurate. We have proposed an approach allowing such systematic inclusion of stochastic protein fluctuation models (Chapter 2) despite incomplete knowledge and have demonstrated its relevance on two examples. In the case of a large model of the TRAIL-induced apoptosis pathway, it enabled to correctly predict subtle measurements of how the cellular state of sister cells diverge from one another as time goes by because of stochastic fluctuations in the level of signaling proteins (Chapter 2). In the case of a synthetic sensing circuit implemented in yeast cells,

our approach permitted to reproduce its induction dynamics with high accuracy, temporal scope and at the single-cell level (Chapter 4).

Models of biochemical reactions that systematically combine stochastic gene expression with protein-protein reaction kinetics at the scale of a full signal transduction pathway (typically for more than 3 proteins) are rare. Perhaps the closest example to our work is the model of $NF\text{-}\kappa B$ used to investigate the digital nature of this pathway activation upon stimulation (Tay et al. 2010; Kellogg & Tay 2015). In this work, stochastic gene expression is modeled for several genes which products feedback on signaling dynamics. However, for several other proteins, their cell-to-cell variability is modeled via static distributions in initial conditions, hence setting a limit to the temporal scope of the model of the order of the mixing time for those proteins, as was the case for the ODE-based model of TRAIL-induced apoptosis that we extended in Chapter 2.

## Linking extended biochemical pathway kinetic models with cellular decisions to simulate cell population dynamics

What can we gain by extending the temporal scope of of biochemical pathway kinetic models by accounting for protein level fluctuations? The main motivation is in fact to enable the simulation of the dynamics of population of cells submitted to given environmental conditions with a level of detail that encompass the molecular scale. Indeed, the behavior of biochemical pathways ultimately governs cell growth, division, death and other cellular phenotypes. In such cases, fluctuations and memory in cells internal state arising from stochastic gene expression are expected to impact population dynamics, because selection effects will apply as soon as different cellular states confer different cell fitness in a given environment. Therefore, integrating extended biochemical pathway kinetic models into cell-based population dynamics models by linking biochemical variables with cellular decisions allows to investigate complex effects such as non-genetic selection operating on fluctuating cellular states. This was a key question we wanted to adress in the context of TRAIL-induced apoptosis (Chapter 2 and 3), as such non-genetic selection could be implicated in transient resistance acquisition observed after treatment (Flusberg et al. 2013). It could also have consequences for the quantitative functioning of the yeast patterning system studied in Chapter 4.

Since several years, the topic of non-genetic selection operating on fluctuating cellular states raised high interest. Theoretical studies considering a single gene which product impact fitness (Sato & Kaneko 2006; Mora & Walczak 2013) showed how the steady-state distribution of that product at steady-state is changed by selection effects. Experimental demonstration of such effects was obtained using synthetic circuits that exhibited random switching between low and high expression for proteins conferring higher metabolic capac-

ity (Acar et al. 2008) or resistance to a drug (Nevozhay et al. 2012). Also, in $E.coli$, fluctuations of a metabolic enzyme in conditions such that its level is highly limiting were shown to generate fluctuations in cell growth, as measured by cross-correlation analysis between expression level and instantaneous growth rate in single cells (Kiviet et al. 2014).

What is the contribution of selection effects into transient resistance acquisition of cancer cells to TRAIL? In Chapter 3, we integrated our TRAIL-induced apoptosis single-cell model into multi-cellular simulations to investigate the long-term response of populations repeatedly treated by TRAIL. We have found a strong increase of resistance in the long-term when comparing killing efficiencies of later treatments with first treatments. Interestingly, a detailed analysis revealed that transient resistance acquisition as predicted by our model involves both selection and 'adaptation' effects. Such 'adaptation' was caused by a cell-level decrease of pro-apoptotic protein levels relatively to anti-apoptotic protein levels, which was mainly mediated by targeted degradation of activated proteins and differential constitutive turnover between pro- and anti- apoptotic proteins. Thus, selection effects are always present but might be minor contributors to non-genetic resistance acquisition. It should be noted that this does not mean that modeling fluctuations was not necessary: 'static' cell-to-cell variability models would have grossly over-estimated selection effects, hence masking the contribution of cell-level resistance acquisition. To our knowledge, this mechanistic explanation for transient resistance acquisition is novel, highlighting the potential of our modeling approach to provide original insights.

Another important modeling work that relates to the question of non-genetic resistance acquisition in cancer cell populations has been recently published (Chisholm et al. 2015). It considers the emergence of drug tolerance in certain type of cancer cell populations treated by a chemotherapeutic agent, which has been experimentally linked to reversible epigenetic changes (Sharma et al. 2010). The model accounts for two phenomenological, stochastically fluctuating cell traits: their survival potential (i.e., drug resistance) and their proliferation potential. Therefore, under drug exposure, cells with low survival potential are eliminated while other can survive. In addition to this selection effect, the model also accounts for cell-level, stress-induced adaptation of the proliferation level: in presence of drug, cells decrease their proliferation. Hence, cell population dynamics is here governed by an interplay between phenotypic fluctuations, drug-mediated selection on those fluctuations, and drug-induced cellular adaptation. This situation is very similar to our simulations of repeated TRAIL treatments. However, a main difference is that our model is 'mechanistic' in the sense that cell response to TRAIL is governed by existing biochemical reactions between proteins. Hence, cell-level and selection-based resistance acquisition is an emerging property of the model and does not result from modeling choices.

'Mechanistic' cell population dynamics models with molecular details are rare but might become more common in the next years. In a recent study (Shokhirev et al. 2015), a multi-

scale model of how $NF\text{-}\kappa B$, cell cycle and apoptosis molecular pathways drives the population dynamics of lymphocytes was constructed and compared to experimental data. Because they relied on ODE models, they modeled protein level variability with static distributions. To solve the inherent problem with that (as already discussed, for example overestimation of selection effects), they decided to remix this 'extrinsic' noise by resampling protein levels from the corresponding distributions each time a cell divides. Our approach provides a less artificial and more mathematically-grounded solution to this problem, and hence is less likely to introduce biases and allows to use temporal decorrelation data to finely calibrate the model.

## From stochastic gene expression in cells to spatially-organized populations

While dissecting the quantitative functioning of cellular pathways in single cells requires precise control of the environment (Spiller et al. 2010), cells in a population are usually not sensing an homogeneous environment. Moreover, they often themselves create environmental heterogeneity by the way they grow and arrange in space or by adopting different phenotypes (Snijder et al. 2009; Hirata et al. 2015). For example, this is the case for tumor spheroids treated by TRAIL (as we studied in Chapter 3) or for cell populations able to send and sense messenger molecules (as for micro-colonies of engineered yeast cells studied in Chapter 4).

In this thesis, we used spatial cell-based modeling to investigate those situations. Spatial cell-based modeling has been increasingly used during the last 20 years to investigate the emerging behavior that results from environment-cell and cell-cell interactions (Graner & Glazier 1992; Drasdo et al. 1995; Galle et al. 2005; Walker et al. 2008; Leeuwen et al. 2009; Ramis-Conde & Drasdo 2012; Jang et al. 2012; Macklin et al. 2012; Schlüter et al. 2015). Most used relatively abstract rules for cellular decisions, while some modeled molecular species (Walker et al. 2008; Ramis-Conde & Drasdo 2012; Schlüter et al. 2015). However, spatial cell-based model that include stochastic gene expression in biochemical pathways are rare (to the best of our knowledge, this work represents the first attempt).

In the context of TRAIL-induced apoptosis, we have found that limited drug penetration within tumor spheroids can exacerbate non-genetic resistance acquisition as predicted for cell populations facing the same drug concentration but is unlikely to cause by itself long-term resistance to repeated treatments. We have focused on a single dimension of heterogeneity in cell micro-environment (drug concentration) but many other environmental factors might play a role, such as mechanical cues that can impact gene expression of many genes (that might include genes involved in the response to TRAIL) via mechanotransduction. For example, it has been found that E-cadherin expression, which is known

to be mechanically regulated, correlates with the expression of TRAIL death receptors (Lu et al. 2014).

In the context of the engineering of synthetic spatial systems, we have shown that our approach could enable the obtention of highly realistic, single-cell resolved predictions of the spatio-temporal response of micro-colonies bearing a sensing circuit and responding to dynamic gradients of messenger molecules (Chapter 4). More traditional modeling approaches used by synthetic biologists (such as ODE models for gene circuit dynamics that are extended into PDEs to represent space (Tabor et al. 2009) or integrated into spatial cell-based models (Gorochowski et al. 2012)) would have failed to reproduce the spatial heterogeneity in the micro-colonies response.

## 6.2 Limitations and perspectives

### Limitations of the single-cell models

One limitation of the single-cell models we used in this thesis (for TRAIL-induced apoptosis and sensing in yeast) is that neither cell size nor progression to the cell cycle is represented. As a result, when we model molecule copy numbers, it does not account for the fact that copy numbers usually increase with cell size (Marguerat & Bähler 2012). Thus our model represents 'effective' copy numbers rather than real copy numbers, and as a consequence when modeling division we do not split randomly the content into the two daughter cells but rather copy the state. This was justified in the case of TRAIL-induced apoptosis as 1) death time distributions were not dependent on cell cycle stage and 2) recently divided sister cells had very high correlation of death times, showing that splitting noise does not have a strong contribution (Spencer et al. 2009). However, it is possible to explicit the dependency with growth and the effect of splitting, as was done for a small network (Volfson et al. 2006). It might be required in conditions where growth can be highly variable between cells and is regulated by one or several molecular factors represented in the model.

Another limitation to those models is that despite their level of detail, many steps are not represented or highly simplified. For example, this is the case for receptor clustering (not modeled) and caspase-8 processing at the DISC (highly simplified) in the kinetic model of TRAIL-induced apoptosis we used (Ho & Harrington 2010; Dickens et al. 2012; Neumann et al. 2010). This might or might not be a problem depending on the purpose of the model and the capacity of the simplified representation used for a given set of reactions to realistically capture the input/output relationship of this set of reactions.

Another concern that apply to our modeling of TRAIL-induced apoptosis relates to the relevance of simulating cell populations for many generations with a model that represents only a single pathway with molecular detail: as time goes by, the risk that cross-talks with

other pathways starts interfering non-trivially with the modeled pathway to determine cell phenotypes increases. Our response to this concern is that it is anyway necessary to take this risk in order to push our understanding forward, but that it might be alleviated by adequate modeling approaches.

## Combining mechanistic and phenomenological models into coarse-grained whole-cell models

Indeed, while efforts spent into building whole-cell models that aim to represent as much molecular detail as possible are promising (Karr et al. 2012), the field of multi-cellular systems biology that address problems such as human diseases or animal development should not wait for the obtention of such models for human or animal cells before attempting to gain a molecular-level, quantitative understanding of the behavior of complex populations evolving in complex environments (such as developping tumors treated by anti-cancer drugs).

