



Deliverable No. 12.3

Report on the development of the Oncosimulator and the utilization of the biomechanism models

Grant Agreement No.: 270089
Deliverable No.: D12.3
Deliverable Name: Report on the development of the Oncosimulator and the utilization
of the biomechanism models
Contractual Submission Date: 31/01/2013
Actual Submission Date: 19/02/2013

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	



COVER AND CONTROL PAGE OF DOCUMENT	
Project Acronym:	p-medicine
Project Full Name:	From data sharing and integration via VPH models to personalized medicine
Deliverable No.:	D 12.3
Document name:	Report on the development of the Oncosimulator and the utilization of the biomechanism models
Nature (R, P, D, O) ¹	R
Dissemination Level (PU, PP, RE, CO) ²	PU
Version:	v1.0
Actual Submission Date:	19/02/2013
Editor:	Georgios S. Stamatakis
Institution:	ICCS (Institute of Communication and Computer Systems – National Technical University of Athens) – <i>In Silico</i> Oncology Group
E-Mail:	gestam@central.ntua.gr

ABSTRACT: This deliverable reports on the current status of the p-medicine Oncosimulator and the special biomechanism models development. The specific biomechanism models could enhance the function of particular submodules of the Oncosimulator such as the ones dealing with the probability of a tumour stem cell to undergo symmetric division and the probability of a tumour cell to undergo apoptosis following chemotherapeutic treatment. Chapter 1 provides an introduction to the overall theme of the document. Chapter 2 deals with aspects of the code parallelization and its Graphics Processing Unit (GPU) adaptation for the special case of the nephroblastoma branch of the Oncosimulator. Chapter 3 referring to the breast cancer branch of the Oncosimulator deals with a dynamic system approach describing vascular tumour growth and the anti-angiogenic treatment effect. It also outlines a discrete state – discrete event tumour growth model. An analysis of the coupling of the dynamic system with a discrete state – discrete event cytokinetic model for the case of combination (antiangiogenic and chemotherapeutic) treatment is presented. Chapter 4 deals with both the cellular and the supercellular as well as the subcellular (molecular) component of the acute lymphoblastic leukemia (ALL) branch of the Oncosimulator. A systems biology approach in conjunction with a discrete entity – discrete event simulation policy is adopted. Chapter 5 refers to special biochemical models at the atomistic level applied inter alia to the JAK2 non-receptor tyrosine kinase. Chapter 6 deals with the modelling of various special molecular biomechanisms including a spatiotemporal and population based model of p53 and its effects on proliferation and differentiation of tumour stem cells and models of the pharmacodynamics of nutlin and the p53 regulated stem cell polarity in breast cancer. Chapter 7 provides some overall conclusions.

¹ R=Report, P=Prototype, D=Demonstrator, O=Other

² PU=Public, PP=Restricted to other programme participants (including the Commission Services), RE=Restricted to a group specified by the consortium (including the Commission Services), CO=Confidential, only for members of the consortium (including the Commission Services)

KEYWORD LIST: in silico oncology, in silico medicine, Oncosimulator, multiscale cancer modelling, Virtual Physiological Human, VPH, cancer simulation, cancer systems biology, systems medicine, personalized medicine, predictive medicine, computational oncology, predictive oncology, clinical trials, Wilms tumour, nephroblastoma, angiogenesis, antiangiogenic treatment, acute lymphoblastic leukemia, ALL, bevacizumab, prednisone, breast cancer, molecular dynamics, JAK2, EGFR, stem cells, proliferative polarity, nutlin, p53

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 270089.

The author is solely responsible for its content, it does not represent the opinion of the European Community and the Community is not responsible for any use that might be made of data appearing therein.

MODIFICATION CONTROL*			
Version	Date	Status	Author
1.0	15/01/2013	Initial Draft	Georgios S. Stamatakos, ICCS-NTUA
2.0	22/01/2013	Revision	Georgios S. Stamatakos, ICCS-NTUA
3.0	28/01/2013	Revision	Dimitra D. Dionysiou, ICCS-NTUA
4.0 = v1.0 submitted	18/02/2013	Revision and Final Draft	Georgios S. Stamatakos, ICCS-NTUA

List of contributors

- G. Stamatakos, ICCS-NTUA
- D. Dionysiou, ICCS-NTUA
- Eleni Georgiadi, ICCS-NTUA
- Marek Blazewicz, PSNC
- Juliusz Pukacki, PSNC
- Katerina Argyri, ICCS-NTUA
- Eleni Kolokotroni, ICCS-NTUA
- Eleftherios Ouzounoglou, ICCS-NTUA
- Shunzhou Wan, UCL
- Alberto d' Onofrio, IEO
- Norbert Graf, USAAR
- Martin Stanulla, CAU

* See Appendix A at the end of this document for a list of affiliation acronyms

Contents

CONTENTS.....	5
EXECUTIVE SUMMARY.....	6
CHAPTER 1 [IN]: INTRODUCTION.....	7
CHAPTER 2 [N]: THE ONCOSIMULATOR: THE NEPHROBLASTOMA BRANCH.....	9
CHAPTER 3 [B]: THE ONCOSIMULATOR: THE BREAST CANCER BRANCH.....	21
CHAPTER 4 [L]: THE ONCOSIMULATOR: THE ACUTE LYMPHOBLASTIC LEUKEMIA (ALL) BRANCH.....	67
CHAPTER 5 [C]: TUMOUR BIOMECHANISM FOCUSED MODELS: VARIOUS BIOCHEMICAL MODELS.....	107
CHAPTER 6 [M]: TUMOUR BIOMECHANISM FOCUSED MODELS: VARIOUS MOLECULAR MODELS.....	120
CHAPTER 7 [CO]: CONCLUSION.....	132
Appendix A - Abbreviations and Acronyms.....	133

Executive Summary

This deliverable outlines the current developmental status of both the p-medicine Oncosimulator and the special biomechanism models. The latter could enhance the function of particular submodules of the Oncosimulator such as the ones dealing with the probability of a tumour stem cell to undergo symmetric division and the probability of a tumour cell to undergo apoptosis following chemotherapeutic treatment. Chapter 1 provides an introduction to the overall theme of the document. Chapter 2 deals with aspects of the code parallelization and its Graphics Processing Unit (GPU) adaptation for the special case of the nephroblastoma branch of the Oncosimulator. It describes the Cactus framework used with its CaKernel. This combination allows performing distributed computations in heterogeneous environment. Chapter 2 also provides information on the installation of Cactus with CaKernel, the software requirements and several compilation and execution aspects. Chapter 3, referring to the breast cancer branch of the Oncosimulator, deals with a dynamic system approach to the modelling of vascular tumour growth and the anti-angiogenic treatment effect. It describes a continuum approach to vascular tumour growth and the effect of bevacizumab targeted treatment. It outlines the implementation and the numerical solution of the formulated mathematical problem and provides indicative results. A parametric analysis is presented whereas several questions of clinical interest are formulated and addressed. It also deals with a discrete state –discrete event tumour growth model. The basic features of the model are presented along with a simulation flowchart. The input parameters are described. A vinorelbin pharmacokinetic model is outlined. An analysis of the coupling of the dynamic system with a discrete state – discrete event cytokinetic model for the case of combination (antiangiogenic and chemotherapeutic) treatment is presented. Chapter 4 deals with both the cellular and supercellular as well as the subcellular (molecular) component of the acute lymphoblastic leukemia (ALL) branch of the Oncosimulator. For the first component a simulation model is described with particular emphasis on the cytokinetic model proposed. An outline of the simulation process is provided along with a delineation of the code implementation. Indicative results are presented and discussed. An local appendix provides the code structure. For the second subcellular component the following major aspects are addressed through the extensive use of systems biology: the precursor-B ALL cell cycle, modelling the mammalian cell cycle sub-cellular dynamics and the pre-B ALL cell cycle subcellular dynamics. The assumptions and modifications that have led to the proposed model are delineated. A model calibration along with simulation results are also presented. Chapter 5 refers to special biochemical models at the atomistic level applied inter alia to the JAK2 non-receptor tyrosine kinase. The biochemical model proposed is outlined and a workflow is provided. Special emphasis is put on the system preparation, the execution of the simulations and the post-process analysis. Further applications which are in progress are briefly outlined. Chapter 6 deals with the modelling of various special molecular biomechanisms. These include a spatiotemporal and population based model of p53 and its effects on proliferation and differentiation of tumour stem cells, a model of the pharmacodynamics of nutlin, a model of the proliferation of stem cells based cellular populations and the effect of nutlin on their dynamics and the wave-pinning and the phase models of the p53 regulated stem cell polarity in breast cancer. Chapter 7 provides some overall conclusions.

Chapter 1 [IN]

INTRODUCTION

This document presents the major advances in the development of both the p-medicine Oncosimulator and the special biomechanism models that have taken place since the submission of Deliverable D12.2. It is noted that the special biomechanism models might refine the function of particular submodules of the Oncosimulator. In the following a brief outline of the following chapters of the document is provided.

Chapter 2 deals with aspects of the code parallelization and its Graphics Processing Unit (GPU) adaptation for the special case of the nephroblastoma branch of the Oncosimulator. It describes the Cactus framework used with its CaKernel. This combination allows performing distributed computations in an heterogeneous environment. Chapter 2 also provides information on the installation of Cactus with CaKernel, the software requirements as well as information on various compilation and execution issues.

Chapter 3, referring to the breast cancer branch of the Oncosimulator, deals with a dynamic system approach to the modelling of vascular tumour growth and the anti-angiogenic treatment effect. It describes a continuum approach to vascular tumour growth and the effect of bevacizumab targeted treatment. It outlines the implementation and the numerical solution of the formulated mathematical problem and provides indicative results. A parametric analysis is presented whereas several questions of clinical interest are formulated and addressed. It also deals with a discrete state –discrete event tumour growth model. The basic features of the model are presented along with a simulation flowchart. The input parameters are described. A vinorelbin pharmacokinetic model is outlined. An analysis of the coupling of the dynamic system with a discrete state – discrete event cytokinetic model for the case of combination (antiangiogenic and chemotherapeutic) treatment is presented.

Chapter 4 deals with both the cellular and supercellular as well as the subcellular (molecular) component of the acute lymphoblastic leukemia (ALL) branch of the Oncosimulator. For the first component a simulation model is described with particular emphasis on the cytokinetic model proposed. An outline of the simulation process is provided along with a delineation of the code implementation. Indicative results are presented and discussed. An local appendix provides the code structure. For the second subcellular component the following major aspects are addressed through the extensive use of systems biology: the precursor-B ALL cell cycle, modelling the mammalian cell cycle sub-cellular dynamics and the pre-B ALL cell cycle subcellular dynamics. The assumptions and modifications that have led to the proposed model are delineated. A model calibration along with simulation results are also presented.

Chapter 5 refers to special biochemical models at the atomistic level applied inter alia to the JAK2 non-receptor tyrosine kinase. The biochemical model proposed is outlined and a workflow is provided. Special emphasis is put on the system preparation, the execution of the simulations and the post-process analysis. Further applications which are in progress are briefly outlined.

Chapter 6 deals with the modelling of various special molecular biomechanisms. These include a spatiotemporal and population based model of p53 and its effects on proliferation and differentiation of tumour stem cells, a model of the pharmacodynamics of nutlin, a model of the proliferation of stem

cells based cellular populations and the effect of nutlin on their dynamics and the wave-pinning and the phase models of the p53 regulated stem cell polarity in breast cancer.

Chapter 7 provides some overall conclusions.

Appendix A provides the meaning of several abbreviations used throughout the deliverable.

Chapter 2 [N]

THE ONCOSIMULATOR: THE NEPHROBLASTOMA BRANCH

Parallelization of the Oncosimulator application code for the case of Wilms (nephroblastoma) tumor and GPU architecture adaptation

N1 Introduction

Since the Oncosimulator simulation model developed by the In Silico Oncology Group, ICCS-NTUA (<http://in-silico-oncology.iccs.ntua.gr/>) for the case of nephroblastoma has been outlined in the p-medicine Deliverables 12.1 and 12.2, this chapter focuses on certain computational aspects [N1-N12]. The simulations for the case of nephroblastoma are performed on a regular rectangular grid. The elements of the grid are called Geometrical Cells or Grid-Cells (GCs) and represent many biological cells. The number of represented biological cells is inversely proportional to the required computational and memory resources. It is worth considering the fact, that each GC is represented by multiple biological and statistical properties. Their number can be greater than 100. This fact makes the particular branch of the p-medicine Oncosimulator a memory-bound application. This requires that the application programmer should put greater effort to access the memory in an efficient manner rather than to perform computations efficiently.

Most of the computations performed over a single GC are independent from those performed on its neighbors. This renders the particular Oncosimulator branch ideal to parallelize. Moreover, this computational pattern along with the data layout and dependencies renders the simulation a subset of stencil computations. Many scientists have been researching and testing different approaches on how to efficiently compute stencils on a variety of different architectures[N7-N9].

The main challenge of stencil computations is to utilize the memory bandwidth efficiently. It can be achieved by fulfilling several computational schemes. The most important one is maximizing cache reuse. This is a difficult task, and dependent on the computing architecture. For instance different CPUs have different caches (L1, L2, L3) with different sizes and length of fetched lines, which requires different blocking techniques and loop traversing. On the other hand GPUs and other dedicated accelerators have specific cache memory types that need to be accessed in an explicit manner. In order to compute efficiently on all architectures, the user has to write a different version of the code for each architecture. This requires additional programming effort and detailed optimizations for each architecture separately.

A solution to the problem might be the use of a certain abstraction allowing for automatic optimizations specific to different architectures. The only requirement is usually to follow some rules about the methods of data access. A framework that is responsible for the handling of such an abstraction usually requires some additional meta-information about the computations that are actually being performed, for instance the size of the boundaries or hints about memory usage.

One such framework that provides mechanisms for this kind of abstraction is Cactus with a plugin handling heterogeneous architectures – CaKernel. Because of the good performance of this framework and its extensibility, it has been utilized in the p-medicine project. Moreover Cactus with CaKernel facilitates the programming and the execution of applications in a distributed environment providing a set of mechanisms for automatic, inter-node synchronization and exchange of neighbouring data.

Chapter 2 or [N] has been structured as follows. The next section describes the utilized frameworks and tools. Subsequently, the installation process is described and the execution process is depicted.

N2 Tools and frameworks

As stated previously, a Cactus framework has been used with its plugin CaKernel. This combination allows performing distributed computations in a heterogeneous environment. In this section both tools and their main features are described.

N2.1 Cactus

The Cactus computational framework [N1, N2, N3] is an open source, modular, highly portable, programming environment for collaborative research using high performance computing. Cactus is distributed with a generic parallel computational toolkit providing parallelization, domain decomposition, coordinates, boundary conditions, interpolators, reduction operators, and efficient I/O in different data formats. However, in the following sections one can find more information about the Cactus computational framework. A detailed documentation is available at: [N1].

N2.1.1 Cactus Architecture

A Cactus application consists of a central piece called the “flesh” and a collection of modules called “thorns” (see Fig. N1). The flesh provides a framework for defining and parsing parameters, scheduling work, interoperation between C, C++, F77, and F90 and interaction with other thorns. A “driver thorn” is required for any Cactus application. It is responsible for a number of important runtime tasks including memory management, synchronization of grid functions, and distribution of data. Thorns are described using a domain specific language (DSL) called the “Cactus Configuration Language” (CCL) [N4]. The information in the CCL files includes the name of the implementation, the declaration of functions and parameters, the schedule of the routines, whether they require synchronization after execution, etc. Details can be found below in Section N2.1.2. Cactus flesh is distributed with the Cactus computational toolkit (CCTK) which includes several arrangements of thorns (i.e. groups of thorns with related functionality) providing basic utilities. A new functionality can be added to Cactus by simply extending

existing thorns or creating new ones. The user may implement their own routines solving a given problem by the same means.

N2.1.2 Cactus Configuration Language

Cactus requires three files in CCL, `interface.ccl`, `schedule.ccl`, and `param.ccl`, from each contributing component to create global data structures, to set runtime parameters, and bind all the C and Fortran subroutines interacting through the schedule tree.

- `interface.ccl` defines the thorn interface and inheritance along with variables and aliased functions. Thorns typically do not define relationships with other specific thorns, nor do they communicate directly with other thorns. Instead they define relationships with an interface, which may be provided by one of multiple thorns. This distinction exists so that thorns providing the same interface may be independently swapped without affecting each other. Interfaces in Cactus are fairly similar to abstract classes in Java or virtual base classes in C++.
- `schedule.ccl`: defines a schedule tree to provide a static workflow which orchestrates a computation in Cactus. Besides controlling when and how scheduled functions provided by thorns should be invoked by the Cactus scheduler, a schedule tree also controls when storage for a so called grid function should be allocated and freed. Application developers are responsible for scheduling their routines at an appropriate place in the tree (e.g. INIT, PRESTEP, EVOL, POSTSTEP, etc.). Cactus ensures that the routines are called at the correct time.
- `param.ccl`: defines parameters which can be specified in a Cactus parameter file and initialized at the start of a Cactus run. Cactus provides tools to parse the parameter file and check the range.

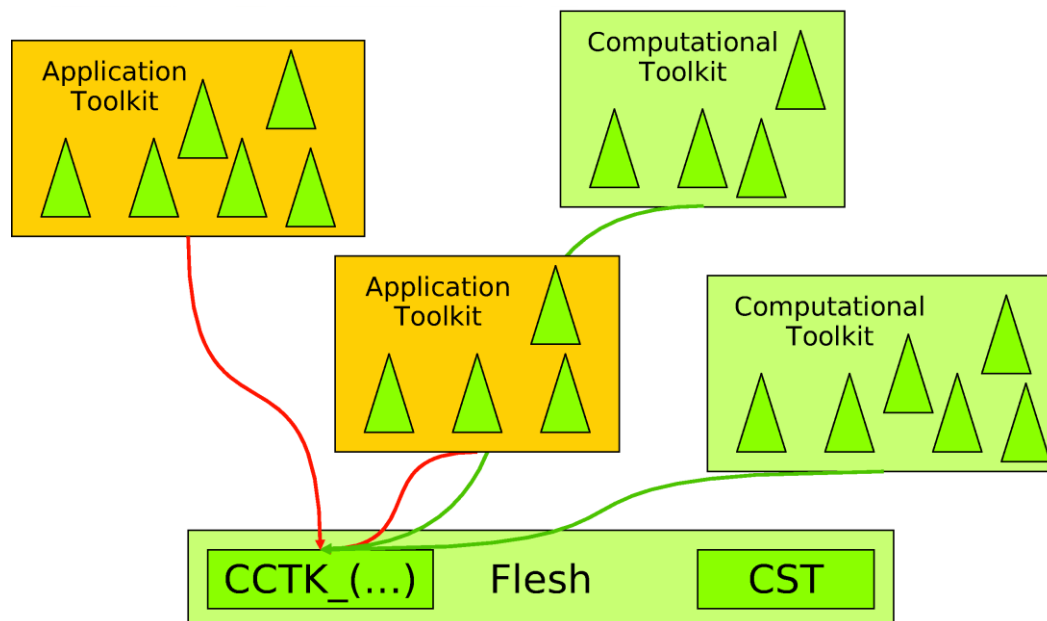


Figure N1: An overview of a typical Cactus application.

Apart from the three required files, there is one more optional file,

- `configuration.ccl`: defines build-time dependencies in terms of provided and required capabilities, e.g. interfaces to Cactus external libraries.

The CCL files will be parsed by the flesh during the compilation stage. After parsing the CCL files, the flesh enables component binding by generating source code to instantiate the different required thorn variables, parameters and functions, as well as checking required thorn dependencies. When Cactus starts up the flesh parses a user provided parameter file which defines which thorns are required and provides key/value pairs of parameter assignments. The flesh then activates only the required thorns, sets the given parameters using default values for parameters which are not specified in the parameter file, and creates the schedule of which functions provided by the activated thorns to run at which time. The Cactus flesh provides the main iteration loop for simulations (although this can be overloaded by any thorn). However, the flesh leaves memory allocation and parallelization to a driver thorn.

N2.2 CaKernel

CaKernel is an extension of CaCUDA kernel abstraction presented in Blazewicz et al. [N5] to implement a programming abstraction in Cactus for heterogeneous architectures with OpenCL. This abstraction enables automatic code generation from a set of highly optimized OpenCL and CUDA templates to simplify the development of scientific applications.

This kernel abstraction makes it easy to write and execute computational kernels, but also makes it possible to optimize the kernel without changing the kernel code itself. The whole optimization process is handled by swapping the templates or adjusting the kernel parameters. In the Kernel Abstraction, there are three major components: Kernel Descriptor, Computation Templates, and Code Generator.

N2.2.1 Kernel Descriptor

The CaKernel kernel descriptor is similar in both format and function to other Cactus Configuration Language (CCL) files. The CaKernel kernel descriptor (`cakernel.ccl`) is used to declare the variables and parameters that will be needed in the computations. The kernel descriptor is defined by:

- `CCTK_KERNEL <name>`: This keyword starts the kernel descriptor definition. It is followed by the kernel name, then by additional parameters of the form `<key>=<value>`. These parameters describe performance characteristics and grid point dependencies. In the current implementation these are:
 - `TYPE`: States on the required neighborhood of each computational grid point. `3DBLOCK` declares that grid points in all directions are required for proper computations. `3DSTENCIL` defines that only grid points in cardinal directions are required. The last value, `BOUNDARY`, is optimized for physical boundary update.
 - `STENCIL`: This parameters defines the number of additional grid points required for computations in each direction.
 - `TILE`: This parameter defines the dimensions of the kernel computational block. In future versions of the framework, it will rather be used as an optimization hint than setting the parameter itself.

- CCKT_KERNEL_VARIABLE: This keyword initiates the declaration of variables required in the current kernel definitions. The pointers to the variables will be automatically provided to the kernel and the data synchronized with the host. Additional parameters of the form <key>=<value> may also be provided.
 - CACHED: When set to YES the variable is automatically fetched and stored in the local kernel's memory, either in shared memory or registers, depending on the current optimization scheme.
 - INTENT: If the intent is IN, then the GPU variable will be updated and synchronized from the CPU before the kernel is executed. If the intent is OUT, the CPU variable will be updated from the GPU variable after the kernel is executed. If it is INOUT, both the effects of IN and OUT are combined. If the intent is SEPERATEINOUT, then two variables are used on the GPU to refer to a single variable on the CPU.
- CCKT_KERNEL_PARAMETER: This keyword initiates the declaration parameters that will be automatically passed to the kernel.

N2.2.2 Computation Templates

Computation templates allow CaKernel to automatically generate OpenCL or CUDA code based on the kernel descriptor and the user's numerical code. Each template is optimized for a given type of computational and communication strategy. A single computational type, i.e. 3DSTENCIL, 3DBLOCK or BOUNDARY, specified in the kernel descriptor corresponds to one or more templates. Templates together with kernel descriptors are passed to the code generator that produces header files containing

definitions of kernels in a form of C-language macros. An example template is shown in Figure N3.

Another advantage of introducing the templates is that the user does not need to be aware of a GPU-specific architecture. Scientists using

```
#define CaKernel_KERNEL_{name}
    _Computations_Begin_s
    for(tmpj = 0; tmpj < tilez_to; tmpj++) \
    {
        __syncthreads();

#define CaKernel_KERNEL_{name}
    _Iterate_Local_Tile_s
    %for_loop(tmpi,
                ) %[
                %var_loop(
                ) %[
                I3D_l(%vname, 0, 0, %var(tmpi)) = \
                I3D_l(%vname, 0, 0, %var(tmpi) + 1); \
                ]%
    ]%
    gk = gk2 + tmpj;

#define CaKernel_KERNEL_{name}
    _Fetch_Front_Tile_To_Cache_s
    %var_loop(
    ) %[
    I3D_l(%vname, 0, 0, stncl_zp) = \
    I3D(%vname, 0, 0, stncl_zp); ]%
    __syncthreads();

#define CaKernel_KERNEL_{name}
    _Limit_Threads_To_Compute_Begin_s \
    /* TODO Add your computations here */ \
    /* TODO Store results to global array */ \

#define CaKernel_KERNEL_{name}
    _Limit_Threads_To_Compute_End_s \
}
#define CaKernel_KERNEL_{name}
    _Computations_End_s \
}
```

Figure N3: A section of a CaKernel template that can be used to generate a local stencil for the variables defined in the kernel descriptor in the shared memory.

```
CCTK_KERNEL_TEMPLATE_GROUP gpu_cuda
{
    CCTK_KERNEL_TEMPLATE mem
    scope=group compile=no
    file="cacuda.vars-template.h" {
        triggers{
            mem$comp, comm$comp, shared
        }
    } "a template that defines the
        necessary variables"

    CCTK_KERNEL_TEMPLATE comm$comp
    scope=group compile=yes
    file="cacuda.comm-template.cu" {
        schedule {
            "schedule CaKernel_CopyToDev in
                CCTK_BASEGRID after
                CaKernel_InitFunctions
            {
                LANG: C
            }" "Allocate memory for variables on
                devices\"
        }
        include { mem, shared }
    } "a template that performs the
        necessary boundary data exchange"
    ...
}
```

Figure N2: A subset of syntax for defining the template hierarchy structure.

CaKernel can concentrate on their domains rather than on technical details of a particular hardware. Moreover, the same computations may be easily launched on other architectures without rewriting the code. CaKernel allows also to specify templates hierarchy. It facilitates inter-template dependencies definition. For example each computational template requires device initialization, memory handling, etc. The templates responsible for these operations are "triggered" to be evaluated if one of the computational schemes is to be used. The templates hierarchy is located in the additional configuration

file `deps.ccl` of which the syntax is similar to the CCL language. The file is read and interpreted during the compilation by the code generator. From the user's point of view the template hierarchy enables compilation of the code for different architectures or different optimization schemes by changing only the name of the computational type in `cakernel.ccl` file. The code generator supports also a small fraction of regular expression syntax to enable multiple architecture and optimization schemes specification within one kernel descriptor definition. The example fragment of the template hierarchy definition syntax is presented in Figure N2.

N2.2.3 Code Generator

```
g.compile("w", "([ \\t\\r\\n]|#.*)*");
g.compile("w1", "([ \\t\\r\\n]|#.*)+");
g.compile("KERNEL", "CCTK_CUDA_KERNEL{-w1
  }{name}{-w1}({key}{-w}={-w}{value}{-w
  })*\\{\\{-w\\}({VAR}{-w}|{PAR}{-w})*\\{\\{-
  w\\}\\}");
g.compile("KERNELS", "~{-w}({KERNEL})*$");
;
g.compile("name", "[A-Za-z0-9_]+");
g.compile("key", "{name}");
g.compile("value", "{name}|{dquote}|{
  squote}");
g.compile("dquote",
  "\"(\\\\\\\\[~]|~\\\\\\\\[~])*\"");
g.compile("squote", "'(\\\\\\\\[~]|~\\\\\\\\[~])
  *'");
g.compile("VAR", "CCTK_KERNEL_VARIABLE({-
  w1}({key}{-w}={-w}{value}{-w})*|)
  \\{\\{-w\\}{name}{-w},{-w}{name}\\}*{-w
  }\\{\\{-w\\}{dquote}\\}");
g.compile("PAR", "CCTK_KERNEL_PARAMETER
  ({-w1}({key}{-w}={-w}{value}{-w})*|)
  \\{\\{-w\\}{name}{-w},{-w}{name}\\}*{-w
  }\\{\\{-w\\}{dquote}\\}");
g.compile("digit", "[0-9]+");
g.compile("any", "[~]*");
g.compile("par", "{key}{-w}={-w}{any}");
```

Figure N5: Piraha PEG grammar for parsing the kernel abstraction

The CaKernel framework uses Piraha [N6] as a parser for the kernel descriptor and as a code generator, automatically creating OpenCL and CUDA-based macros from cakernel.ccl and templates. The Piraha Parsing Expression Grammar

```
#include <CaKernel_Update_Vel_3DBlock.h>
CAKERNEL_Update_Vel_Begin
/* users temporary variables */
CCTK_REAL tmpf = 0, v_sum, v, va, vb;
CAKERNEL_Update_Vel_Computations_Begin
/* more user's code above */
v = (I3D_l(vx,0,0,0) + I3D_l(vx,-1,0,0)
  + I3D_l(vx,0,1,0) + I3D_l(vx
  ,-1,1,0)) / 4.0;
va = I3D_l(vy,1,0,0) - I3D_l(vy,0,0,0);
vb = I3D_l(vy,0,0,0) - I3D_l(vy,-1,0,0)
  ;
tmpf = params.cagh_dx * 2;
v_sum -= v / tmpf * (va + vb + COPYSIGN
  (0.2, v) * (vb - va));
/* more user's code below */
CAKERNEL_Update_Vel_Computations_End
CAKERNEL_Update_Vel_End
```

Figure N4: A section of a scientific code responsible for explicit time integration. The macros in the code are defined in an automatically generated header file.

(PEG) is designed to resemble the Java regular expression API and syntax, but at the same time it provides the full power of a grammar parser. Figure N5 presents the grammar defined in Piraha and used to parse the kernel definitions. Both OpenCL and CUDA implementations share the same parser and grammar, only the templates are separated. As a result, CaKernel is able to generate OpenCL as well as CUDA code with the same set of tools from the same kernel descriptor and user's code.

Additionally, the code generator gives an opportunity to optimize OpenCL and CUDA code during the evaluation of a template by performing loop unrolling. Although recommended on GPUs, the compilers often are not capable to perform this optimization. This enhancement allows to speed up applications.

The code generator makes it possible to solve a particular problem with the preferred algorithm for stencil-based computations. The user only needs to write code for a single element in the computational domain and the automatically generated macros loop around all the points that are necessary to carry out local computations. Figure N4 contains an example of the code written for single grid-point in CaKernel environment.

N3 Installation

Cactus with CaKernel is installed locally in the user's home directory. The installation process is identical to the one for Cactus, it only requires some additional modules responsible for CaKernel and the

Oncosimulator. The user might download and install Cactus either from repositories in order to obtain the newest version of the software, or utilize a specially created for this purpose package. The requirements for both methods of obtaining the software as well as its compilation are described in the following sections.

N3.1 Software requirements

For a proper compilation and execution of the nephroblastoma Oncosimulator within the Cactus/CaKernel environment certain software conditions have to be met.

- A unix-like environment is expected; this might be either Linux, Windows (with Cygwin) or OS X (with MacPorts);
- Compilers installed: C/C++, F90, NVIDIA CUDA (nvcc);
- Libraries: MPI;
- Language interpreters: Java, Perl;
- Applications: GNU Make, svn, git, mercurial;

However the application should work in selected environments, it has been tested only on Linux. In specific on: Fedora, Ubuntu and Debian.

N3.2 Getting the software from a repository

This method allows to download the latest version of the software. It also allows to update the once downloaded software. The Cactus installation with the use of this method is described in [N10].

After having downloaded the required GetComponents script allowing to download the software from the repositories, the user has to also specify the file with the modules (thorns) to be fetched. The list of thorns and repositories being part of the Oncosimulator implemented in Cactus/CaKernel are available at [N11]. However some repositories are private, and permission to access them needs to be granted. After having collected all required credentials one should execute the command:

```
./GetComponents <ThornList name>
```

This will fetch all necessary software sources into the Cactus directory.

N3.3 Getting the software from the package

Obtaining the software from the compressed package is less complicated. One just needs to fetch the file located under link [N12] and decompress it using their favorite archiving software. For instance under Linux, one should execute the command:

```
7z x <archive name>
```

As a result one will obtain a directory analogous to the one in the previous section with the only difference that the software might be slightly outdated.

N3.4 Compilation

In general each Cactus compilation is called configuration. It is a set of thorns to be compiled into a single executable along with all flags and variables required for a proper compilation. The data required for each compilation is stored separately in different directories: `configs/<configuration name>`. In order to compile one has to follow the following set of instructions:

- go to Cactus directory;
- hit `make <configuration name>` in the cactus directory for the first time;
- the framework will ask whether it should create a configuration named `<configuration name>`; type yes; the flags and the required variables will be fetched from `~/cactus/config` file which should have the same syntax as the standard *Makefile* file; if the file is empty, or does not exist, defaults would be taken;
- now one should choose a list of thorns to be compiled; this can be done either by selecting the thorns manually (hit `make <configuration name>` one more time) or by selecting the same `ThornList`, that could be used to fetch the thorns from the repository; the `ThornList` is located at `arrangements/Oncosimulator/OncoWilmsTumour/par/ThornList`; the `ThornList` should be copied into `configs/<configuration name>/` directory;
- finally one should hit `make <configuration name>` to initiate the actual compilation;
- as a result, a file `exe/cactus_<configuration name>` is created.

More information on the discussed topic is available in the Cactus User Guide [N2], where the compilation is very well documented.

N4 Running

Each executed Cactus application requires a parameter file. Each parameter file consists of lines:

`<thorn>::<name> = <value>`

where `<thorn>`, `<name>` and `<value>` stands for the name of the thorn the parameter was declared in, the name of the parameter and its value respectively. The only exception is the parameter named `ActiveThorns`, which defines which thorns should be activated during the execution. This includes the allocation of variables, the setting up of parameters and the execution of scheduled functions. The example parameter file for the Oncosimulator simulation is located at `arrangements/Oncosimulator/OncoWilmsTumour/par/wilm.par`. In order to run the Oncosimulator in the Cactus environment, one should execute the binary file with a path to the parameter file as an argument of the execution:

`./cactus_<config name> <path to parameter file>`

The most important parameters are the following.

Name	Description
driver::global_n[xyz]	dimensions of the simulation grid in 3 directions;
Cactus::cctk_final_time	end simulation after n hours;
Carpet::output_timer_tree_every	output performance timings every n iterations;
OncoCellBucket::InputImageFileName	the initial tumour state; the file consists of unsigned chars and its dimensions needs to be equal to those specified by driver::global_n[xyz] parameters; otherwise an error occurs;
OncoWilmsTumour::Mode	specifies the simulation mode, which might be Therapy (with drug applied) or Growth, corresponding to free tumor growth;
OncoCellBucket::Reconstruction	Specifies whether reconstruction should be applied;
OncoCellBucket::Drug	Defines which drug should be applied; The possible values are: 0 (VCR – Vincristine), 1 (ACT – Dactinomycin), 2 (COMBI - Combination: VCR+ACT);

If a parameter is not specified in the parameter file, it is set to its default value. In order to check all possible parameters and their default values one should check *param.ccl* files of the Oncosimulator thorns, which are located in arrangements/Oncosimulator.

N5 Summary

An accelerated version of the Oncosimulator code concerning nephroblastoma response to chemotherapy has been implemented in Cactus with CaKernel plugin. This allows executing the particular code in a heterogeneous environment.

N6 References

- [N1] Cactus Computational Toolkit. <http://www.cactuscode.org>.
- [N2] G. Allen, T. Goodale, G. Lanfermann, T. Radke, D. Rideout, and J. Thornburg. <http://cactuscode.org/documentation/UsersGuide.pdf>.
- [N3] T. Goodale, G. Allen, G. Lanfermann, J. Masso, T. Radke, E. Seidel, and J. Shalf. The Cactus framework and toolkit: Design and applications. In High Performance Computing for Computational Science - VECPAR 2002, 5th International Conference, Porto, Portugal, June 26-28, 2002, pages 197–227, Berlin, 2003. Springer.
- [N4] G. Allen, T. Goodale, F. Loeffler, D. Rideout, E. Schnetter, and E.L. Seidel. Component Specification in the Cactus Framework: The Cactus Configuration Language. In Grid2010: Proceedings of the 11th IEEE/ACM International Conference on Grid Computing, 2010. (arXiv:1009.1341).
- [N5] M. Blazewicz, S.R. Brandt, P. Diener, D.M. Koppelman, K. Kurowski, F. Loeffler, E. Schnetter, and J. Tao. A massive data parallel computational framework on petascale/exascale hybrid computer systems. In International Conference on Parallel Computing, Ghent, Belgium, 2011.
- [N6] S.R. Brandt and G. Allen. Piraha: A simplified grammar parser for component languages. In 2010 11th ACM/IEEE International Conference on Grid Computing, 2011.
- [N7] M. Christen, O. Schenk, and H. Burkhart, “PATUS : A Code Generation and Autotuning Framework For Parallel Iterative Stencil Computations on Modern Microarchitectures.”
- [N8] K. Datta, M. Murphy, V. Volkov, S. Williams, J. Carter, L. Oliker, D. Patterson, J. Shalf, and K. Yelick, “Stencil computation optimization and auto-tuning on state-of-the-art multicore architectures,” 2008 SC - International Conference for High Performance Computing, Networking, Storage and Analysis, no. November, pp. 1–12, Nov. 2008.
- [N9] S. M. F. Rahman, Q. Yi, and A. Qasem, “Understanding Stencil Code Performance On MultiCore Architectures.”
- [N10] Cactus installation process: <http://cactuscode.org/download/>
- [N11] Oncosimulator ThornList
<http://apps.man.poznan.pl/trac/onco-cactus/attachment/wiki/WikiStart/ThornList>
- [N12] Oncosimulator in Cactus, full package (password to the archive: p-medicine-onco):
<http://apps.man.poznan.pl/trac/onco-cactus/downloads/1>

Chapter 3 [B]

THE ONCOSIMULATOR: THE BREAST CANCER BRANCH

B1 The dynamical system approach describing vascular tumour growth and the anti-angiogenic treatment effect

B1.1 The continuum approach of vascular tumour growth

A model chosen to serve as a basis of our continuum approach is the one introduced by (Poleszczuk et al., 2011). This model is a member of a family of approaches based on the work of (Hahnfeldt et al., 1999).

The biological phenomena incorporated into the continuum simulation model are listed below:

- Proliferation of tumour cells
- Apoptosis of tumour cells
- Postvascular dormancy where angiogenesis stimulation and inhibition has come into balance
- Endothelial cell death
- Spontaneous loss of functional vasculature
- Excretion of endogenous pro-angiogenic factors (such as vascular endothelial growth factor, fibroblast growth factors, platelet-derived growth factor, angiopoietin-1 etc.)
- Excretion of endogenous anti-angiogenic factors (angiostatin, endostatin, angiopoietin-2 etc.)
- Anti-angiogenic treatment – induced endothelial cell death and resulting tumour cell death

The underlying assumptions made in the context of the basic framework of the model are stated below:

- the tumour is a three dimensional spheroid
- the diffusion process is in a quasi-stationary state i.e. the tumour growth rate as well as the rate of change of drug concentration are relatively small compared to the rate of distribution of angiogenesis stimulators
- the concentration of the stimulator is a radially symmetric function i.e.

$$\|x\| = \|y\| \Rightarrow n(x) = n(y)$$

where $n(\cdot)$ is the function that gives the concentration of the stimulators.

The dynamical system of Poleszczuk et al. consists of two differential equations that reflect the interplay between tumour volume and carrying capacity.

1st equation: It describes the tumour volume rate of change and, as Hahnfeldt et al., the authors also adopt the Gompertzian growth equation

$$\frac{dV}{dt} = -\lambda_1 \cdot V \cdot \ln\left(\frac{V}{K}\right) \quad (B1)$$

2nd equation: It describes the rate of change of the carrying capacity K

$$\frac{dK}{dt} = -\lambda_2 \cdot K + c \cdot \frac{(\beta + V^p)}{\alpha \cdot (\beta + V^p) + I(t)} - d \cdot K \cdot V^{2/3} \quad (B2)$$

The parameters and variables involved in the equations are explained in Table B1. Most parameters involved in the model are not pure numbers but parameters that stand for physical quantities and as such, they have units that are also listed below.

Table B1: Parameters and variables involved in Poleszczuk et al. model

	Mathematical Symbol	Code Symbol	Units	Description
Independent variable	t	t	days	time
Dependent Variable	V	$y(1)$	vol	tumour volume
Dependent Variable	K	$y(2)$	vol	tumour capacity
Parameter	λ_1	lamda1	1/time	proportionality constant
Parameter	λ_2	lamda2	1/time	proportionality constant
Parameter	c	c	$1/(\text{time} \cdot \text{vol}^p) \cdot \text{conc}$	proportionality constant
Parameter	d	d	$1/(\text{time} \cdot \text{vol}^{2/3})$	proportionality constant
Parameter	α	alpha	$\text{conc} / \text{vol}^p$	constant
Parameter	β	beta	vol^p	constant
Parameter	p	p	≥ 0	Hill coefficient
Parameter	I	I	conc	drug concentration in plasma

B1.2 Incorporating the Bevacizumab treatment effect: the pharmacokinetic model

This part of the deliverable presents the computation of drug concentration $I(t)$ at a given time-point t , a function that is involved in equation (B2).

In order to simulate bevacizumab mono-therapy in a reliable way, we have studied the equations governing the Bevacizumab pharmacokinetic properties according to the two-compartmental pharmacokinetic model proposed by (Bertrand and Mentre, 2008) by taking into account that bevacizumab is only administered via the intravenous route. The pharmacokinetics module has been incorporated into the model of Poleszczuk et al.

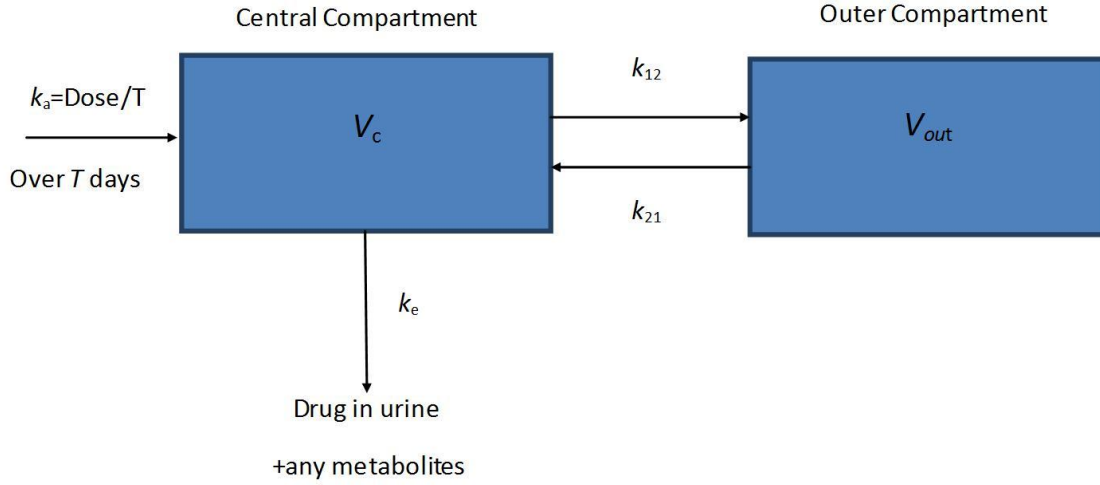


Figure B1: The two – compartmental pharmacokinetic model.

The central compartment represents blood and all highly perfused tissues i.e. vital organs that are in rapid equilibrium with blood such as lungs, kidney and liver. Drug elimination is supposed to occur from the central compartment according to the first order transfer rate k_e , since kidney and liver are the two most important clearing organs. The outer compartment corresponds to poorly perfused tissues.

Taking into account the fact that the concentrations of drugs cannot usually be measured at the effector site, the concentration of drug in plasma is typically assumed to represent the concentration in the target tissue. Hence, the equation characterizing the central compartment is the one to be introduced in the vascular tumour growth model.

If T stands for the duration of infusion, D for dose and t_D for the time-point of the administration, the function that describes bevacizumab concentration in plasma, following one infusion, is the following:

While the infusion takes place, namely while $t - t_D \leq T$:	
$I(t) = \frac{D}{T} \cdot \left[\frac{A}{a} (1 - e^{-a(t-t_D)}) + \frac{B}{b} (1 - e^{-b(t-t_D)}) \right]$	(B3)
After the infusion, namely while $T < t - t_D$:	
$I(t) = \frac{D}{T} \cdot \left[\frac{A}{a} (1 - e^{-aT}) \cdot e^{-a(t-t_D-T)} + \frac{B}{b} (1 - e^{-bT}) \cdot e^{-b(t-t_D-T)} \right]$	(B4)

where,

$a = \frac{k_{21} \cdot k_e}{b}$	(B5)
$b = 0.5 \cdot \left(k_{12} + k_{21} + k_e - \sqrt{(k_{12} + k_{21} + k_e)^2 - 4k_{21} \cdot k_e} \right)$	(B6)
$A = \frac{1}{V_p} \frac{a - k_{21}}{a - b}$	(B7)
$B = \frac{1}{V_p} \frac{b - k_{21}}{b - a}$	(B8)

Table B2: Parameters and variables involved in the two-compartmental pharmacokinetic model.

	Mathematical Symbols	Code Symbols	Units	Description
Independent variable	t	t	days	time
Dependent variable	I	I	conc (mg/ml)	concentration of administered inhibitor in blood
Parameter	-	n	-	number of infusions
Parameter	D	dose	mg	dose
Parameter	T	inf_dur	day	duration of infusion
Parameter	V_p	Vp	volume (ml)	volume of central compartment
Parameter	k_{12}	k12	1/day	inter-compartmental rate constant (from central compartment to outer compartment)
Parameter	k_{21}	k21	1/day	inter-compartmental rate constant (from outer compartment to central compartment)
Parameter	k_e	ke	1/day	elimination rate constant
Parameter	-	dosage	mg/kg	dosage
Parameter	-	w	kg	patient's weight
Parameter	T_D	T1,...,Tn	days	administration time-points of anti-angiogenic treatment

B1.3 Implementation, numerical solution and indicative results

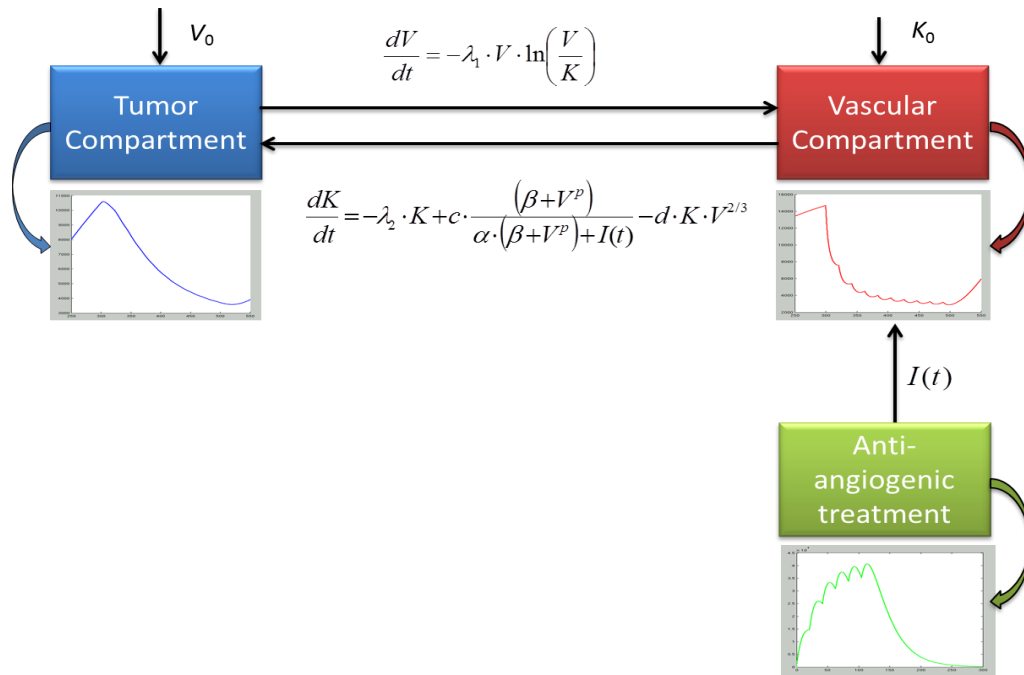


Figure B2: The structure of the model along with the two equations describing vascular tumour growth.

The m-files involved in the implementation of the model are listed and explained below.

DrugConcentration.m: This function implements the two-compartmental pharmacokinetic model for administration via intravenous infusion (as it is the case for the bevacizumab related clinical trials involved in p-medicine). It calculates the concentration of anti-angiogenic agent at each time point taking into account all previous infusions.
Input: time - point t
Input parameters: number of infusions n , duration of infusions (inf_dur), volume of distribution (V_c), transfer constants (k_{12}, k_{21}, k_e), dosage ($dosage$), patient's weight (w), administration time points (T_i 's)
Output: the concentration of anti-angiogenic agent at the given time point

VascularTumorGrowth.m: This function computes the derivatives involved in solving the Poleszczuk et al. equations which describe the vascular tumour growth.
Input: time point t , initial values (V_0, K_0)
Input parameters: $\lambda_1, \lambda_2, c, d, p, \alpha, \beta$
Output: -

VascularTumorGrowth_main.m: Given initial values (V_0, K_0) and a time interval, the specific script file resolves the problem with the solver ode45 which implements a Runge Kutta method with a variable time step for efficient computation. It also plots the variables V and K (representing tumour volume and tumour capacity respectively) as functions of time in common axial system as well as the phase – plane of the system.

Multiple code executions have been conducted testing the code with different parameter values, time span, initial values and code module (free growth, constant treatment, intermittent treatment). Below, several indicative results are presented, adopting the classical graph representation.

Code execution 1: First of all, the drug concentration for the case of intermittent bevacizumab monotherapy was plotted. The administration of 15 mg/kg in ten 90-minute intravenous infusions given every three weeks was simulated and steady – state is obtained at approximately the 100th day i.e. in ((bevacizumab terminal half – life) × 5) days as the relevant literature suggests (McCormack and Keam, 2008). The simulated treatment scheme which has been applied in the context of various clinical trials (Lu et al., 2008) is characterized by the parameter values shown in Table B3 (code execution 3).

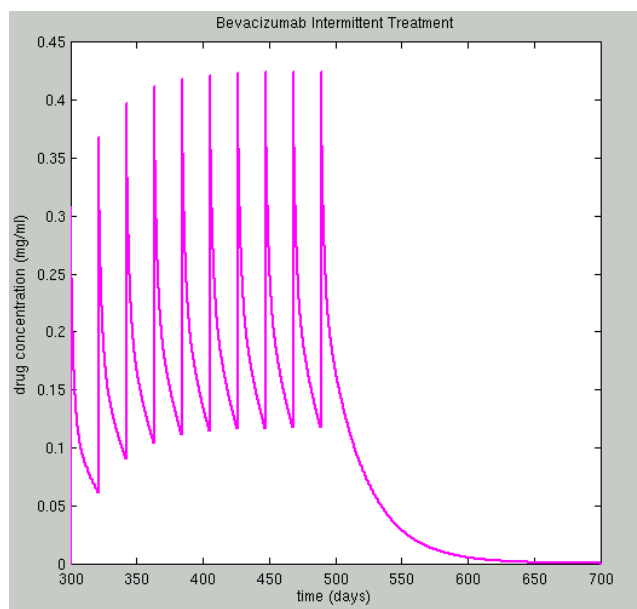


Figure B3: The time course of bevacizumab concentration in plasma. The details of the simulated scheme are shown in table B3 (code execution 3).

Code Execution 2: The time evolution of untreated tumour growth where λ_1 was set equal to 0.003465 1/day in order to reflect a real life value of human breast cancer doubling time i.e. 200 days (Peer et al., 1993) and initial values equal to $(V_0, K_0) = (8000\text{mm}^3, 9750\text{mm}^3)$. A plateau of the tumour volume is obtained approximately after 3000 days. This plateau is just a theoretically set point value reached beyond the life of a patient. The value of the plateau is determined by the point where angiogenesis stimulators and inhibitors come into balance. The tumour is characterized by the parameter values listed in Table B3.

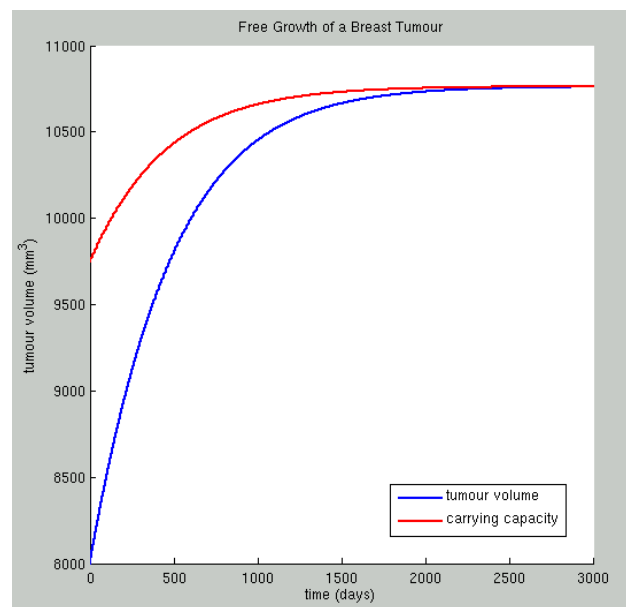


Figure B4: Simulation results for an untreated breast tumour characterized by the parameter values shown in Table B3.

Code Execution 3: The application of an intermittent bevacizumab mono-therapy treatment scheme consists of ten 90-minute intravenous infusions of 15 mg/kg (Figure B3) on the breast tumour. The dosage, route of administration and duration of infusion are adopted from a scheme used in clinical practice (Lu et al., 2008; http://www.gene.com/download/pdf/avastin_prescribing.pdf, Visited: 1/2013).

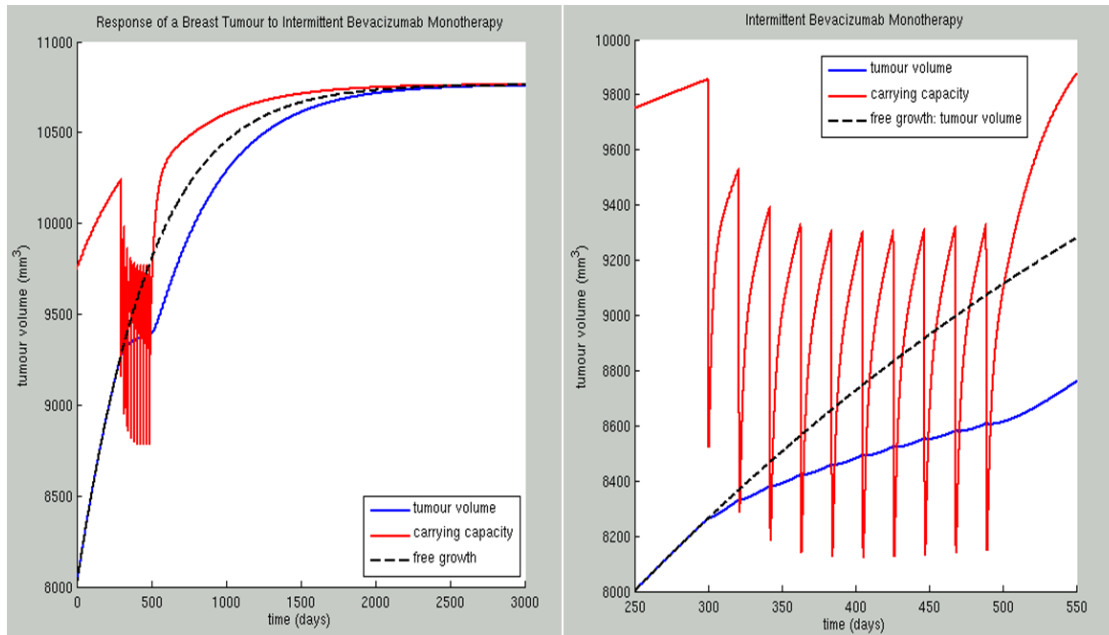


Figure B5: Simulation results for a breast tumour treated with bevacizumab according to a scheme used in clinical practice (Lu et al., 2008) and characterized by the parameter values shown in Table B2

The time course of volume V and carrying capacity K for a breast tumour treated with Bevacizumab is shown in Figure B5. The tumour and the administered treatment are characterized by the parameter values shown in Table B3. Treatment begins on the 300th day when the tumour size is already equal to 8000 mm³ i.e. when the radius of the tumour is equal to 20 mm. Before the 300th day an overshoot of tumour volume (V) is observed due to the fact that the tumour is untreated. With the start of the administration regimen a sharp decrease of tumour capacity is obvious and then the response of tumour volume follows. Indeed, anti-angiogenic treatment influences directly the vascular compartment and through that, the tumour compartment response. However, the specific tumour is not very responsive to the treatment scheme applied as the tumour volume does not decrease but increases with a slower growth rate comparing to the one for the case of free growth (black dotted line). Another interesting observation is that each time the curve $K(t)$ intersects curve $V(t)$, the monotonicity of the function $V(t)$ changes. This is reasonable if one keeps in mind that tumour capacity is actually defined as the maximal tumour volume that the current vascular system can support.

Code Execution 4: The time evolution of untreated tumour growth where λ_1 has been set equal to 0.0038115 1/day in order to reflect a realistic value of human breast cancer doubling time i.e. 188 days (Peer et al., 1993) and initial values as equal to $(V_0, K_0) = (8000\text{mm}^3, 9750\text{ mm}^3)$. The time-point 0 corresponds to a tumour radius equal to $8000^{1/3} \approx 20\text{ mm}$, namely to a vascularised tumour. The growth slowdown is obvious and the asymptotic limit obtained is approximately equal to 9200 mm^3 after approximately 2200 days.

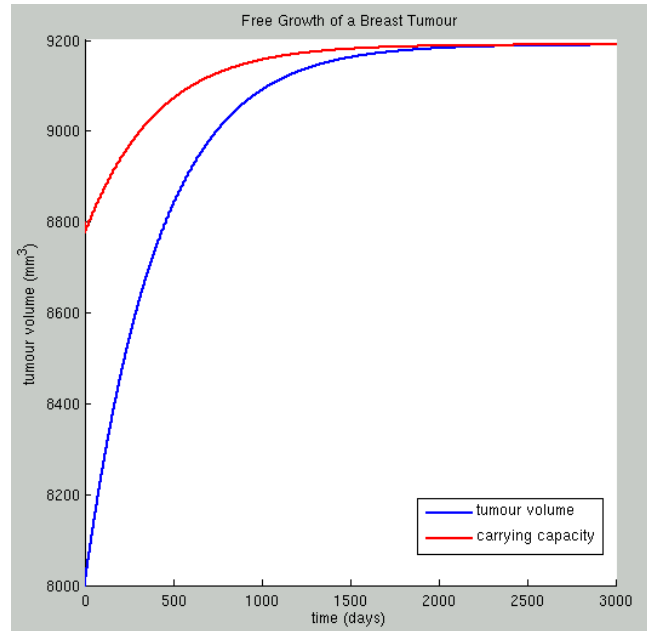


Figure B6: Simulation results for an untreated breast tumour characterized by the parameter values shown in table B3.

Code Execution 5: An intermittent bevacizumab monotherapy treatment scheme consisting of twenty 90-minute intravenous infusions of 20 mg/kg and administered every 21 days is applied on a breast tumour characterized by the parameter values shown in table B3. The specific scheme details such as dosage, frequency of infusions and duration of infusions have been clinically used in the context of the assessment of Bevacizumab pharmacokinetic profile via a population analysis of 491 cancer patients (http://www.gene.com/download/pdf/avastin_prescribing.pdf, Visited: 1/2013).

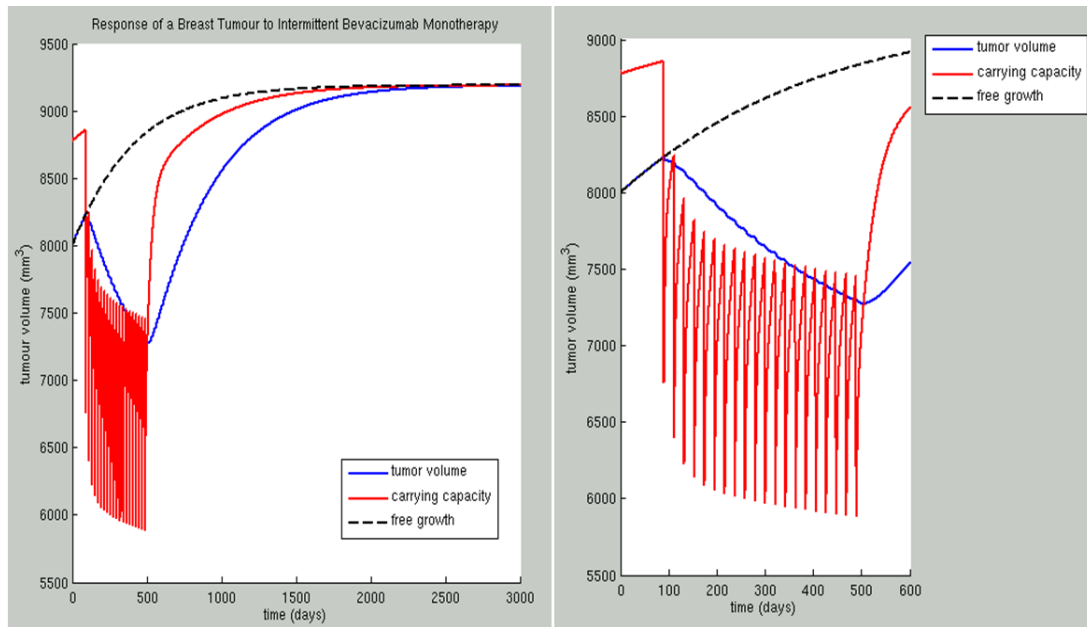


Figure B7: Simulation results for a breast tumour treated with intermittent Bevacizumab mono-therapy scheme. The tumour and treatment scheme are characterized by the parameter values shown in Table B3

The time course of volume V and carrying capacity K for a breast tumour responsive to the specific bevacizumab treatment, is shown in Figure B7. Treatment begins on the 90th day when the tumour size is already equal to 8000 mm³ i.e. when the radius of the tumour is equal to 20 mm. In analogy with the results of the previous treatment, it is obvious that before the start of treatment, i.e. before the 90th day, one can observe the time course of an untreated tumour. Following the administration of inhibitor, the vasculature regresses and the tumour volume decreases. As expected, the treatment has also delayed the reaching of the tumour plateau.

Table B3: Parameter values used in the demonstrated code executions in the context of presentation of indicative results

Parameters	Code execution 2 Free Growth	Code execution 3 Intermittent Bevacizumab Monotherapy	Code execution 4 Free Growth	Code execution 5 Intermittent Bevacizumab Monotherapy	Units
Vascular Tumour Growth Model					
lamda1	0.003465	0.0003465	0.0038115	0.0038115	1/day
lamda2	0	0	0	0	1/day
c	5.85	5.85	5.265	5.265	mg/(ml·day · mm ^{3p})
alpha	1	1	1	1	mg/(mm ^{3p} · ml)
beta	1	1	0.001	0.001	mm ^{3p}
p	0	0	0	0	-
d	0.012	0.012	0.012	0.012	1/(day · mm ²)
Initial Values					
[v0,k0]	[8000,9750]	[8000,9750]	[8000,8800]	[8000,8800]	[mm ³ , mm ³]
Simulation Time Span					
tspan	[0,3000]	a. [0,3000] b. [250,550]	[0,3000]	a. [0,3000] b. [0,600]	days

Therapy – Related Parameters					
k12	n/a	0.223	n/a	0.223	1/day
k21	n/a	0.215	n/a	0.215	1/day
kel	n/a	0.0779	n/a	0.0779	1/day
Vp	n/a	2660	n/a	2660	ml
dose	n/a	dosage*weight	n/a	dosage*weight	mg
duration of infusion	n/a	90	n/a	90	min
dosage	n/a	15	n/a	20	mg/kg
weight	n/a	50	n/a	50	kg
number of infusions	n/a	10	n/a	20	-
administration time-points	n/a	T1=300, T2=321, T3=342, ..., T40=489	n/a	T1=90, T2=111, T3=132, ..., T20=489	days

B1.4 Parametric Analysis

In order to investigate the impact of the input factors on the model and to identify the model parameters exerting the greatest influence on the model predictions, a thorough sensitivity analysis is being currently conducted. In order to avoid the impractically high complexity of a complete factorial experimental design, firstly a *varying one parameter at a time* approach was selected so as to study the dynamic behaviour of a breast tumour. Thus, the effect of the perturbation of a factor when all others are kept constant at the reference values adopted in Hahnfeldt et al. has been studied. The only exception was λ_1 , the value of which has been adapted so that it reflects a real-life breast tumour doubling time. After completion of the analysis on parameter sensitivity the most influential of the parameters involved in the model will be determined. All parameters related to tumour free growth, except for λ_2 which is the proportionality constant involved in the omitted term in equation (B2), were varied by a uniform 5, 10 and 20 percent either side of the baseline value.

λ_1

The specific parameter is a proportionality constant with units of 1/time which is involved in the equation describing the rate of change of tumour volume (equation (B1)).

λ_1 is actually the parameter that implicitly represents the doubling time of the tumour. Indeed, with the biologically justifiable simplification of exponential tumour growth, the following equation holds:

$$\frac{dV}{dt} = \lambda_1 \cdot V$$

This implies that

$$V = V_0 \cdot e^{\lambda_1 t}$$

For the tumour doubling time t_d the following holds:

$$2V_0 = V_0 \cdot e^{\lambda_1 t_d}$$

$$\text{Finally, } t_d = \frac{\ln 2}{\lambda_1} \quad (B3)$$

Code Execution 6: All parameters were kept constant while lamda1 (λ_1) was perturbed by a uniform 5, 10 and 20 percent either side of the base value (blue line) which was set equal to 0.003465 1/day in order to reflect a doubling time of 200 days. Values 0.003465, 0.00329175, 0.0031185, 0.002772, 0.00363825, 0.0038115 and 0.004158 were assigned to lamda1. With the use of equation (B3) the respective values of tumour doubling time are 200 days, 210.52 days, 222.22 days, 250 days, 190.47 days, 181.81 days and 166.66 days.

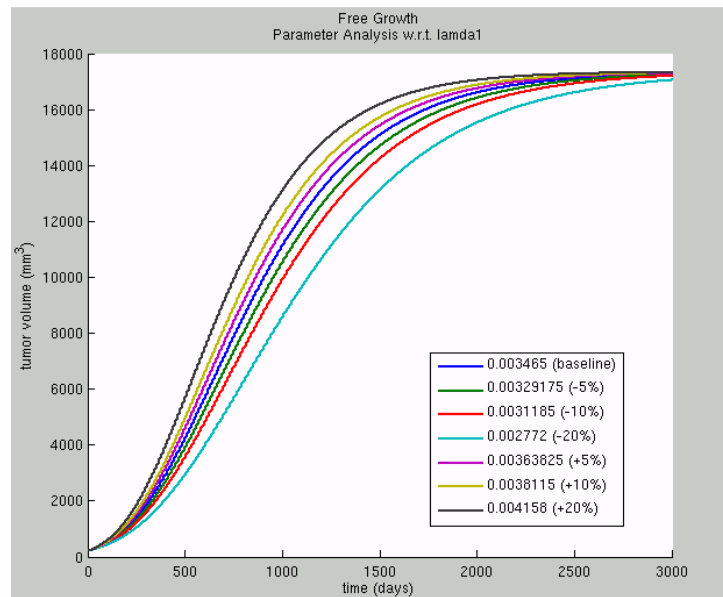


Figure B8: Effect of input parameter lamda1 (λ_1) on breast tumour volume. The values assigned to the rest of the model input parameters are reported in Table B4.

Taking equation (B3) into account, it is expected that higher values of lamda1 (λ_1) would result in a tumour with lower doubling time and therefore to a more aggressive tumour. Indeed, increasing the value of lamda1 (λ_1) leads to a faster growing tumour while decreasing the value of lamda1 (λ_1) has the opposite effect, namely leads to a less aggressive tumour.

C

Parameter c is a proportionality constant involved in equation (B2) and specifically in the term representing the stimulatory capacity of the tumour upon the inducible vasculature. Its unit is $1/(\text{time} \cdot \text{vol}^p) \cdot \text{conc.}$

Code Execution 7: All parameters were kept constant while c was perturbed by a uniform 5, 10 and 20 percent either side of the base value (blue line) which was extracted from (Hahnfeldt et al., 1999). Values 5.85, 5.265, 4.68, 6.435 and 7.02 were assigned to c .

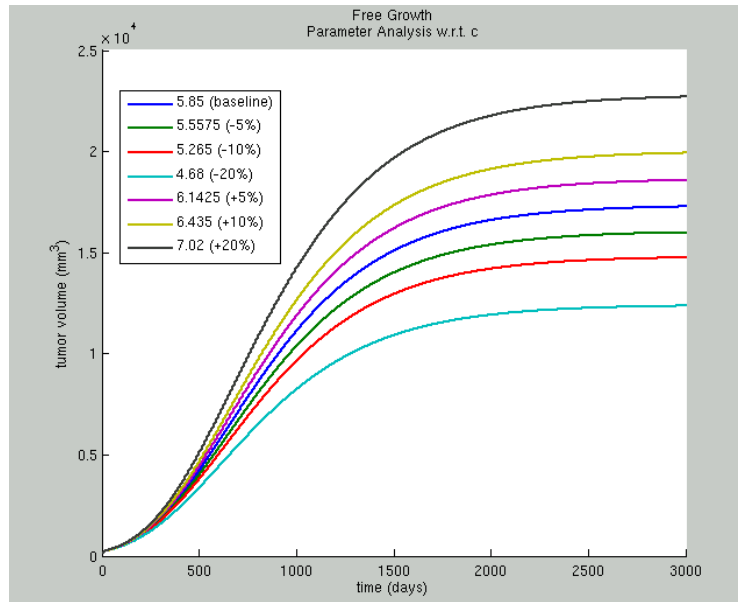


Figure B9: Effect of input parameter c on breast tumour volume. The values assigned to the rest of the model input parameters are reported in Table B4.