This imply to make choices regarding which aspects of cellular environmental states should be understood with molecular resolution and which aspects can be represented phenomenologically via abstract sub-models. The resulting models will thus necessarily mix different levels of detail/abstraction. For example, in our simulations of repeated TRAIL treatments on multi-cellular spheroids, the decision to enter apoptosis is modeled with molecular details while cell growth and division are modeled in a very phenomenological and simple manner. Such models will maybe fall short on explaining some observations, motivating either an improvement of the abstract sub-models (for example, accounting for memory effects in the cell cycle duration by using stochastic processes instead of sampling a cell cycle duration in the same distribution at each division) or the replacement of abstract/phenomenological sub-models with mechanistic, molecular level representations. While such models might seem un-elegant and somehow messy, they will probably be necessary for realizing the potential of multi-cellular systems biology to quantitatively explain complex biological phenomena with molecular detail.

Recently, a coarse-grained whole-cell model of *E. coli* has been proposed (Weisse et al. 2015). This model has been constructed to account for three fundamental cellular trade-offs that link gene expression, growth and metabolism: making proteins costs energy, mRNAs compete for translation, and the total proteome is constant (in steady-state growth). It correctly predicts empirical relationships between growth rate, nutrient levels and the ribosomal protein fraction. Thus, this model could be used as a 'chassis' into which mechanistic, molecular-level models are integrated in order to realistically simulate how a specific process interacts with the global physiology of cells. However, because this model abstracts progression within the cell cycle and cell division, a 'single-cell' version should be constructed first to enable the meaningful integration with single-cell models of biochemical

reaction pathways.

## Using synthetic circuits to study the population consequences of cellular decisions

Building quantitative models in order to realistically simulate the dynamics of cell populations with molecular detail is still very difficult and is currently rarely attempted. Much more progress is needed to establish methods and concepts that are robust enough to make such modeling rewarding in terms of biological insights. We believe that multi-cellular synthetic biology can play a decisive role towards this goal. This could be realized by the combined experimental and theoretical investigation of how synthetic gene circuits (that are designed, well-characterized and tunable) can interact with, and eventually control, the dynamics of cell populations. The gain for systems biology research could be double: in addition to establishing modeling methods and concepts, it is likely to provide insights on the 'design principles' by which natural gene circuits implement various, complex and robust functions at the population level.

# Appendix 1: Derivation of the auto-correlation function for the stochastic protein turnover model

We now present the derivation of the autocorrelation function. The first step is to obtain the differential equations describing the temporal evolution of the first moments. Such equations are obtained by using the master equation, which gives how the joint probability distribution $p^t(G=g, m=x, P=y)$ is changed during an infinitesimal time interval $dt$ because of the reactions that can occur:

$$\frac{dp^{t+dt}(g,x,y)}{dt} = \left[k_{off}(1-g)+k_{on}g\right]p^t(1-g,x,y)+k_{sm}g\,p^t(g,x-1,y)$$
$$+\gamma_m(x+1)p^t(g,x+1,y)+k_{sp}x\,p^t(g,x,y-1)+\gamma_p(y+1)p^t(g,x,y+1)$$
$$-\left[k_{on}(1-g)+k_{off}g+k_{sm}g+\gamma_m x+k_{sp}x+\gamma_p y\right]p^t(g,x,y)$$

For example, if one is interested in the temporal evolution of $\mathbb{E}[G]$:

$$\frac{d\mathbb{E}[G]}{dt} = \frac{d\left(\sum_{g,x,y} g p^t(g,x,y)\right)}{dt}$$
$$= \sum_{g,x,y} g\frac{d\left(p^t(g,x,y)\right)}{dt}$$
$$= \sum_{g,x,y} g\left[k_{off}(1-g)+k_{on}g\right]p^t(1-g,x,y)+k_{sm}g^2p^t(g,x-1,y)$$
$$+g\gamma_m(x+1)p^t(g,x+1,y)+gk_{sp}x\,p^t(g,x,y-1)+g\gamma_p(y+1)p^t(g,x,y+1)$$

$$-g\left[k_{on}\left(1-g\right)+k_{off}g+k_{sm}g+\gamma_m x+k_{sp}x+\gamma_p y\right]p^t\left(g,x,y\right)$$

Noticing that $g\left(1-g\right)=0$, $g^2=g$ and that the probability distribution cannot charge states where $x=-1$ or $y=-1$, one gets:

$$\frac{d\mathbb{E}\left[G\right]}{dt}=k_{on}\sum_{g,x,y}gp^t\left(1-g,x,y\right)-k_{off}\sum_{g,x,y}gp^t\left(g,x,y\right)$$

$$=k_{on}\left(1-\mathbb{E}\left[G\right]\right)-k_{off}\mathbb{E}\left[G\right]$$

$$=k_{on}-\gamma_g\mathbb{E}\left[G\right]$$

With similar calculations, one also obtains the two following equations:

$$\frac{d\mathbb{E}\left[m\right]}{dt}=k_{sm}\mathbb{E}\left[G\right]-\gamma_m\mathbb{E}\left[m\right]$$

$$\frac{d\mathbb{E}\left[P\right]}{dt}=k_{sp}\mathbb{E}\left[m\right]-\gamma_p\mathbb{E}\left[P\right]$$

This system of autonomous ordinary differential equations is linear and can be solved analytically. In particular, we have:

$$\mathbb{E}\left[P\right]\left(t+T\right)=f\left(t,\mathbb{E}\left[G\right]\left(T\right),\mathbb{E}\left[m\right]\left(T\right),\mathbb{E}\left[P\right]\left(T\right)\right)$$

$$=\mathbb{E}\left[G\right]\left(T\right)c_G\left(t\right)+\mathbb{E}\left[m\right]\left(T\right)c_m\left(t\right)+\mathbb{E}\left[P\right]\left(T\right)c_P\left(t\right)+c_0\left(t\right)$$

where:

$$c_P\left(t\right)=e^{-\gamma_p t}$$

$$c_m\left(t\right)=\frac{\mathbb{E}_{eq}\left[P\right]}{\mathbb{E}_{eq}\left[m\right]}\left\{e^{-\gamma_m t}-e^{-\gamma_p t}\right\}$$

$$c_G\left(t\right)=\frac{\mathbb{E}_{eq}\left[P\right]}{\mathbb{E}_{eq}\left[G\right]}\frac{\gamma_m\gamma_p}{\left(\gamma_p-\gamma_m\right)\left(\gamma_m-\gamma_g\right)\left(\gamma_p-\gamma_g\right)}\left\{\left(\gamma_m-\gamma_g\right)e^{-\gamma_p t}+\left(\gamma_g-\gamma_p\right)e^{-\gamma_m t}+\left(\gamma_p-\gamma_m\right)e^{-\gamma_g t}\right\}$$

$$c_0\left(t\right)=\mathbb{E}_{eq}\left[P\right]-\mathbb{E}_{eq}\left[P\right]c_P\left(t\right)-\mathbb{E}_{eq}\left[m\right]c_m\left(t\right)-\mathbb{E}_{eq}\left[G\right]c_G\left(t\right)$$

This particular ordering has been chosen to facilitate the next step of the auto-correlation function derivation, which is to find an appropriate expression for the quantity $\mathbb{E}_{eq}\left[P\left(t\right)P\left(0\right)\right]$:

$$\mathbb{E}_{eq}\left[P(t)P(0)\right]=\sum_{g_0,x_0,y_0}\sum_{g,x,y}yy_0\mathbb{P}\left(\begin{array}{l|l} G(t)=g & G(0)=g_0 \\ m(t)=x & m(0)=x_0 \\ P(t)=y & P(0)=y_0 \end{array}\right)\mathbb{P}_{eq}(g_0,x_0,y_0)$$

letting the expectation of $P(t)$ when starting from a dirac in $(g_0,x_0,y_0)$ at $t=0$ appear:

$$=\sum_{g0,x0,y0}y_0\mathbb{E}\left(\begin{array}{l|l} P(t) & G(0)=g_0 \\ & m(0)=x_0 \\ & P(0)=y_0 \end{array}\right)\mathbb{P}_{eq}(g_0,x_0,y_0)$$

$$=\sum_{g0,x0,y0}y_0 f(t,g_0,x_0,y_0)\mathbb{P}_{eq}(g_0,x_0,y_0)$$

$$=\sum_{g0,x0,y0}\left[y_0g_0c_G(t)+y_0x_0c_m(t)+{y_0}^2c_P(t)+y_0c_0(t)\right]\mathbb{P}_{eq}(g_0,x_0,y_0)$$

letting crossed second order moments of the steady-state distribution appear:

$$=\mathbb{E}_{eq}\left[GP\right]c_G(t)+\mathbb{E}_{eq}\left[mP\right]c_m(t)+\mathbb{E}_{eq}\left[P^2\right]c_P(t)+\mathbb{E}_{eq}\left[P\right]c_0(t)\mathbb{E}_{eq}\left[P\right]$$

We thus have:

$$\mathbb{E}_{eq}\left[P(t)P(0)\right]-\mathbb{E}_{eq}\left[P\right]^2=Var(P)c_P(t)+Cov(m,P)c_m(t)+Cov(G,P)c_G(t)$$

and therefore:

$$A\left[P\right](t)=\frac{\mathbb{E}_{eq}\left[P(t)P(0)\right]-\mathbb{E}_{eq}\left[P\right]^2}{Var(P)}=c_P(t)+\frac{Cov(m,P)}{Var(P)}c_m(t)+\frac{Cov(G,P)}{Var(P)}c_G(t)$$

Finally, after having plugged the expressions for the covariances (given in [39]) and for the functions $c_G(t),c_m(t)$ and $c_P(t)$, we get a final expression (i.e. containing only terms for which expression in terms of rate parameters has been given already) for the autocorrelation function:

$$A[P](t) = e^{-\gamma_p t} + h_m(\gamma_m, \gamma_p, t) + \frac{1}{CV[P]^2} \frac{1 - \mathbb{E}[G]}{\mathbb{E}[G]} \frac{\gamma_p \gamma_m}{(\gamma_m + \gamma_g)(\gamma_p + \gamma_g)} h_g(\gamma_g, \gamma_m, \gamma_p, t)$$

with:

$$h_m(\gamma_m, \gamma_p, t) = \frac{\gamma_p}{\gamma_p - \gamma_m} \left\{ e^{-\gamma_m t} - e^{-\gamma_p t} \right\}$$

$$h_g(\gamma_g, \gamma_m, \gamma_p, t) = \frac{\gamma_p \gamma_m}{(\gamma_p - \gamma_m)(\gamma_m - \gamma_g)(\gamma_p - \gamma_g)} \left\{ (\gamma_m - \gamma_g) e^{-\gamma_p t} + (\gamma_p - \gamma_m) e^{-\gamma_g t} + (\gamma_g - \gamma_p) e^{-\gamma_m t} \right\}$$

The exactness of this expression was verified by Monte-Carlo simulations using Gillespie algorithm, for a parameter set carefully chosen such that each term in the expression has a non-negligible contribution to the function value. One can note how the three time-scales $\gamma_g, \gamma_m, \gamma_p$, corresponding to the promoter, the mRNA and the protein respectively, are convoluted together in this expression.