According to equation (B2) an increased value of parameter c would result in higher angiogenic stimulatory capacity of the tumour and thus to a faster growing breast tumour. Similarly, decreasing c would result in a slower growing breast tumour, a conclusion that is also corroborated in a graphical way (figure B9).

alpha (α)

A constant involved in equation (B2) and specifically in the term representing the stimulatory capacity of the tumour upon the inducible vasculature. Its unit is $\text{conc} / \text{vol}^p$.

Code Execution 8: All parameters were kept constant while alpha (α) was perturbed by a uniform 5, 10 and 20 percent either side of the base value (blue line) which was extracted from (Poleszczuk et al., 2011). Thus, values 1, 0.95, 0.9, 0.8, 1.05, 1.1 and 1.2 were assigned to alpha.

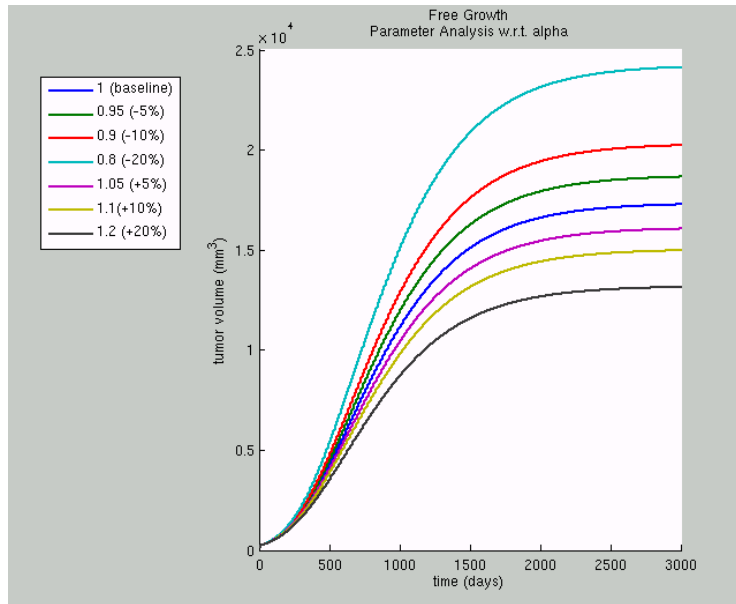


Figure B10: Effect of input alpha (α) on breast tumour volume. The values assigned to the rest of the model input parameters are reported in Table B4.

Since the specific parameter is involved in equation (B2) as a constant in the denominator of the term that represents angiogenic stimulatory capacity of the tumour, higher values of alpha (α) are associated with weaker stimulatory capacity and consequently to less aggressive tumours. In a similar way, decreasing the value of alpha (α) will lead to a more aggressive tumour, a pattern that can also be graphically observed (figure B10).

beta (β)

Parameter beta (β) is a constant involved in equation (B2) and specifically in the term representing the stimulatory capacity of the tumour upon the inducible vasculature. It has units of vol^p . This parameter is eliminated from equation (B2) for the case of free tumour growth due to the form of the specific term that is involved and thus, different values will be assigned to it for the case of constant treatment.

Code Execution 9: All parameters were kept constant while beta (β) was perturbed by a uniform 5, 10 and 20 percent either side of the base value (blue line) which was extracted from (Poleszczuk et al., 2011). Values 1, 0.95, 0.9, 1.05 and 1.1 were assigned to beta.

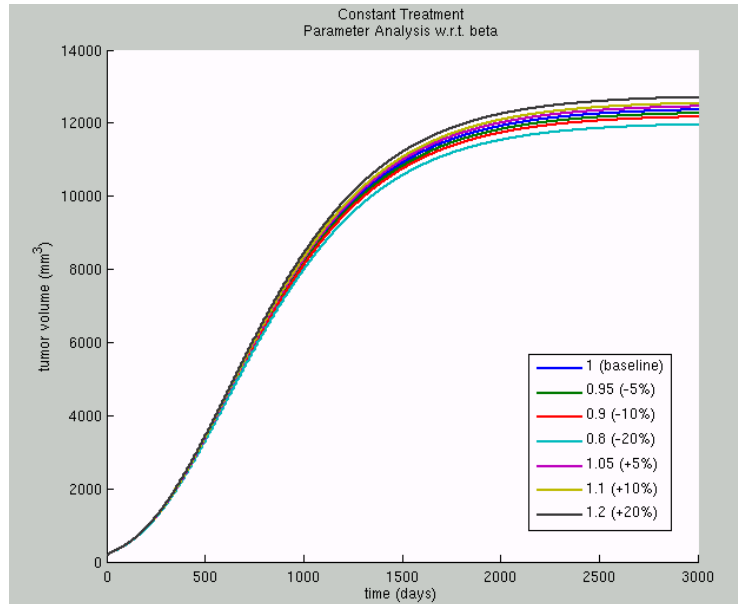


Figure B11: Effect of input parameter beta (β) on breast tumour volume. The values assigned to the rest of the model input parameters are reported in Table B4.

Higher values of parameter beta (β) result in faster growing tumours while low values of beta (β) are associated with slower growing tumours. A parameter involved in equation (B2) and specifically in the term representing the stimulatory capacity of the tumour upon the inducible vasculature.

Code Execution 10: This parameter is actually a Hill coefficient which is eliminated from the equation for the case of free tumour growth and thus, different integer values (any other choice would contradict the binding of n ligands that it is supposed to reflect) will be assigned to it for the case of constant treatment. The values 0, 1, 2, 3 and 4 were assigned to p .

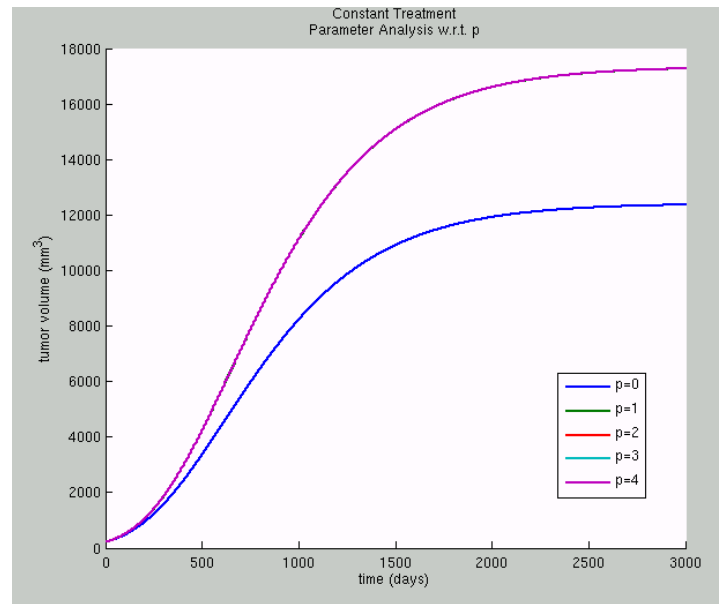


Figure B12: Effect of input parameter p on breast tumour volume. The values assigned to the rest of the model input parameters are reported in Table B4.

Based on figure 12, for $p=1$, $p=2$, $p=3$ and $p=4$ we derived a common treatment response. The tumour is obviously more aggressive than the one for $p=0$.

d

Parameter d is a proportionality constant involved in equation (B2) and specifically in the term representing endogenous inhibition of previously generated vasculature. Its unit is $1/(\text{time} \cdot \text{vol}^{2/3})$.

Code Execution 11: All parameters were kept constant while d was perturbed by a uniform 5, 10 and 20 percent either side of the base value (blue line) which was extracted from (Hahnfeldt et al., 1999). The values 0.00873, 0.0082935, 0.007857, 0.0091665 and 0.009603 were assigned to d .

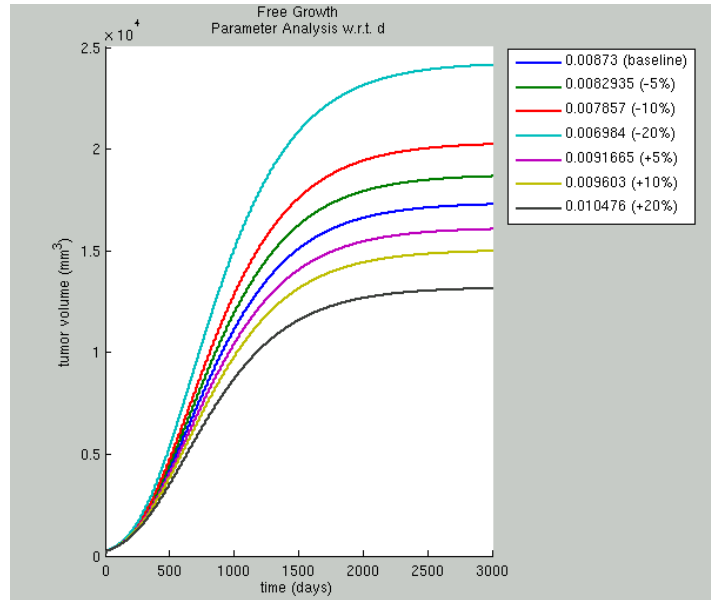


Figure B13: Effect of input parameter d on breast tumour volume. The values assigned to the rest of the model input parameters are reported in Table B4.

High values of parameter d lead to tumours with high endogenous inhibition of previously generated vasculature and faster growing tumours. In a similar way, lower values of d correspond to tumours with small concentrations of angiogenesis inhibitors and thus, to slower growing tumours.

According to code executions 6 – 11, where all parameters but one were held constant on each graph, the parameters d , α , and c seem to be those exerting the greatest influence on the model predictions. In order for this conclusion to become clearer, we proceeded to the demonstration of the effect of perturbing each parameter at a time by a uniform +10% (for $p=0$ and $p=1$) and -10% on breast tumour volume in common axial system (for $p=0$ and $p=1$). The baseline values of the parameters are shown in Table B4.

Indeed, based on Figure B14, it becomes obvious that out of all model parameters that affect the tumour volume the most critical ones are d , α , and c (light blue, magenta and red line respectively).

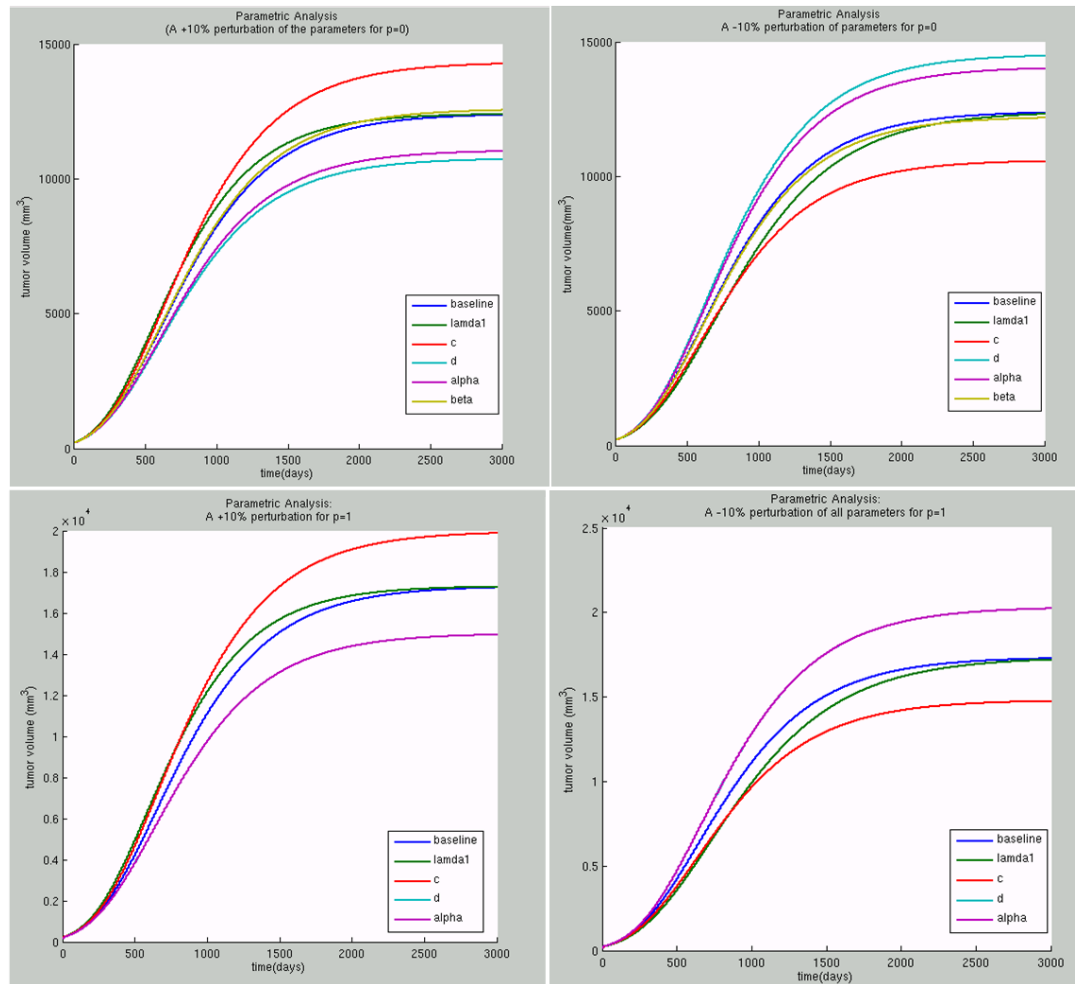


Figure B14: Depicting the effect of perturbing each parameter by a uniform +10% and -10% on breast tumour volume in a common axial system, for p=0 and p=1.

Table B4: Parameter values used for the code executions in the context of parametric analysis of the model.

Parameters	Baseline Values	Code Execution 5 (perturbing lamda1)	Code Execution 6 (perturbing c)	Code Execution 7 (perturbing alpha)	Code Execution 8 (perturbing beta)	Code Execution 9 (perturbing p)	Code Execution 10 (perturbing d)
lamda1 (1/time)	0.003465	0.003465, 0.00329175, 0.0031185, 0.002772, 0.00363825, 0.0038115, 0.004158	0.003465	0.003465	0.003465	0.003465	0.003465
c (1/(day · mm ^{3p}) · mg/ml))	5.85	5.85	5.85, 5.5575, 5.265, 4.68, 6.1425, 6.435, 7.02	5.85	5.85	5.85	5.85
alpha (mg / (ml·mm ^{3p}))	1	1	1	1, 0.95, 0.9 0.8, 1.05, 1.1, 1.2	1	1	1
beta (mm ^{3p})	1	1	1	1	1, 0.95, 0.9 0.8, 1.05, 1.1, 1.2	1	1
p	0	0	0	0	0	0,1,2,3,4	0
d (1/(day · mm ²))	0.00873	0.00873	0.00873	0.00873	0.00873	0.00873, 0.0082935, 0.007857, 0.006984, 0.0091665, 0.009603, 0.010476	0.00873

B1.5 Questions of clinical interest

The goal of *in silico* oncology is to construct a reliable decision support system to enable clinicians to select the optimal treatment scheme, based on patient-specific clinical data. Therefore, the modeler needs to focus on issues of possible clinical interest, so as to validate the results that are derived by the model as well as to gain insight into possible answers to the questions posed by a clinician. To this end, several code executions have been performed in order to investigate the impact of drug dosage, the drug administration starting point and the frequency of drug administrations on the breast tumour volume response for the case of intermittent Bevacizumab mono-therapy.

Q1. How does varying the administered dosage affect the response of a breast tumour to intermittent bevacizumab monotherapy?

Code Execution 12: An intermittent bevacizumab monotherapy scheme was applied to the breast tumour characterized by the parameter values shown in Table 5, administering though different dosage (5 mg/kg, 10 mg/kg, 15 mg/kg, 20 mg/kg) each time, while keeping all other therapy – related parameters constant. The effect of varying bevacizumab dosage on the response of the breast tumour is revealed in Figure B15.

Based on the derived results the following observations can be made. While dosage increases, the breast tumour has a better response to treatment, exactly as it would be expected based on clinical experience. More specifically, dosage values 5 mg/kg, 10 mg/kg and 15 mg/kg may not be capable of decreasing the breast tumour volume, but they do achieve a decrease in its growth rate. On the contrary, a dosage equal to 20 mg/kg results in decreasing breast tumour volume. Thus, by the end of the simulated therapy, all four schemes proved to be beneficial to the patient when compared to the results for the case of free growth (black dotted line), since dosage values 5 mg/kg, 10 mg/kg, 15 mg/kg and 20 mg/kg decreased the free growth tumour volume to 97%, 95%, 92.5% and 90% (at the 500th day). However, according to Figure B15, only the administration of a dosage equal to 20 mg/kg (light blue line) achieved a decrease of the initial volume (before therapy) for the specific breast tumour.

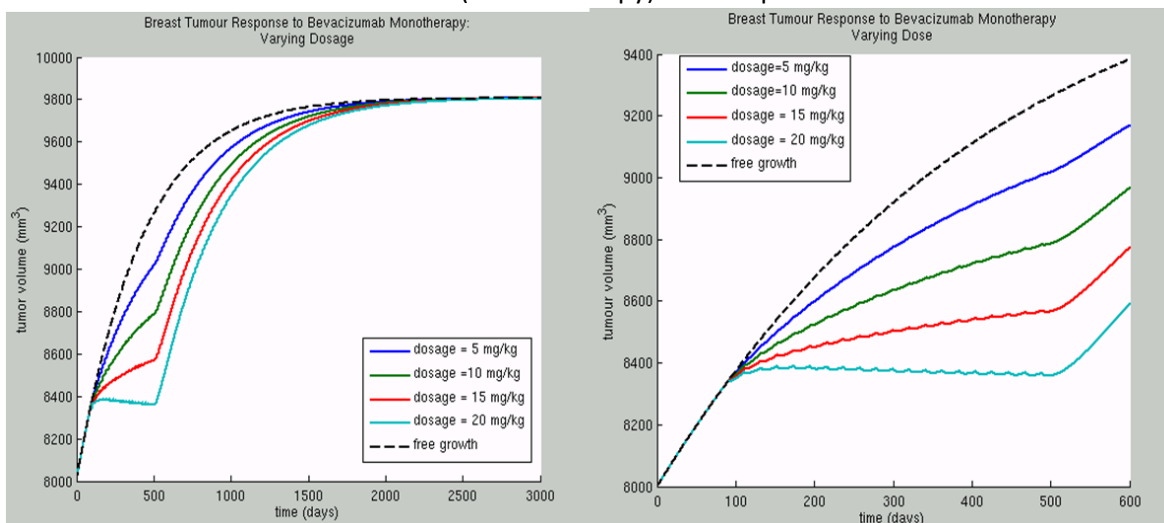


Figure B15: Testing the response of a breast tumour to different dosage values of bevacizumab. The values assigned to the rest of the model input parameters are reported in Table B5.

Q2. What is the impact of varying the starting point of therapy to the response of a breast tumour to intermittent Bevacizumab monotherapy?

Code Execution 13: We have performed five code executions in order to simulate the application of the same intermittent Bevacizumab monotherapy scheme to a breast tumour characterized by the parameter values shown in table B5, but each time having selected a different therapy starting point.

Results shown in Figure B16 suggest that the lowest value of tumour volume is succeeded when therapy starts at the 50th day (blue line), i.e. when the regimen is administered at the earliest starting point. The more the therapy delays, the higher the tumour volume by the end of the treatment. However, it is worth noting that the results also indicate an increasing percentage of decrease as the starting point delays. The specific behaviour which suggests an increasing time dependent sensitivity to anti-angiogenic treatment could be due to the fact that in a fully vascularised tumour each vessel may be supporting more tumour cells and thus, destroying a vessel via antiangiogenic treatment could cause the death of more tumour cells. Nevertheless, further investigation is needed in order to check the validity of these conclusions.

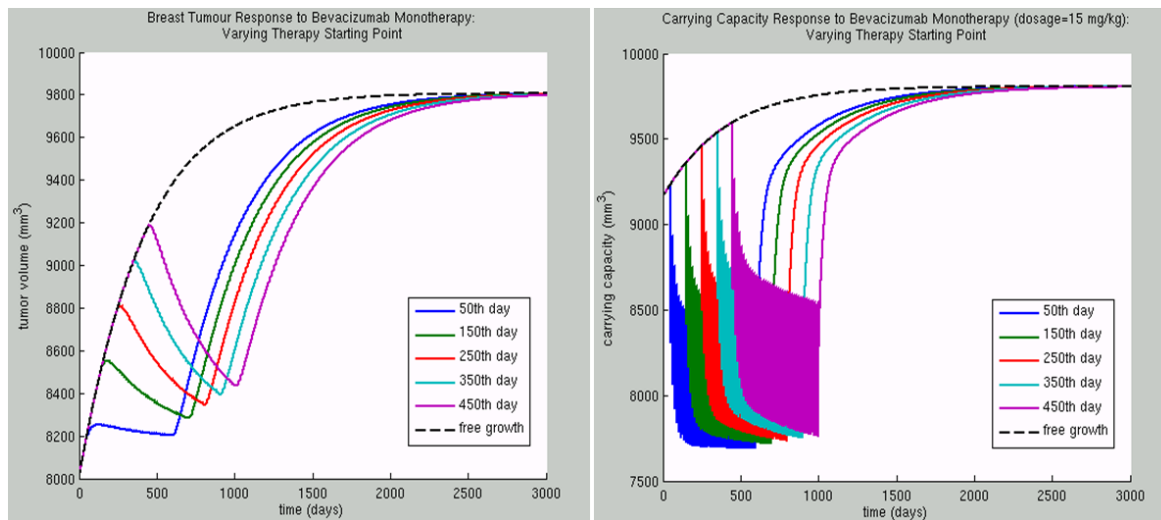


Figure B16: Testing the response of a breast tumour to different Bevacizumab treatment starting points. The values assigned to the rest of the model input parameters are reported in Table B5.

Q3. How does the variation of the frequency of administered doses of Bevacizumab affects the response of a breast tumour to therapy?

Code Execution 14: We have simulated the application of three intermittent Bevacizumab monotherapy schemes to a breast tumour characterized by the parameter values shown in Table B5, keeping always constant a daily dosage equal to 1 mg/kg/day. Specifically, administering 7 mg/kg once a week (blue line), 14 mg/kg once every 14 days (green line), and 21 mg/kg (red line) once every 21 days, while keeping all other tumour and therapy related parameters constant are the selected treatment schemes.

The effect of varying the frequency of administered doses of bevacizumab on the response of the breast tumour characterized by the parameter values shown in Table B5 is revealed in Figures B17 and B18. Based on the predictions, it can be observed that intravenous bevacizumab mono-therapy of 21 mg/kg every three weeks seems to be at least as efficient as administering 14 mg/kg every two weeks, a treatment scheme which in a similar way seems to be at least as efficient as administering 7 mg/kg once a week.

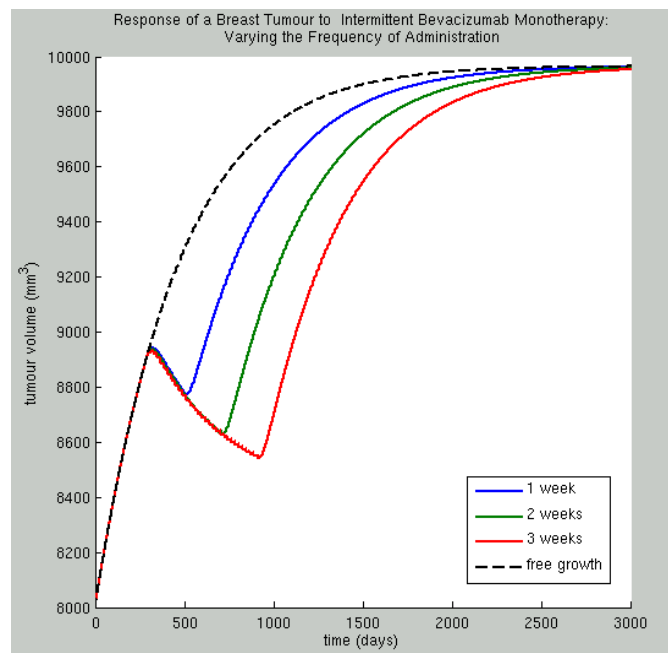


Figure B17: Testing the response of a breast tumour while varying the frequency of bevacizumab administration. The parameter values characterizing both the tumour and the treatment schemes are reported in Table B5

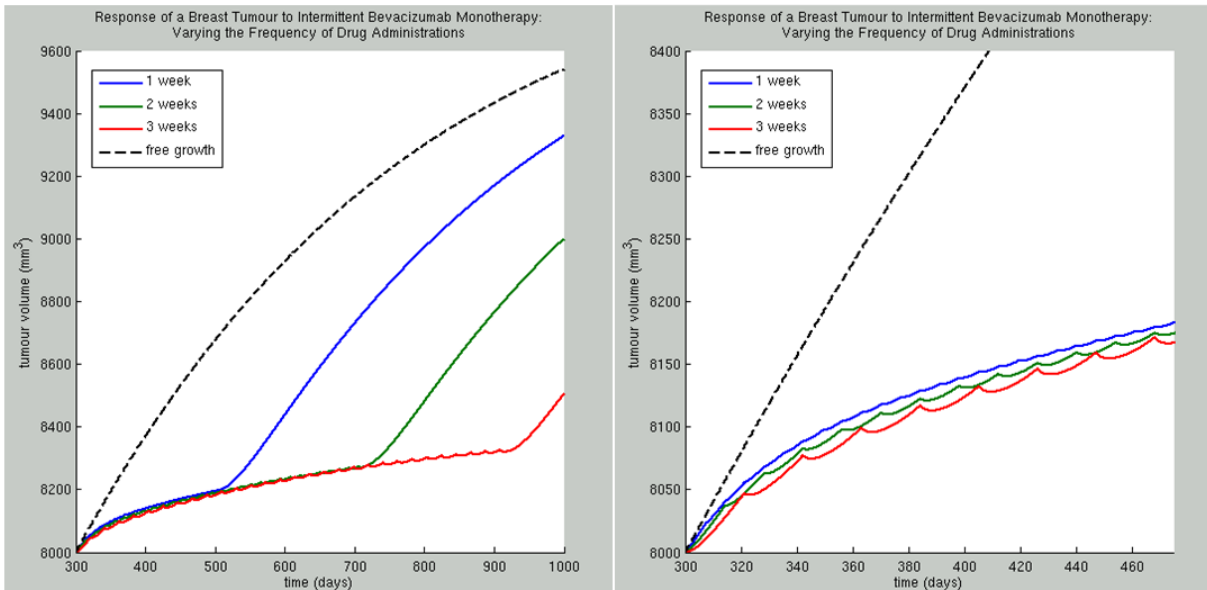


Figure B18: Testing the response of a breast tumour while varying the frequency of bevacizumab administration. The parameter values characterizing both the tumour and the treatment schemes are reported in Table B5

Code Execution 15: The recommended dosage of bevacizumab for various kinds of cancers (metastatic colorectal cancer, non-squamous non-small cell lung cancer, glioblastoma and metastatic renal cell carcinoma) is either 10 mg/kg once every 14 days or 15 mg/kg given once every 21 days, given as an IV infusion (http://www.gene.com/download/pdf/avastin_prescribing.pdf, Visited: 1/2013). In the three following figures (figures B19, B20), the simulation results of the two aforementioned treatment schemes (keeping all other tumour and therapy – related parameters as in table 5, code execution 14) are also demonstrated and seem to be in agreement with the conclusions drawn so far. Indeed, administering 15 mg/kg once every 21 days seems to be at least as effective as administering 10 mg/kg once every 14 days (Figure B20).

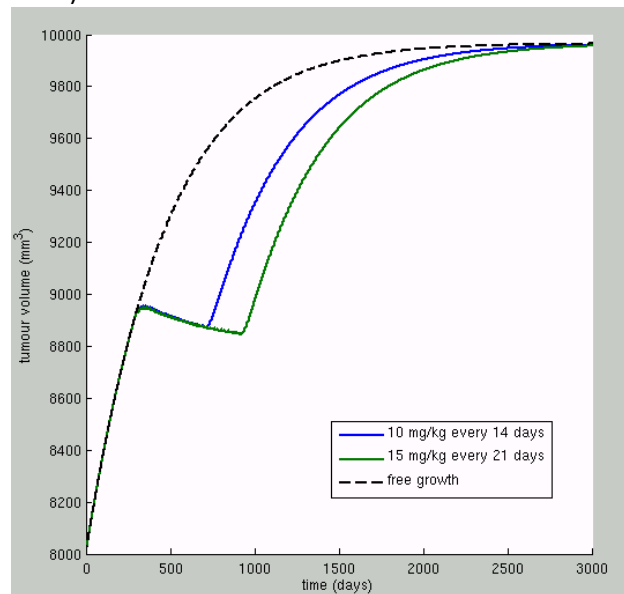


Figure B19: Comparing two Bevacizumab monotherapy treatment schemes applied in clinical practice. The parameter values characterizing the tumour are reported in Table B5, code execution 14.

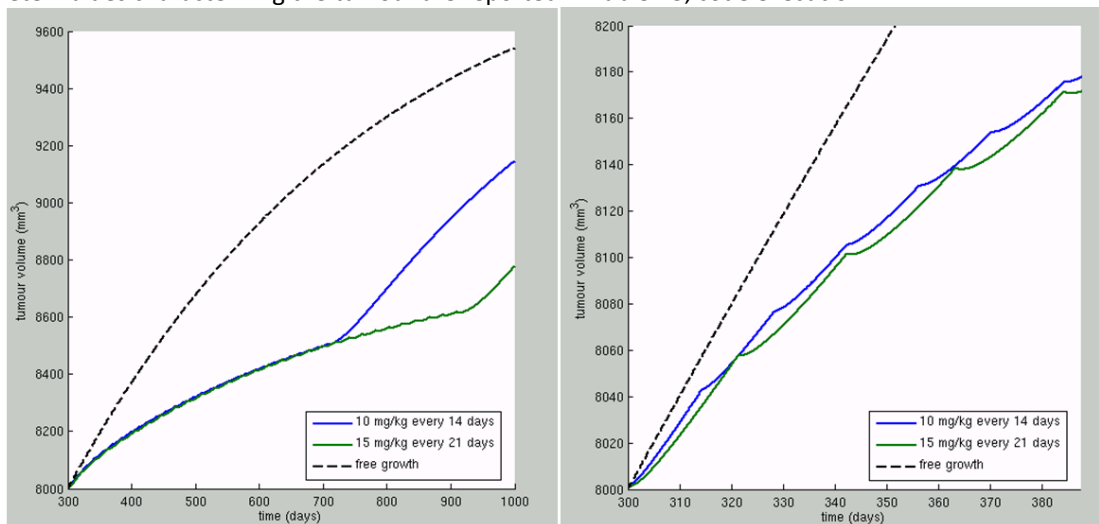


Figure B20: Comparing two Bevacizumab monotherapy treatment schemes applied in clinical practice. The parameter values characterizing the tumour are reported in table B5, code execution 14.

Table B5: Parameter values used in demonstrated code executions in the context of investigating clinician – related research questions.

Parameters	Code execution 12	Code execution 13	Code execution 14	Units
Vascular Tumour Growth Model				
lamda1	0.0038115	0.0038115	0.003465	1/day
lamda2	0	0	0	1/day
c	4	4	5.5575	mg/ (ml·day· mm ^{3p})
alpha	1	1	1	mg / (mm ^{3p} · ml)
beta	1	1	1	mm ^{3p}
p	0	0	0	-
d	0.00873	0.00873	0.012	1/(day · mm ²)
Initial Values				
[V0,K0]	[8000,9200]	[8000,9200]	[8000,9200]	[mm ³ , mm ³]
Simulation Time Span				
tspan	a. [0,3000] b. [0,600],	[0,3000]	a. [0,3000] b. [300,1000]	days
Therapy – Related Parameters				
k12	0.223	0.223	0.223	1/day

k21	0.215	0.215	0.215	1/day
ke	0.0779	0.0779	0.0779	1/day
Vp	2660	2660	2660	ml
dose	dosage*weight	dosage*weight	dosage*weight	mg
duration of infusion	90	90	90	min
dosage	0,5,10,15,20	15	7,14,21	mg/kg
weight	50	50	50	kg
number of infusions	20	40	30	-
administration time-points	T1=90,	T1=50,100,150,200,250	T1=300,	days
frequency of infusions	21	14	7,14,21	days

B2 The discrete event – discrete state tumour growth model

B2.1 Basic features of the model

The model simulates the spatiotemporal response of breast cancer clinical tumours to chemotherapeutic treatment with the single agent vinorelbine. It is based on the consideration of a discrete time and space stochastic cellular automaton, representing the tumour region. More specifically, the tumour region can be considered as a grid of “geometrical cells” (GCs, the elementary volume of the grid). Each GC corresponds to a cluster of heterogeneous cells found in various states. Specific rules regulate the transition between these states, as well as cell movement throughout the tumour volume; the aim is a realistic, conformal to the initial shape of the tumour, simulation of spatial evolution.

Free Growth: The adopted cytokinetic model incorporates the biological mechanisms of cell cycling, quiescence, differentiation and loss. Stem, Limited Mitotic Potential (LIMP), differentiated (DIFF), apoptotic and necrotic cells represent the distinct cell categories of the model. More specifically, tumour sustenance is attributed to the presence of a cell population that exhibits stem cell like properties. Specifically, cancer stem cells have the ability to preserve their own population, as well as give birth to cells of limited mitotic potential (LIMP cells) that follow the path towards terminal differentiation (DIFF cells). A proliferating tumour cell (stem or LIMP) passes through the successive cell cycle phases. Phases within or out of the cell cycle (G1, S, G2, M, G0) constitute different states in which cells may be found. After the completion of mitosis a fraction of newborn cells will enter the dormant phase, whereas the rest will continue to cycle. Transition to quiescence (dormant, G0, phase) and “awakening” of dormant cells are regulated by local metabolic conditions. All cell categories may die through spontaneous apoptosis. However, for dormant and differentiated cells necrosis is the main cell loss mechanism caused by inadequate nutrients’ and oxygen supply.