# Appendix 2: Parameters values of the single-cell TRAIL-induced apoptosis model

**Standard stochastic protein turnover models**

| | Constraints | | | | Rate values | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Protein | Switching speed $k_{on}+k_{off}$ ($hrs^{-1}$) *fixed* | Mean mRNA level *fixed* | Mean protein level *fixed* | Protein level Coefficient of Variation *fixed* | Off to on promoter rate $k_{on}$ ($hrs^{-1}$) *derived* | On to off promoter rate $k_{off}$ ($hrs^{-1}$) *derived* | $k_{sm}$ ($min^{-1}$) *derived* | mRNA deg. rate $\gamma_m$ ($hrs^{-1}$) *fixed* | Protein synth. rate $k_{sp}$ ($min^{-1}$) *derived* | Protein deg. rate $\gamma_p$ ($hrs^{-1}$) *fixed* |
| Receptor | 10.44 | 17 | 1000 | 0.25 | 0.388 | 10.05 | 0.587 | 0.077 | 0.025 | 0.0257 |
| Caspase-8 | | | 10000 | 0.25 | 0.388 | 10.05 | 0.587 | | 0.25 | |
| Bar | | | 1000 | 0.25 | 0.388 | 10.05 | 0.587 | | 0.025 | |
| Caspase-3 | | | 10000 | 0.282 | 0.289 | 10.15 | 0.789 | | 0.25 | |
| Caspase-6 | | | 10000 | 0.25 | 0.388 | 10.05 | 0.587 | | 0.25 | |
| XIAP | | | 100000 | 0.288 | 0.275 | 10.17 | 0.829 | | 2.52 | |
| PARP | | | 1000000 | 0.25 | 0.388 | 10.05 | 0.587 | | 25.2 | |
| Bid | | | 60000 | 0.288 | 0.275 | 10.17 | 0.829 | | 1.51 | |
| Bax | | | 80000 | 0.271 | 0.318 | 10.12 | 0.717 | | 2.01 | |
| Bcl-2 | | | 30000 | 0.294 | 0.268 | 10.18 | 0.871 | | 0.76 | |
| Pore | | | 500000 | 0.25 | 0.388 | 10.05 | 0.587 | | 12.6 | |
| CytoC_m | | | 500000 | 0.25 | 0.388 | 10.05 | 0.587 | | 12.6 | |
| Smac | | | 100000 | 0.25 | 0.388 | 10.05 | 0.587 | | 2.51 | |
| Apaf | | | 100000 | 0.25 | 0.388 | 10.05 | 0.587 | | 2.51 | |
| Caspase-9 | | | 100000 | 0.25 | 0.388 | 10.05 | 0.587 | | 2.51 | |

**Specific stochastic protein turnover models for Flip and Mcl1: "non-fitted" model**

| | Constraints | | Rate values | | | | | |
|---|---|---|---|---|---|---|---|---|
| Protein | Mean mRNA level | Mean protein level | $k_{on}$ ($hrs^{-1}$) | $k_{off}$ ($hrs^{-1}$) | $k_{sm}$ ($min^{-1}$) | $\gamma_m$ ($hrs^{-1}$) | $k_{sp}$ ($min^{-1}$) | $\gamma_p$ ($hrs^{-1}$) |
| Flip | 17 | 2000 | 0.388 | 10.05 | 2.6186 | 0.3466 | 2.7182 | 1.39 |
| Mcl-1 | 17 | 20000 | 0.388 | 10.05 | 2.6186 | 0.3466 | 27.182 | 1.39 |

**Specific stochastic protein turnover models for Flip and Mcl1: "fitted" model**

| | Constraints | | Rate values | | | | | |
|---|---|---|---|---|---|---|---|---|
| Protein | Mean mRNA level | Mean protein level | $k_{on}$ ($hrs^{-1}$) | $k_{off}$ ($hrs^{-1}$) | $k_{sm}$ ($min^{-1}$) | $\gamma_m$ ($hrs^{-1}$) | $k_{sp}$ ($min^{-1}$) | $\gamma_p$ ($hrs^{-1}$) |
| Flip | 17 | 2000 | 0.0417 | 0.0625 | 0.4910 | 0.6931 | 3.3978 | 1.73 |
| Mcl-1 | 17 | 20000 | 0.0417 | 0.0625 | 0.4910 | 0.6931 | 33.9780 | 1.73 |

**Degradation of non-native forms**

| Protein | Half-life (hours) |
|---|---|
| TRAIL | 9 |
| Pore* | 1.9 |
| Flip:Receptor | same as Flip |
| Mcl-1:tBid | same as Mcl-1 |
| All others | 5 |

## Model equations

Here we give all the differential equations governing the evolution of protein levels. The mRNA levels are governed by stochastic reactions as described earlier.

$$
\begin{aligned}
\frac{dR}{dt} &= k_{sp}^{R}\ mRNA_{R} - \gamma_{p}^{R}\ R - k_1\ TRAIL\ R + k_2\ TRAIL : R \\
&+ k_4\ R^{*} \\
\frac{dpC8}{dt} &= k_{sp}^{pC8}\ mRNA_{pC8} - \gamma_{p}^{pC8}\ pC8 - k_5\ pC8\ R^{*} + k_6\ pC8 : R^{*} \\
&- k_{14}\ pC8\ C6 + k_{15}\ pC8 : C6 \\
\frac{dpC3}{dt} &= k_{sp}^{pC3}\ mRNA_{pC3} - \gamma_{p}^{pC3}\ pC3 - k_8\ pC3\ C8 + k_9\ pC3 : C8 \\
&- k_{69}\ pC3\ Apop + k_{70}\ pC3 : Apop \\
\frac{dpC6}{dt} &= k_{sp}^{pC6}\ mRNA_{pC6} - \gamma_{p}^{pC6}\ pC6 - k_{11}\ pC6\ C3 + k_{12}\ pC6 : C3 \\
\frac{dPARP}{dt} &= k_{sp}^{PARP}\ mRNA_{PARP} - \gamma_{p}^{PARP}\ PARP - k_{17}\ PARP\ C3 + k_{18}\ PARP : C3 \\
\frac{dflip}{dt} &= k_{sp}^{flip}\ mRNA_{flip} - \gamma_{p}^{flip}\ flip - k_{20}\ flip\ R^{*} + k_{21}\ flip : R^{*} \\
\frac{dBar}{dt} &= k_{sp}^{Bar}\ mRNA_{Bar} - \gamma_{p}^{Bar}\ Bar - k_{22}\ Bar\ C8 + k_{23}\ Bar : C8 \\
\frac{dXIAP}{dt} &= k_{sp}^{XIAP}\ mRNA_{XIAP} - \gamma_{p}^{XIAP}\ XIAP - k_{24}\ C3\ XIAP + k_{25}\ C3 : XIAP \\
&+ k_{26}\ C3 : XIAP - k_{50}\ XIAP\ Apop + k_{51}\ XIAP : Apop \\
&- k_{62}\ smac\ XIAP + k_{63}\ smac : XIAP \\
\frac{dBid}{dt} &= k_{sp}^{Bid}\ mRNA_{Bid} - \gamma_{p}^{Bid}\ Bid - k_{27}\ Bid\ C8 + k_{28}\ Bid : C8 \\
\frac{dBax}{dt} &= k_{sp}^{Bax}\ mRNA_{Bax} - \gamma_{p}^{Bax}\ Bax - k_{30}\ Bax\ tBid + k_{31}\ Bax : tBid \\
\frac{dPore}{dt} &= k_{sp}^{Pore}\ mRNA_{Pore} - \gamma_{p}^{Pore}\ Pore - k_{39}\ Pore\ Bax_{m,4}^{*} + k_{40}\ Pore : Bax_{m,4}^{*} \\
\frac{dsmac_m}{dt} &= k_{sp}^{smac_m}\ mRNA_{smac_m} - \gamma_{p}^{smac_m}\ smac_m - k_{55}\ smac_m\ Pore^{*} + k_{56}\ smac_m : Pore^{*} \\
\frac{dCyC_m}{dt} &= k_{sp}^{CyC_m}\ mRNA_{CyC_m} - \gamma_{p}^{CyC_m}\ CyC_m - k_{52}\ CyC_m\ Pore^{*} + k_{53}\ CyC_m : Pore^{*} \\
\frac{dpC9}{dt} &= k_{sp}^{pC9}\ mRNA_{pC9} - \gamma_{p}^{pC9}\ pC9 - k_{67}\ Apaf^{*}\ pC9 + k_{68}\ Apop \\
\frac{dApaf}{dt} &= k_{sp}^{Apaf}\ mRNA_{Apaf} - \gamma_{p}^{Apaf}\ Apaf - k_{64}\ Apaf\ CyC + k_{65}\ Apaf : CyC \\
\frac{dMcl1}{dt} &= k_{sp}^{Mcl1}\ mRNA_{Mcl1} - \gamma_{p}^{Mcl1}\ Mcl1 - k_{42}\ Mcl1\ tBid + k_{43}\ Mcl1 : tBid \\
\frac{dBcl2}{dt} &= k_{sp}^{Bcl2}\ mRNA_{Bcl2} - \gamma_{p}^{Bcl2}\ Bcl2 - k_{44}\ Bcl2\ Bax_{m}^{*} + k_{45}\ Bcl2 : Bax_{m}^{*} \\
&- k_{46}\ Bcl2\ Bax_{m,2}^{*} + k_{47}\ Bcl2 : Bax_{m,2}^{*} - k_{48}\ Bcl2\ Bax_{m,4}^{*} \\
&+ k_{49}\ Bcl2 : Bax_{m,4}^{*}
\end{aligned}
$$

$$
\begin{aligned}
\frac{dTRAIL}{dt} &= -\gamma_p^{TRAIL}\ TRAIL - k_1\ TRAIL\ R + k_2\ TRAIL : R \\
&\quad + k_4\ R^* \\
\frac{dTRAIL : R}{dt} &= -\gamma_p^{TRAIL:R}\ TRAIL : R + k_1\ TRAIL\ R - k_2\ TRAIL : R \\
&\quad - k_3\ TRAIL : R \\
\frac{dR^*}{dt} &= -\gamma_p^{R^*}\ R^* + k_3\ TRAIL : R - k_4\ R^* \\
&\quad - k_5\ pC8\ R^* + k_6\ pC8 : R^* \\
&\quad + k_7\ pC8 : R^* - k_{20}\ flip\ R^* \\
&\quad + k_{21}\ flip : R^* \\
\frac{dpC8 : R^*}{dt} &= -\gamma_p^{pC8:R^*}\ pC8 : R^* + k_5\ pC8\ R^* - k_6\ pC8 : R^* \\
&\quad - k_7\ pC8 : R^* \\
\frac{dC8}{dt} &= -\gamma_p^{C8}\ C8 + k_7\ pC8 : R^* - k_8\ pC3\ C8 \\
&\quad + k_9\ pC3 : C8 + k_{10}\ pC3 : C8 \\
&\quad + k_{16}\ pC8 : C6 - k_{22}\ Bar\ C8 \\
&\quad + k_{23}\ Bar : C8 - k_{27}\ Bid\ C8 \\
&\quad + k_{28}\ Bid : C8 + k_{29}\ Bid : C8 \\
\frac{dpC3 : C8}{dt} &= -\gamma_p^{pC3:C8}\ pC3 : C8 + k_8\ pC3\ C8 - k_9\ pC3 : C8 \\
&\quad - k_{10}\ pC3 : C8 \\
\frac{dC3}{dt} &= -\gamma_p^{C3}\ C3 + k_{10}\ pC3 : C8 - k_{11}\ pC6\ C3 \\
&\quad + k_{12}\ pC6 : C3 + k_{13}\ pC6 : C3 \\
&\quad - k_{17}\ PARP\ C3 + k_{18}\ PARP : C3 \\
&\quad + k_{19}\ PARP : C3 - k_{24}\ C3\ XIAP \\
&\quad + k_{25}\ C3 : XIAP + k_{71}\ pC3 : Apop \\
\frac{dpC6 : C3}{dt} &= -\gamma_p^{pC6:C3}\ pC6 : C3 + k_{11}\ pC6\ C3 - k_{12}\ pC6 : C3 \\
&\quad - k_{13}\ pC6 : C3 \\
\frac{dC6}{dt} &= -\gamma_p^{C6}\ C6 + k_{13}\ pC6 : C3 - k_{14}\ pC8\ C6 \\
&\quad + k_{15}\ pC8 : C6 + k_{16}\ pC8 : C6 \\
\frac{dpC8 : C6}{dt} &= -\gamma_p^{pC8:C6}\ pC8 : C6 + k_{14}\ pC8\ C6 - k_{15}\ pC8 : C6 \\
&\quad - k_{16}\ pC8 : C6 \\
\frac{dPARP : C3}{dt} &= -\gamma_p^{PARP:C3}\ PARP : C3 + k_{17}\ PARP\ C3 - k_{18}\ PARP : C3 \\
&\quad - k_{19}\ PARP : C3
\end{aligned}
$$

$$
\begin{aligned}
\frac{dcPARP}{dt} &= -\gamma_p^{cPARP}\ cPARP + k_{19}\ PARP : C3 \\
\frac{dflip : R^*}{dt} &= -\gamma_p^{flip:R^*}\ flip : R^* + k_{20}\ flip\ R^* - k_{21}\ flip : R^* \\
\frac{dBar : C8}{dt} &= -\gamma_p^{Bar:C8}\ Bar : C8 + k_{22}\ Bar\ C8 - k_{23}\ Bar : C8 \\
\frac{dC3 : XIAP}{dt} &= -\gamma_p^{C3:XIAP}\ C3 : XIAP + k_{24}\ C3\ XIAP - k_{25}\ C3 : XIAP \\
&\quad - k_{26}\ C3 : XIAP \\
\frac{dC3 : ub}{dt} &= -\gamma_p^{C3:ub}\ C3 : ub + k_{26}\ C3 : XIAP \\
\frac{dBid : C8}{dt} &= -\gamma_p^{Bid:C8}\ Bid : C8 + k_{27}\ Bid\ C8 - k_{28}\ Bid : C8 \\
&\quad - k_{29}\ Bid : C8 \\
\frac{dtBid}{dt} &= -\gamma_p^{tBid}\ tBid + k_{29}\ Bid : C8 - k_{30}\ Bax\ tBid \\
&\quad + k_{31}\ Bax : tBid + k_{32}\ Bax : tBid \\
&\quad - k_{42}\ Mcl1\ tBid + k_{43}\ Mcl1 : tBid \\
\frac{dBax : tBid}{dt} &= -\gamma_p^{Bax:tBid}\ Bax : tBid + k_{30}\ Bax\ tBid - k_{31}\ Bax : tBid \\
&\quad - k_{32}\ Bax : tBid \\
\frac{dBax^*}{dt} &= -\gamma_p^{Bax^*}\ Bax^* + k_{32}\ Bax : tBid - k_{33}\ Bax^* \\
&\quad + k_{34}\ Bax_m^* \\
\frac{dBax_m^*}{dt} &= -\gamma_p^{Bax_m^*}\ Bax_m^* + k_{33}\ Bax^* - k_{34}\ Bax_m^* \\
&\quad - 2k_{35}\ Bax_m^*\ Bax_m^* + 2k_{36}\ Bax_{m,2}^* \\
&\quad - k_{44}\ Bcl2\ Bax_m^* + k_{45}\ Bcl2 : Bax_m^* \\
\frac{dBax_{m,2}^*}{dt} &= -\gamma_p^{Bax_{m,2}^*}\ Bax_{m,2}^* + k_{35}\ Bax_m^*\ Bax_m^* - k_{36}\ Bax_{m,2}^* \\
&\quad - 2k_{37}\ Bax_{m,2}^*\ Bax_{m,2}^* + 2k_{38}\ Bax_{m,4}^* \\
&\quad - k_{46}\ Bcl2\ Bax_{m,2}^* + k_{47}\ Bcl2 : Bax_{m,2}^*
\end{aligned}
$$

$$\frac{dBax^*_{m,4}}{dt} = -\gamma_p^{Bax^*_{m,4}}\ Bax^*_{m,4} + k_{37}\ Bax^*_{m,2}\ Bax^*_{m,2} - k_{38}\ Bax^*_{m,4}$$
$$- k_{39}\ Pore\ Bax^*_{m,4} + k_{40}\ Pore : Bax^*_{m,4}$$
$$- k_{48}\ Bcl2\ Bax^*_{m,4} + k_{49}\ Bcl2 : Bax^*_{m,4}$$
$$\frac{dPore : Bax^*_{m,4}}{dt} = -\gamma_p^{Pore:Bax^*_{m,4}}\ Pore : Bax^*_{m,4} + k_{39}\ Pore\ Bax^*_{m,4} - k_{40}\ Pore : Bax^*_{m,4}$$
$$- k_{41}\ Pore : Bax^*_{m,4}$$
$$\frac{dPore^*}{dt} = -\gamma_p^{Pore^*}\ Pore^* + k_{41}\ Pore : Bax^*_{m,4} - k_{52}\ CyC_m\ Pore^*$$
$$+ k_{53}\ CyC_m : Pore^* + k_{54}\ CyC_m : Pore^*$$
$$- k_{55}\ smac_m\ Pore^* + k_{56}\ smac_m : Pore^*$$
$$+ k_{57}\ smac_m : Pore^*$$
$$\frac{dMcl1 : tBid}{dt} = -\gamma_p^{Mcl1:tBid}\ Mcl1 : tBid + k_{42}\ Mcl1\ tBid - k_{43}\ Mcl1 : tBid$$
$$\frac{dBcl2 : Bax^*_m}{dt} = -\gamma_p^{Bcl2:Bax^*_m}\ Bcl2 : Bax^*_m + k_{44}\ Bcl2\ Bax^*_m - k_{45}\ Bcl2 : Bax^*_m$$
$$\frac{dBcl2 : Bax^*_{m,2}}{dt} = -\gamma_p^{Bcl2:Bax^*_{m,2}}\ Bcl2 : Bax^*_{m,2} + k_{46}\ Bcl2\ Bax^*_{m,2} - k_{47}\ Bcl2 : Bax^*_{m,2}$$
$$\frac{dBcl2 : Bax^*_{m,4}}{dt} = -\gamma_p^{Bcl2:Bax^*_{m,4}}\ Bcl2 : Bax^*_{m,4} + k_{48}\ Bcl2\ Bax^*_{m,4} - k_{49}\ Bcl2 : Bax^*_{m,4}$$
$$\frac{dApop}{dt} = -\gamma_p^{Apop}\ Apop - k_{50}\ XIAP\ Apop + k_{51}\ XIAP : Apop$$
$$+ k_{67}\ Apaf^*\ pC9 - k_{68}\ Apop$$
$$- k_{69}\ pC3\ Apop + k_{70}\ pC3 : Apop$$
$$+ k_{71}\ pC3 : Apop$$
$$\frac{dXIAP : Apop}{dt} = -\gamma_p^{XIAP:Apop}\ XIAP : Apop + k_{50}\ XIAP\ Apop - k_{51}\ XIAP : Apop$$
$$\frac{dCyC_m : Pore^*}{dt} = -\gamma_p^{CyC_m:Pore^*}\ CyC_m : Pore^* + k_{52}\ CyC_m\ Pore^* - k_{53}\ CyC_m : Pore^*$$
$$- k_{54}\ CyC_m : Pore^*$$
$$\frac{dCyC : r}{dt} = -\gamma_p^{CyC:r}\ CyC : r + k_{54}\ CyC_m : Pore^* - k_{58}\ CyC : r$$
$$+ k_{59}\ CyC$$

$$\frac{dsmac_m : Pore^*}{dt} = -\gamma_p^{smac_m:Pore^*}\ smac_m : Pore^* + k_{55}\ smac_m\ Pore^* - k_{56}\ smac_m : Pore^*$$
$$- k_{57}\ smac_m : Pore^*$$
$$\frac{dsmac : r}{dt} = -\gamma_p^{smac:r}\ smac : r + k_{57}\ smac_m : Pore^* - k_{60}\ smac : r$$
$$+ k_{61}\ smac$$
$$\frac{dCyC}{dt} = -\gamma_p^{CyC}\ CyC + k_{58}\ CyC : r - k_{59}\ CyC$$
$$- k_{64}\ Apaf\ CyC + k_{65}\ Apaf : CyC$$
$$+ k_{66}\ Apaf : CyC$$
$$\frac{dsmac}{dt} = -\gamma_p^{smac}\ smac + k_{60}\ smac : r - k_{61}\ smac$$
$$- k_{62}\ smac\ XIAP + k_{63}\ smac : XIAP$$
$$\frac{dsmac : XIAP}{dt} = -\gamma_p^{smac:XIAP}\ smac : XIAP + k_{62}\ smac\ XIAP - k_{63}\ smac : XIAP$$
$$\frac{dApaf : CyC}{dt} = -\gamma_p^{Apaf:CyC}\ Apaf : CyC + k_{64}\ Apaf\ CyC - k_{65}\ Apaf : CyC$$
$$- k_{66}\ Apaf : CyC$$
$$\frac{dApaf^*}{dt} = -\gamma_p^{Apaf^*}\ Apaf^* + k_{66}\ Apaf : CyC - k_{67}\ Apaf^*\ pC9$$
$$+ k_{68}\ Apop$$
$$\frac{dpC3 : Apop}{dt} = -\gamma_p^{pC3:Apop}\ pC3 : Apop + k_{69}\ pC3\ Apop - k_{70}\ pC3 : Apop$$
$$- k_{71}\ pC3 : Apop$$

# Appendix 3: Simulation of the single-cell model of TRAIL-induced apoptosis and comparison with experimental data

## Simulating TRAIL-induced apoptosis with stochastic protein turnover

In general, the TRAIL signaling protein-protein reactions are taking place concurrently with stochastic protein turnover. When the noise in signaling reactions is neglected due to high protein copy number, those reactions can be simulated using ODEs. However, rates of protein synthesis are in our model stochastic, as they follow mRNA fluctuations. Promoter activity and mRNA fluctuations were simulated using an implementation of the Gillespie algorithm in C++ (Numerical Recipes). Messenger RNA trajectories were computed and stored in advance because protein levels do not affect the rates of promoter state switches, mRNA production and degradation (see Chapter 5). The ODEs governing evolution of all protein levels were then simulated using the Semi-Implicit Extrapolation method implemented in C++ ((Deuflhard 1985), Numerical Recipes). This method was significantly faster than a more standard Runge-Kutta method (Dormand-Prince, C++, Numerical Recipes) but gave identical results (see Chapter 5, section *Simulating combined signal transduction and stochastic gene expression*).