Treatment: The model addresses the case of vinorelbine agent. When a tumour is chemotherapeutically treated, a fraction of cancerous proliferating cells are lethally hit by the drug. These cells enter a rudimentary cell cycle that leads to apoptotic death through a cell cycle phase depending each time on the specific chemotherapeutic agent. In the simulation model the case of vinorelbine agent that primarily binds to tubulin and inhibits microtubule assembly, resulting in cell mitosis disruption and eventual cell death is addressed. The effect of the drug is considered instantaneous at the time of its administration.

B2.2 Simulation flowchart

At each time step, the discretizing mesh covering the anatomical region of interest is virtually scanned in order to apply the basic metabolic, cytokinetic, and mechanical rules that govern the spatiotemporal evolution of the tumour system (e.g. Stamatakis et al.2010) . The time unit considered in this paper is taken one hour since this is approximately the duration of mitosis, the shortest cell cycle phase (Bast et

al., 2000). A simplified flowchart of the proposed and adopted algorithm pertaining to free growth is described below (e.g. Stamatakos et al. 2010):

1. The full version of the model supports 3D tumours of arbitrary shape. The first step is the definition of the mesh and the characterization of the GCs (occupied – non occupied) based on the available patient specific imaging data. In case of unavailability of the required data for the 3D reconstruction of the tumour (e.g. MRI, CT scans), for example when only the maximum sizes are given from ultrasound measurements, the tumour can be approximated by a spheroid. The values given to the model input parameters are defined for the occupied GCs based on the metabolic region they belong. Histological data such as differentiation grade, as well as tumour specific data, can be incorporated for a further refinement of the values attributed to the model input parameters. In case of macroscopically homogeneous tumours all occupied GCs are characterized by the same values.

2. In order to avoid an abnormal free growth behaviour at the beginning of the simulation, the relative population (expressed as the fraction of the total tumour cell population) of each equivalence class and its equivalence subclasses used for the population initialization of the occupied GCs (step 3 below) must be adjusted to the values attributed to the cell category and cell phase transition rates.

3. The biological cells residing within each geometrical cell of the mesh are distributed into the five cell categories i.e. the stem, LIMP, differentiated, apoptotic and necrotic categories based on the fractions calculated during the previous step. Each stem or limp cell can be either proliferating residing in any of the cell cycle phases (G1, S, G2, M) or dormant. The initial distribution of the proliferating cells throughout the cell cycle phases is assumed analogous to the corresponding cell cycle phase durations. For any given time instant the biological cells belonging to the same cell category and cell cycle phase within a given GC are assumed synchronized. However, biological cells belonging to different GCs or to different categories and cell cycle phases within the same GC are not assumed synchronized. The mean time spent by the biological cells of a given equivalence subclass in the same subclass is initialized using a random number generator (Monte Carlo technique). The time under consideration can vary between 0 and the maximum time of the corresponding phase.

4. The first scan of the mesh aims at updating the state of each GC according to the proposed and adopted cytokinetic model. Initially spontaneous apoptosis induced cell loss is calculated for the mitotic potential cell categories (stem, LIMP, DIFF) based on the spontaneous apoptotic rates assumed, whereas the transition of differentiated cells to the necrotic phase is implemented based on the relevant necrotic rate. The products of the cell loss mechanisms (apoptotic and necrotic cell classes) that are removed from the tumour bulk are calculated and subtracted from the relevant cell class. The time registers of the proliferating subclasses within each geometrical cell are updated by subtracting one time unit (1 hour). When the mean time that the clustered cells have spent in the corresponding phase becomes equal to the phase duration then the cells enter the next phase and therefore the equivalence subclass ($G1 \rightarrow S$, $S \rightarrow G2$, $G2 \rightarrow M$, $M \rightarrow G1$ or $M \rightarrow G0$). The dormant cells are removed their class with a rate equal to the reciprocal of the G0 phase duration and transferred to the necrotic cell class or G1 subclass with a probability determined by the relevant model parameter.

As far as chemotherapy is concerned, at any time instant corresponding to vinorelbine administration, the numbers of proliferating belonging to each one of the phases G1, S, G2 or M and to each one of the

stem or LIMP mitotic potential categories that are designated as hit by the drug are computed. The latter is achieved through the utilization of the cell kill rate (CKR) parameter. In terms of the simulation model's parameters the cell kill rate is the percentage of LIMP and stem cells hit by the chemotherapeutic agent after each drug administration. In the model, the fraction of cells marked as hit by the chemotherapeutic agent is considered the same for all proliferating cell cycle phases. The above mentioned cell numbers are added to the corresponding cell numbers of the drug affected equivalence subclasses, designated as "phase"hit. In the simulation model the cytotoxic effect has been modeled by assuming that the apoptotic death of drug affected cells occurs at the end of M phase.

5. The second scan aims to preserve a constant cell density throughout the tumour volume allowing a variation of 10% around the mean value considered. To this end, adequately shaped morphological rules are introduced, which may lead to tumour expansion, as is the case in free tumour growth, or no change in tumour volume or tumour shrinkage as is usually the case after treatment administration. These rules regulate (i) the unloading of the overloaded geometrical cells due to mitoses and the creation of new tumour GCs (leading to differentiated tumour expansion) and (ii) the freeing of geometrical cells with few biological cells (leading to differentiated tumour shrinkage). During the above process artificial tumour fragmentation may occur and a special procedure has been devised in order to achieve macroscopic tumour cohesion. Random directions for differentiated growth expansion and shrinkage are used.

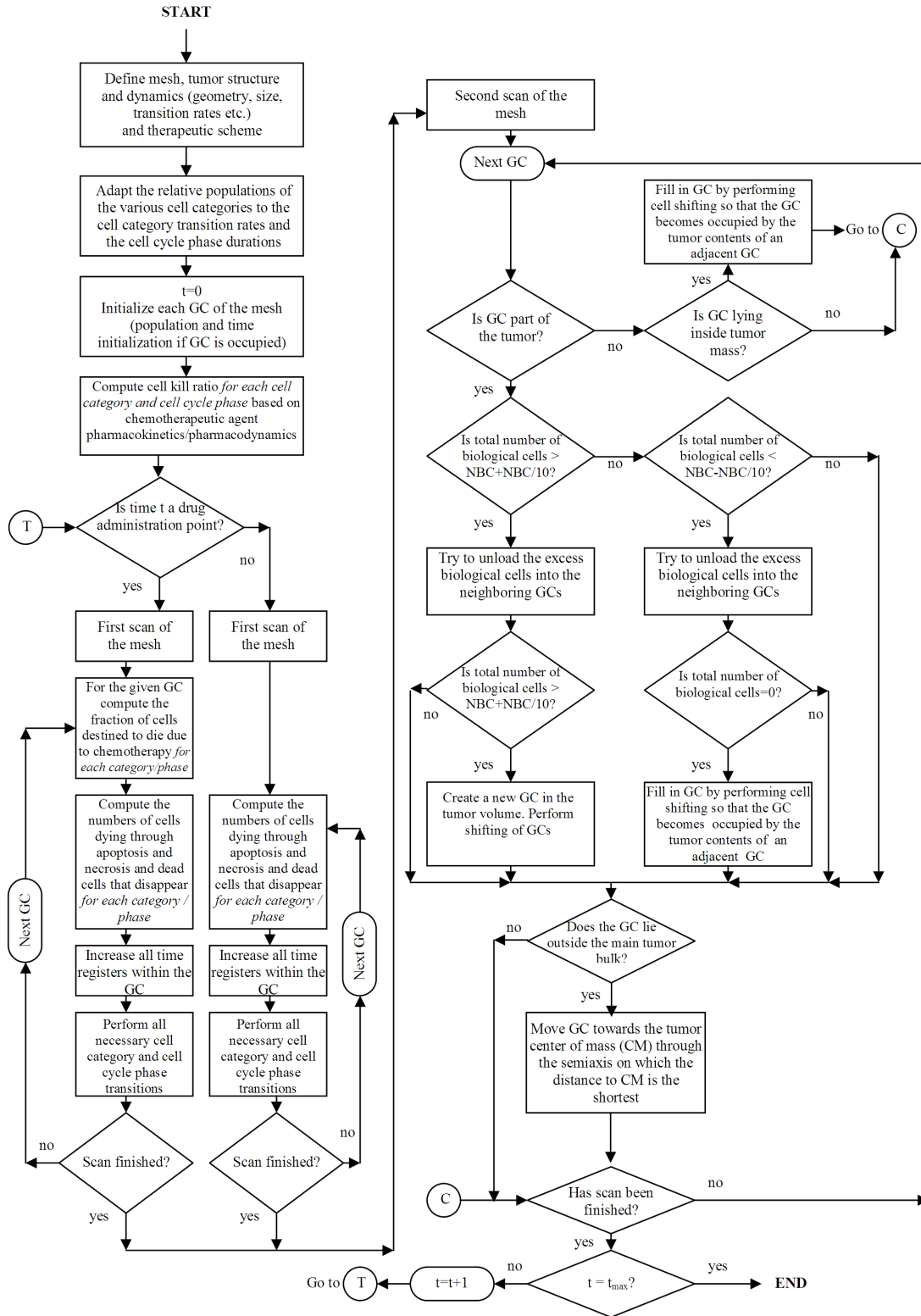


Figure 21: Simplified flow chart of the simulation procedure for a solid tumor of arbitrary shape. Abbreviations: GC: geometrical cell, t: simulation time step.

B2.3 Input parameter description

Table B6 presents the simulation model input parameters and their range of values according to pertinent literature or based on logical assumptions supported by basic science or clinical experience in case of lack of literature data. In the following a summarizing description of the model parameters related to free growth kinetics and their adopted values is provided for reasons of clarity:

(i) Cell density

This parameter is related to the number of biological cells initially occupying each geometrical cell (GC) belonging to the tumour. In order to preserve a roughly constant mean cell density throughout the tumour volume, the number of biological cells is allowed to fluctuate between a minimum and a maximum value. Therefore, as described in previous paragraphs, morphological rules governing the spatial evolution of the tumour have been introduced. A typical value of tumour cell density found in literature is 10^6 biological cells/mm³ (Steel, 1997).

(ii) Cell cycle phases duration

According to literature the duration of the various cell cycle phases follows a normal distribution. Although this observation has been explicitly taken into account in an in vitro tumour growth and treatment response model by our group (Zacharaki et al., 2004), in the present model cell cycle phase durations are considered constant and equal to their literature based mean values in order to accelerate executions. It is pointed out however that algorithmically constructing a normal distribution of phase durations is a pretty simple task.

Based on literature, the cell cycle duration may vary from approximately 20h to 96h (Sterin et al., 2001; Cos et al., 1996; Descamps et al., 1998; Barnes et al., 2001; Meyer et al., 1984). The duration of mitosis is considered constant and equal to 1h (Bast et al., 2000). The rest of the cell cycle phases durations are computed based on Salmon et al. (2001), after having taken into consideration the above assumption regarding the constant duration of mitosis. More specifically the following equations are used: $TG1 = TS = 0.41(Tc-TM)$, $TG2 = 0.18(Tc-TM)$, $TM = 1h$.

(iii) Duration of G0, duration of apoptosis and necrosis

According to literature dormant cells resting in G0 phase can survive under hypoxic conditions for 4-10 days (Maseide et al., 2000). Tumour apoptotic cells are generally considered to be rapidly phagocytosed in vivo (Dewey et al., 1995), contrary to the time-consuming process of necrosis products removal.

(iv) Cell category/phase transition rates/fractions

The values of these parameters have been selected based on both qualitative or semi-quantitative information and the dictates of the accumulated basic science and clinical experience.

(v) The Number of mitotic divisions that LIMP (committed progenitor) cells undergo before they become terminally differentiated (also called LIMP mitotic stages)

By varying the number of LIMP mitotic stages (assumed range 1 to 10) tumours characterized by different differentiation degrees can be simulated. To the best of our knowledge, there exists no specific information regarding precise values for this parameter in the pertinent literature. Nevertheless, the choice of its value is governed by the need to achieve specific relative percentages of stem and LIMP cells (which is also a reflection of a tumour's degree of differentiation). A larger number in this interval leads to lower percentages of stem cells.

Table B6: Code input parameters

Parameter symbol	Description	Value range	References
CELL PHASE DURATIONS			
$T_c[\text{class}]$ $\text{class} \in \{\text{stem}, \text{limp}^{\dagger}\}$	Cell cycle duration*	20 h – 96 h	Sterin et al., 2001; Cos et al., 1996 ; Descamps et al., 1998 ; Barnes et al., 2001; Meyer et al., 1984
$T_{G0}[\text{class}]$ $\text{class} \in \{\text{stem}, \text{limp}^{\dagger}\}$	G0 (dormant phase) duration* i.e. time interval before a dormant cell enters necrosis	96 h – 240 h	Maseide et al., 2000
T_N	Time needed for both necrosis to be completed and its lysis products to be removed from the tumour [‡]	20 h (0 h-100h)	Ginsberg T.,1996 (similar value range)
T_A	Time needed for both apoptosis to be completed and its products to be removed from the tumour [‡]	6 h (0 h – 25 h)	Ginsberg T.,1996 (similar value range), Dewey et al., 1995
CELL CATEGORY/PHASE TRANSITION RATES			
R_A	Apoptosis rate of living stem and limp tumour cells (fraction of cells dying through apoptosis per h)	$0.0\text{h}^{-1} - 1.0\text{h}^{-1}$	
R_{NDiff}	Necrosis rate of differentiated tumour cells	$0.0\text{h}^{-1} - 1.0\text{h}^{-1}$	
R_{ADiff}	Apoptosis rate of differentiated tumour cells	$0.0\text{h}^{-1} - 1.0\text{h}^{-1}$	
$P_{G0toG1}[\text{class}]$ $\text{class} \in \{\text{stem}, \text{limp}^{\dagger}\}$	Fraction of dormant (stem and LIMP) cells that re-enter cell cycle per time unit [‡]	$0.0\text{h}^{-1} - 1.0\text{h}^{-1}$	
P_{sleep}	Fraction of cells entering the G0 phase following mitosis [‡]	0.0 - 0.5	
P_{sym}	Fraction of stem cells that perform symmetric division [‡]	0.0 – 1.0	
MISCELLANEOUS PARAMETERS			
Cell density	Number of biological cells normally contained within a voxel/geometrical cell (GC) of the mesh	10^6	Steel et al.,1997
Voxel dimension	Dimension of voxel/GC edge in mm	1 or 2mm	
N_{LIMP}	Number of mitoses performed by LIMP cells before becoming differentiated	1 - 10	
$x_{dim}, y_{dim}, z_{dim}$	Number of geometrical cells along the x, y, z axis respectively		

tumour_length	Dimensions of the three tumour axes in mm in the case a triaxial ellipsoidal tumour is considered	
tumour_breadth		
tumour_width		
Execution time	Execution stop time (in days) after initialization	
Spatial evolution?	Will geometric spatial evolution be included in the simulations? "True": yes. The functions for tumour differential expansion/shrinkage are enabled "False": no. The functions for tumour differential expansion/shrinkage are disabled	0 or 1
Input image?	true: consider the input image file false: no input image file	0 or 1
Image filename	Name of the input image file	
Output directory	Name of the directory where the output files are stored	

CHEMOTHERAPY AND EFFECT

T _{1st,adm}	Time point of the first drug administration since pretreatment scan (simulation initialization) in days	Depends on scheme
T _{2nd,adm}	Time point after initialization (in days) when the 2 nd drug administration takes place (=last administration if total administrations <2)	Depends on scheme
T _{3rd,adm}	Time point after initialization (in days) when the 3 rd drug administration takes place (=last administration if total administrations <3)	Depends on scheme
T _{4th,adm}	Time point after initialization (in days) when the 4 th drug administration takes place (=last administration if total administrations <4)	Depends on scheme
T _{5th,adm}	Time point after initialization (in days) when the 5 th drug administration takes place (=last administration if total administrations <5)	Depends on scheme
T _{6th,adm}	Time point after initialization (in days) when the 6 th drug administration takes place (=last administration if total administrations <6)	Depends on scheme
CKR	Cell Kill Ratio of Vinorelbine	0-1
Cell kill factor	Factor adapting cell killing probability to stem or LIMP cells	0-1
[class]		
class ∈ {stem, limp [†] }		

* Cell phase durations can be defined separately for the stem and the LIMP tumour cell category.

B2.4 Vinorelbine Pharmacokinetic Model

This section presents the computation of vinorelbine concentration $C(t)$ in plasma. According to existing literature, the majority of vinorelbine data have been analysed using three - compartment methods (Gregory and Smith, 2000; Gauvin et al., 2000; Rezai et al., 2011).

The equations governing the pharmacokinetic properties of vinorelbine have been implemented according to the three-compartmental pharmacokinetic model proposed by (Bertrand and Mentre, 2008) for the case of administration via intravenous route.

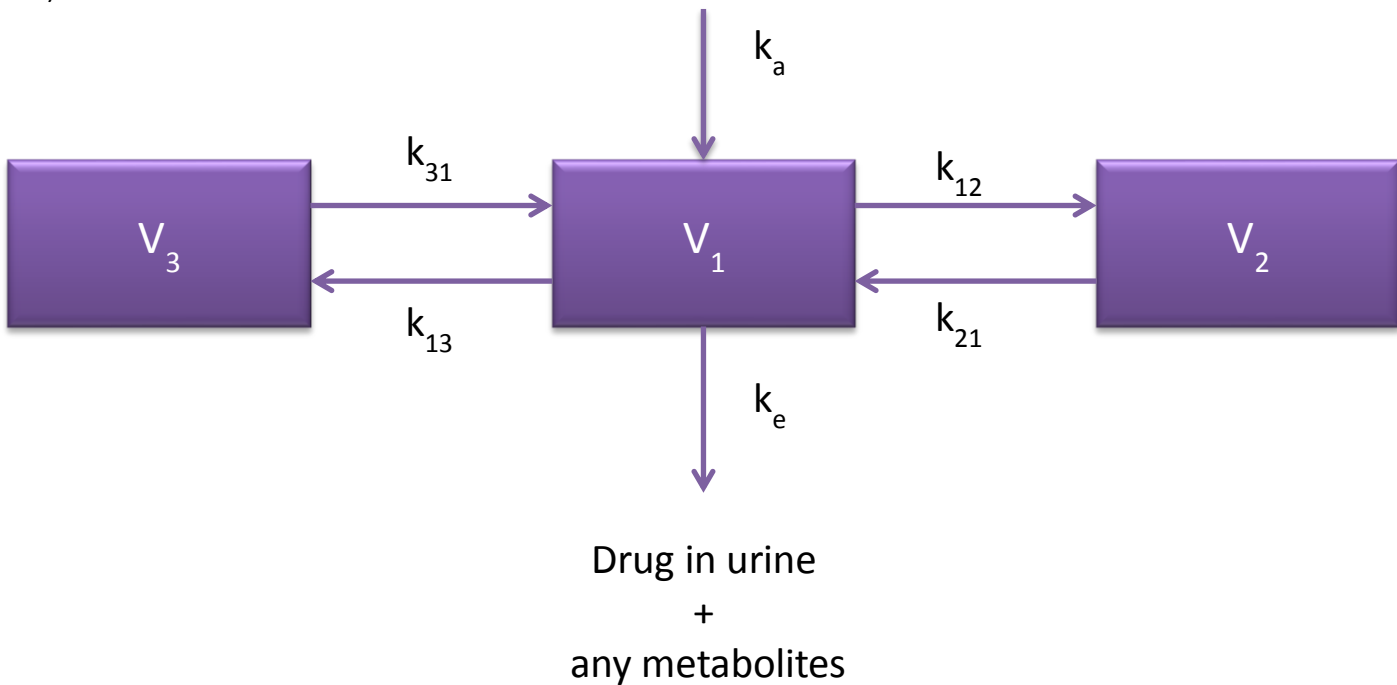


Figure B22: The three – compartmental mammillary pharmacokinetic model.

The central compartment has a volume V_1 , the tissue compartment volume V_2 and the deep tissue compartment volume V_3 .

The central compartment represents blood and all highly perfused tissues i.e. vital organs that are in rapid equilibrium with blood such as lungs, kidney and liver. Drug elimination is supposed to occur from the central compartment according to the first order transfer rate k_e since kidney and liver are the two most important clearing organs. The tissue compartment corresponds to poorly perfused tissues and the deep tissue compartment is very poor in tissue including sites such as bone and fat.

Taking into account the fact that concentrations of drugs cannot usually be measured at the effector site, the concentration of drug in plasma is typically assumed to represent the concentration in the target tissue. Hence, the equation characterizing the central compartment is the one to be linked with vinorelbine pharmacodynamics.

If T stands for the duration of infusion, D for dose and t_D for the time-point of the administration, the function that describes vinorelbine concentration in plasma following one infusion, is the following:

While the infusion takes place, namely while $t - t_D \leq T$:	
$C(t) = \frac{D}{T} \cdot \left[\frac{A}{a} (1 - e^{-a(t-t_D)}) + \frac{B}{b} (1 - e^{-b(t-t_D)}) + \frac{C}{c} (1 - e^{-c(t-t_D)}) \right]$	(B4)
After the infusion, namely while $T < t - t_D$:	
$C(t) = \frac{D}{T} \cdot \left[\frac{A}{a} (1 - e^{-aT}) \cdot e^{-a(t-t_D-T)} + \frac{B}{b} (1 - e^{-bT}) \cdot e^{-b(t-t_D-T)} + \frac{C}{c} (1 - e^{-cT}) \cdot e^{-c(t-t_D-T)} \right]$	(B5)

where,

$a_0 = k_e \cdot k_{21} \cdot k_{31}$	(B6)
$a_1 = k_e \cdot k_{31} + k_{21} \cdot k_{31} + k_{21} \cdot k_{13} + k_e \cdot k_{21} + k_{31} \cdot k_{12}$	(B7)
$a_2 = k_e + k_{12} + k_{13} + k_{21} + k_{31}$	(B8)
$p_1 = a_1 - a_2^2 / 3$	(B9)
$q = 2a_2^3 / 27 - a_1 \cdot a_2 / 3 + a_0$	(B10)
$r_1 = \sqrt[3]{-p_1 / 27}$	(B11)
$r_2 = 2r_1^{1/3}$	(B12)
$\phi = \arccos\left(-\frac{q}{2r_1}\right) / 3$	(B13)
$a = -(\cos(\phi) \cdot r_2 - a_2 / 3)$	(B14)
$b = -(\cos(\phi + 2\pi/3) \cdot r_2 - a_2 / 3)$	(B15)
$c = -(\cos(\phi + 4\pi/3) \cdot r_2 - a_2 / 3)$	(B16)
$A = \frac{1}{V_1} \cdot \frac{k_{21} - a}{a - b} \cdot \frac{k_{31} - a}{a - c}$	(B17)
$B = \frac{1}{V_1} \cdot \frac{k_{21} - b}{b - a} \cdot \frac{k_{31} - b}{b - c}$	(B18)
$C = \frac{1}{V_1} \cdot \frac{k_{21} - c}{c - b} \cdot \frac{k_{31} - c}{c - a}$	(B19)

Table B7: Parameters and variables involved in the three-compartmental pharmacokinetic model

	Mathematical Symbols	Code Symbols	Units	Description
Independent variable	t	t	days	time
Dependent variable	C	C	conc (mg/ml)	concentration of vinorelbine in blood
Parameter	-	n	-	number of infusions
Parameter	D	dose	mg	dose
Parameter	T	inf_dur	day	duration of infusion
Parameter	V_1	V1	volume (ml)	volume of central compartment
Parameter	k_{12}	k12	1/day	inter-compartmental rate constant
Parameter	k_{21}	k21	1/day	inter-compartmental rate constant
Parameter	k_{13}	k13	1/day	inter-compartmental rate constant
Parameter	k_{31}	k31	1/day	inter-compartmental rate constant
Parameter	k_e	ke	1/day	elimination rate constant
Parameter	-	dosage	mg/kg	dosage
Parameter	-	w	kg	patient's weight
Parameter	T_D	T1,...,Tn	days	administration time-points of anti-angiogenic treatment

Implementation and indicative results

In order to calculate vinorelbine concentration in plasma, the following m-file has been implemented:

VinorelbineConcentration.m: This function implements the three-compartmental pharmacokinetic model for vinorelbine administration via intravenous infusions. It calculates the concentration of vinorelbine at each time point taking into account all previous infusions.
Input: time - point t
Input parameters: number of infusions (n) , duration of infusions (inf_dur), volume of distribution (V_c), transfer constants ($k_{12}, k_{21}, k_{13}, k_{31}, k_e$), dosage ($dosage$), patient's weight (w), administration time points (T_i 's)
Output: the concentration of Vinorelbine at the given time point

Code Execution 17: The administration of 1.2 mg/kg in eight 10-minute intravenous infusions given once a week was simulated (Wong et al., 2006). The simulated treatment scheme is characterized by parameter values shown in Table B8.

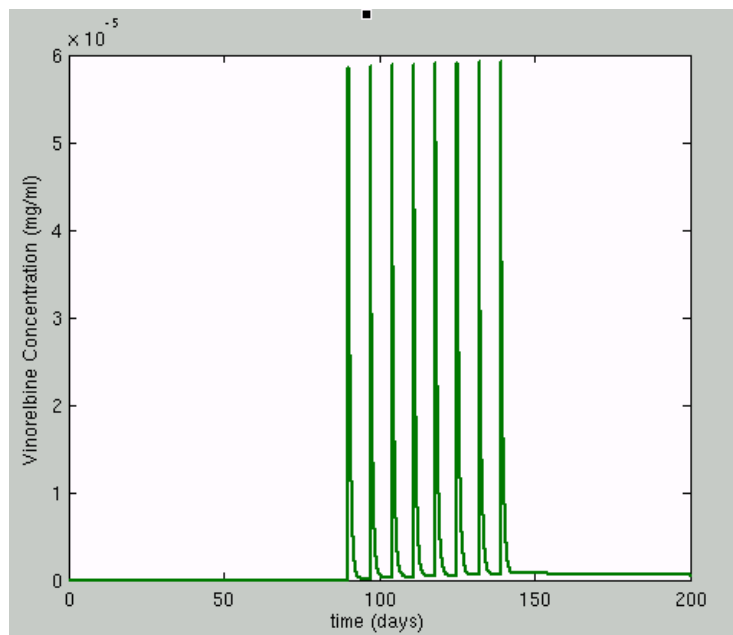


Figure B23: The time course of vinorelbine concentration in plasma. The details of the simulated scheme are shown in Table B8

Table B8: Parameters values used in order to derive the results of code execution 17.

Parameters	Code Execution 21	Units
k12	8.05	1/h
k21	0.36	1/h
k13	0.64	1/h
k31	0.00026	1/h
ke	1.3347	1/h
V1	24. 8	lt
dose	dosage*weight	mg
duration of infusion	1/6	h
dosage	1.2	mg/kg
weight	50	kg
number of infusions	8	-
administration time-points	T1=90, T2=97, T3=104, ..., T8=139	days

The therapy – related parameter values shown in the Table above were extracted from (Wong et al., 2006).

B3 Coupling of the dynamical system with the discrete state – discrete event cytokinetic model: The case of combination treatment

Since the scientific problem addressed consists not only of the simulation of bevacizumab mono-therapy but also of the simulation of combination treatment of vinorelbine administered with bevacizumab, the ultimate goal of our modelling approach, which is in process, is to couple the two modeling approaches outlined in the context of the current deliverable. To this end, a linking point between the vascular tumour growth dynamical system and the discrete entity discrete state tumour growth model, needs to be determined.

A first approach, would be to use the vascular tumour growth dynamical system in order to perturb a key-parameter involved in the discrete state-discrete event cytokinetic model. Among many candidate parameters, a satisfying choice would be P_{sleep} , i.e. the fraction of cells entering G0 phase following mitosis, a parameter that already takes into account tumor induced angiogenesis in an implicit way (e.g. through the modification of the probability of a newborn cell to re-enter the cell cycle in relation to its value in the largely necrotic layer of the tumor) (Stamatakis et al., 2010; Dionysiou et al., 2004; Stamatakis et al., 2002).

Indeed, tumour angiogenesis, i.e. the tumour-induced growth of new capillary blood vessels in the body from already existing vasculature, is a key process in tumour development, as new vessels provide the tumour mass with nutrients and oxygen which are clearly of vital importance for its survival and extra growth. Therefore, it can be assumed that poorly vascularised tumour areas are more likely to consist of tumour cells with higher probability to enter G0 phase following mitosis due to the lack of oxygen and nutrients, compared to sufficiently vascularised tumour areas.

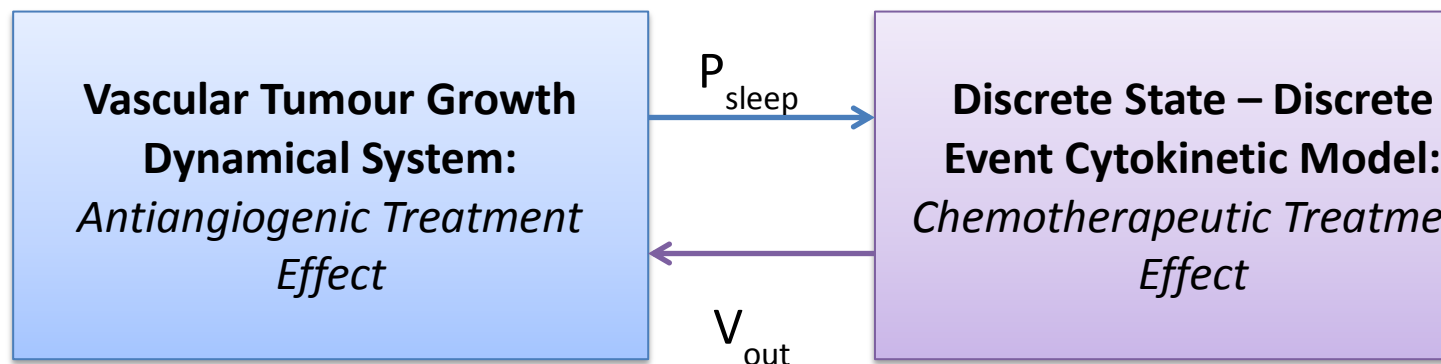


Figure B24: A diagram describing the coupling of the dynamical system simulating anti- angiogenic effect with the discrete state – discrete event cytokinetic model reflecting chemotherapeutic effect.

Figure B24 depicts the basic idea of the coupling of the two models for the case of combination treatment. Input V_{input} is created from the data provided by the clinical partners, such as DICOM imaging files which is fed into the system following a preprocessing procedure. At each time-step, the DS-DE cytokinetic model simulates the response of breast cancer clinical tumours to chemotherapeutic treatment consisting of the single agent vinorelbine. When the DS-DE cytokinetic model reaches a bevacizumab administration time-point, it provides as output the volume of the treated tumour V_{out} which then serves as an input to the vascular tumour growth model.

Subsequently, the dynamical system computes the tumour capacity K for the given volume (V_{out}) based on the details of the anti-angiogenic treatment scheme and it provides as output the adjusted probability of tumour cells entering the G0 phase following mitosis. Finally, P_{sleep} is fed to DS-DE cytokinetic model. This process can obviously be repeated as suggested by the administered treatment scheme.

A next step, besides thoroughly studying the behaviour and testing the performance of the overall hybrid system, would be a modification of the adopted approach, so as to also account for the synergistic effect of the parallel administration of anti-angiogenic and chemotherapeutic treatment. Indeed, it is of paramount importance that the system in its final form reflects the improved penetration of chemotherapeutic drug due to the normalization of the surviving tumour vasculature, a key-mechanism to the anti-neoplastic effect of bevacizumab.

B4 References

Barnes, J. A., Dix, D. J., Collins, B. W., Luft, C., Allen J. W., 2001. Expression of inducible Hsp70 enhances the proliferation of MCF-7 breast cancer cells and protects against the cytotoxic effects of hyperthermia. *Cell Stress & Chaperones* 6 (4), 316–325.

Bast, R.C., Kufe, D.W., Pollock, R.E., Weichelbaum, R.R., Holland, J.F., Frei E., Eds. 2000. *Cancer Medicine* 5th ed., B.C. Decker Inc, Hamilton, Ontario, Canada

Bertrand J, Mentré F. Mathematical Expressions of the Pharmacokinetic and Pharmacodynamic Models implemented in the MONOLIX software (<http://www.monolix.org>). 2008.

Cos, S., Recio, J., Sanchez-Barcelo, E. J., 1996. Modulation of the length of cell cycle time of MCF-7 human breast cancer cells by melatonin. *Life Sci.* 58, 811-816.

Descamps, S., Lebourhis, X., Delehedde, M., Boilly, B., Hondermarck H., 1998. Nerve Growth Factor Is Mitogenic for Cancerous but Not Normal Human Breast Epithelial Cells. *J. Biol. Chem.* 273 (27), 16659–16662.

Dewey, W.C., Ling, C.C., Meyn, R.E., 1995. Radiation-induced apoptosis: relevance to radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 33 (4), 781–796.

D.D.Dionysiou, G.S. Stamatakis, N.K.Uzunoglu, K.S.Nikita and A. Marioli , “A Four Dimensional In Vivo Model of Tumour Response to Radiotherapy: Parametric Validation Considering Radiosensitivity, Genetic Profile and Fractionation”, *J. theor. Biol.*, vol. 230, pp. 1-20, 2004.

Gauvin A, Pinguet F, Culine S, Astre C, Bresolle R Gomeni. Bayesian estimate of vinorelbine pharmacokinetic parameters in elderly patients with advanced metastatic cancer. *Clin Cancer Res.* 2000;6:2690–2695

Ginsberg, T., 1996. Modellierung und Simulation der Proliferationsregulation und Strahlentherapie normaler und maligner Gewebe, *Fortschritt-Berichte, VDI Verlag: Duesseldorf, Reihe 17: Biotechnik, Nr 140*, 103-107.

Gregory RK, Smith IE. Vinorelbine a clinical review. *Br J Cancer* 2000;82:1907±13.

P. Hahnfeldt, D. Panigrahy, J. Folkman and L. Hlatky, Tumour development under angiogenic signaling: A dynamical theory of tumour growth, treatment response and postvascular dormancy, *Cancer Res.*, **59** (1999), 4770-4775.

Highlights of prescribing information, Retrieved 1/2013 from
http://www.gene.com/download/pdf/avastin_prescribing.pdf

Lu Jian-Feng, Rene Bruno, Steve Eppler, William Novotny, Bert Lum, Jacques Gaudreault, Clinical pharmacokinetics of bevacizumab in patients with solid tumours, *Cancer Chemother Pharmacol* (2008) 62:779–786

Maseide, K. Rofstad, E.K. 2000. Mathematical modeling of chronic hypoxia in tumors considering potential doubling time and hypoxic cell lifetime. *Radiother. Oncol.* 54, 171-177.

Meyer, J.S., McDivitt, R., Stone, K., Prey, M., Bauer, W., 1984. Practical breast carcinoma cell kinetics: review and update. *Breast Cancer Res. Treat.* 4, 79-88.

McCormack PL, Keam SJ. *Drugs*. 2008;68(4):487–506.

Peer PG, van Djick JA, Hendricks JH, Holland R, Verbeek AL. Age-dependent growth of primary breast cancer. *Cancer*. 1993;713547-3551]

Poleszczuk J, Bodnar M, Foryś U., New approach to modelling of antiangiogenic treatment on the basis of Hahnfeldt et al. model, *Math Biosci Eng*. 2011 Apr;8(2):591-603. doi: 10.3934/mbe.2011.8.591.

Reynolds, L. P., Killilea, S. D. & Redmer, D. A. (1992) Angiogenesis in the female reproductive cycle. *FASEB J.*, **6**, 886–892.

Rezai K, Urien S, Isambert N, Roche H, Dieras V, Berille J, Bonnetterre J, Brain E, Lokiec F, Pharmacokinetic evaluation of the vinorelbine-lapatinib combination in the treatment of breast cancer patients, *Cancer Chemother Pharmacol*. 2011 Apr 26.

Salmon, S.E., Sartorelli, A.C., 2001. Cancer Chemotherapy. In: Katzung, B.G. (Eds.), *Basic & Clinical Pharmacology*. Lange Medical Books/McGraw-Hill: International Edition, pp. 923-1044.

Sterin, M., Cohen, J.S., Mardor, Y., Berman, E., Ringel, I., 2001. Levels of phospholipid metabolites in breast cancer cells treated with antimitotic drugs. *Cancer Res*. **61**, 7536-7543.

G.S.Stamatakis, E.A.Kolokotroni, D.D.Dionysiou, E.Ch.Georgiadi, C.Desmedt. An advanced discrete state - discrete event multiscale simulation model of the response of a solid tumor to chemotherapy: Mimicking a clinical study. *Journal of Theoretical Biology* 266, 124-139, 2010

G.S. Stamatakis, D.D. Dionysiou, E.I. Zacharaki, N.A. Mouravliansky, K.S.Nikita, N.K. and Uzunoglu , “In silico radiation oncology: combining novel simulation algorithms with current visualization techniques” , *IEEE Proceedings: Special Issue on Bioinformatics: Advances and Challenges* , vol. 90, no. 11 , pp. 1764-1777 , 2002.

Steel, G.G. 1997. Clonogenic cells and the concept of cell survival. In: Steel, G.G. (Ed.), *Basic Clinical Radiobiology*, second ed. Arnold: London, UK.

Sterin, M., Cohen, J.S., Mardor, Y., Berman, E., Ringel, I., 2001. Levels of phospholipid metabolites in breast cancer cells treated with antimitotic drugs. *Cancer Res*. **61**, 7536-7543.

Wong M, Balleine RL, Blair EY, et al: Predictors of vinorelbine pharmacokinetics and pharmacodynamics in patients with cancer. *J Clin Oncol* 24:2448-2455, 2006

Zacharaki, E.I., Stamatakis, G.S., Nikita, K.S., Uzunoglu, N.K., 2004. Simulating growth dynamics and radiation response of avascular tumour spheroid model validation in the case of an EMT6/Ro multicellular spheroid. *Comput. Methods Programs Biomed*. **76**, 193-206.

Chapter 4 [L]

THE ONCOSIMULATOR: THE ACUTE LYMPHOBLASTIC LEUKEMIA (ALL) BRANCH

L1 ALL: The Cellular and Supercellular Component

L1.1 The simulation model

L1.1.1 Introduction

The non spatial model developed for the cellular and supercellular simulation of the temporal evolution and response to therapy of non solid cancers, such as acute lymphoblastic leukemia (ALL) in the context of p-medicine, constitutes a modification of previously developed core simulation algorithms and codes of the *In Silico* Oncology Group (e.g. Stamatakis et al., 2010). However, this new non spatial approach can also be used for the simulation of both metabolically homogeneous solid tumours and tumours comprising regions of different metabolic activity (e.g. proliferating, necrotic), when no emphasis is put on the spatial features of a tumour's evolution.

The new model critical features are (a) omission of the simulation of the spatial evolution of the tumour and (b) consideration of more compartments for proliferating cells. In the classical spatial models (e.g. Stamatakis et al. 2010), (Stamatakis et al 2011)), the tumour region is treated as a grid (or “mesh”) of “Geometrical Cells” (GCs, the elementary volume of the grid). The morphological rules, which are introduced to govern the deletion or creation of new GCs, result in a shrinkage or expansion of the tumour conformal to its initial shape. The handling of the spatial features of a tumour evolution is a computationally demanding task, particularly so for large tumour sizes. By ignoring these spatial features in the non spatial algorithm, a more detailed modeling of the cytokinetic properties of the proliferating cells becomes feasible, through the consideration of more distinct proliferating cell compartments. More specifically, the cycling and dormant (if applicable) cancerous cells are distributed over a number of classes/compartments that equals the duration of the relevant cell phase (in hours). Each compartment corresponds to a time interval of one hour, as described analytically in the next section.

This handling presents the advantage of eliminating quantization errors related to the existence of only five proliferative cell sub-classes corresponding to the five phases of the cell cycle (G1, S, G2, M), as is the case in the classical spatial compartmental code (see the results section) . On the other hand, the exclusion of the spatial evolution of the tumour does not affect the temporal evolution of the various cancerous cell categories and the total cell population.

L1.2 Cytokinetic model

In the case of acute lymphoblastic leukemia (ALL) three types of cancer cell compartments can be distinguished: the bone marrow, the circulating blood and the extramedullary blast cell compartments. At a first approximation, the cytokinetic model adopted here, considers only the compartment of cancer origin, that is the bone marrow cell compartment. The inclusion of the circulating blood and the extramedullary cell compartments will be considered in future versions of the code.

The cytokinetic model (Figure L1.1) incorporates the biological mechanisms of cell cycling, quiescence (if applicable), recruitment, differentiation and death. Cancer sustenance is attributed to the presence of the cancer stem cells, which have the ability to preserve their own population, as well as give birth to cells that follow the path towards terminal differentiation. More specifically, five cancerous cell categories can be identified in the model:

- i. Stem cells: cells with unlimited proliferating potential,
- ii. LIMP cells: limited mitotic potential cells (LIMP) able to perform a limited number of divisions before terminal differentiation,
- iii. Differentiated cells: terminally differentiated cells
- iv. Apoptotic cells: cells that have died through apoptosis
- v. Necrotic cells: cells that have died through necrosis

Stem, LIMP and differentiated cells constitute three categories with distinct mitotic potential. From the mathematical standpoint each cell category defines an equivalence class. Each stem or LIMP cell can be either proliferating or dormant. Proliferation or dormancy creates another level of cell population partitioning. A finer partitioning of proliferating cells (stem and LIMP) into subclasses is introduced in the compartmental code by considering a number of compartments that equals the duration of the cell cycle. Each compartment corresponds to each hour of the total duration of the G1 (gap 1), S (synthesis), G2 (gap 2), M (mitosis) phases and therefore contains the biological cells residing in the given hour. Similarly dormant cells (stem and LIMP) are distributed in a number of compartments that correspond to each hour of the duration of the G0 phase. A further partitioning in the case of therapeutic intervention is treatment hitting i.e. a boolean variable denoting whether a biological cell has been hit by treatment.

A proliferating cancer cell (stem or LIMP) passes through the successive cell cycle phases. After the completion of mitosis, a fraction of newborn cells will enter the dormant phase, due to insufficient nutrient supply and oxygenation, whereas the rest will continue to cycle. Under conditions of lack of nutrients, dormant cells are assumed to survive only for a limited time length. After the expiration of this time, dormant cells die through necrosis, unless the local conditions of nutrient and oxygen supply have been reinstated, allowing the re-entering of the dormant cells into the active cell cycle. Any cell may die through spontaneous apoptosis. Differentiated cells may die through apoptosis or necrosis. Table L1.1 lists the model parameters and the corresponding biological mechanisms as described above.

When cancer is chemotherapeutically treated, a fraction of cancer cells are lethally hit by the drug or its metabolites. Lethally hit cycling cancer cells enter a rudimentary cell cycle that leads to apoptotic death via a specific phase dictated by the action mechanism of the chemotherapeutic agent used. Similarly, in

the case of cell cycle non specific drugs, lethally hit dormant (G0) cells enter the G0hit phase. Marking of a cell as hit by the drug is assumed to take place at the instant of drug administration although its actual time of death is dictated by the specific pharmacokinetics and pharmacodynamics of the drug. It is pointed out however that cell cycle phase specific drugs can be readily modeled by the cytokinetic model shown in Figure L1.1 by appropriately selecting the “Chemo” induced exit from the normal cell cycle for both cases of stem and LIMP cells.

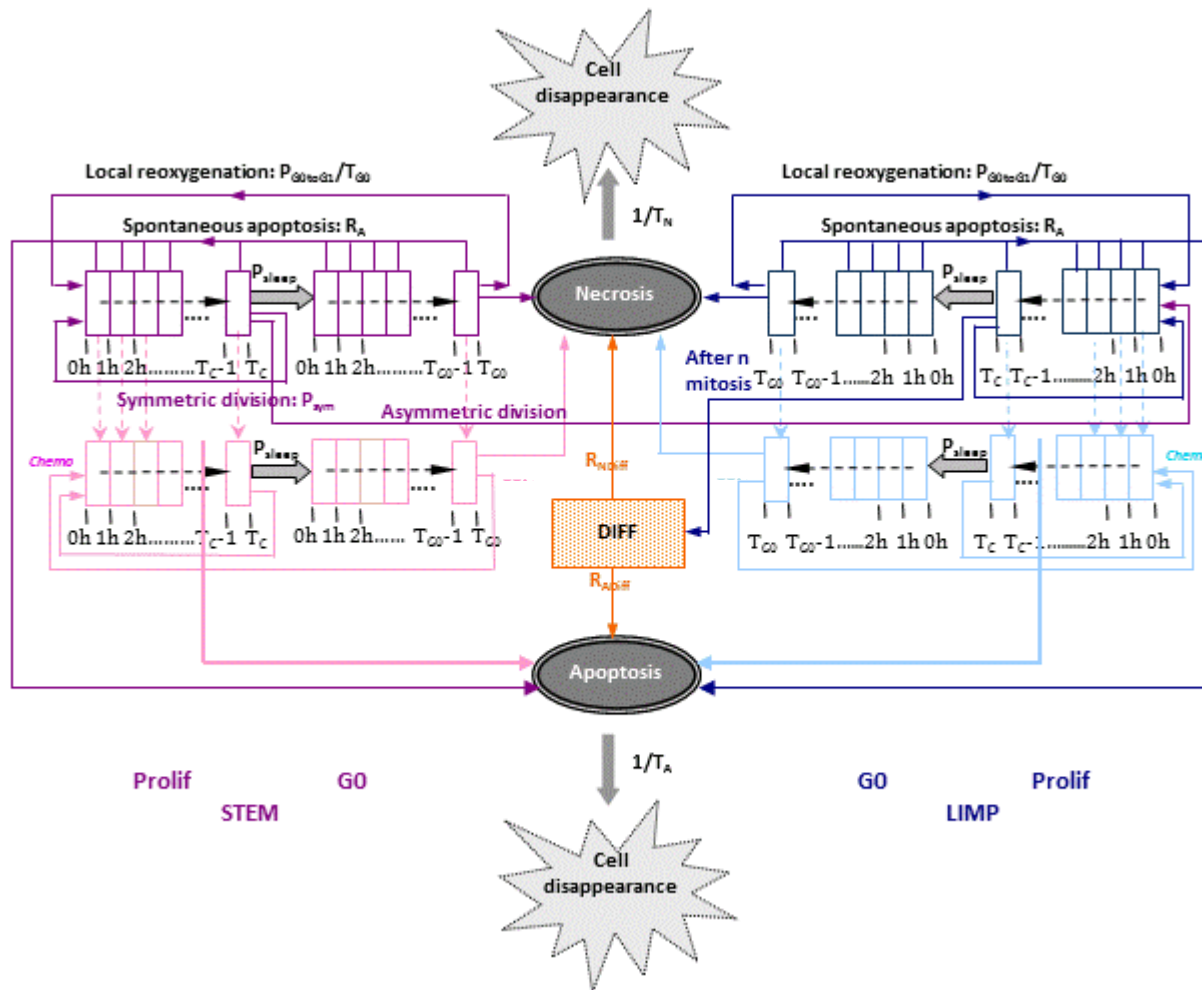


Figure L1.1: Generic cytokinetic model of bone marrow cell compartment (cell category / phase transition diagram) for cancer response to chemotherapy.

Abbreviations: STEM: stem cell. LIMP: Limited Mitotic Potential cancer cell (also called committed or restricted progenitor cell). DIFF: terminally DIFFerentiated cancer cell. Prolif: Proliferating cell. G0: dormant cell. Chemo: chemotherapeutic treatment. Hit: lethally hit cells by the drug

L1.3 Simulation outline

The modeling approach is discrete in time. The model incorporates individualized clinical data such as histopathologic (e.g., type of cancer, grade) and the genetic data (e.g., p53 status, if available). A simplified flowchart of the simulation algorithm pertaining to free growth and therapy is depicted in Figure L1.2 and described below.

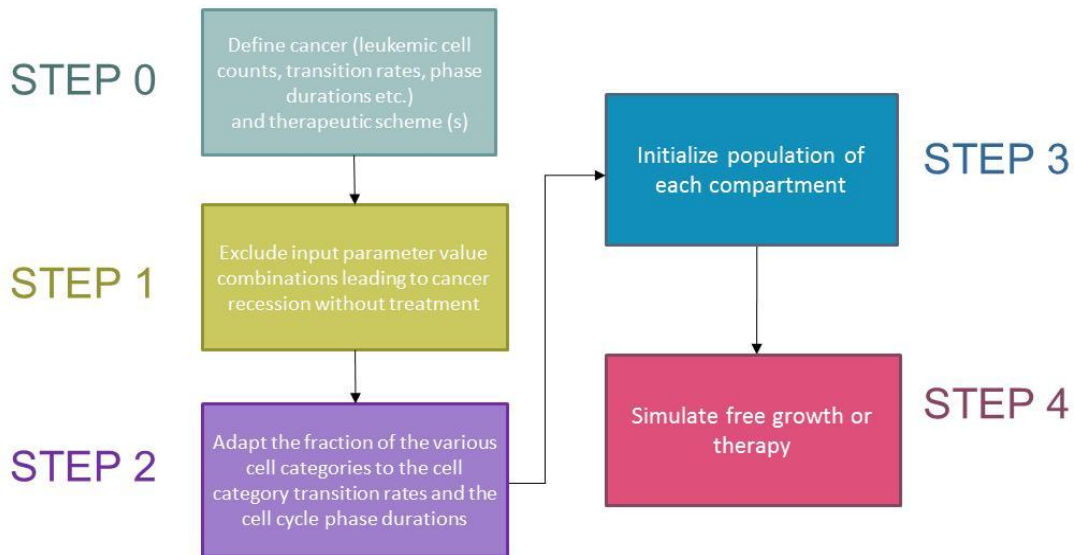


Figure L1.2 Simplified flowchart of the simulation procedure.

L1.3.1 STEP 0: Cancer definition

The first step is the definition of the clinical cancer considered based on the available patient-specific data. For non solid cancers, such as ALL, the model accepts as an input cell counts. The values assigned to the model input parameters (see paragraph L0) are defined. Histological data such as differentiation grade, as well as other cancer specific data, can be incorporated for further refinement of the values attributed to the model input parameters.

L1.3.2 STEP 1: Free growth check

Based on the cytokinetic model described previously, stem cells are responsible for sustaining the cancer and their behavior plays a determinant role for cancer free growth evolution. Depending on the values assigned to the model parameters that describe the life course of stem cells, it is possible to simulate cancer with variable degree of aggressiveness in terms of growth rate. Furthermore, there exist certain 'forbidden' value combinations of these parameters that lead to biologically irrelevant cancers, i.e. cancers that diminish over time of their own accord, unable to sustain growth.

A condition is applied to check whether the value combination of input parameters leads to a growing or

self-diminishing cancer. The condition has been derived (Kolokotroni et al. 2011) from an analytical treatment of model assumptions following the methodology of Bertuzzi et al. (1997).

$$(1 + P_{sym}) \left(1 - P_{sleep} + P_{sleep} \frac{P_{G0toG1}/T_{G0}}{R_A + 1/T_{G0}} \right) e^{-R_A T_C} \geq 1$$

In order to take into account any divergence between the simulating results and the above condition we assume that for values of the left side of the above inequality lower than 0.9 cancer free growth cannot be sustained, whereas for values above 1.1 free growth is ensured. In the middle value range the cancer free growth is checked based on simulation results during tumorigenesis process. The flowchart of the above procedure is depicted in Figure L1.3a.

L1.3.3 STEP 2: Tumour Initialization (Tumorigenesis)

In order to avoid an abnormal free growth behaviour at the beginning of the simulation, an automatic tumour initialization methodology has been developed, according to which the initial relative population (expressed as a fraction of the total cancer cell population) of each equivalence class and its equivalence subclasses is adjusted to the values attributed to the cell category and cell phase transition rates. The technique consists of the following processes. An initial number of proliferating stem cells e.g. 10000 and stem G0 cells e.g. 1000 are distributed throughout the cell cycle and dormant phase respectively. The distribution of these initial stem cells throughout the relevant phase is an exponential decay of the hour they reside in, with a decay constant equal to the apoptosis rate of stem and LIMP cells, R_A :

$$N_{S, Prolif}[m] = \frac{10000}{\sum_{i=0}^{T_C-1} e^{-R_A \cdot i}} e^{-R_A \cdot m}, \quad m = 0, 1, 2, \dots, T_C - 1$$

$$N_{S, G0}[n] = \frac{1000}{\sum_{i=0}^{T_{G0}-1} e^{-R_A \cdot i}} e^{-R_A \cdot n}, \quad n = 0, 1, 2, \dots, T_{G0} - 1$$

where m and n the hour of the duration of cell cycle and dormant phase that the proliferating $N_{S, Prolif}$ and dormant $N_{S, G0}$ stem cells reside in respectively. The sum quantity at the denominator ensures that the total initial stem proliferating and dormant cells are equal to 10000 and 1000 respectively.

The system is left to evolve according to the cytokinetic model of Figure L1.1 and produce the rest of the cell category populations. The code execution has to continue until equilibrium is reached and the various cell categories population percentages have been stabilized. The equilibrium condition applied is described by the following inequality:

$$\left| \frac{\sum_{\text{time step } i=n-N+1}^n \frac{\text{cell category rel. population at time step } i}{N}}{\sum_{\text{time step } j=(n-M)-N+1}^{n-M} \frac{\text{cell category rel. population at time step } j}{N}} - \frac{\sum_{\text{time step } j=(n-M)-N+1}^{n-M} \frac{\text{cell category rel. population at time step } j}{N}}{\sum_{\text{time step } j=(n-M)-N+1}^{n-M} \frac{\text{cell category rel. population at time step } j}{N}} \right| \leq \varepsilon \quad (\text{L1.1})$$

for l consecutive averages

In words the average of N consecutive values of the relative population of cells clustered in a given mitotic potential category is calculated every M time steps (hours). The variation of the average is also calculated. If the variation of the average is less than ε , where ε is a small positive real number, for l consequent average values equilibrium has been reached. The condition must be true for all cell category relative populations. The flowchart of the above procedure is depicted in Figure L1.3.b.

L1.3.4 STEP 3: Population initialization

The biological cells are distributed over the five cell categories (i.e. stem, LIMP, differentiated, apoptotic, necrotic), based on the fractions calculated during the previous step. The initial distribution of the proliferating and dormant cells throughout the relevant phase is assumed to be an exponential decay of the time they have spent since entering the relevant phase (time is discrete). More specifically, the number of stem or LIMP proliferating cells having already spent m hours since entering cell cycle is:

$$N_{Prolif}[m] = \frac{N_{Total} * F_{Prolif}}{\sum_{i=0}^{T_C-1} e^{-[(k+R_A)*i]}} e^{-[(k+R_A)*m]}, \quad m = 0, 1, 2, \dots, T_C - 1$$

where R_A the apoptosis rate of stem and LIMP cells, F_{Prolif} the fraction of proliferating cells, k is the growth rate of the cancer, N_{Total} the total number of cancerous proliferating cells of the relevant cell compartment (stem or LIMP). F_{Prolif} and k are calculated during tumorigenesis process. The sum quantity at the denominator ensures that the total initial proliferating cells are equal to $N_{Total} * F_{Prolif}$ for each stem or LIMP cell compartment. Likewise, the number of stem or LIMP G0 cells having already spent n hours since entering dormant phase is:

$$N_{G0}[n] = \frac{N_{Total} * F_{G0}}{\sum_{i=0}^{T_{G0}-1} e^{-[(k+R_A)*i]}} e^{-[(k+R_A)*n]}, \quad n = 0, 1, 2, \dots, T_{G0} - 1$$

where F_{G0} the fraction of stem dormant cells calculated during tumorigenesis process.

The flowchart of the above procedure is depicted in Figure L1.3.c

L1.3.5 STEP 4: Tumour evolution

Every hour the state of the tumour is updated according to the proposed and adopted cytokinetic model of Figure L1.1 as follows. The spontaneous apoptosis induced cell loss from the class of differentiated

cells and from each non treatment and treatment perturbed cell cycle phase and the G0 phase is calculated for each cell category, based on the spontaneous apoptotic rates assumed. The necrotic and apoptotic cells destined to disappear are computed based on the reciprocal of the duration of the relevant phase. The progression of stem and LIMP cells throughout and between the cell cycle and dormant phase occurs with the following order. The proliferating cells (stem or LIMP) residing in the last hour of cell cycle duration (mitosis phase, see paragraph L0) are duplicated and the number of cells to be transferred to the dormant phase or reenter the cell cycle are computed. In the case of stem cells symmetric transitions to stem subclasses and asymmetric transitions to LIMP subclasses will take place. The proliferating cells residing in each hour of the cell cycle are transferred to the next one starting from the penultimate compartment and proceeding successively to the first. The first compartment (first hour) is filled with the newborn cells destined to enter cell cycle. Subsequently, the dormant cells that die through necrosis or reenter cell cycle are computed and the relevant transitions are implemented. The progression of dormant cells throughout the phase takes place in a similar manner as before, starting from the penultimate compartment and proceeding successively to the first.

Summarizing, the following transitions between the various cell categories take place. For stem and LIMP cells: Progression through each cell cycle phase, G0→G1, G0→Necrosis or G0→Apoptosis. For differentiated cells: Differentiated→Necrosis, Differentiated→Apoptosis. For dead cells of any mitotic potential category: Apoptosis→Cell disappearance, Necrosis→Cell disappearance. The corresponding rates are parameters of the model.

As far as chemotherapy is concerned, at any time instant corresponding to drug administration, the numbers of proliferating and dormant cells belonging to each hour of the relevant phase and to each one of the stem or LIMP mitotic potential categories that are designated as hit by the drug are computed. The latter is achieved through the utilization of the cell kill rate (CKR) parameter that corresponds to the drug and dose (per m² of the patient surface) considered. In terms of the simulation model parameters the cell kill rate is the percentage of LIMP and stem cells hit by the chemotherapeutic agent after each drug administration. The above mentioned cell numbers are added to the corresponding cell numbers of the drug affected equivalence subclasses, designated as “phase”hit. The progression of proliferating and dormant cells throughout the cycling and dormant phase and their removal is dictated by the action mechanism of the chemotherapeutic agent.

The flowchart of the above procedure is depicted in Fig.L1.3d.

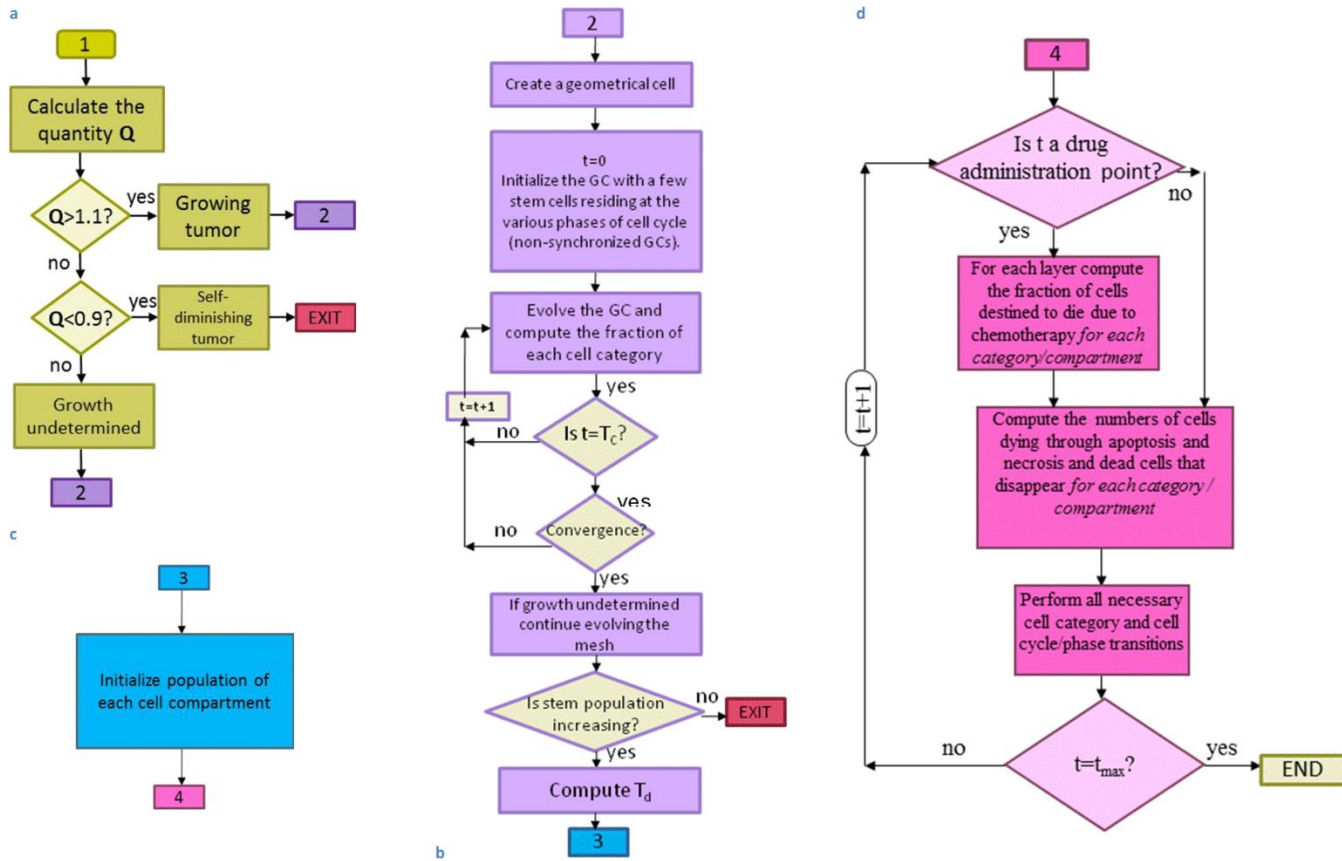


Figure L1.3 Flowcharts of the following simulation procedures (a) free growth check (b) tumorigenesis (c) population initialization (d) cancer evolution

L1.4 Code implementation

L1.4.1 Code input

Table L1.1 presents the simulation model input parameters and their range of values according to pertinent literature or based on logical assumptions supported by basic science or clinical experience in case of lack of literature data. In the following a description of the model parameters related to free growth and their adopted values is provided:

(i) Cell cycle duration

Based on literature, even though the cell cycle duration in cell lines in vitro is found to be around one day, in human cancers the cell cycle duration may vary considerably, even among cancers of the same histological type. In vivo studies have indicated that the cell cycle duration vary from 15 h to more than 100 h, with a mean value around 2-3 days (Steel et al., 2002). The duration of the mitosis phase is considered constant and equal to 1 h (Bast et al., 2000). The rest of the cell cycle phases durations are computed based on Salmon et al. (2001), after having taken into consideration the above assumption regarding the constant duration of mitosis. More specifically the following equations are used: $T_{G1} = T_S = 0.41(T_c - T_M)$, $T_{G2} = 0.18(T_c - T_M)$, $T_M = 1h$. The duration of the cell cycle phases is used to determine the time point that hit cells die through apoptosis depending on the action mechanism of the drug.

(ii) Duration of G0, duration of apoptosis and necrosis

According to literature dormant cells resting in G0 phase can survive under hypoxic conditions for 4-10 days (Maseide et al., 2000). Cancer apoptotic cells are generally considered to be rapidly phagocytosed in vivo (Dewey et al., 1995), contrary to the time-consuming process of necrosis products removal.

(iii) Cell category/phase transition rates/fractions

The values of these parameters have been selected based on both qualitative or semi-quantitative information as well as on the accumulated basic science and clinical experience. Systematic use of clinical trial data is expected to permit a quantitative refinement of the initial assumptions.

(v) Number of mitotic divisions that LIMP (committed progenitor) cells undergo before they become terminally differentiated

This parameter allows us to simulate tumours of different differentiation grades.

Table L1.1: MODEL PARAMETERS

Parameter symbol	Description	Value	References
CELL PHASE DURATIONS			
$T_c[\text{class}]$ $\text{class} \in \{\text{stem}, \text{LIMP}^{\dagger}\}$	Cell cycle duration in hours *	20 h – 96 h	Steel et al. 2002
$T_{G0}[\text{class}]$ $\text{class} \in \{\text{stem}, \text{LIMP}^{\dagger}\}$	G0 (dormant phase) duration* i.e. time interval before a dormant cell enters necrosis or re-enters cell cycle in hours	96 h – 240 h	Maseide et al., 2000
T_N	Time (in hours) needed for both necrosis to be completed and its lysis products to be removed	(0-several days)	Ginsberg T.,1996 (similar value range)
T_A	Time (in hours) needed for both apoptosis to be completed and its products to be removed ‡	(0 – 25 h)	Ginsberg T.,1996 (similar value range) Dewey et al., 1995
CELL CATEGORY/PHASE TRANSITION RATES			
R_A	Apoptosis rate of living stem and LIMP^{\dagger} tumour cells [fraction of cells dying through apoptosis per hour]	$0.0 - 1.0 \text{ h}^{-1}$	
R_{NDiff}	Necrosis rate of differentiated tumour cells [fraction of cell number per hour]	$0.0 - 1.0 \text{ h}^{-1}$	
R_{ADiff}	Apoptosis rate of differentiated tumour cells [fraction of cell number per hour]	$0.0 - 1.0 \text{ h}^{-1}$	
$P_{G0toG1}[\text{class}]$ $\text{class} \in \{\text{stem}, \text{LIMP}^{\dagger}\}$	Fraction of dormant (stem and LIMP^{\dagger}) cells that have just left dormant phase and re-enter cell cycle	0.0 – 1.0	
P_{sleep}	Fraction of cells entering the G0 phase following mitosis	0.0 – 1.0	
P_{sym}	Fraction of stem cells that perform symmetric ‡	0.0 – 1.0	
MISCELLANEOUS PARAMETERS			
N_{LIMP}	Number of mitoses performed by LIMP^{\dagger} cells before becoming differentiated	1 - 10	
Cell kill factor[class] $\text{class} \in \{\text{stem}, \text{LIMP}^{\dagger}\}$	Factor adapting cell killing probability to stem or LIMP^{\dagger} cells	0.0 – 1.0	
N_{cells}	Total leukemic cell count		
DRUG ADMINISTRATION PARAMETERS			
$T_{1st,combi}$	Time point after initialization (in days)		

$T_{2nd, combi}$	when the 1st combination drug administration takes place Time interval (in days) between 1st and 2nd combination drug administration	depends on chemotherapeutic schedule
$T_{3rd, combi}$	Time interval (in days) between 2nd and 3rd combination drug administration	depends on chemotherapeutic schedule
$T_{4th, combi}$	Time interval (in days) between 3rd and 4th combination drug administration	depends on chemotherapeutic schedule
$T_{5th, combi}$	Time interval (in days) between 4th and 5th combination drug administration (=0 if the chemotherapeutic scheme simulated comprises of less than 5 sessions)	depends on chemotherapeutic schedule
$T_{6th, combi}$	Time interval (in days) between 5th and 6th combination drug administration (=0 if the chemotherapeutic scheme simulated comprises of less than 6 sessions)	depends on chemotherapeutic schedule
$T_{1st, single}$	Time point after initialization (in days) when the 1st single drug administration takes place	depends on chemotherapeutic schedule
$T_{2nd, single}$	Time interval (in days) between 1st and 2nd single drug administration	depends on chemotherapeutic schedule
$T_{3rd, single}$	Time interval (in days) between 2nd and 3rd single drug administration	depends on chemotherapeutic schedule
$T_{4th, single}$	Time interval (in days) between 3rd and 4th single drug administration	depends on chemotherapeutic schedule
$T_{5th, single}$	Time interval (in days) between 4th and 5th single drug administration (=0 if the chemotherapeutic scheme simulated comprises of less than 5 sessions)	depends on chemotherapeutic schedule
$T_{6th, single}$	Time interval (in days) between 5th and 6th single drug administration (=0 if the chemotherapeutic scheme simulated comprises of less than 6 sessions)	depends on chemotherapeutic schedule
T_{stop}	Execution stop time (in days) after initialization	
CKR_A	Cell kill ratio for drug A	0.0 – 1.0
CKR_B	Cell kill ratio for drug B	0.0 – 1.0

[†] A LIMP tumour cell denotes a limited mitotic potential tumour cell (also referred to as LIMP or committed progenitor tumour cell). It leads to a terminally differentiated tumour cell.

* Phase durations can be defined separately for the stem and the LIMP tumour cell category.

L1.4.2 Code output

The code output is the following:

1. tumor_evolution.dat: A dat file containing the various total cell categories populations that constitute the tumour for each time step. The cell categories classes reported are stem proliferating, stem G0, LIMP proliferating, LIMP G0, differentiated, apoptotic, necrotic, stem (proliferating + G0), LIMP (proliferating + G0), proliferating (stem +LIMP) ,G0 (stem+LIMP), living, dead, and total.
2. tumor_dynamics.dat: A dat file containing the free growth dynamic characteristics of the simulated tumour i.e. the blast cell counts doubling time and the relative population of the various cell categories. The cell categories classes reported are stem proliferating, stem G0, LIMP proliferating having undergone 1 mitosis, LIMP proliferating having undergone 2 mitosis,..., LIMP proliferating having undergone N_{LIMP} mitosis, LIMP G0 having undergone 1 mitoses, LIMP G0 having undergone 2 mitoses,..., LIMP G0 having undergone N_{LIMP} mitoses differentiated, apoptotic and necrotic. Parameter N_{LIMP} denotes the number of mitoses performed by LIMP cells before becoming terminally differentiated.
3. oncosim-parameters.log: A log file containing the values assigned to the model input parameters.