## In-silico sister cells experiment

To sample the state (promoter activity, mRNA and protein levels) of the mother cells, all stochastic protein turnover models were simulated during 25 days (starting for the mean state) for each of the $10^4$ (or $10^5$ for Figures 2.8, 2.9) mother cells. This duration was verified by comparison with analytical results to be sufficient to reach the steady-state distribution. Sister cells were simply constructed by duplication of the mother cell state. Because in experiments from (Spencer et al. 2009), the distribution of durations between division and treatment was not uniform (see Figures S5-b) and 1-g) in this reference), we applied a sampling algorithm to approximately reproduce those distributions. The overall impact on

correlation curves was generally low compared to results obtained with assuming a fully uniform distribution of division times in the pre-stimulus recording interval. MOMP was considered to have occurred when half of mitochondrial Smac has been released.

## In-silico repeated TRAIL experiment

A naïve population of $10^4$ cells was obtained as in the sister cells experiment. Each cell was assigned a random time of next division. To account for the fact that the distribution of next division times is not uniform in growing cell populations, we used a distribution obtained by simulating simple growth. New cells were attributed a next division time according to a cell cycle duration normally distributed with 27 hours mean and 3 hours standard deviation. Cells in which cPARP levels exceeded $10^5$ were considered dead as in (Gaudet et al. 2012). To closely mimic the experimental protocol used in (Flusberg et al. 2013), we accounted for the effect of passing cells by checking population size each day and if needed, removing randomly cells until $10^4$ were left. Resistance gain is computed as $RG = \frac{Resistance^{survivors}}{Resistance^{naive\,cells}}$ where $Resistance = \frac{AliveCellNumber\,(treatment+8hours)}{AliveCellNumber\,(treatment)}$, similarly to (Flusberg et al. 2013).

## Quantification of model-data agreement

For the estimation of FLIP/Mcl-1 model rates based on cell fate variability experimental data; for the validation against transient cell fate inheritance data and for robustness analysis (Figure 2.12), it was needed to quantify the agreement/discrepancy between each model tested and the observed data. Such quantification was performed as follows.

**MOMP time distribution** Data was extracted from (Spencer et al. 2009) (Figures S4 b) and c)). It consists in MOMP time histograms (number of cells which did MOMP in a given 20 minutes time interval between 0 and 8 hours after treatment, 24 intervals in total). It was transformed in MOMP time frequencies by dividing by the total cell number. The same MOMP time frequencies were computed from simulated results. An agreement cost was then computed as the squared deviation between the two sets of frequencies, which respectively represent the empirical/model MOMP time distributions.

**Surviving fraction** An agreement cost for surviving fractions was simply computed as the squared difference between the surviving fractions observed experimentally and in simulations 8 hours after treatment.

**Sister cell MOMP time correlation curve** Spencer et al. (2009) quantified the transient inheritance of MOMP times by computing a curve of sister cells MOMP time correlation as follows: pairs of sister cells for which both cells did MOMP before 8 hours were sorted as a function of the average time between division and MOMP, and linear regression correlation coefficients were computed for all groups obtained by sliding a window of constant size along the sorted pairs. For each group, mean time between division and MOMP was also computed, thus providing the abscissa of the corresponding point in the curve. From this data (Figure S5 d) in (Spencer et al. 2009)), twenty representative points were extracted. To compute a comparable curve from simulations results, we applied the same quantification of sister cell MOMP time correlations. The group size was chosen such that fraction of total pairs in each group is 10%, similarly to what was originally done. One should note that the correlation values are available at different time points between the experimental and simulated curves. Thus, to permit a quantification of the agreement cost, each point in the experimental curve was mapped to the point in the simulated curve for which time points are the closest. The cost then penalizes, for each pair of points, a difference in the correlations but also in the time. Formally, if for each point $i$ in the data curve, the point $j = closestDivToMOMP_i$ in the simulated curve such that $DivToMOMP_j^{simulation}$ is the closest to $DivToMOMP_j^{data}$, the cost is then computed as:

$$\begin{aligned} cost = &\sum \left( Corr_i^{data} - Corr_{closestDivToMOMP(i)}^{simulation} \right)^2 \\ &+ \sum \left( DivToMOMP_i^{data} - DivToMOMP_{closestDivToMOMP(i)}^{simulation} \right)^2 \end{aligned}$$

## Comparison between the different types of data

To permit comparison between the different types of data, in each case a threshold for the cost defining agreement/disagreement was manually set by visual comparison of experimental and simulated data. For visualization purposes, panels A,B,E of Figure 2.12 represent a linearly normalized cost such that the threshold value correspond to 0.5 (represented in yellow). Normalized costs above 1 are capped to 1 and represented in red. The threshold costs used were 0.01/0.01, 0.01/0.01, and 2/2.5 for MOMP time distributions, surviving fractions and sister cell correlations respectively (TRAIL+CHX condition/TRAIL alone condition). To assign an agreement cost for MOMP time distribution AND surviving fractions (Figure 2.12, panel C), the maximum of the two normalized costs is taken.

# Appendix 4: List of software tools used in this thesis

- Code from Numerical Recipes (c++ edition) for solving ordinary differential equations and continuous time markov chains (i.e., the Gillespie algorithm), and sometimes scalar optimization (used in Chapter 4)
- Matlab for data analysis, fitting and plotting
- Qt and Qt Creator for helping with the compilation and editing of c++ code
- NumPy was sometimes used for scalar optimization
- Mayavi (python library), POVRAY or matplotlib (python) for visualization of spatially organized cell populations

# References


Acar, M. et al., 2008. Stochastic switching as a survival strategy in fluctuating environments. *Nature Genetics*, 40(4), pp.471–475.

Aebersold, R., 2005. Molecular Systems Biology: a new journal for a new biology? *Molecular Systems Biology*, 1, p.2005.0005.

Albeck, J.G. et al., 2008. Quantitative analysis of pathways controlling extrinsic apoptosis in single cells. *Molecular Cell*, 30(1), pp.11–25.

Albeck, J.G. et al., 2008. Modeling a snap-action, variable-delay switch controlling extrinsic cell death. *PLoS Biology*, 6(12), pp.2831–2852.

Aldridge, B.B. et al., 2011. Lyapunov exponents and phase diagrams reveal multifactorial control over TRAIL-induced apoptosis. *Molecular Systems Biology*, 7, p.553.

Alfonsi, A. et al., 2005. Adaptive simulation of hybrid stochastic and deterministic models for biochemical systems. *ESAIM: Proceedings*, 14, pp.1–13.

Anderson, A.R.A. & Quaranta, V., 2008. Integrative mathematical oncology. *Nature reviews. Cancer*, 8(3), pp.227–234.

Ashkenazi, A. & Dixit, V.M., 1998. Death receptors: signaling and modulation. *Science (New York, N.Y.)*, 281(5381), pp.1305–1308.

Ashkenazi, A. et al., 1999. Safety and antitumor activity of recombinant soluble Apo2 ligand. *The Journal of clinical investigation*, 104(2), pp.155–162.

Baker, J.H.E. et al., 2008. Direct visualization of heterogeneous extravascular distribution of trastuzumab in human epidermal growth factor receptor type 2 overexpressing xenografts. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 14(7), pp.2171–2179.

Balaban, N.Q. et al., 2004. Bacterial persistence as a phenotypic switch. *Science (New York, N.Y.)*, 305(5690), pp.1622–1625.

Balagaddé, F.K. et al., 2008. A synthetic Escherichia coli predator-prey ecosystem. *Molecular Systems Biology*, 4, p.187.

Balázsi, G. et al., 2011. Cellular decision making and biological noise: from microbes to mammals. *Cell*, 144(6), pp.910–925.

Barbone, D. et al., 2011. The Bcl-2 repertoire of mesothelioma spheroids underlies ac-

quired apoptotic multicellular resistance. *Cell death & disease*, 2, p.e174.

Batchelor, E., Loewer, A. & Lahav, G., 2009. The ups and downs of p53: understanding protein dynamics in single cells. *Nature reviews. Cancer*, 9(5), pp.371–377.

Bendall, S.C. et al., 2014. Single-Cell Trajectory Detection Uncovers Progression and Regulatory Coordination in Human B Cell Development. *Cell*, 157(3), pp.714–725.

Bendall, S.C. et al., 2011. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science (New York, N.Y.)*, 332(6030), pp.687–696.

Benner, S.A. & Sismour, A.M., 2005. Synthetic biology. *Nature reviews. Genetics*, 6(7), pp.533–543.

Bertaux, F. et al., 2014. Modeling Dynamics of Cell-to-Cell Variability in TRAIL-Induced Apoptosis Explains Fractional Killing and Predicts Reversible Resistance. *PLoS computational biology*, 10(10), p.e1003893.

Bhola, P.D. & Simon, S.M., 2009. Determinism and divergence of apoptosis susceptibility in mammalian cells. *Journal of cell science*, 122(Pt 23), pp.4296–4302.

Blainey, P.C. & Quake, S.R., 2014. Dissecting genomic diversity, one cell at a time. *Nature methods*, 11(1), pp.19–21.

Blake, W.J. et al., 2006. Phenotypic Consequences of Promoter-Mediated Transcriptional Noise. *Molecular Cell*, 24(6), pp.853–865.

Bodenmiller, B. et al., 2012. Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nature Biotechnology*, 30(9), pp.858–867.

Braeuer, S.J. et al., 2006. Constitutively activated nuclear factor-kappaB, but not induced NF-kappaB, leads to TRAIL resistance by up-regulation of X-linked inhibitor of apoptosis protein in human cancer cells. *Molecular cancer research : MCR*, 4(10), pp.715–728.

Breitschopf, K., Zeiher, A.M. & Dimmeler, S., 2000. Ubiquitin-mediated degradation of the proapoptotic active form of bid. A functional consequence on apoptosis induction. *The Journal of biological chemistry*, 275(28), pp.21648–21652.

Brown, C.R. et al., 2013. Linking stochastic fluctuations in chromatin structure and gene expression. *PLoS Biology*, 11(8), p.e1001621.

Brown, W. & Johnsen, R.M., 1981. Diffusion in polyacrylamide gels. *Polymer*, 22(2), pp.185–189.

Cao, Y., Gillespie, D.T. & Petzold, L.R., 2007. Adaptive explicit-implicit tau-leaping method with automatic tau selection. *The Journal of Chemical Physics*, 126(22), p.224101.

Chalfie, M. et al., 1994. Green fluorescent protein as a marker for gene expression. *Science (New York, N.Y.)*, 263(5148), pp.802–805.

Chaudhary, P.M. et al., 1997. Death receptor 5, a new member of the TNFR family, and

DR4 induce FADD-dependent apoptosis and activate the NF-kappaB pathway. *Immunity*, 7(6), pp.821–830.

Chen, M.-T., 2008. *Artificial Cell-Cell Communication and Multi-cellular Pattern Formation in Yeast Saccharomyces cerevisiae*. PhD thesis. MIT.