L1.5 Indicative results

A virtual clinical case of acute lymphoblastic leukemia is considered, with an initial bone marrow blast cell count of 512.000.000. The values attributed to the various input model parameters are presented in Table L1.2. Figures L1.4 and L1.5 present the simulation outcome of the above described non spatial code and the spatial code described in Stamatakis et al. (2010). The two predictions are compared. Two kinds of executions have been performed simulating the tumour free growth and its response to chemotherapy, respectively. In the case of therapy, the drug is assumed to be administrated orally (bolus administration) according to a hypothetical fractionation scheme. More specifically, the chemotherapeutic agent is assumed to be administrated on the 5th, 12th, 26th and 33rd day of the simulated period. Figure L1.4 and Figure L1.5 show the time course of various cell populations for free growth and therapy respectively. In the case of therapy, the stem and LIMP populations include both surviving cells and cells affected by the drug (and therefore destined to die) but not yet dead. Agreement of the predictions produced by both models is observed. Furthermore, the oscillating pattern present in the simulation results of the spatial code is absent from the ones of the non spatial code.

Table L1.2: MODEL INPUT PARAMETER VALUES USED

Parameter	Value
Cell cycle duration of stem and LIMP cells	90 h
Duration of dormant phase	96 h
Duration of necrosis phase	120 h
Duration of apoptosis phase	6 h
LIMP stages number	7
Apoptosis rate of stem and LIMP cells	0.001 h ⁻¹
Apoptosis rate of differentiated cells	0.001 h ⁻¹
Necrosis rate of differentiated cells	0.01 h ⁻¹
Fraction of dormant cells re-entering cell cycle	0.01
Fraction of newborn cells entering dormant phase after mitosis	0.159
Symmetric division fraction	0.4
Factor adapting cell kill ratio to stem and LIMP cells	1
Cell counts	512.000.000
Factor adapting cell kill ratio to proliferating layer	1
Factor adapting cell kill ratio to necrotic layer	1
Cell kill ratio	0.4

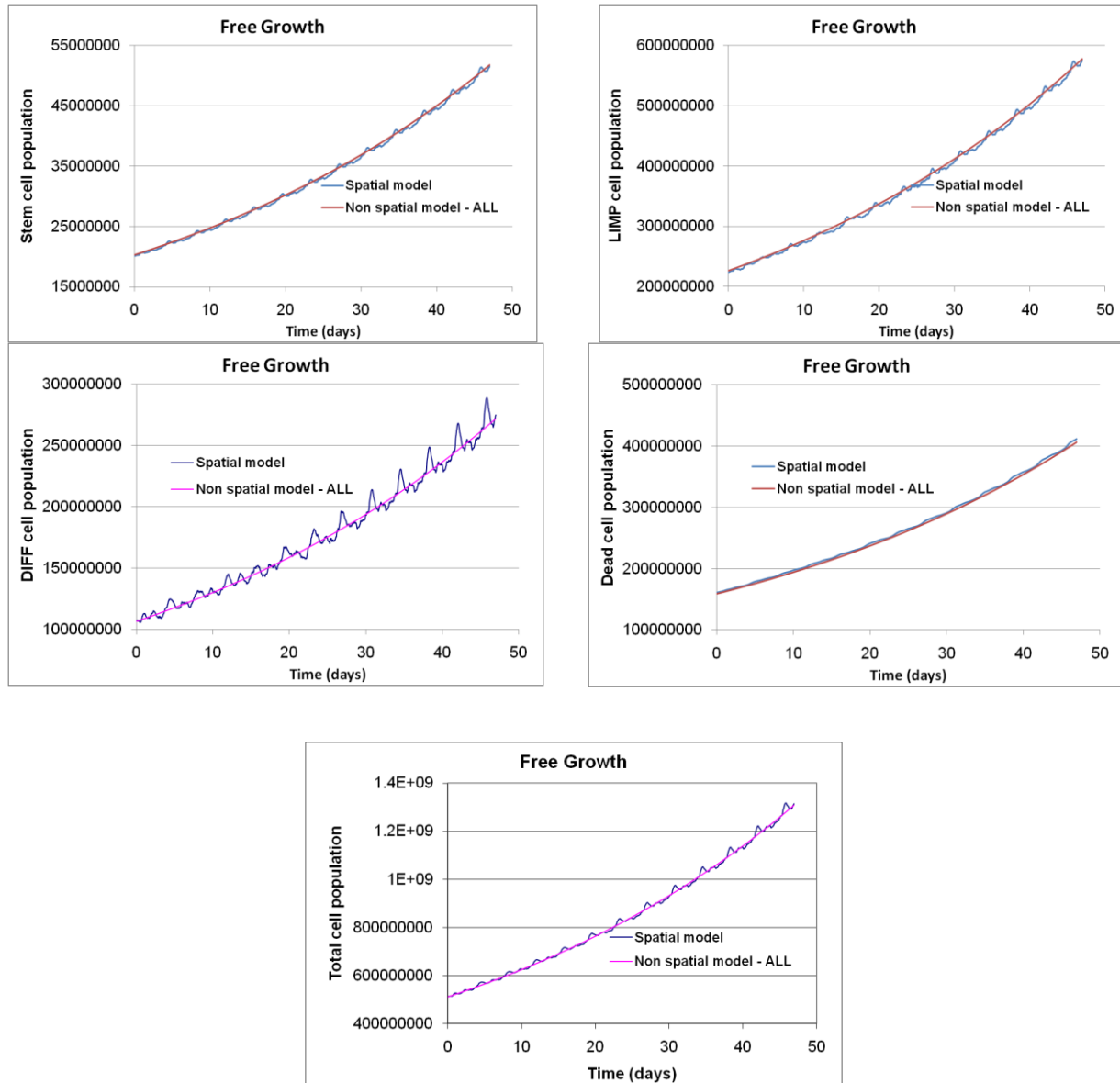


Figure L1.4: Comparison of the non spatial model (ALL modeling) and the spatial model of Stamatakis et al. (2010) for free cancer growth. The time course of various cell categories is depicted for each model. The values assigned to model parameters are reported in Table .

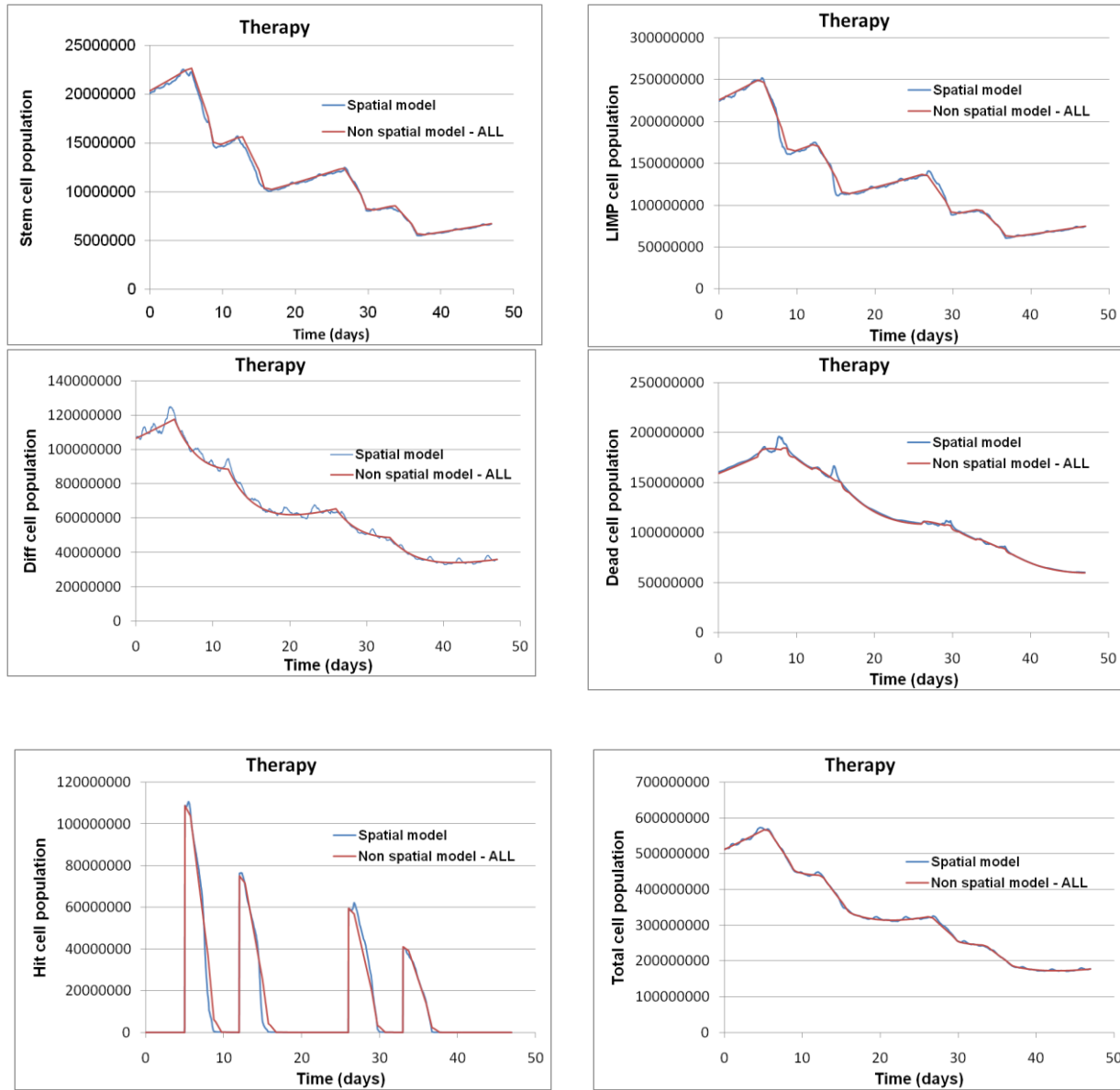


Figure L1.5 : Comparison of the non spatial model (ALL modeling) and the spatial model of Stamatakis et al. (2010) for response to chemotherapy. The time course of various cell categories is depicted for each model. The values assigned to model parameters are reported in Table .

L1.6 REFERENCES

- Bast RC, Kufe DW, Pollock RE, Weichelbaum RR, Holland JF, Frei E, Cancer Medicine 5th ed. (2000) Decker B.C. Inc, Hamilton, Ontario, Canada.
- Bertuzzi, A., Gandolfi, A., Sinisgalli, C., 1997. Steel's Potential Doubling Time and Its Estimation in Cell Populations Affected by Nonuniform Cell Loss. *Mathematical Biosciences* 143, 61-89
- Dewey WC, Ling CC, Meyn RE. 'Radiation-induced apoptosis: relevance to radiotherapy'. *Int J Radiat Oncol Biol Phys* (1995), 33(4):781–796.
- Ginsberg T. 'Modellierung und Simulation der Proliferationsregulation und Strahlentherapie normaler und maligner Gewebe'. *Fortschritt-Berichte*, VDI Verlag: Duesseldorf, Reihe 17: Biotechnik (1996), 140:103–107.
- Kolokotroni E.A., Dionysiou D.D., Uzunoglu N.K., Stamatakis G.S., "Studying the growth kinetics of untreated clinical tumors by using an advanced discrete simulation model", *Mathematical and Computer Modelling*, in press 2011
- Maseide K, Rofstad EK. 'Mathematical modeling of chronic hypoxia in tumors considering potential doubling time and hypoxic cell lifetime'. *Radiother Oncol* (2000), 54:171–177.
- Salmon SE and Sartorelli AC. 'Cancer Chemotherapy'. In: Katzung BG (Eds.), *Basic & Clinical Pharmacology* Lange Medical Books/McGraw-Hill: International Edition, 2001. pp. 923-1044.
- Stamatakis G.S., Kolokotroni E.A., Dionysiou D.D., Georgiadi E.C., Desmedt C., An advanced discrete state-discrete event multiscale simulation model of the response of a solid tumor to chemotherapy: Mimicking a clinical study. *J Theor Biol.* 266(1) (2010) 124–139.
- Stamatakis G.S., Georgiadi E.Ch., Graf N, Kolokotroni E.A., and Dionysiou D.D., "Exploiting Clinical Trial Data Drastically Narrows the Window of Possible Solutions to the Problem of Clinical Adaptation of a Multiscale Cancer Model", *PLOS ONE* 6(3), e17594 2011
- Steel G.G. (Editor), *Basic Clinical Radiobiology*, third ed., Arnold, London, 2002.

L1.7 APPENDIX: Code structure

The non spatial code consists of the following classes:

Class `virtual_tumor`

Simulates Chemotherapy or Free growth

Attributes

Attribute	Description
<code>plegma mesh</code>	Geometrical mesh

Constructor

Constructor	Description	Parameters
<code>virtual_tumor</code> (<code>int no_comp</code>)	Passes the number of metabolic compartments, in the case of solid cancers, to the constructor of class <code>plegma</code> .	<code>no_comp</code> Number of metabolic compartments, (one for non solid cancer, one or more for solid cancer)

Functions

Function	Description	Parameters	Return value
<code>void therapy()</code>	Simulates response to therapy, for intervals between sessions simulate growth of tumour	None	None
<code>void grow()</code>	Simulates free growth of cancer	None	None

Class `cell_bucket`

Sets a metabolic compartment (one for non solid cancer, one or more for solid cancer)

Attributes

Attribute	Description
<code>double stem_prolif</code> [<code>MAX_Tc</code>]	Number of proliferating/cycling stem cells having spent 0h, 1h, 2h,..., T_c since entering cell cycle. T_c denotes the cell cycle duration of stem cells.

int stem_G0[MAX_TG0]	Number of dormant stem cells having spent 0h, 1h, 2h,..., T_{G0} since entering dormant, G0, phase. T_{G0} denotes the duration of dormant phase of stem cells.
double stem_hit_prolif[N_DRUGS][MAX_Tc]	Number of proliferating/cycling stem cells hit by each involved drug having spent 0h, 1h, 2h,..., T_c since entering cell cycle.
double stem_hit_G0[N_DRUGS][MAX_TG0]	Number of dormant stem cells hit by each involved drug having spent 0h, 1h, 2h,..., T_{G0} since entering dormant, G0, phase.
double LIMP_prolif[max_LIMP_stages_number][MAX_Tc]	Number of proliferating/cycling LIMP cells having spent 0h, 1h, 2h,..., T_c since entering cell cycle. T_c denotes the cell cycle duration of LIMP cells.
double LIMP_G0[max_LIMP_stages_number][MAX_TG0]	Number of dormant LIMP cells having spent 0h, 1h, 2h,..., T_{G0} since entering dormant, G0, phase. T_{G0} denotes the duration of dormant phase of LIMP cells.
double LIMP_hit_prolif[N_DRUGS][MAX_Tc]	Number of proliferating/cycling LIMP cells hit by each involved drug having spent 0h, 1h, 2h,..., T_c since entering cell cycle.
double LIMP_hit_G0[N_DRUGS][MAX_TG0]	Number of LIMP cells hit by each involved drug having spent 0h, 1h, 2h,..., T_{G0} since entering dormant, G0, phase.
double a_pop	Number of apoptotic cells
double n_pop	Number of necrotic cells
double diff_pop	Number of terminally differentiated cells
int tumor_layer	It defines the cancer layer that the compartment belongs =0 necrotic layer for solid cancer/ =1 proliferating layer for solid and non solid cancer

Functions

Function	Description	Parameters	Return value
void tumor_init(int layer)	Initialization of population of compartment	layer Metabolic layer that the compartment belongs (necrotic or proliferating)	None
double total_population()	Calculation of total cell population of compartment	None	Total cell population

void advance_cycle()	Update of the state of the compartment according to the cytokinetic model	None	None
void chemo_action (float drugA_cell_kill, float drugB_cell_kill);	Simulation of drug administration. Computes number of cells that will die due to therapy and updates the populations.	drugA_cell_kill Cell kill ratio of drug A drugB_cell_kill Cell kill ratio of drug B	None

Class plegma

Sets the Geometrical Mesh (GM)

Attributes

Attribute	Description
<code>cell_bucket *c</code>	Compartment
<code>int no_comp</code>	Number of compartments

Constructor

Constructor	Description	Parameters
<code>plegma(int no_layers)</code>	Creates the mesh of compartments as an array of objects of class <code>cell_bucket</code> .	<code>no_layers</code> number of metabolic layers

Destructor

Destructor	Description	Parameters
<code>~plegma()</code>	Destroys the mesh of compartments by deleting the array of objects of class <code>cell_bucket</code>	None

Functions

Function	Description	Parameters	Return value
<code>void get_number_cells(double *cells_a)</code>	Returns the amount of each kind of cells	<code>cells_a</code> array where results will be written. We assume below each cell categories in this order: [0] total cells [1] stem proliferating [2] stem G0 [3] stem hit proliferating [4] stem hit G0 [5] LIMP	None

		proliferating [6] LIMP G0 [7] LIMP hit proliferating [8] LIMP hit G0 [9] differentiated [10] apoptotic [11] necrotic	
void free_growth_condition()	Checks whether the value combination of input parameters leads to a growing or self-diminishing cancer based on an analytical condition derived according to “Bertuzzi, A., Gandolfi, A., Sinisgalli, C., 1997. Steel's Potential Doubling Time and Its Estimation in Cell Populations Affected by Nonuniform Cell Loss. Mathematical Biosciences 143, 61-89”.	None	None
void tumorigenesis(int layer)	Calculates the fraction of each cell category that the cancerous system tends to establish at equilibrium for specific code parameters. These values are fed as an input during the tumour initialization process.	layer Metabolic area considered (proliferative or necrotic)	None

L2 ALL: Cell cycle modelling at the sub-cellular scale

L2.1 Introduction

The cell cycle duration (t_{cc}) is one of the most important parameters of the p-medicine Oncosimulator with obvious implications on treatment personalization. Although an extensive literature review might result in the extraction of a t_{cc} value proper for the tumor type and some patient characteristics, a detailed model simulating the biochemical dynamics of the cell cycle progression at the sub-cellular scale could allow a considerably more accurate prediction of t_{cc} based on the patient's specific genomic and/or proteomic status. Moreover, the utilization of such a model could allow the testing of numerous hypotheses concerning the cell cycle of tumor cells and its regulation by anti-cancer therapy. For example, glucocorticoids which play a central role in the treatment of Acute Lymphoblastic Leukemia (ALL) are known to cause G1 cell cycle arrest; however, the exact drug effects leading to this outcome have not yet been fully understood (Renner et al., 2003).

Searching the pertinent literature shows that numerous models for the subcellular dynamics of the cell cycle have been proposed. Representative examples include (Conradie et al., 2010; Haberichter et al., 2007; Iwamoto et al., 2011; Novák and Tyson, 2004; Pfeuty, 2012; Swat et al., 2004; Yao et al., 2008). Models take into account either the whole cell cycle machinery or specific phases and the transitions between them. However, the majority of these models are structured and calibrated in a way so as to fit the general case of the mammalian cell cycle. They do not normally refer to any specific tissues or cell categories. Moreover, although many of these modeling studies present simulations of cell cycle deregulations (gene mutation/deletion/overexpression) commonly found in cancer cells, to our knowledge, no modeling study has been published dealing with the simulation of cell cycle sub-cellular dynamics of leukemic cells. Therefore, the objectives of the present section are firstly to explore the cell cycle kinetics and the related sub-cellular biochemical dynamics of leukemic cells, by searching the literature; secondly to collect the existing mathematical models of the biochemical regulation of the cell cycle at the molecular scale from the literature and model repositories and finally to take the first steps in order to appropriately modify one of these models so as to simulate the ALL cell cycle dynamics reported in literature.

In the following sub-sections of the chapter the literature reported findings for the ALL cell cycle, both at the cellular and subcellular scale, as well as the proposed model are presented. It is noted that at this stage, only the case of precursor-B ALL (pre-B) has been taken into account since this ALL subtype is the most common among leukemias (Reiter et al., 1994). Another reason is the comparatively large collection of studies dealing with this subtype at present, in contrast with other ALL subtypes.

L2.2 Precursor-B ALL cell cycle

L2.2.1 Precursor-B ALL cell cycle kinetics

Primary Pre-B ALL cells show a t_{cc} value of 112.5 (± 46.8) hours compared with the 65.5 (± 46.8) hours for normal cells of the same type (Hirt et al., 1992). Interestingly the diversity on this value has been correlated with the acuteness of the symptoms of the disease (Hirt et al., 1997b). What it is believed to characterize the acuteness of the disease is the resistance of ALL cells to differentiation stimuli

(Greaves, 1986) and the very low rate of programmed cell death (apoptosis) execution, where apoptosis reduces the leukemic cell mass by 4% per day in contrast with an average 10-11% expansion per day caused by cell birth (Hirt et al., 1997b).

As far as the distribution of pre-B ALL cells in cell cycle phases is concerned, it has been shown that the majority of cells reside in the G1 phase (more than 80%) and only a minor proportion of cells can be found in S (~7%) and G2 phases (Hirt et al., 2009; Schmitz et al., 2006, 2005). Moreover the percentage of quiescent cells (found in true G0) is really low (~3%) (Hirt et al., 1997a). Although this accumulation of cells in the G1 phase, being the cell cycle phase where a decision for differentiation or proliferation is made (Pardee, 1989), ALL cells are differentiation resistant. This phenomenon has been explained by showing that almost every ALL cell in the G1 phase resides in the late-G1 sub-phase of the cell cycle (Hirt et al., 2009; Leibundgut et al., 1999; Schmitz et al., 2006, 2005) at or beyond the restriction point (Zetterberg et al., 1995). After this point cells have already decided not to differentiate in a way that will be presented in more detail in the following sections.

In accordance with experimental findings it is believed that in many cancers a subpopulation of cells commonly showing very slow rates of replication have the properties of cancer stem cells. However, in the case of pre-B ALL such a hypothesis is highly controversial, at least as far as oncogenic (in the case of leukemia, leukemogenic) properties of a distinct population are concerned. More specifically it has been recently shown that cells from almost every maturation stage could lead to leukemogenesis in xenografts (Rehe et al., 2013). Moreover, there is evidence that a maturation hierarchy, known to exist in the case of Acute Myeloid Leukemia (AML), does not exist in pre-B ALL (McClellan and Majeti, 2013). Nevertheless, and irrespectively from the leukemogenic properties, subpopulations with very different cell cycle kinetics have been found in bone marrow samples from patients with pre-B ALL (Hirt et al., 2011). More specifically, based on the identification of ALL cells by the levels of the CD19 marker expression two distinct populations have been found: the CD19⁻ population with frequency of appearance of ~6% and almost no or minimal proliferation activity (BrdU-LI % incorporation <1%) and the CD19⁺ population, consisting the ~93% of the whole population with significantly higher proliferation activity (~7%). The latter was further sub-clustered into CD19^{low} and CD19^{high} cells, with almost identical frequency (45.2% and 47.9% respectively) but with significantly different proliferation activity (2.9% and 8.9% BrdU-LI incorporation respectively). In the analysis presented in the following subsections and for the steps taken in order to formulate the proposed model, it is assumed that all the information extracted from the literature refers to CD19⁺ cells, since this is the dominant population in pre-B ALL samples.

L2.2.1 Precursor-B ALL cell cycle at sub-cellular scale

As already mentioned the majority of pre-B ALL cells reside in late-G1 phase of the cell cycle. This has been shown by measuring the levels of phosphorylation of the retinoblastoma protein, which has been found to be almost exclusively hyper-phosphorylated in pre-B ALL extracts (Hirt et al., 2009; Leibundgut et al., 1999; Schmitz et al., 2006, 2005). However this phosphorylation state for the retinoblastoma protein is known to be found only when cells have started passing the restriction point and initiate the G1/S transition. The broadly accepted model for the biochemical dynamics of the restriction point in mammalian cells is depicted in the following figure (taken from (Takahashi-Yanaga and Sasaguri, 2008)).

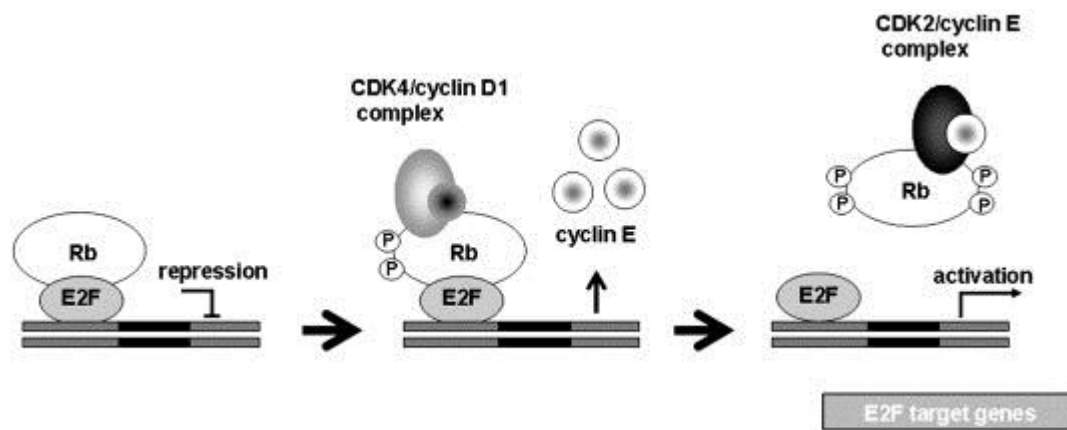


Figure L2.1 A simplified depiction of the concept of the Restriction Point in G1-Phase of the cell cycle (taken from (Takahashi-Yanaga and Sasaguri, 2008))

Briefly, the progression through the restriction point (R-point) is controlled by the sequential phosphorylation of retinoblastoma protein (pRb) or its family member proteins p107 and p130 by at least two different cyclin dependent kinase (Cdk) complexes, Cyclin D/Cdk4,6 and cyclin E/cdk2 leading to its inactivation. The inactivation of pRb protein allows the activation of E2F transcription factors which are crucial for the initiation of DNA replication. One of the known targets of E2F is Cyclin A.

Although the hyper-phosphorylated status of the retinoblastoma protein dictates an irreversible and rapid transition to S phase, this step is not taken in the expected rate by the pre-B ALL cells (only ~7% of cells are found in S phase). Therefore, a deregulation in the cell cycle machinery, and more specifically in the late-G1 phase and its progress towards G1/S transition should be caused in pre-B ALL due to their cancerous transformation. However, it should be mentioned that the G1/S transition, as far as the expression/activation of CyclinA is concerned, has been found to be undisturbed (Hirt et al., 2009). It should be pointed that the levels of Cyclin A have been found highly correlated with the percentage of pre-B ALL cells in the S phase (Hirt et al., 2009). Therefore, a deregulation that directly refers to cyclinA is excluded and something else happening earlier in the execution of cell cycle should be altered.

(Schmitz et al, 2006) have shown that the sequential phosphorylation of pRb known to happen during the execution of the restriction point may be deregulated in pre-B ALL. It has been found that the redundancy in phosphorylation of pRb by CDK2 and CDK4 is limited. In more detail, it has been shown that CDK4 phosphorylates CDK2 preferred phosphorylation sites of the retinoblastoma protein. Moreover, in the same study it is mentioned that the pRb, in pre-B ALL cells, even hyper-phosphorylated, could partially maintain its nuclear tethering ability and continue inhibiting E2F transcription factors. This finding is also mentioned in (Hirt et al., 2009).

L2.3 Models for the mammalian cell cycle sub-cellular dynamics

As mentioned earlier, numerous models aiming at simulating the biochemical dynamics that leads to the regulation of the cell cycle are available in literature. These models follow diverse modeling approaches spanning from fully quantitative to fully qualitative modeling and focus differently on several aspects of the cell cycle machinery. For the present task we focus on certain well-known

models, frequently cited in many studies appearing in Table L2.1. However, in the next developmental steps of the ALL cell cycle model, an even broader review of the models available in literature will be undertaken.

In order to choose the generic model that would serve as the basis for the development of the proposed specific ALL model, concrete criteria have been formulated. They are included in Table L2.1 along with the level at which each existing and retrieved model satisfies them.

Table L2.1 Existing generic cell cycle models that might serve as the basis of the specific ALL model under development

Model Name (based on first author's name in the corresponding publication)	Publication	Biomodels database entry (http://www.ebi.ac.uk/biomodels-main/)	Whole cell cycle modelling	Detailed restriction point modelling	Detailed addressing of each pRb phosphorylation step separately	Cyclin A modelling
Conradie et al.	(Conradie et al., 2010)	BIOMD0000000265	YES	YES	NO	YES
Yao et al.	(Yao et al., 2008)	BIOMD0000000318	NO	YES	NO	NO
Swat et al.	(Swat et al., 2004)	BIOMD0000000228	NO	YES	NO	NO
Iwamoto et al.	(Iwamoto et al., 2011)	No entry	YES	YES	YES	YES
Haberichter et al.	(Haberichter et al., 2007)	BIOMD0000000109	NO	YES	YES	YES

L2.4 Pre-B ALL cell cycle sub-cellular dynamics proposed model

L2.4.1 The basic model selected

Based on the characteristics of the Haberichter et al. model and more specifically on

1. The detail by which the G1 phase progression and the G1/S transition are modelled (for example, no pRb phosphorylation steps are summarized where one of the major deregulations of ALL cell cycle is believed to happen (Schmitz et al., 2006))
2. The focus that the model shows on the restriction point
3. The modelling of the activation of cyclin A, since this has been shown to be correlated with the passage of pre-B ALL cells into the S-phase (Hirt et al., 2009)
4. The availability of the model in curated SBML file by the access free Biomodels database,

the Haberichter et al. model has been selected in order to serve as the basis for the development of the proposed specific model of the sub-cellular dynamics of the pre-ALL cell cycle.

Although a full description of the model is available in the related publication and in the Biomodels Database entry (<http://www.ebi.ac.uk/biomodels-main/BIOMD0000000109>), a brief presentation of the model is provided in Figure L2.2 and Tables L2.2 L2.3 and L2.4 where the reaction diagram and the respective molecular entities and biochemical reactions that are modeled are shown.

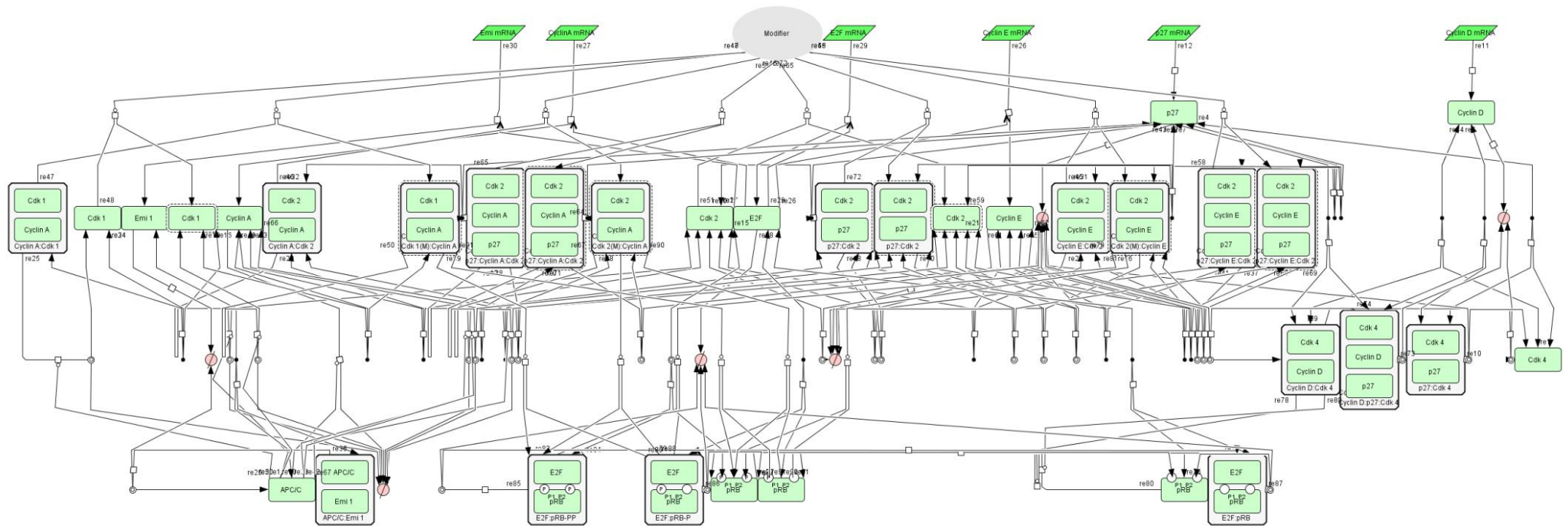


Figure L2.2 The reaction diagram of Haberichter's model (taken from: (Haberichter et al., 2007))

Table L2.2 Proteins and their complexes that are modeled in the Haberichter et al. model
(taken from: (Haberichter et al., 2007))

Proteins and complexes in the model	Meaning
CyclinD	Cyclin D
Cdk4[00]	Cdk 4
Cdk4[10]	p27:Cdk 4
Cdk4[01]	Cyclin D:Cdk 4
Cdk4[11]	p27:Cyclin D:Cdk 4
p27	p27
CyclinE	Cyclin E
CyclinA	Cyclin A
Cdk2[000]	Cdk 2
Cdk2[100]	p27:Cdk2
Cdk2[010]	Cdk2(M)
Cdk2[110]	p27:Cdk2(M)
Cdk2[001]	Cyclin E:Cdk2
Cdk2[101]	p27:Cyclin E:Cdk2
Cdk2[011]	Cyclin E:Cdk2(M)
Cdk2[111]	p27:Cyclin E:Cdk2(M)
Cdk2[002]	Cyclin A:Cdk2
Cdk2[102]	p27:Cyclin A:Cdk2
Cdk2[012]	Cyclin A:Cdk2(M)
Cdk2[112]	p27:Cyclin A:Cdk2(M)
Cdk1[00]	Cdk1
Cdk1[10]	Cdk1(M)
Cdk1[01]	Cyclin A:Cdk1
Cdk1[11]	Cyclin A:Cdk1(M)
pRb[00]	pRb
pRb[10]	pRb-P
pRb[20]	pRb-PP
pRb[01]	E2F:pRb
pRb[11]	E2F:pRb-P
pRb[21]	E2F:pRb-PP
E2F	E2F
Emi1	Emi1
APC/C	APC/C
APC/C:Emi1	APC/C:Emi1

Table L2.3 Enzyme-substrate complexes that are modeled in the Haberichter et al. model
(taken from: (Haberichter et al., 2007))

Enzyme-substrate complexes in the model	Enzyme	Substrate	Product
Cdk4[01]_pRb[00]_pRb[10]_Int	Cyclin D:Cdk 4	pRb	pRb-P
Cdk4[01]_pRb[01]_pRb[11]_Int	Cyclin D:Cdk 4	E2F:pRb	E2F:pRb-P
Cdk2[011]_pRb[10]_pRb[20]_Int	Cyclin E:Cdk2(M)	pRb-P	pRb-PP
Cdk2[011]_pRb[11]_pRb[21]_Int	Cyclin E:Cdk2(M)	E2F:pRb-P	E2F:pRb-PP
Cdk2[012]_pRb[10]_pRb[20]_Int	Cyclin A:Cdk2(M)	pRb-P	pRb-PP

Cdk2[012]_pRb[11]_pRb[21]_Int	Cyclin A:Cdk2(M)	E2F:pRb-P	E2F:pRb-PP
Cdk1[11]_pRb[10]_pRb[20]_Int	Cyclin A:Cdk1(M)	pRb-P	pRb-PP
Cdk1[11]_pRb[11]_pRb[21]_Int	Cyclin A:Cdk1(M)	E2F:pRb-P	E2F:pRb-PP
APC/C_CyclinA_Int	APC/C	Cyclin A	---
APC/C_Cdk2[000]_Cdk2[002]_Int	APC/C	Cyclin A:Cdk2	Cdk 2
APC/C_Cdk2[100]_Cdk2[102]_Int	APC/C	p27:Cyclin A:Cdk2	p27:Cdk2
APC/C_Cdk2[010]_Cdk2[012]_Int	APC/C	Cyclin A:Cdk2(M)	Cdk2(M)
APC/C_Cdk2[110]_Cdk2[112]_Int	APC/C	p27:Cyclin A:Cdk2(M)	p27:Cdk2(M)
APC/C_Cdk1[00]_Cdk1[01]_Int	APC/C	Cyclin A:Cdk1	Cdk1
APC/C_Cdk1[10]_Cdk1[11]_Int	APC/C	Cyclin A:Cdk1(M)	Cdk1(M)

Table L2.4 Biochemical reactions that are modeled in the Haberichter et al. model (taken from: (Haberichter et al., 2007)). The textlined reactions are referred to in the text by their textlining colour .