Chen, M.-T. & Weiss, R., 2005. Artificial cell-cell communication in yeast Saccharomyces cerevisiae using signaling elements from Arabidopsis thaliana. *Nature Biotechnology*, 23(12), pp.1551–1555.

Cheng, J. et al., 2006. Multiple mechanisms underlie resistance of leukemia cells to Apo2 Ligand/TRAIL. *Molecular Cancer Therapeutics*, 5(7), pp.1844–1853.

Chinnaiyan, A.M. et al., 2000. Combined effect of tumor necrosis factor-related apoptosis-inducing ligand and ionizing radiation in breast cancer therapy. *Proceedings of the National Academy of Sciences of the United States of America*, 97(4), pp.1754–1759.

Chisholm, R.H. et al., 2015. Emergence of drug tolerance in cancer cell populations: an evolutionary outcome of selection, nongenetic instability, and stress-induced adaptation. *Cancer research*, 75(6), pp.930–939.

Christian, P.A., Thorpe, J.A. & Schwarze, S.R., 2009. Velcade sensitizes prostate cancer cells to TRAIL induced apoptosis and suppresses tumor growth in vivo. *Cancer biology & therapy*, 8(1), pp.73–80.

Colman-Lerner, A. et al., 2005. Regulated cell-to-cell variation in a cell-fate decision system. *Nature*, 437(7059), pp.699–706.

Conway, J., 1970. The game of life. *Scientific American*.

Crivat, G. & Taraska, J.W., 2012. Imaging proteins inside cells with fluorescent tags. *Trends in biotechnology*, 30(1), pp.8–16.

Dada, J.O. & Mendes, P., 2011. Multi-scale modelling and simulation in systems biology. *Integrative biology : quantitative biosciences from nano to macro*, 3(2), pp.86–96.

Dar, R.D. et al., 2012. Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proceedings of the National Academy of Sciences of the United States of America*, 109(43), pp.17454–17459.

Deisboeck, T.S. et al., 2011. Multiscale cancer modeling. *Annual Review of Biomedical Engineering*, 13, pp.127–155.

Del Vecchio, D., 2015. Modularity, context-dependence, and insulation in engineered biological circuits. *Trends in biotechnology*, 33(2), pp.111–119.

Delbrück, M., 1945. The burst size distribution in the growth of bacterial viruses (bacteriophages). *Journal of bacteriology*, 50(2), pp.131–135.

DeRosier, L.C. et al., 2007. Combination treatment with TRA-8 anti death receptor 5 antibody and CPT-11 induces tumor regression in an orthotopic model of pancreatic cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 13(18

Pt 2), pp.5535s–5543s.

Deuflhard, P., 1985. Recent Progress in Extrapolation Methods for Ordinary Differential Equations. *SIAM Review*, 27(4), pp.505–535.

Deveraux, Q.L. et al., 1998. IAPs block apoptotic events induced by caspase-8 and cytochrome c by direct inhibition of distinct caspases. *The EMBO journal*, 17(8), pp.2215–2223.

Dickens, L.S. et al., 2012. A death effector domain chain DISC model reveals a crucial role for caspase-8 chain assembly in mediating apoptotic cell death. *Molecular Cell*, 47(2), pp.291–305.

Dimberg, L.Y. et al., 2013. On the TRAIL to successful cancer therapy? Predicting and counteracting resistance against TRAIL-based therapeutics. *Oncogene*, 32(11), pp.1341–1350.

Drasdo, D. & Höhme, S., 2005. A single-cell-based model of tumor growth in vitro: monolayers and spheroids. *Physical biology*, 2(3), pp.133–147.

Drasdo, D., Hoehme, S. & Block, M., 2007. On the Role of Physics in the Growth and Pattern Formation of Multi-Cellular Systems: What can we Learn from Individual-Cell Based Models? *Journal of Statistical Physics*, 128(1-2), pp.287–345.

Drasdo, D., Hoehme, S. & Hengstler, J.G., 2014. How predictive quantitative modelling of tissue organisation can inform liver disease pathogenesis. *Journal of Hepatology*, 61(4), pp.951–956.

Drasdo, D., Kree, R. & McCaskill, J., 1995. Monte Carlo approach to tissue-cell populations. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, 52(6), pp.6635–6657.

Du, C. et al., 2000. Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition. *Cell*, 102(1), pp.33–42.

Ehrhardt, H. et al., 2003. TRAIL induced survival and proliferation in cancer cells resistant towards TRAIL-induced apoptosis mediated by NF-kappaB. *Oncogene*, 22(25), pp.3842–3852.

Elowitz, M.B. et al., 2002. Stochastic gene expression in a single cell. *Science (New York, N.Y.)*, 297(5584), pp.1183–1186.

Fallahi-Sichani, M. et al., 2013. Metrics other than potency reveal systematic variation in responses to cancer drugs. *Nature Chemical Biology*, 9(11), pp.708–714.

Falschlehner, C. et al., 2007. TRAIL signalling: decisions between life and death. *The international journal of biochemistry & cell biology*, 39(7-8), pp.1462–1475.

Ferraro, E. et al., 2008. Apoptosome-deficient cells lose cytochrome c through proteasomal degradation but survive by autophagy-dependent glycolysis. *Molecular biology of the cell*, 19(8), pp.3576–3588.

Flusberg, D. & Sorger, P.K., 2013. Modulating cell-to-cell variability and sensitivity to

death ligands by co-drugging. *Physical biology*, 10(3), p.035002.

Flusberg, D. et al., 2013. Cells surviving fractional killing by TRAIL exhibit transient but sustainable resistance and inflammatory phenotypes. *Molecular biology of the cell*, 24(14), pp.2186–2200.

Fraser, H.B. et al., 2004. Noise Minimization in Eukaryotic Gene Expression. *PLoS Biology*, 2(6), pp.e137–5.

Fulda, S., 2011. Novel insights into the synergistic interaction of Bortezomib and TRAIL: tBid provides the link. *Oncotarget*, 2(5), pp.418–421.

Fulda, S. et al., 2002. Smac agonists sensitize for Apo2L/TRAIL- or anticancer drug-induced apoptosis and induce regression of malignant glioma in vivo. *Nature medicine*, 8(8), pp.808–815.

Galle, J., Hoffmann, M. & Aust, G., 2008. From single cells to tissue architecturea bottom-up approach to modelling the spatio-temporal organisation of complex multi-cellular systems. *Journal of Mathematical Biology*, 58(1-2), pp.261–283.

Galle, J., Loeffler, M. & Drasdo, D., 2005. Modeling the effect of deregulated proliferation and apoptosis on the growth dynamics of epithelial cell populations in vitro. *Biophysj*, 88(1), pp.62–75.

Gaudet, S. et al., 2012. Exploring the contextual sensitivity of factors that determine cell-to-cell variability in receptor-mediated apoptosis. *PLoS computational biology*, 8(4), p.e1002482.

Geva-Zatorsky, N. et al., 2006. Oscillations and variability in the p53 system. *Molecular Systems Biology*, 2, p.2006.0033.

Gonzalvez, F. et al., 2012. TRAF2 Sets a threshold for extrinsic apoptosis by tagging caspase-8 with a ubiquitin shutoff timer. *Molecular Cell*, 48(6), pp.888–899.

Gorochowski, T.E. et al., 2012. BSim: An Agent-Based Tool for Modeling Bacterial Populations in Systems and Synthetic Biology. *PLoS ONE*, 7(8), pp.e42790–9.

Graner, F. & Glazier, J.A., 1992. Simulation of biological cell sorting using a two-dimensional extended Potts model. *Physical review letters*, 69(13), p.2013.

Gupta, P.B. et al., 2011. Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell*, 146(4), pp.633–644.

Gygi, S.P. et al., 1999. Correlation between protein and mRNA abundance in yeast. *Molecular and cellular biology*, 19(3), pp.1720–1730.

Hairer, E., NΓErsett, S.P. & Wanner, G., 2009. Solving Ordinary Differential Equations I: Nonstiff Problems (Springer Series In Computational Mathematics) Author: Ernst.

Hart, Y. et al., 2014. Paradoxical Signalingby a Secreted Molecule Leads to Homeostasis

of Cell Levels. *Cell*, 158(5), pp.1022–1032.

Haseltine, E.L. & Rawlings, J.B., 2005. On the origins of approximations for stochastic chemical kinetics. *The Journal of Chemical Physics*, 123(16), p.164115.

Hellwig, C.T. et al., 2008. Real Time Analysis of Tumor Necrosis Factor-related Apoptosis-inducing Ligand/Cycloheximide-induced Caspase Activities during Apoptosis Initiation. *Journal of Biological Chemistry*, 283(31), pp.21676–21685.

Hirata, E. et al., 2015. Intravital Imaging Reveals How BRAF Inhibition Generates Drug-Tolerant Microenvironments with High Integrin β1/FAK Signaling. *Cancer Cell*, 27(4), pp.574–588.

Hirschhaeuser, F. et al., 2010. Multicellular tumor spheroids: an underestimated tool is catching up again. *Journal of biotechnology*, 148(1), pp.3–15.

Hlavacek, W.S. et al., 2006. Rules for modeling signal-transduction systems. *Science's STKE : signal transduction knowledge environment*, 2006(344), pp.re6–re6.

Ho, K.L. & Harrington, H.A., 2010. Bistability in apoptosis by receptor clustering. *PLoS computational biology*, 6(10), p.e1000956.

Hoehme, S., 2010. *Agent-based modeling of growing cell populations and the regenerating liver based on image processing*. PhD thesis.

Huet, H.A. et al., 2014. Multivalent nanobodies targeting death receptor 5 elicit superior tumor cell killing through efficient caspase induction. *mAbs*, 6(6), pp.1560–1570.

Iyer, V. & Struhl, K., 1996. Absolute mRNA levels and transcriptional initiation rates in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, 93(11), pp.5208–5212.

Jagiella, N., 2012. *Parameterization of Lattice-Based Tumor Models from Data.* PhD thesis. Inria: UPMC.

Jang, S.S. et al., 2012. Specification and Simulation of Synthetic Multicelled Behaviors. *ACS synthetic biology*, 1(8), pp.365–374.

Johnstone, R.W., Frew, A.J. & Smyth, M.J., 2008. The TRAIL apoptotic pathway in cancer onset, progression and therapy. *Nature reviews. Cancer*, 8(10), pp.782–798.

Kallenberger, S.M. et al., 2014. Intra- and Interdimeric Caspase-8 Self-Cleavage Controls Strength and Timing of CD95-Induced Apoptosis. *Science Signaling*, 7(316), pp.ra23–ra23.

Karr, J.R. et al., 2012. A whole-cell computational model predicts phenotype from genotype. *Cell*, 150(2), pp.389–401.

Kellogg, R.A. & Tay, S., 2015. Noise Facilitates Transcriptional Control under Dynamic Inputs. *Cell*, 160(3), pp.381–392.

Khmelinskii, A. et al., 2012. Tandem fluorescent protein timers for in vivo analysis of

protein dynamics. *Nature Biotechnology*, 30(7), pp.708–714.

Kim, H. et al., 2006. Hierarchical regulation of mitochondrion-dependent apoptosis by BCL-2 subfamilies. *Nature cell biology*, 8(12), pp.1348–1358.

Kim, J. et al., 2012. A programmable microfluidic cell array for combinatorial drug screening. *Lab on a chip*, 12(10), pp.1813–1822.

Kiviet, D.J. et al., 2014. Stochasticity of metabolism and growth at the single-cell level. *Nature*, 514(7522), pp.376–379.

Klumpp, S. & Hwa, T., 2014. Bacterial growth: global effects on gene expression, growth feedback and proteome partition. *Current opinion in biotechnology*, 28, pp.96–102.

Klumpp, S., Zhang, Z. & Hwa, T., 2009. Growth Rate-Dependent Global Effects on Gene Expression in Bacteria. *Cell*, 139(7), pp.1366–1375.

Lalaoui, N. et al., 2011. TRAIL-R4 promotes tumor growth and resistance to apoptosis in cervical carcinoma HeLa cells through AKT. *PLoS ONE*, 6(5), p.e19679.

Landau, L.D. et al., 2012. *Theory of Elasticity*, Elsevier.

Lane, D. et al., 2006. Acquired resistance to TRAIL-induced apoptosis in human ovarian cancer cells is conferred by increased turnover of mature caspase-3. *Molecular Cancer Therapeutics*, 5(3), pp.509–521.

Lapidus, S., Han, B. & Wang, J., 2008. Intrinsic noise, dissipation cost, and robustness of cellular networks: the underlying energy landscape of MAPK signal transduction. *Proceedings of the National Academy of Sciences of the United States of America*, 105(16), pp.6039–6044.

Lasken, R.S. & McLean, J.S., 2014. Recent advances in genomic DNA sequencing of microbial species from single cells. *Nature Publishing Group*, 15(9), pp.577–584.

Laussmann, M.A. et al., 2012. Proteasome Inhibition Can Impair Caspase-8 Activation upon Submaximal Stimulation of Apoptotic Tumor Necrosis Factor-related Apoptosis Inducing Ligand (TRAIL) Signaling. *Journal of Biological Chemistry*, 287(18), pp.14402–14411.

Le Novère, N., 2015. Quantitative and logic modelling of molecular and gene networks. *Nature reviews. Genetics*, 16(3), pp.146–158.

Lee, S. et al., 2010. Optimized combination therapy using bortezomib, TRAIL and TLR agonists in established breast tumors. *Cancer immunology, immunotherapy : CII*, 59(7), pp.1073–1081.

Leeuwen, I.M.M. van et al., 2009. An integrative computational model for intestinal tissue renewal. *Cell proliferation*, 42(5), pp.617–636.

Lemke, J. et al., 2014. Selective CDK9 inhibition overcomes TRAIL resistance by concomitant suppression of cFlip and Mcl-1. *Cell death and differentiation*, 21(3), pp.491–502.

Lemke, J. et al., 2014. Getting TRAIL back on track for cancer therapy. *Cell death and*

*differentiation*, 21(9), pp.1350–1364.

Leverkus, M. et al., 2003. Proteasome Inhibition Results in TRAIL Sensitization of Primary Keratinocytes by Removing the Resistance-Mediating Block of Effector Caspase Maturation. *Molecular and cellular biology*, 23(3), pp.777–790.

Li, B. & Dou, Q.P., 2000. Bax degradation by the ubiquitin/proteasome-dependent pathway: involvement in tumor survival and progression. *Proceedings of the National Academy of Sciences of the United States of America*, 97(8), pp.3850–3855.

Lipniacki, T. et al., 2006. Stochastic regulation in early immune response. *Biophysical journal*, 90(3), pp.725–742.

Lipniacki, T. et al., 2007. Single TNFalpha trimers mediating NF-kappaB activation: Stochastic robustness of NF-kappaB signaling. *BMC Bioinformatics*, 8(1), p.376.

Liu, C. et al., 2011. Sequential establishment of stripe patterns in an expanding cell population. *Science (New York, N.Y.)*, 334(6053), pp.238–241.

Lopez, C.F. et al., 2013. Programming biological models in Python using PySB. *Molecular Systems Biology*, 9, pp.1–19.

Loriaux, P.M. & Hoffmann, A., 2013. A Protein Turnover Signaling Motif Controls the Stimulus-Sensitivity of Stress Response Pathways. *PLoS computational biology*.

Lu, M. et al., 2014. E-Cadherin Couples Death Receptorsto the Cytoskeleton to Regulate Apoptosis. *Molecular Cell*, 54(6), pp.987–998.

Luo, X. et al., 1998. Bid, a Bcl2 interacting protein, mediates cytochrome c release from mitochondria in response to activation of cell surface death receptors. *Cell*, 94(4), pp.481–490.

Macklin, P. et al., 2012. Patient-calibrated agent-based modelling of ductal carcinoma in situ (DCIS): from microscopic measurements to macroscopic predictions of clinical progression. *Journal of Theoretical Biology*, 301, pp.122–140.

Marguerat, S. & Bähler, J., 2012. Coordinating genome expression with cell size. *Trends in genetics : TIG*, 28(11), pp.560–565.

McAdams, H.H. & Arkin, A.P., 1997. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 94(3), pp.814–819.

Mehling, M. & Tay, S., 2014. Microfluidic cell culture. *Current opinion in biotechnology*, 25, pp.95–102.

Menke, C. et al., 2011. Distinct TRAIL resistance mechanisms can be overcome by proteasome inhibition but not generally by synergizing agents. *Cancer research*, 71(5), pp.1883–1892.

Merchant, M.S. et al., 2004. Interferon gamma enhances the effectiveness of tumor necrosis factor-related apoptosis-inducing ligand receptor agonists in a xenograft model of Ew-

ing's sarcoma. *Cancer research*, 64(22), pp.8349–8356.

Miller, C. et al., 2011. Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. *Molecular Systems Biology*, 7(1), pp.458–458.

Miller, M. et al., 2012. Modular Design of Artificial Tissue Homeostasis: Robust Control through Synthetic Cellular Heterogeneity. *PLoS computational biology*, 8(7), p.e1002579.

Mirams, G.R. et al., 2013. Chaste: an open source C++ library for computational physiology and biology. *PLoS computational biology*.

Molina, N. et al., 2013. Stimulus-induced modulation of transcriptional bursting in a single mammalian gene. *Proceedings of the National Academy of Sciences of the United States of America*, 110(51), pp.20563–20568.

Mora, T. & Walczak, A.M., 2013. Effect of phenotypic selection on stochastic gene expression. *The journal of physical chemistry. B*, 117(42), pp.13194–13205.

Munsky, B., Trinh, B. & Khammash, M., 2009. Listening to the noise: random fluctuations reveal gene network parameters. *Molecular Systems Biology*, 5, p.318.

Naumann, I. et al., 2011. Bortezomib primes neuroblastoma cells for TRAIL-induced apoptosis by linking the death receptor to the mitochondrial pathway. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 17(10), pp.3204–3218.

Neal, M.L. et al., 2014. A reappraisal of how to build modular, reusable models of biological systems. *PLoS computational biology*, 10(10), p.e1003849.

Neuert, G. et al., 2013. Systematic identification of signal-activated stochastic gene regulation. *Science (New York, N.Y.)*, 339(6119), pp.584–587.

Neumann, L. et al., 2010. Dynamics within the CD95 death-inducing signaling complex decide life and death of cells. *Molecular Systems Biology*, 6, p.352.

Nevozhay, D. et al., 2012. Mapping the Environmental Fitness Landscape of a Synthetic Gene Circuit. *PLoS computational biology*, 8(4), pp.e1002480–17.

Nijhawan, D. et al., 2003. Elimination of Mcl-1 is required for the initiation of apoptosis following ultraviolet irradiation. *Genes & development*, 17(12), pp.1475–1486.

Oltvai, Z.N., Milliman, C.L. & Korsmeyer, S.J., 1993. Bcl-2 heterodimerizes in vivo with a conserved homolog, Bax, that accelerates programmed cell death. *Cell*, 74(4), pp.609–619.

Palani, S. & Sarkar, C.A., 2011. Synthetic conversion of a graded receptor signal into a tunable, reversible switch. *Molecular Systems Biology*, 7, pp.1–7.

Pampaloni, F., Reynaud, E.G. & Stelzer, E.H.K., 2007. The third dimension bridges the gap between cell culture and live tissue. *Nature reviews. Molecular cell biology*, 8(10), pp.839–845.

Paszek, P., 2007. Modeling stochasticity in gene regulation: characterization in the terms

of the underlying distribution function. *Bulletin of mathematical biology*, 69(5), pp.1567–1601.

Payne, S. et al., 2013. Temporal control of self-organized pattern formation without morphogen gradients in bacteria. *Molecular Systems Biology*, 9, pp.1–10.

Pitti, R. et al., 1996. Induction of apoptosis by Apo-2 ligand, a new member of the tumor necrosis factor cytokine family. *The Journal of biological chemistry*, 271(22), pp.12687–12690.

Poukkula, M. et al., 2005. Rapid turnover of c-FLIPshort is determined by its unique C-terminal tail. *The Journal of biological chemistry*, 280(29), pp.27345–27355.

Press, W.H., 2007. *Numerical Recipes 3rd Edition*, Cambridge University Press.

Primeau, A.J. et al., 2005. The distribution of the anticancer drug Doxorubicin in relation to blood vessels in solid tumors. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 11(24 Pt 1), pp.8782–8788.

Purnick, P.E.M. & Weiss, R., 2009. The second wave of synthetic biology: from modules to systems. *Nature reviews. Molecular cell biology*, 10(6), pp.410–422.

Raj, A. et al., 2006. Stochastic mRNA synthesis in mammalian cells. *PLoS Biology*, 4(10), p.e309.

Ramis-Conde, I. & Drasdo, D., 2012. From genotypes to phenotypes: classification of the tumour profiles for different variants of the cadherin adhesion pathway. *Physical biology*, 9(3), p.036008.

Ramis-Conde, I. et al., 2008. Modeling the influence of the E-cadherin-beta-catenin pathway in cancer cell invasion: a multiscale approach. *Biophysical journal*, 95(1), pp.155–165.

Rando, O.J. & Verstrepen, K.J., 2007. Timescales of genetic and epigenetic inheritance. *Cell*, 128(4), pp.655–668.

Raser, J.M. & O'Shea, E.K., 2005. Noise in gene expression: origins, consequences, and control. *Science (New York, N.Y.)*, 309(5743), pp.2010–2013.

Regot, S. et al., 2014. High-sensitivity measurements of multiple kinase activities in live single cells. *Cell*, 157(7), pp.1724–1734.

Rehm, M. et al., 2006. Systems analysis of effector caspase activation and its control by X-linked inhibitor of apoptosis protein. *The EMBO journal*, 25(18), pp.4338–4349.

Rehm, M. et al., 2009. Dynamics of outer mitochondrial membrane permeabilization during apoptosis. *Cell death and differentiation*, 16(4), pp.613–623.

Ricicova, M. et al., 2013. Dissecting genealogy and cell cycle as sources of cell-to-cell variability in MAPK signaling using high-throughput lineage tracking. *Proceedings of the National Academy of Sciences of the United States of America*, 110(28), pp.11403–11408.