Reactions	Rate constants [in order of appearance; forward before backward reactions]
<--> CyclinD	ks_CyclinD, kd_CyclinD
Cdk4[01] --> Cdk4[00]	kd_CyclinD
Cdk4[11] --> Cdk4[10]	kd_CyclinD
<--> p27	ks_p27, kd_p27
Cdk4[10] --> Cdk4[00]	kd_p27
Cdk4[11] --> Cdk4[01]	kd_p27
Cdk2[100] --> Cdk2[000]	kd_p27
Cdk2[101] --> Cdk2[001]	kd_p27
Cdk2[102] --> Cdk2[002]	kd_p27
Cdk2[110] --> Cdk2[010]	kd_p27
Cdk2[111] --> Cdk2[011]	kd1_p27
Cdk2[112] --> Cdk2[012]	kd_p27
<--> CyclinE	ks_CyclinE, kd_CyclinE
Cdk2[001] --> Cdk2[000]	kd_CyclinE
Cdk2[101] --> Cdk2[100]	kd_CyclinE
Cdk2[011] --> Cdk2[010]	kd_CyclinE
Cdk2[111] --> Cdk2[110]	kd_CyclinE
<--> CyclinA	ks_CyclinA, kd_CyclinA
Cdk1[01] --> Cdk1[00]	kd_CyclinA
Cdk1[11] --> Cdk1[10]	kd_CyclinA
Cdk2[002] --> Cdk2[000]	kd_CyclinA
Cdk2[102] --> Cdk2[100]	kd_CyclinA
Cdk2[012] --> Cdk2[010]	kd_CyclinA
Cdk2[112] --> Cdk2[110]	kd_CyclinA
<--> E2F	ks_E2F, kd0_E2F
pRb[01] --> pRb[00]	kd_E2F
pRb[11] --> pRb[10]	kd_E2F
pRb[21] --> pRb[20]	kd_E2F
<--> Emi	ks_Emi, kd_Emi

APC/C:Emi --> APC/C	kd_Emi
Cdk4[00] + CyclinD <--> Cdk4[01]	kb_CyclinD__Cdk4, ku_CyclinD__Cdk4
Cdk4[10] + CyclinD <--> Cdk4[11]	kb_CyclinD__Cdk4, ku_CyclinD__Cdk4
Cdk4[00] + p27 <--> Cdk4[10]	kb_p27__Cdk4, ku_p27__Cdk4
Cdk4[01] + p27 <--> Cdk4[11]	kb_p27__Cdk4, ku_p27__Cdk4
Cdk2[000] + p27 <--> Cdk2[100]	kb_p27__Cdk2, ku_p27__Cdk2
Cdk2[001] + p27 <--> Cdk2[101]	kb_p27__Cdk2, ku_p27__Cdk2
Cdk2[002] + p27 <--> Cdk2[102]	kb_p27__Cdk2, ku_p27__Cdk2
Cdk2[010] + p27 <--> Cdk2[110]	kb_p27__Cdk2, ku_p27__Cdk2
Cdk2[011] + p27 <--> Cdk2[111]	kb_p27__Cdk2, ku_p27__Cdk2
Cdk2[012] + p27 <--> Cdk2[112]	kb_p27__Cdk2, ku_p27__Cdk2
Cdk2[000] --> Cdk2[010]	k_act
Cdk2[100] --> Cdk2[110]	k_act
Cdk2[001] --> Cdk2[011]	k_act
Cdk2[101] --> Cdk2[111]	k_act
Cdk2[002] --> Cdk2[012]	k_act
Cdk2[102] --> Cdk2[112]	k_act
Cdk2[000] + CyclinE <--> Cdk2[001]	kb_CyclinE__Cdk2, ku_CyclinE__Cdk2
Cdk2[100] + CyclinE <--> Cdk2[101]	kb_CyclinE__Cdk2, ku_CyclinE__Cdk2
Cdk2[010] + CyclinE <--> Cdk2[011]	kb_CyclinE__Cdk2, ku_CyclinE__Cdk2
Cdk2[110] + CyclinE <--> Cdk2[111]	kb_CyclinE__Cdk2, ku_CyclinE__Cdk2
Cdk2[000] + CyclinA <--> Cdk2[002]	kb_CyclinA__Cdk2, ku_CyclinA__Cdk2
Cdk2[100] + CyclinA <--> Cdk2[102]	kb_CyclinA__Cdk2, ku_CyclinA__Cdk2
Cdk2[010] + CyclinA <--> Cdk2[012]	kb_CyclinA__Cdk2, ku_CyclinA__Cdk2
Cdk2[110] + CyclinA <--> Cdk2[112]	kb_CyclinA__Cdk2, ku_CyclinA__Cdk2
Cdk1[00] --> Cdk1[10]	k_act
Cdk1[01] --> Cdk1[11]	k_act
Cdk1[00] + CyclinA <--> Cdk1[01]	kb_CyclinA__Cdk1, ku_CyclinA__Cdk1
Cdk1[10] + CyclinA <--> Cdk1[11]	kb_CyclinA__Cdk1, ku_CyclinA__Cdk1
pRb[00] + Cdk4[01] <--> Cdk4[01]_pRb[00]_pRb[10]_Int --> pRb[10] + Cdk4[01]	kb_D4__pRb, ku_D4__pRb, kup_D4__pRb
pRb[01] + Cdk4[01] <--> Cdk4[01]_pRb[01]_pRb[11]_Int --> pRb[11] + Cdk4[01]	kb_D4__pRb, ku_D4__pRb, kup_D4__pRb
pRb[10] + Cdk2[011] <--> Cdk2[011]_pRb[10]_pRb[20]_Int --> pRb[20] + Cdk2[011]	kb_E2__pRb, ku_E2__pRb, kup_E2__pRb
pRb[11] + Cdk2[011] <--> Cdk2[011]_pRb[11]_pRb[21]_Int --> pRb[21] + Cdk2[011]	kb_E2__pRb, ku_E2__pRb, kup_E2__pRb
pRb[10] + Cdk2[012] <--> Cdk2[012]_pRb[10]_pRb[20]_Int --> pRb[20] +	kb_A2__pRb, ku_A2__pRb, kup_A2__pRb

Cdk2[012]	
pRb[11] + Cdk2[012] <--> Cdk2[012]_pRb[11]_pRb[21]_Int --> pRb[21] + Cdk2[012]	kb_A2__pRb, ku_A2__pRb, kup_A2__pRb
pRb[10] + Cdk1[11] <--> Cdk1[11]_pRb[10]_pRb[20]_Int --> pRb[20] + Cdk1[11]	kb_A1__pRb, ku_A1__pRb, kup_A1__pRb
pRb[11] + Cdk1[11] <--> Cdk1[11]_pRb[11]_pRb[21]_Int --> pRb[21] + Cdk1[11]	kb_A1__pRb, ku_A1__pRb, kup_A1__pRb
pRb[20] --> pRb[00]	kt_pRb__Dephos
pRb[21] --> pRb[01]	kt_pRb__Dephos
pRb[00] + E2F <--> pRb[01]	kb_E2F__pRb, ku_E2F__pRb
pRb[10] + E2F <--> pRb[11]	kb_E2F__pRb, ku_E2F__pRb
pRb[21] --> pRb[20] + E2F	ku_E2F__pRb
APC/C + Emi <--> APC/C:Emi	kb_Emi__APCC, ku_Emi__APCC
CyclinA + APC/C <--> APC/C_CyclinA_Int --> APC/C	kb_APCC__CyclinA, ku_APCC__CyclinA, kud_APCC__CyclinA
Cdk2[002] + APC/C <--> APC/C_Cdk2[000]_Cdk2[002]_Int --> Cdk2[000] + APC/C	kb_APCC__CyclinA, ku_APCC__CyclinA, kud_APCC__CyclinA
Cdk2[102] + APC/C <--> APC/C_Cdk2[100]_Cdk2[102]_Int --> Cdk2[100] + APC/C	kb_APCC__CyclinA, ku_APCC__CyclinA, kud_APCC__CyclinA
Cdk2[012] + APC/C <--> APC/C_Cdk2[010]_Cdk2[012]_Int --> Cdk2[010] + APC/C	kb_APCC__CyclinA, ku_APCC__CyclinA, kud_APCC__CyclinA
Cdk2[112] + APC/C <--> APC/C_Cdk2[110]_Cdk2[112]_Int --> Cdk2[110] + APC/C	kb_APCC__CyclinA, ku_APCC__CyclinA, kud_APCC__CyclinA
Cdk1[01] + APC/C <--> APC/C_Cdk1[00]_Cdk1[01]_Int --> Cdk1[00] + APC/C	kb_APCC__CyclinA, ku_APCC__CyclinA, kud_APCC__CyclinA
Cdk1[11] + APC/C <--> APC/C_Cdk1[10]_Cdk1[11]_Int --> Cdk1[10] + APC/C	kb_APCC__CyclinA, ku_APCC__CyclinA, kud_APCC__CyclinA

The initial amount values for the biomolecular entities, the kinetic rate laws and the respective kinetic parameter values are available at (<http://www.ebi.ac.uk/biomodels-main/BIOMD0000000109>).

In the following Figures, the simulation results of the original model using the SBML compatible software COPASI 4.8 (Hoops et al., 2006) are given. A comparison with the results given by the online simulator of biomodels database has been made in order to validate the results returned by COPASI). The molecular entities chosen to be presented are those dictated by literature that the literature dictates (section L2.2) in order to show deregulation in pre-B ALL case either in their levels or in the time needed to reach adequate levels to execute their function (hyper-phosphorylated pRb, hypo-phosphorylated pRb, cyclinA). The simulated time for the following results has been chosen to

be 10 hours (600 minutes) since this has been the time during which the cells pass into the S phase according to the experiments used for the calibration of the original model (Haberichter et al., 2007).

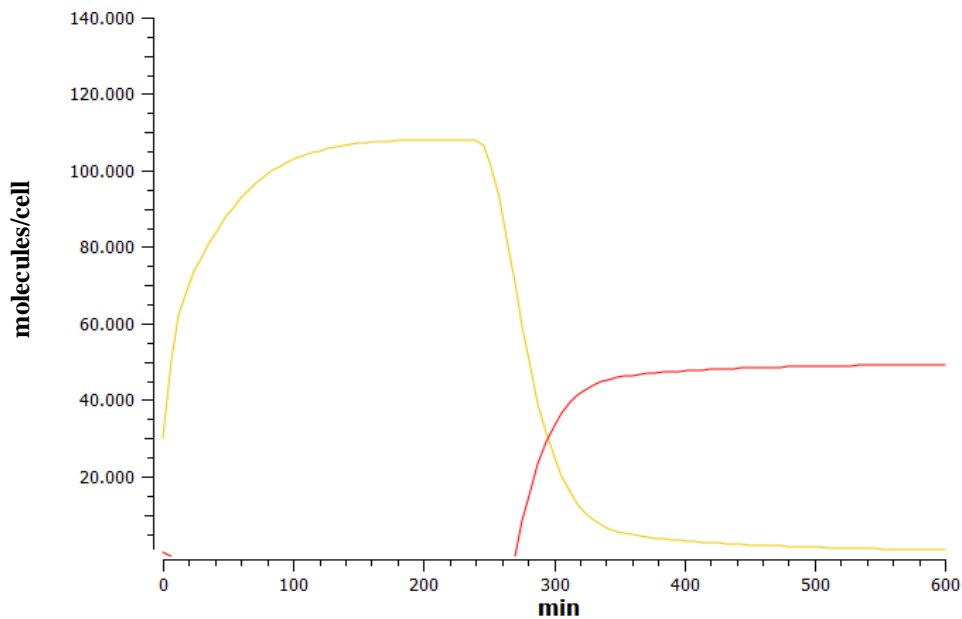


Figure L2.3 Simulation results of the original model (Haberichter et al. model) for 600 minutes (10 hours) showing the levels (molecules/cell) of hypo-phosphorylated pRb (yellow line) and the hyper-phosphorylated pRb (red line). *It is noted that in the following graphs of Chapter 4 the vertical axis generally represents a sample of the cell and the modelules under consideration for reference/calibration purposes.*

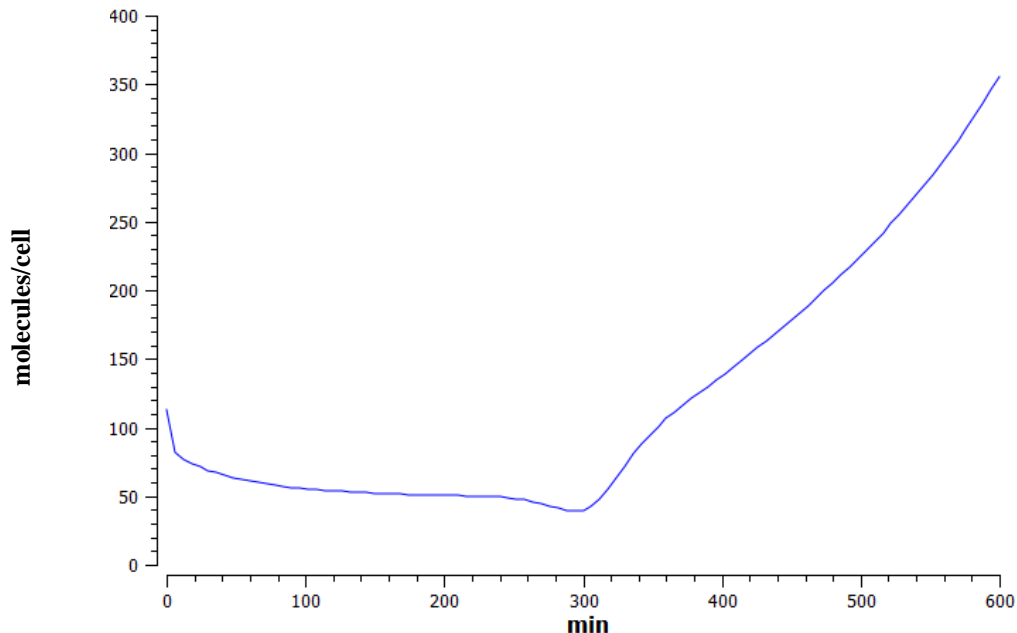


Figure L2.4 Simulation results of the original model (the Haberichter et al. model) for 600 minutes (10 hours) showing the levels (molecules/cell) of CyclinA (blue line).

L2.4.2 Modifications leading to the proposed model

Based on the observed deregulations in pre-B ALL cell cycle sub-cellular biochemical dynamics presented in section L2.1 and the findings regarding the direct regulation of cyclinA by pRB, specific

modifications have been made in the Haberichter et al. model referring to both the model structure and some kinetic rate laws and reaction parameters. These modifications are summarized as follows. *It is noted that a more detailed description of each intervention is provided in the subsequent sub-sections*

1. The retinoblastoma protein can reach the hyper-phosphorylated state not only by the sequential phosphorylation by CycD/Cdk4 and CycE/Cdk2 complexes but also directly by CycD/Cdk4 alone. The retinoblastoma proteins that are hyper-phosphorylated by only the contribution of Cyc/Cdk4 (pRb-pp_by_cdk4) are distinct from the total hyper-phosphorylated retinoblastoma protein since it is hypothesized that only these show the aberrant functionality.
2. The rate by which the E2F transcription factors are liberated from the inhibition of pRb-pp_by_cdk4 protein is very low compared to the relative rate that the E2F TFs are freed from the normal hyper-phosphorylated pRb.
3. The pRb-pp_by_cdk4 can form complexes with free E2F transcription factors

L2.4.2.1 Hyper-phosphorylation of pRB also by Cdk4/CyclinD

Based on the findings of (Schmitz et al., 2006) presented in detail in section L2.2, we hypothesize that the Cdk4/CyclinD complexes can contribute not only to the hypo-phosphorylation of retinoblastoma protein but also to its hyper-phosphorylation. This phenomenon could explain the hyper-phosphorylated status of pRb from the very first hours of the G1 phase execution since CyclinD and Cdk4 have been adequately expressed since the initiation of the G1 phase in contrast with Cdk2 and CyclinE. Therefore, in order to model this phenomenon the following reactions have been introduced into the model (the newly defined molecular entities are in bold):

- New_React1: pRb[10] + Cdk4[01] -> **Cdk4[01]pRb[10]pRb[20]Int**
- New_React2: pRb[11] + Cdk4[01] -> **Cdk4[01]pRb[11]pRb[21]Int**
- New_React3: **Cdk4[01]pRb[10]pRb[20]Int** -> pRb[20]cdk4 + Cdk4[01]
- New_React4: **Cdk4[01]pRb[10]pRb[20]Int** -> pRb[10] + Cdk4[01]
- New_React5: **Cdk4[01]pRb[11]pRb[21]Int** -> pRb[21]cdk4 + Cdk4[01]
- New_React6: **Cdk4[01]pRb[11]pRb[21]Int** -> pRb[11] + Cdk4[01]
- New_React7: **pRb[21]cdk4** -> **pRb[20]cdk4** (leading to the degradation of E2F and not to its liberation)
- New_React8: **pRb[20]cdk4** -> **pRb[00]**
- New_React9: **pRb[21]cdk4** -> **pRb[01]**

As far as the kinetic rate laws and the related parameters for these reactions are concerned, it has been assumed that they have exactly the same form and values with the respective reactions for:

- The hypo-phosphorylation step (textlined in light red in table L4, reactions rxnY085, rxnY086, rxnY087 in SBML version of the model)
 - for New_React1 and New_React2, mass action rate law with parameter value: $3.1e - 05$ (1/molecules * min)
 - for New_React3 and New_React5, mass action rate law with parameter value: 1.69466 (1/min)
 - For New_React4 and New_React6, mass action rate law with parameter value: 0.1 (1/min)

- The degradation of E2F bound to normal pRb[21] (textlined in light blue in table L4, reaction rxnY033)
 - for New_React7, mass action rate law with parameter value: 0.006465 (1/min)
- The desphosphorylation of pRb[20] and pRB[21] (textlined in light orange in table L4, reactions rxnY109 and rxnY110)
 - For New_React8 and New_React9, mass action rate law with parameter value: 0.023194 (1/min)

L2.4.2.2 Inhibition of E2F by hyper-phosphorylated pRb

As the findings of (Hirt et al., 2009; Schmitz et al., 2006) already presented in section L2.2 dictates, there is evidence that even the hyper-phosphorylated form of pRb could contribute to the inhibition of E2F transcription factors. We hypothesize that this inhibition could be the result of a weaker disintegration of E2F/pRb-pp_by_cdk4 complexes together with the acquirement of the ability for pRb-pp_by_cdk4 to de novo inhibit free E2F. To model such hypotheses two specific modifications of the model have been made.

- A new reaction similar to rxnY115 referring to pRb-pp_by_Cdk4 has been introduced:
 - New_React10: **pRb[21]cdk4** -> **pRb[20]cdk4** + E2F, with a mass action kinetic rate law and parameter $k_{pRbppCdk4E2Ffree}$, which has been estimated by manual fine tuning as described in the following sections
- The reverse of the above mentioned reaction having the following form has been introduced:
 - New_React11: **pRb[20]cdk4** + E2F -> **pRb[21]cdk4**

and its kinetic rate law has been selected to be a mass action with parameter value $k_{pRbppCdk4E2FInh}$. The latter has also been estimated by manual fine tuning

L2.4.2.3 Delay in the modifier activation time

In the Haberichter et al. model, in order for the modelling effort to be consistent with the existing evidence about a linkage between cell growth and regulation of proliferation [inhibition of metabolic processes results in a G1 arrest, suggesting that metabolism plays a key role in G1 progression], an unspecified activating 'modifier' pathway that responds to cellular growth and results in activation of cyclin E/Cdk2 complexes, allowing cells to progress through the cell cycle, has been modelled by adding a time triggered switch which induces the activation of Cdk2. In the original version of the model the activation of the switch happens at the end of early G1 and the time point that this is done has been estimated based on experimental data. However, in the case of pre-B ALL, where the early G1 is omitted and the execution of the G1 phase is prolonged, we hypothesize that this activation is postponed, so the parameter ModifierTime (time point of the switch activation) has been set to candidate for changes during the calibration of the model though manual fine tuning. It should be mentioned that genes associated with glucose metabolism [which is crucial part in cancer cells metabolism due to the Warburg effect (López-Lázaro, 2008)] are differentially expressed between two distinct populations of pre-B ALL patients (prednisolone-sensitive and prednisolone-resistant) (Holleman et al., 2004).

L2.4.3 Proposed model calibration and simulation results

L2.4.3.1 Calibration of the model for the mean S-phase transition time point

As described above, at this early stage of the development of the model for the sub-cellular dynamics of the pre-B ALL cell cycle, the calibration of the model parameters has been chosen to be done by manual fine tuning. Taking into account that there has not been any quantitative data available from pre-B ALL cells, the current objective is to adequately fit the observed experimental findings in a semi-qualitative way.

A fundamental question is whether it is possible to recalibrate the model in such a way that although the hyper-phosphorylated form of retinoblastoma protein almost dominates the total levels of the protein, the transition into the S-phase happens at a time point that is consistent with the observed cell cycle kinetics in pre-B ALL as presented in section L2.1.

In order to calculate the earliest, the mean and the latest time point at which the transition into the S-phase can take place in pre-B ALL cells we take into account the following values of the duration of the cell cycle and the related phases, since the G1-phase is thought to be the most tunable phase of the cell cycle (Robert C Bast et al., 2000):

- $t_{cc} = 112.5 \text{ h } (\pm 46.8)$ (Hirt et al., 1992)
- $t_s = 18 \text{ h}$ (duration of the S-phase in pre-B ALL primary cells) (Hirt et al., 1992)
- $t_M = 1 \text{ h}$ (Robert C Bast et al., 2000)
- $t_{G2} = 6 \text{ h}$ (Robert C Bast et al., 2000)

Therefore, by taking the lower, mean and higher values of t_{cc} we can estimate the minimal, mean and maximal duration of the G1 phase (equivalently the time point of transition into the S-phase) which is

- Earliest time point for the S-phase transition ($t_{(G1/S)_{min}}$): ~2400 minutes (40 hours)
- Mean time point for the S-phase transition ($t_{(G1/S)_{mean}}$): ~5200 minutes (86,6 hours)
- Latest time point for the S-phase ($t_{(G1/S)_{max}}$): ~8000 minutes (133,3 hours)

As a first step in the the model calibration we have set the objective to predict the transition into the S-phase at the mean time point for pre-B ALL cells. The criterion for the passage into the S-phase has been the increase of CyclinA to levels similar to those reached in the original model during the time needed for the corresponding cellular system to enter the S-phase (350 cyclinA molecules/cell in 600 minutes).

By introducing the new reactions to the model (New_React 1-11) and by manually tuning the parameters mentioned in sections L2.4.2.2 and L2.4.2.3 in such a way that the levels of cyclinA will reach the criterion levels approximately at 5200 minutes (86,6 hours), the simulation results presented in Figures L5 and L6 have been acquired. It should be mentioned that in order to effectively check the fulfillment of the criterion, two specific actions have been introduced into the model. The first action registers the time point at which the CyclinA reaches the criterion levels (a variable with the name `transition_time` is used). The second action returns the value of the above register to zero if the rate of CyclinA levels ($d\text{CyclinA}/dt$) is lower than or equal to zero after CyclinA reaches the threshold. More specifically, the criterion is fulfilled only if the threshold is reached and CyclinA is still increasing (since the levels of CyclinA rise especially during S-phase). The following results have been achieved by setting:

- $\text{kpRbppCdk4E2Ffree} = 0.0001 \text{ (1/min)}$
- $\text{kpRbppCdk4E2FInh} = 5.204e^{-5} \text{ (1/molecules * min)}$
- $\text{ModifierTime} = 2200 \text{ minutes (36,6 hours)}$

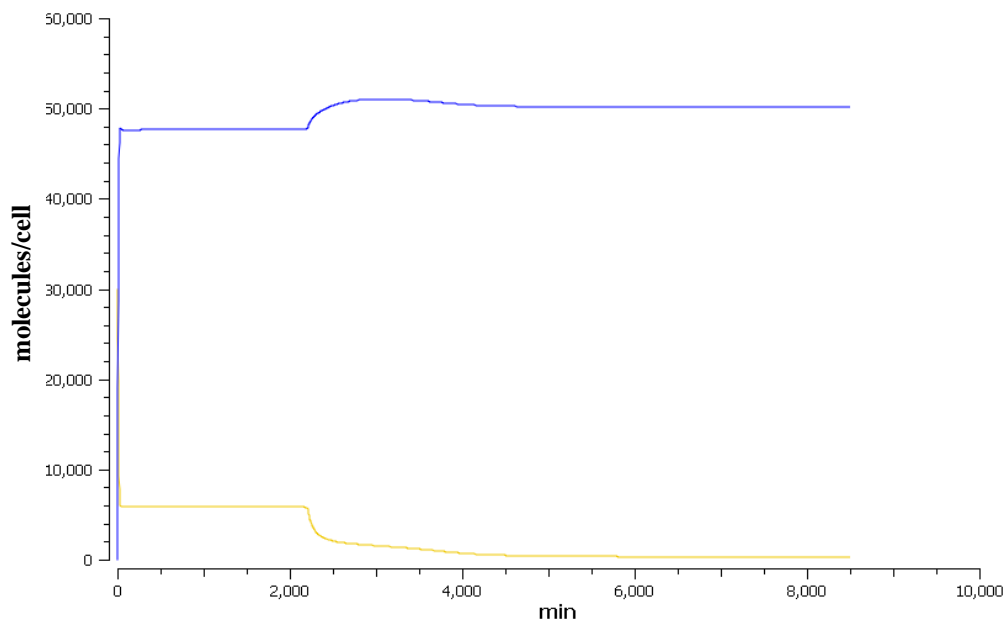


Figure L2.5 Simulation results of the proposed model (extension of the Haberichter et al. model) for 8500 minutes (~141 hours) showing the levels (molecules/cell) of hypo-phosphorylated pRb (yellow line) and the hyper-phosphorylated pRb (blue line), by selecting the parameter values in such a way that the S-phase transition happens approximately at 5200 minutes (86,6 hours).

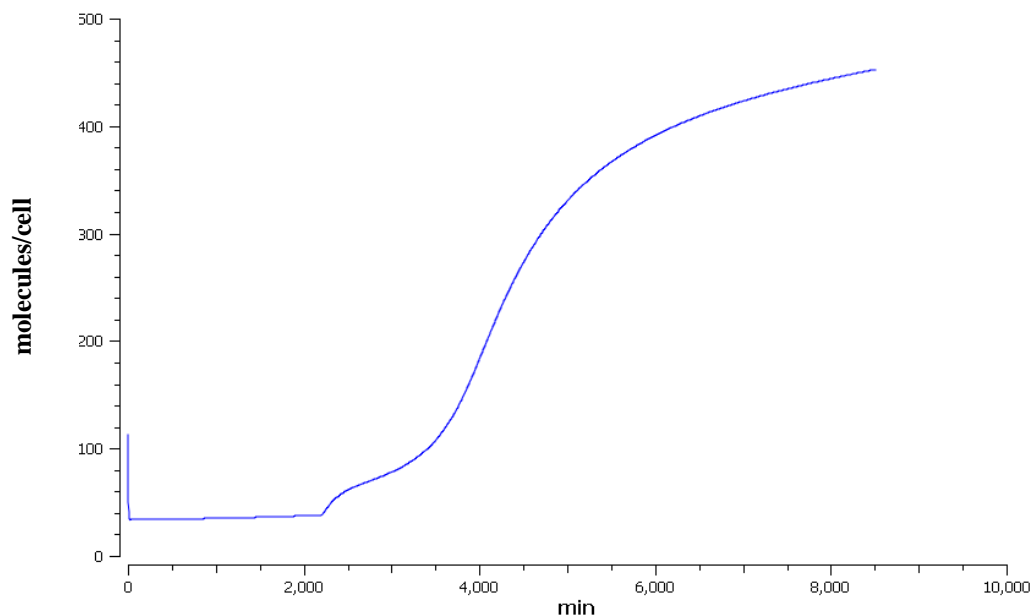


Figure L2.6 Simulation results of the proposed model (extension of the Haberichter et al. model) for 8500 minutes (~141 hours) showing the levels (molecules/cell) of cyclinA (blue line) by selecting the parameter values in such a way that the S-phase transition happens at approximately 5200 minutes (86,6 hours).

According to Figure L2.5, the hyper-phosphorylated form of the retinoblastoma protein rapidly reaches sufficiently high levels in contrast with the hypo-phosphorylated form which is almost cleared after some hours of hours of the G1-phase execution. However, according to Figure L2.6,

although the cyclinA is gradually rising, the time needed to reach the hypothesized threshold for S-phase passage is significantly high, reaching 5247.326095 minutes (almost identical to the value calculated to be the mean time point for the S-phase transition in pre-B ALL cells. These preliminary results are thought to be consistent with the previously presented literature findings about pre-B ALL cell cycle characteristics, concerning the unusual co-existence of hyper-phosphorylated retinoblastoma protein with the delayed passage into the G1-phase.

L2.4.3.2 Calibration of the model to meet the range of S-phase transition time points in pre-B ALL

Following the calibration of the model for the mean t_{cc} (and the relative $t_{(G1/S)_{mean}}$), the question has arisen as to whether the model is capable of simulating the whole range of the S-phase transition time points observed in pre-B ALL (presented in L2.4.3.1). In order to explore this we have searched for the parameters of the model to which the value of the variable `transition_time` is highly sensitive. This has been done by utilizing the Sensitivity Analysis and Parameter Scan functionalities included in COPASI. The methods implemented are available in (De la Fuente et al., 2002). For this initial phase of the development and testing of the model of the sub-cellular biochemical dynamics of the cell cycle in pre-B ALL, we have requested that a rather small number of parameters should be perturbed for simplicity reasons. Moreover, it should be mentioned that we have pre-selected the `ModifierTime` parameter, since during the initial calibration of the model presented in the previous sub-section it has been shown to significantly affect the transition time point. In this context we have searched for two parameters, one from the set of kinetic parameters and one from the set of the initial levels of biomolecules that in conjunction with the `ModifierTime` parameter would be able, when perturbed, to lead into S-phase transition time points distributed over the entire calculated range.

In order to identify these parameters, two distinct sensitivity analysis steps have been taken: one calculating the sensitivity of the `transition_time` to all kinetic parameters and one calculating the sensitivity of the `transition_time` to all initial levels of the biomolecules modeled in the system. Following the sensitivity analysis execution, the parameters with the highest or lowest impact on sensitivity were selected for further analysis using the parameter scan functionality. Multiple simulations of the model, each time with different values of the selected parameters based on specific bounds and intervals have been performed. Each time a combination of three parameters (one kinetic parameter, one initial level value and the `ModifierTime`) was tested for their appropriateness to generate the tunability observed at the S-phase transition time point. After numerous tests we have selected the following combination of parameters and the corresponding bounds of their values:

- Degradation rate of p27 cyclin dependent kinase inhibitor (`kdYp27`): 0.0005-0.0016 (1/min)
- Anaphase-Promoting Complex (APCC) initial levels: 10000-35000 (molecules/cell)
- `ModifierTime`: 2000-7600 (minutes) (33.3 – 126.6 hours)

The above mentioned values have been obtained through manual fine tuning during the parameter scan procedure. The objective has been to end up with simulations that would predict that the transition into the S-phase happens at time points between 2400 and 8000 minutes (40-133.3 hours) and that no parameter values combination leads to G1-arrest (no transition into the S-phase). The results of the parameter scan (5 intervals for each parameter) for the above parameters are presented in Figure L2.7.