Roux, J. et al., 2015. Fractional killing arises from cell-to-cell variability in overcoming a caspase activity threshold. *Molecular Systems Biology*, 11(5), p.803.

Salis, H. & Kaznessis, Y., 2005. Accurate hybrid stochastic simulation of a system of

coupled chemical or biochemical reactions. *The Journal of Chemical Physics*, 122(5), p.054103.

Sato, K. & Kaneko, K., 2006. On the distribution of state values of reproducing cells. *Physical biology*, 3(1), pp.74–82.

Schleich, K. et al., 2012. Stoichiometry of the CD95 death-inducing signaling complex: experimental and modeling evidence for a death effector domain chain model. *Molecular Cell*, 47(2), pp.306–319.

Schlüter, D.K., Ramis-Conde, I. & Chaplain, M.A.J., 2015. Multi-scale modelling of the dynamics of cell colonies: insights into cell-adhesion forces and cancer invasion from in silico simulations. *Journal of the Royal Society, Interface / the Royal Society*, 12(103).

Schnell, S., Grima, R. & Maini, P., 2007. Multiscale Modeling in Biology. *American Scientist*.

Schwanhäusser, B. et al., 2011. Global quantification of mammalian gene expression control. *Nature*, 473(7347), pp.337–342.

Selimkhanov, J., Hasty, J. & Tsimring, L.S., 2012. Recent advances in single-cell studies of gene regulation. *Current opinion in biotechnology*, 23(1), pp.34–40.

Senecal, A. et al., 2014. Transcription Factors Modulate c-Fos Transcriptional Bursts. *Cell reports*.

Shahrezaei, V. & Swain, P.S., 2008. Analytical distributions for stochastic gene expression. *Proceedings of the National Academy of Sciences*, 105(45), pp.17256–17261.

Shalek, A.K. et al., 2014. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature*, 510(7505), pp.363–369.

Shanker, A. et al., 2008. Treating Metastatic Solid Tumors With Bortezomib and a Tumor Necrosis Factor-Related Apoptosis-Inducing Ligand Receptor Agonist Antibody. *JNCI Journal of the National Cancer Institute*, 100(9), pp.649–662.

Sharma, S.V. et al., 2010. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell*, 141(1), pp.69–80.

Sharova, L.V. et al., 2009. Database for mRNA Half-Life of 19 977 Genes Obtained by DNA Microarray Analysis of Pluripotent and Differentiating Mouse Embryonic Stem Cells. *DNA Research*, 16(1), pp.45–58.

Sheridan, J.P. et al., 1997. Control of TRAIL-induced apoptosis by a family of signaling and decoy receptors. *Science (New York, N.Y.)*, 277(5327), pp.818–821.

Sherman, M.S. & Cohen, B.A., 2014. A computational framework for analyzing stochasticity in gene expression. *PLoS computational biology*, 10(5), p.e1003596.

Shibata, T. & Fujimoto, K., 2005. Noisy signal amplification in ultrasensitive signal transduction. *Proceedings of the National Academy of Sciences of the United States of America*, 102(2),

pp.331–336.

Shirley, S. & Micheau, O., 2013. Targeting c-FLIP in cancer. *Cancer letters*, 332(2), pp.141–150.

Shokhirev, M.N. et al., 2015. A multi-scale approach reveals that NF-$\kappa$B cRel enforces a B-cell decision to divide. *Molecular Systems Biology*, 11, p.783.

Sigal, A. et al., 2006. Variability and memory of protein levels in human cells. *Nature*, 444(7119), pp.643–646.

Singh, A. et al., 2012. Dynamics of protein noise can distinguish between alternate sources of gene-expression variability. *Molecular Systems Biology*, 8, p.607.

Snijder, B. et al., 2009. Population context determines cell-to-cell variability in endocytosis and virus infection. *Nature*, 461(7263), pp.520–523.

Son, J.K., Varadarajan, S. & Bratton, S.B., 2010. TRAIL-activated stress kinases suppress apoptosis through transcriptional upregulation of MCL-1. *Cell death and differentiation*, 17(8), pp.1288–1301.

Song, J.J. et al., 2006. Evidence for Two Modes of Development of Acquired Tumor Necrosis Factor-related Apoptosis-inducing Ligand Resistance: INVOLVEMENT OF Bcl-xL. *The Journal of biological chemistry*, 282(1), pp.319–328.

Spencer, S.L. et al., 2009. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature*, 459(7245), pp.428–432.

Spiller, D.G. et al., 2010. Measurement of single-cell dynamics. *Nature*, 465(7299), pp.736–745.

Starruß, J. et al., 2014. Morpheus: a user-friendly modeling environment for multiscale and multicellular systems biology. *Bioinformatics (Oxford, England)*, 30(9), pp.1331–1332.

Stoma, S. et al., 2013. STL-based analysis of TRAIL-induced apoptosis challenges the notion of type I/type II cell line classification. *PLoS computational biology*, 9(5), p.e1003056.

Sun, B.K. et al., 2011. TRAIL-induced caspase/p38 activation is responsible for the increased catalytic and invasive activities of Akt. *International journal of oncology*, 38(1), pp.249–256.

Sun, C. et al., 2014. Reversible and adaptive resistance to BRAF(V600E) inhibition in melanoma. *Nature*, 508(7494), pp.118–122.

Sun, J., 2014. *Towards Synthetic Ecology: Genetically Programmable 4-Module Population Control System in Yeast*. PhD thesis. MIT.

Suter, D.M. et al., 2011. Mammalian genes are transcribed with widely different bursting kinetics. *Science (New York, N.Y.)*, 332(6028), pp.472–474.

Suzuki, Y., Nakabayashi, Y. & Takahashi, R., 2001. Ubiquitin-protein ligase activity of X-linked inhibitor of apoptosis protein promotes proteasomal degradation of caspase-3

and enhances its anti-apoptotic effect in Fas-induced cell death. *Proceedings of the National Academy of Sciences of the United States of America*, 98(15), pp.8662–8667.

Sütterlin, T. et al., 2013. Bridging the scales: semantic integration of quantitative SBML in graphical multi-cellular models and simulations with EPISIM and COPASI. *Bioinformatics (Oxford, England)*, 29(2), pp.223–229.

Tabor, J.J. et al., 2009. A Synthetic Genetic Edge Detection Program. *Cell*, 137(7), pp.1272–1281.

Tang, F. et al., 2009. mRNA-Seq whole-transcriptome analysis of a single cell. *Nature methods*, 6(5), pp.377–382.

Tanouchi, Y. et al., 2012. Programming stress-induced altruistic death in engineered bacteria. *Molecular Systems Biology*, 8, pp.1–11.

Tawa, P. et al., 2004. Catalytic activity of caspase-3 is required for its degradation: stabilization of the active complex by synthetic inhibitors. *Cell death and differentiation*, 11(4), pp.439–447.

Tay, S. et al., 2010. Single-cell NF-kappaB dynamics reveal digital activation and analogue information processing. *Nature*, 466(7303), pp.267–271.

Thorne, B.C., Bailey, A.M. & Peirce, S.M., 2007. Combining experiments with multi-cell agent-based modeling to study biological tissue patterning. *Briefings in bioinformatics*, 8(4), pp.245–257.

Thorpe, J.A., Christian, P.A. & Schwarze, S.R., 2008. Proteasome inhibition blocks caspase-8 degradation and sensitizes prostate cancer cells to death receptor-mediated apoptosis. *The Prostate*, 68(2), pp.200–209.

Thurber, G.M., Schmidt, M.M. & Wittrup, K.D., 2008. Factors determining antibody distribution in tumors. *Trends in pharmacological sciences*, 29(2), pp.57–61.

Toni, T. & Tidor, B., 2013. Combined Model of Intrinsic and Extrinsic Variability for Computational Network Design with Application to Synthetic Biology. *PLoS computational biology*, 9(3), p.e1002960.

Turing, A.M., 1952. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London B*, (237), pp.37–72.

Viñuelas, J. et al., 2013. Quantifying the contribution of chromatin dynamics to stochastic gene expression reveals long, locus-dependent periods between transcriptional bursts. *BMC Biology*, 11(1), p.15.

Volfson, D. et al., 2006. Origins of extrinsic variability in eukaryotic gene expression. *Nature*, 439(7078), pp.861–864.

Vucic, D., Dixit, V.M. & Wertz, I.E., 2011. Ubiquitylation in apoptosis: a post-translational modification at the edge of life and death. *Nature reviews. Molecular cell biology*, 12(7), pp.439–

452.

Walczak, H. et al., 1999. Tumoricidal activity of tumor necrosis factor-related apoptosis-inducing ligand in vivo. *Nature medicine*, 5(2), pp.157–163.

Walker, D.C., Georgopoulos, N.T. & Southgate, J., 2008. From pathway to population - a multiscale model of juxtacrine EGFR-MAPK signalling. *BMC systems biology*, 2(1), p.102.

Wang, D. & Bodovitz, S., 2010. Single cell analysis: the new frontier in 'omics'. *Trends in biotechnology*, 28(6), pp.281–290.

Wang, X. et al., 2008. Akt-mediated eminent expression of c-FLIP and Mcl-1 confers acquired resistance to TRAIL-induced cytotoxicity to lung cancer cells. *Molecular Cancer Therapeutics*, 7(5), pp.1156–1163.

Weens, W., 2012. *Mathematical modeling of liver tumor*. PhD thesis. UPMC.

Weisse, A.Y. et al., 2015. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proceedings of the National Academy of Sciences*, 112(9), pp.E1038–47.

Wiley, S.R. et al., 1995. Identification and characterization of a new member of the TNF family that induces apoptosis. *Immunity*, 3(6), pp.673–682.

Wilkinson, D.J., 2009. Stochastic modelling for quantitative description of heterogeneous biological systems. *Nature reviews. Genetics*, 10(2), pp.122–133.

Wilt, L.H.A.M. de et al., 2013. Bortezomib and TRAIL: a perfect match for apoptotic elimination of tumour cells? *Critical reviews in oncology/hematology*, 85(3), pp.363–372.

Wolkenhauer, O. et al., 2014. Enabling multiscale modeling in systems medicine. *Genome medicine*, 6(3), p.21.

Yamada, K.M. & Cukierman, E., 2007. Modeling tissue morphogenesis and cancer in 3D. *Cell*, 130(4), pp.601–610.

Yoshida, T. et al., 2009. Repeated treatment with subtoxic doses of TRAIL induces resistance to apoptosis through its death receptors in MDA-MB-231 breast cancer cells. *Molecular cancer research : MCR*, 7(11), pp.1835–1844.

You, L. et al., 2004. Programmed population control by cell-cell communication and regulated killing. *Nature*, 428(6985), pp.868–871.

Youle, R.J. & Strasser, A., 2008. The BCL-2 protein family: opposing activities that mediate cell death. *Nature reviews. Molecular cell biology*, 9(1), pp.47–59.

Youn, Y.S. et al., 2007. Biological and physicochemical evaluation of the conformational stability of tumor necrosis factor-related apoptosis-inducing ligand (TRAIL). *Biotechnology Letters*, 29(5), pp.713–721.

Zechner, C. et al., 2012. Moment-based inference predicts bimodality in transient gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 109(21), pp.8340–8345.