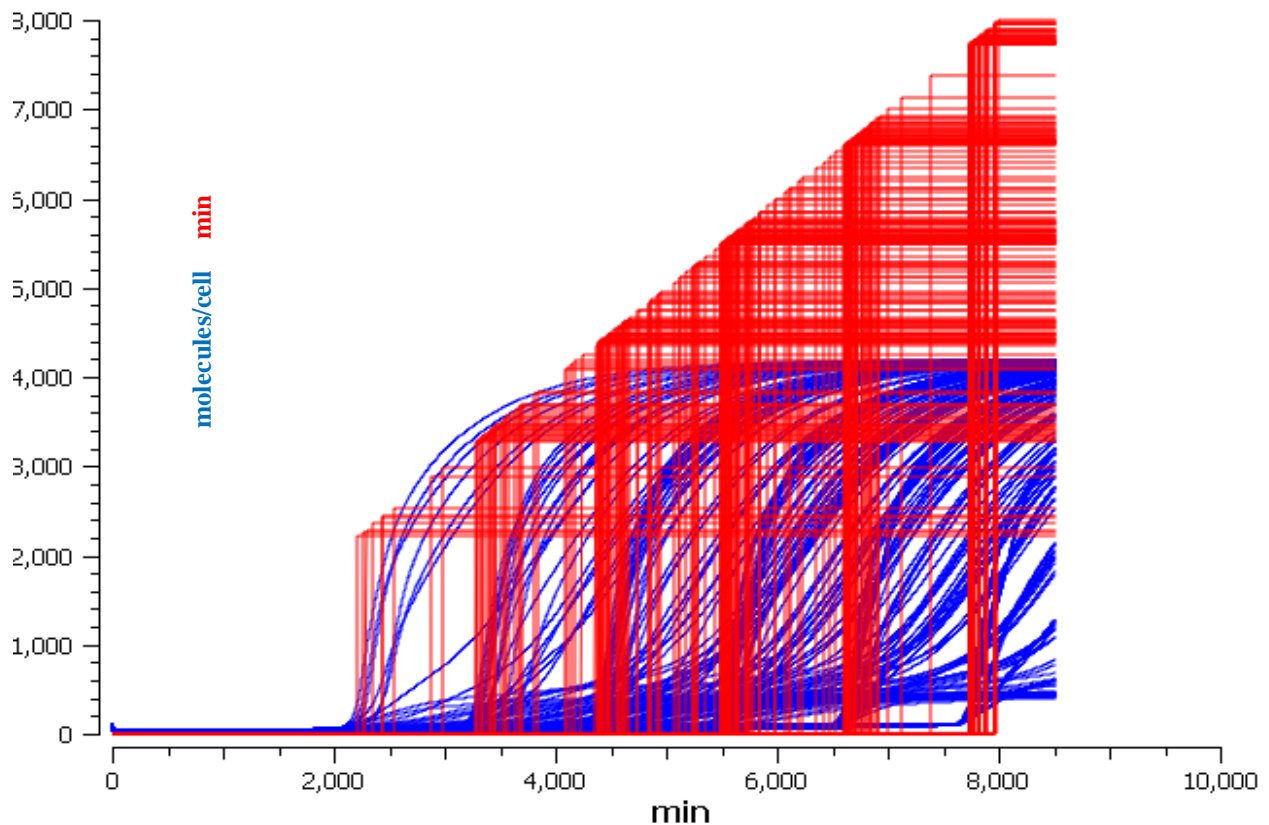


Figure L2.7 Parameter scan results of the proposed model (extension of the Haberichter et al. model) for 8500 min following the calibration for the mean S-phase transition time. The parameters scanned have been $kdYp27$, APCC initial levels and ModifierTime. The blue line shows the levels (molecules/cell) of CyclinA and the red line shows the status of the transition_time variable, indicating the time point where the S-phase transition happens for every combination of parameters values. *It is noted that the vertical axis generally represents a sample of the cell and the modelules under consideration for reference/calibration purposes.*

According to Figure L2.7 by choosing a specific combination of the values of the three above mentioned parameters, the S-phase transition can be predicted to happen in a broad range of time points similar to the observed in the pre-B ALL case (see section L2.4.3.1). The selected parameters are critical indeed for the execution of the cell cycle. p27 is one of the central cyclin dependent kinase inhibitors. APCC, except for promoting the Anaphase (not modeled here), regulates the degradation of CyclinA as can be seen in the reactions of the model. ModifierTime controls the activation of one of the central components of the G1 phase machinery. However, it should be noted that in the next phases of the model development a detailed literature search for the factors leading to the observed variability in the S-phase transition time point is to be undertaken. In that way a more precise prediction of the cell cycle duration specific for every patient and based on their own clinical and/or genomic-proteomic status is expected to be achieved.

In order to demonstrate the above points we have used a random sample from three uniform distributions having bounds identical to the bounds of the above three parameters in order to create a virtual case for which a prediction for the cell cycle would be made. The sample has been the following

- $kdYp27$ (1/min) = 0.00124842
- APCC initial levels (molecules/cell) = 34919.5

- ModifierTime (minutes) = 4383.64 (~73 hours)

The values imply that for this specific case the cell is characterized by reduced p27 degradation rate generally inhibiting the cell cycle progression, increased levels of APCC which could lead to increased CyclinA degradation and significantly postponed activation of the switch activating Cdk2 which may indicate a slow rate of cell growth/metabolism. Therefore, a delayed transition into the S-phase compared with the mean case (L2.4.3.1) is expected. By introducing these parameter values into the model, a transition to S-phase at 6491.563667 minutes (108.2 hours) is predicted which is consistent with the above mentioned expectation. The results of the simulation for this case are shown in the following figure:

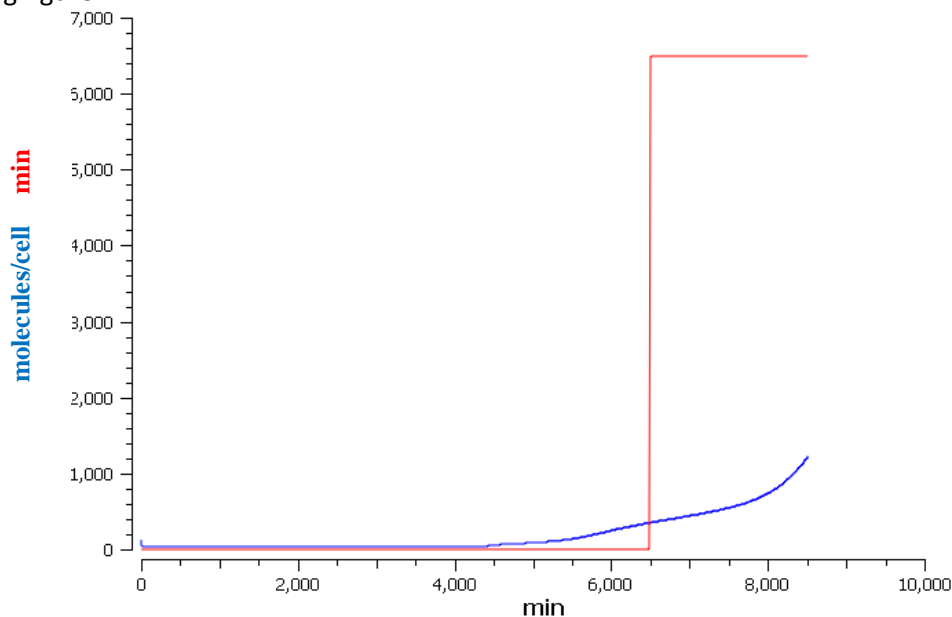


Figure L2.8 Simulation results of the proposed model (extension of the Haberichter et al. model) for 8500 min (~141 hours) showing the levels (molecules/cell) of cyclinA (blue line) by using randomly selected parameter values for kd_{Yp27} , APCC initial levels and ModifierTime. The predicted time point for the transition to S-phase is shown by the red line.

L2.5 Conclusions

The first steps to create a model capable of simulating the sub-cellular biochemical dynamics of the pre-B ALL cell cycle have been taken. The central objective has been to test whether it is possible to create a model that would predict a delayed, although undisturbed, passage to the S-phase even with the domination of the hyper-phosphorylated form of the retinoblastoma protein, a phenomenon observed in pre-B ALL. Based on the results presented, it seems that a more detailed and more precise model could emerge in the future by possibly integrating further cell cycle models that could serve as a basis for the consideration of additional biological/biochemical phenomena that might contribute to the observed deregulation of the cell cycle in pre-B ALL and in ALL in general. For example a model like (Iwamoto et al., 2011), simulating the whole cell cycle machinery, might enable us to study in more detail not only the deregulations of the G1/S transition observed in ALL, but also the possible contribution of other phases to the observed cell cycle kinetics. Moreover, such a model might enable the modelling of cell cycle phase specific drug effects. Finally, as far as other biological phenomena deregulating the leukemic cell cycle are concerned, an hypothesis regarding the direct regulation of CyclinA by pRb (Angus et al., 2002) has been planned to be tested.

L2.6 References

Angus, S.P., Fribourg, A.F., Markey, M.P., Williams, S.L., Horn, H.F., DeGregori, J., Kowalik, T.F., Fukasawa, K., Knudsen, E.S., 2002. Active RB elicits late G1/S inhibition. *Exp. Cell Res.* 276, 201–213.

Conradie, R., Bruggeman, F.J., Ciliberto, A., Csikász-Nagy, A., Novák, B., Westerhoff, H.V., Snoep, J.L., 2010. Restriction point control of the mammalian cell cycle via the cyclin E/Cdk2:p27 complex. *FEBS J.* 277, 357–367.
De la Fuente, A., Snoep, J.L., Westerhoff, H.V., Mendes, P., 2002. Metabolic control in integrated biochemical systems. *European Journal of Biochemistry* 269, 4399–4408.

Greaves, M.F., 1986. Differentiation-linked leukemogenesis in lymphocytes. *Science* 234, 697–704.

Haberichter, T., Mädge, B., Christopher, R.A., Yoshioka, N., Dhiman, A., Miller, R., Gendelman, R., Aksenov, S.V., Khalil, I.G., Dowdy, S.F., 2007. A systems biology dynamical model of mammalian G1 cell cycle progression. *Mol. Syst. Biol.* 3, 84.

Hirt, A., Antic, V., Wang, E., Lüthy, A.R., Leibundgut, K., Von der Weid, N., Tobler, A., Wagner, H.P., 1997a. Acute lymphoblastic leukaemia in childhood: cell proliferation without rest. *Br. J. Haematol.* 96, 366–368.

Hirt, A., Leibundgut, K., Lüthy, A.R., Von der Weid, N., Wagner, H.P., 1997b. Cell birth and death in childhood acute lymphoblastic leukaemia: how fast does the neoplastic cell clone expand? *Br. J. Haematol.* 98, 999–1001.

Hirt, A., Schmid, A.-M., Ammann, R.A., Leibundgut, K., 2011. In pediatric lymphoblastic leukemia of B-cell origin, a small population of primitive blast cells is noncycling, suggesting them to be leukemia stem cell candidates. *Pediatr. Res.* 69, 194–199.

Hirt, A., Schmid, A.M., Julmy, F., Schmitz, N.M.R., Leibundgut, K., 2009. Expression of cyclin A in childhood acute lymphoblastic leukemia cells reveals undisturbed G1-S phase transition and passage through the S phase. *Leukemia* 23, 414–417.

Hirt, A., Werren, E.M., Luethy, A.R., Gerdes, J., Wagner, H.P., 1992. Cell cycle analysis in lymphoid neoplasia of childhood: differences among immunologic subtypes and similarities in the proliferation of normal and leukaemic precursor B cells. *Br. J. Haematol.* 80, 189–193.

Holleman, A., Cheok, M.H., Den Boer, M.L., Yang, W., Veerman, A.J.P., Kazemier, K.M., Pei, D., Cheng, C., Pui, C.-H., Relling, M.V., Janka-Schaub, G.E., Pieters, R., Evans, W.E., 2004. Gene-Expression Patterns in Drug-Resistant Acute Lymphoblastic Leukemia Cells and Response to Treatment. *New England Journal of Medicine* 351, 533–542.

Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., Singhal, M., Xu, L., Mendes, P., Kummer, U., 2006. COPASI--a COMplex PATHway Simulator. *Bioinformatics* 22, 3067–3074.

Iwamoto, K., Hamada, H., Eguchi, Y., Okamoto, M., 2011. Mathematical modeling of cell cycle regulation in response to DNA damage: exploring mechanisms of cell-fate determination. *BioSystems* 103, 384–391.

Leibundgut, K., Schmitz, N., Tobler, A., Lüthy, A.R., Hirt, A., 1999. In childhood acute lymphoblastic leukemia the hypophosphorylated retinoblastoma protein, p110RB, is diminished, as compared with normal CD34+ peripheral blood progenitor cells. *Pediatr. Res.* 45, 692–696.

López-Lázaro, M., 2008. The warburg effect: why and how do cancer cells activate glycolysis in the presence of oxygen? *Anticancer Agents Med Chem* 8, 305–312.

McClellan, J.S., Majeti, R., 2013. The cancer stem cell model: B cell acute lymphoblastic leukaemia breaks the mould. *EMBO Mol Med* 5, 7–9.

Novák, B., Tyson, J.J., 2004. A model for restriction point control of the mammalian cell cycle. *J. Theor. Biol.* 230, 563–579.

Pardee, A.B., 1989. G1 events and regulation of cell proliferation. *Science* 246, 603–608.

Pfeuty, B., 2012. Strategic cell-cycle regulatory features that provide mammalian cells with tunable G1 length and reversible G1 arrest. *PLoS ONE* 7, e35291.

Rehe, K., Wilson, K., Bomken, S., Williamson, D., Irving, J., Den Boer, M.L., Stanulla, M., Schrappe, M., Hall, A.G., Heidenreich, O., Vormoor, J., 2013. Acute B lymphoblastic leukaemia-propagating cells are present at high frequency in diverse lymphoblast populations. *EMBO Mol Med* 5, 38–51.

Reiter, A., Schrappe, M., Ludwig, W.D., Hiddemann, W., Sauter, S., Henze, G., Zimmermann, M., Lampert, F., Havers, W., Niethammer, D., 1994. Chemotherapy in 998 unselected childhood acute lymphoblastic leukemia patients. Results and conclusions of the multicenter trial ALL-BFM 86. *Blood* 84, 3122–3133.

Renner, K., Ausserlechner, M.J., Kofler, R., 2003. A conceptual view on glucocorticoid-induced apoptosis, cell cycle arrest and glucocorticoid resistance in lymphoblastic leukemia. *Curr. Mol. Med.* 3, 707–717.

Robert C Bast, J., Kufe, D.W., Pollock, R.E., Weichselbaum, R.R., Holland, J.F., Emil Frei, I.I.I., 2000. Holland-Frei Cancer Medicine [WWW Document]. URL <http://www.ncbi.nlm.nih.gov/books/NBK20777/>

Schmitz, N.M.R., Hirt, A., Aebi, M., Leibundgut, K., 2006. Limited redundancy in phosphorylation of retinoblastoma tumor suppressor protein by cyclin-dependent kinases in acute lymphoblastic leukemia. *Am. J. Pathol.* 169, 1074–1079.

Schmitz, N.M.R., Leibundgut, K., Hirt, A., 2005. CDK2 catalytic activity and loss of nuclear tethering of retinoblastoma protein in childhood acute lymphoblastic leukemia. *Leukemia* 19, 1783–1787.

Swat, M., Kel, A., Herzog, H., 2004. Bifurcation analysis of the regulatory modules of the mammalian G1/S transition. *Bioinformatics* 20, 1506–1511.

Takahashi-Yanaga, F., Sasaguri, T., 2008. GSK-3 β regulates cyclin D1 expression: a new target for chemotherapy. *Cell. Signal.* 20, 581–589.

Yao, G., Lee, T.J., Mori, S., Nevins, J.R., You, L., 2008. A bistable Rb-E2F switch underlies the restriction point. *Nat. Cell Biol.* 10, 476–482.

Zetterberg, A., Larsson, O., Wiman, K.G., 1995. What is the restriction point? *Curr. Opin. Cell Biol.* 7, 835–842.

Chapter 5 [C]

TUMOUR BIOMECHANISM FOCUSED MODELS: VARIOUS BIOCHEMICAL MODELS

C1 Introduction

In this chapter, we will describe one component of the multilevel modelling undertaken within p-medicine: modelling cancer at the atomistic level. The atomistic level modelling provides a valuable tool for investigating the links between our current wealth of genotypic data and therapeutic medicine (C1-C29). In the move toward personalized medicine, this model offers the potential to gain insights, both qualitative and quantitative, at the smallest length and time scales. Here we briefly summary the workflow of the biochemical model and its application in cancer research. The basic methodologies described are generic and can be applied to a variety of pathologies.

A simulator has been built (see Chapter 5 in deliverable D12.2) to study the structural and energetic properties of biomolecular systems. The simulator could generate qualitative and quantitative insights into the causation of disease, and assist in clinical decision making for drug selection as well as future drug development. The simulator would provide a quantitative, predictive modelling for drug efficacy/resistance on a patient with specific genomics. The detailed protocols of the simulator have been described in Deliverable D12.2. For the integrity of the document, in this document we briefly overview the method and workflow of the simulator.

In this deliverable a biochemical model for the dimerization and activity of Janus kinase 2 (JAK2) is outlined. An asymmetric dimer of pseudokinase and kinase domains is built, and molecular dynamics (MD) simulation is performed on the proposed dimer system. Comparisons are made between our form of dimer and a previous suggested anti-symmetric model. Our results show that the asymmetric model is more reasonable to explain the mutation effect on its biological activity. The biochemical model has also been applied to other proteins like epidermal growth factor receptors (EGFRs) and Kirsten rat sarcoma (KRAS).

C2 The biochemical model

The model has been described in detail in Deliverable D12.2. Here we only briefly summarize the methodologies applied to the model and its workflow.

C2.1 Methodologies applied to the model

The model employs large scale MD techniques(C1) to study the structural and energetic properties of biomolecular systems in atomistic detail. Molecular dynamics is a computer simulation technique where the time evolution of a set of interacting atoms is followed by integrating their equations of motion (equation C1). MD simulations have been used successfully in binding free-energy calculations (C2) and rational design of drugs. They can help in visualizing and understanding structures and dynamics at an atomistic level when combined with molecular graphics programs.

The classical molecular dynamics simulation method is based on Newton's equation of motion (equation C1). Numerical integration of the differential equations of motion yields a trajectory that describes the evolution of the system's positions and momenta through time. From this trajectory, the average values of properties can be determined.

$$\mathbf{F} = m \frac{d^2 \mathbf{r}}{dt^2} \quad (\text{C1})$$

Various features can be analysed throughout the trajectory to gain structural and dynamical insights of a system, to predict meaningful statistical properties, and to relate structure to function. Both time-dependent and time-averaged properties can be extracted from the production run. Time dependent properties can be displayed graphically, where one of the axes corresponds to time and the other to the quantity of interest, such as energy, distance, root-mean-square deviation (RMSD), etc. The averaged data is usually presented tabularly. In equation C2, the average is calculated for a quantity of interest Q over a trajectory with N frames

$$\langle Q \rangle = \frac{1}{N} \sum_{i=1}^N Q_i \quad (\text{C2})$$

Quantities of interest can be related to structure such as distances, angles, RMSDs, or related to energy such as pair-pair interactions, potential energies and free energies. Some of the averaged properties can be compared with experimental data. For example, the average structures are compared with these from x-ray crystallography and/or NMR experiments.

C2.2 Workflow of the model

To perform modelling and calculation with optimal efficiency, a highly automated molecular simulation based workflow tool has been adapted into this project. The simulator decomposes the workflow into three main components: (a) the model builder which prepares all necessary files and generates a simulation-ready molecular model, (b) the simulation component which performs large scale molecular dynamics runs on high performance computing resource and (c) the post-production analysis which calculates the averaged or time-dependent properties of the system.

The workflow of the simulator consists of a sequence of connected steps, including performing of the individual components and data transferring between various computational resources (details in deliverable D12.2). The discrete tasks can be performed through a series of scripting command. These scripts are used to automate the process of setting up an MD simulation and creating the necessary input data sets on a local computing resource, running a series of chained MD simulations on a powerful grid resource, post processing the simulation trajectories on local cluster or national grid services, and transferring the data between various resources. The whole process can be performed automatically by a single command, or the discrete tasks be executed separately, on various computational resources. When executed separately, the preparation, simulation and analysis jobs need to be submitted manually using qsub command after logging into various computational resources, and data be transferred manually among various resources.

C3 Application of the simulator to JAK2

JAK2, a member of the Janus kinase (JAK) family of non-receptor tyrosine kinases, participates in the evolutionarily conserved JAK/STAT signalling pathway in almost all eukaryotic organisms. The kinase consists of a kinase and a pseudokinase domain, and possesses but also negatively regulates the kinase activity. The JAK/STAT pathway regulates numerous aspects of haematopoiesis and the immune response, and its abnormal activation has been implicated in the promotion of cancers and immune disorders. JAK2's importance is further underscored by its constitutive activation (i.e. it is always turned on) via cancer mutations, which contributes to its pro-proliferation activities. Most JAK2 mutations are found in the Janus homology 2 (JH2) domain and the connecting region between JH2 and JH3 domains (Figure C1)(C3). JH1 and JH2 are homologous phosphate-transferring domains; JH1, the kinase domain, exhibits tyrosine kinase catalytic activity, while JH2, the pseudokinase domain, is essential for regulation of tyrosine phosphorylation but itself lacking enzymatic activity. The JH2 domain has been proven to act as allosteric regulator of other protein kinase domains.

JAK2 is also an important molecular in ALL-BFM study for clinical trials (WP9). Some driver mutations in ALL are JAK2 activating mutations. JAK2 inhibitors, either already on the market or in clinical trials, are the most reasonable choice for ALL. Indeed, modelling of JAK2 was suggested by our p-medicine partner, Prof Martin Stanulla, in Christian-Albrechts-Universitaet Zu Kiel (CAU).

C3.1 System Preparation

Preparation of protein structures. The structure of the JAK2 JH2 domain was constructed using Modeller (C4), a homology modelling software tool for predicting the three-dimensional structures of proteins. The x-ray structure of EGFR (PDB id: 1M14)(5) was used as the template structure. The kinase domains of JAK2 and EGFR have ~21% sequence identity and another ~41% similarity. The highly homologous catalytic kinase core predominates the high degree of structural conservation for JAK2 and EGFR. The homology model of the JH2 domain closely resembles most x-ray structures of tyrosine kinase domains. The structure of the JH1–JH2 dimer was constructed based on the same EGFR structure, in which the JH1 coordinates were taken from an x-ray structure (PDB id: 3E64)(C6). In the dimer, JH1 was located as the activator and JH2 as the receiver. Choosing the same EGFR structure as template for both the JH2 and JH1–JH2 structures may make the JH1–JH2 interface less sterically hindered. The V617F mutant JH2 was obtained by modification of valine at 617 to phenylalanine within the JH2 structure. Four molecular models were constructed for WT and V617F mutant JAK2 systems which are designated as WT-JH2, V617F-JH2, JH1-WT-JH2 and JH1-V617F-JH2 respectively. Figure C1 shows our asymmetric model of JH1-JH2, and a previous published anti-symmetric model with which we compared in our study.

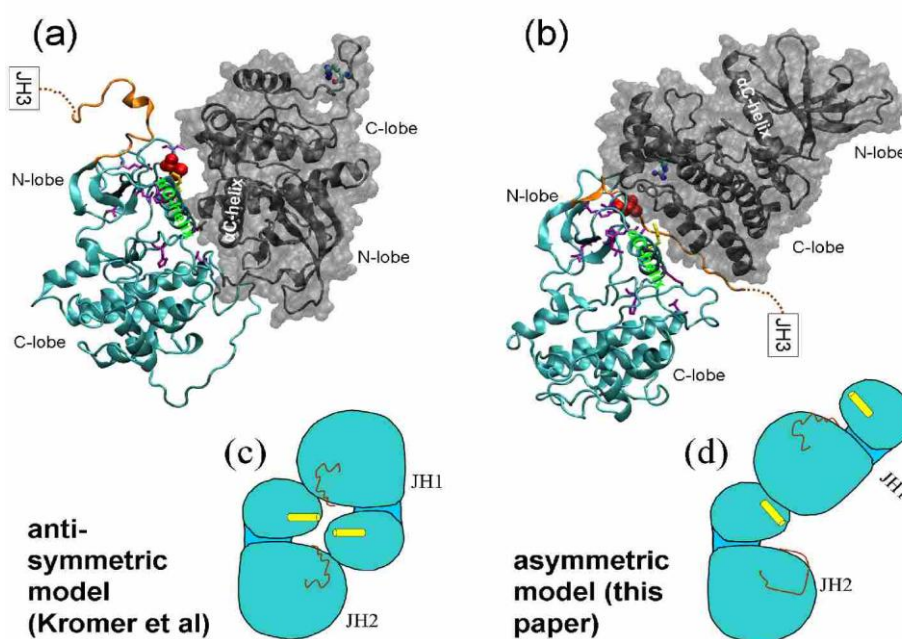


Figure C1. The structures of the JH1-JH2 dimer in antisymmetric (a,c) and asymmetric (b,d) modes. The proteins are shown in cartoon representation (a,b), with JH1 in gray and JH2 in cyan. JH1 is also shown in surface representation for better visualization of the interacting surface between the two domains. The V617F mutation in JH2 is shown in a ball and stick model (red), F595 in bond (yellow), R1063 in JH1 domain in ball and stick (cyan for carbon, blue for nitrogen, and red for oxygen), other JH2 mutations in bond (purple), and the link between JH2 and JH3 in cartoon representation (orange). The schematic representations (c,d) highlight the distinct arrangements of the two domains and the conformational differences of the α C-helix (yellow) and A-loop (orange) within the JH2 domain.

Simulation-ready models. Initial preparation of all four models was implemented using our simulator(C7). Protonation states were assigned for all ionisable residues to their default values at neutral pH. The proteins were solvated using TIP3P water molecules(8) and then electrostatically neutralized by adding counterions (Na^+). The distance between the edges of the solvent box and the closest atom in the protein was 14Å. The AMBER ff03 all-atom force field(9) was used to describe the protein parameters. This procedure resulted in a final atom count of approximately 52,000 and 96,000 atoms for the JH2 domain and JH1-JH2 domain models respectively (Figure C2).

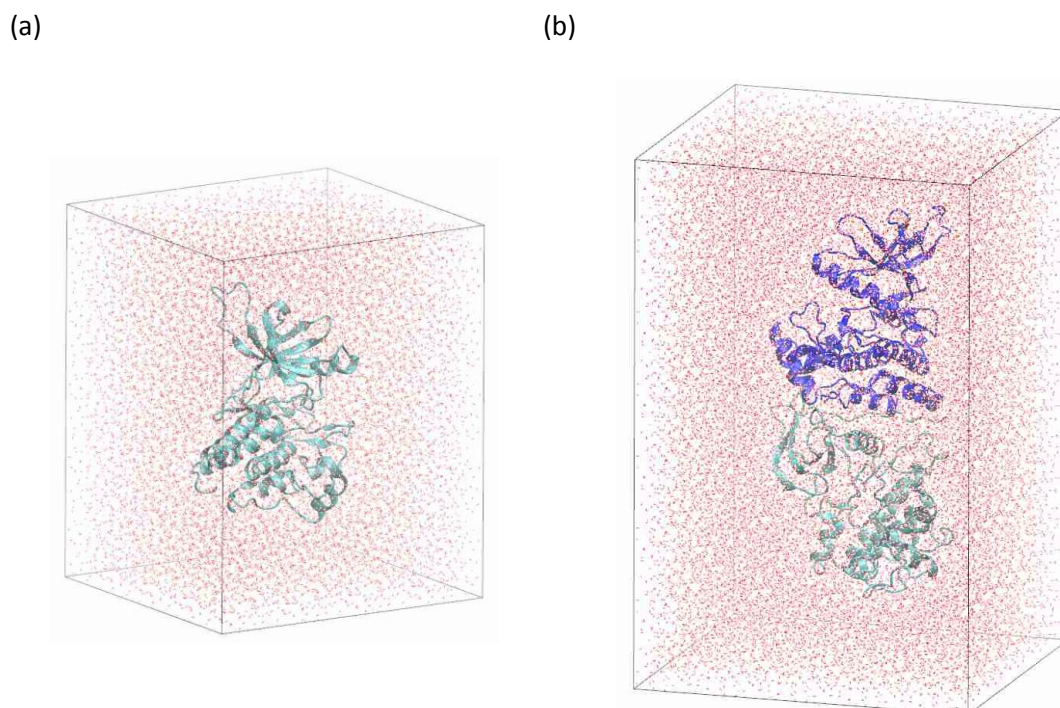


Figure C2. Simulation-ready models for JH2 (a) and JH1-JH2 (b) domains. The proteins (in cartoon representation) are solvated in boxes of water molecules (red).

C3.2 Simulations

Molecular dynamics (MD) simulations were performed on the four separate JAK2 models that differ in terms of sequence (wild-type or V617F mutation) or size (JH2 domain alone, or JH1–JH2 domains). The MD package NAMD2.7 (C10) was used throughout the equilibration and production simulations. For each molecular system, energy minimizations were first performed with heavy protein atoms restrained at their initial positions. Then a series of short simulations (each of 25 ps duration) were conducted, while the restraints on heavy atoms were gradually reduced. Finally, 100 ns production simulations were run and coordinates were recorded every 10 ps. Periodic boundary conditions were imposed in all three spatial dimensions. The systems were maintained at a temperature of 300 K and a pressure of 1 bar in the isothermal-isobaric ensemble (NPT). The simulations were performed on 96 cores (four 24-core nodes) for the JH2 system, and 192 (eight 24-core nodes) for JH1–JH2, on HECToR (Cray XE6), the UK national supercomputer based in Edinburgh. Each individual MD simulation took approximately 2 h/ns on HECToR.

C3.3 Post-process analyses

Analysis of the root mean square deviation (RMSD) was conducted for the backbone atoms, based on comparison with the initial structures. Thermodynamic analysis was performed using the molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) method (C11) within the AMBER package (C12). The MM/PBSA method is an efficient way to estimate the binding free energy of a complex from a classical MD simulation. The JH1–JH2 binding free energy can be evaluated, following the MM/PBSA approach, as:

$$\Delta G_{binding} = G^{dimer} - G^{JH1} - G^{JH2} \quad (C3)$$

where

$$G^i = \langle E_{MM} \rangle + \langle G_{solv} \rangle - TS_{conf} \quad (C4)$$

Here, $\langle \dots \rangle$ denotes an average for a set of structures from MD simulations, while E_{MM} is the molecular mechanics energy of species i (the index i corresponding to either dimer, JH1 or JH2 domain) in the gas phase. E_{MM} is the sum of bonded energy, van der Waals (E_{vdW}) and electrostatic (E_{elec}) interactions. G_{solv} is the solvation free energy; it is estimated as the sum of the electrostatic solvation free energy (G_{PB}) calculated using the Poisson-Boltzmann equation and the nonpolar solvation free energy (G_{SA}) calculated from the solvent accessible surface area. $-TS_{conf}$ is the contribution from the configurational entropy S_{conf} at temperature T . As the JH1 and JH2 domains bind non-covalently, the bonded terms cancel exactly in the binding free energy calculation (Equation C3). Hence, the binding free energy can be written as:

$$\begin{aligned} \Delta G_{binding} &= \langle \Delta E_{MM} \rangle + \langle \Delta G_{solv} \rangle - T \langle \Delta S_{conf} \rangle \\ &= \langle \Delta E_{vdW} \rangle + \langle \Delta E_{elec} \rangle + \langle \Delta G_{PB} \rangle + \langle \Delta G_{SA} \rangle - T \langle \Delta S_{conf} \rangle \end{aligned} \quad (C5)$$

Here, the free energy difference upon binding $\Delta G_{binding}$ is decomposed into ΔE_{MM} and ΔG_{solv} , and further into ΔE_{vdW} , ΔE_{elec} , ΔG_{PB} and ΔG_{SA} . The $-T\Delta S_{conf}$ term denotes the configurational entropy change ΔS_{conf} at temperature T . The entropy contribution ΔS_{conf} was not included in this study. Previous study has shown that inclusion of the configurational entropy does not make any significant contributions to estimation of protein stability for the Abelson and EGFR kinases (C13). The structures used for the thermodynamic calculations were extracted from the last 80ns MD trajectories of the JH1–JH2 systems.

C3.4 Results

Locations of cancer-related mutations. Negative-stain electron microscopy experiments have demonstrated that a mutation at the asymmetric interface eliminates the association of the dimeric EGFRs, and hence suppresses EGFR kinase catalytic activity (C14). For the regulation-related mutations in the JH2 domain of JAK2, it is reasonable to assume that most are located at or near the JH1–JH2 interface. The mutations either strengthen or weaken the interactions between JH1 and JH2 domains, and hence affect regulation of the JAK2 kinase activation.

From a visual inspection of the asymmetric JH1–JH2 structure constructed here and the previously built anti-symmetric model (C15), it is apparent that the mutations are in closer proximity to the interface in the asymmetric dimer than they are in the anti-symmetric form (Figure C1). For example, V617, the more commonly mutated residue of JAK2, is closer to the kinase domain JH1 in the asymmetric model than it is in the anti-symmetric one, with closest heavy atom distances of 3.2Å and 7.9Å in the two models. Mutations are present at the link region between JH2 and JH3 in patients with myeloproliferative neoplasm and polycythemia vera. This region has a close contact with the JH1 domain in the asymmetric model, but no direct interactions with JH1 in the anti-symmetric model (Figure C1). R1063H, a mutation in JH1 domain identified in some polycythemia vera patients, is also near the JH1–JH2 interface in the asymmetric model with a closest distance of 4.7Å to the JH2 domain, while the residue is far away from the interface in the anti-symmetric conformation with a distance of 34.0Å.

Initial separation of the two domains in the V617F mutant JAK2. Figure C3 shows the root-mean-squared deviation (RMSD) of backbone atoms for JH2 and JH1–JH2 domains, with respect to the

structures from which the simulations are initiated. The simulations of JH2 alone were initiated from a structure built by homology modelling. It is not surprising that its RMSD displays a rapid increase in the first 20 ns simulation, both for the wild-type and V617F mutant forms (Figure C3a,b) because of the uncertainties in the modelled structures (C16). The V617F mutant JH2 has a smaller RMSD than the wild-type after 20ns simulations, probably due, at least in part, to the presence of F617 which stabilizes the α C-helix in the N-lobe (C17). The interactions between JH1 and JH2 stabilize the N-lobe of the JH2 (Figure C1), and are attributed to the decreased JH2 flexibility. The wild-type simulation of JH1–JH2 displays RMSD values of ca. 4Å for the whole dimer after approximately 20 ns, and remains as such for about 55 ns when the value begins to increase. As the RMSDs for the JH1 or JH2 domain alone remain stable throughout the last 80 ns simulations, the rapid increase of the RMSD for JH1–JH2 is due to the relative movement between the two domains, possibly implicating incipient JH1–JH2 separation. The V617F simulation, by contrast, does not exhibit any stable period before it reaches equilibrium at about 40ns, with a much larger RMSD than that of the wild-type at its stable state. Although the wild-type and mutant JH1–JH2 have comparable RMSD at the last 20ns simulations, they adopt different orientations of JH1 and JH2. At the end of the simulations, for example, the two structures differ by an RMSD larger than 7.5Å for all backbone atoms, mainly due to the differences in the interdomain orientations. The mutation is likely to make the JH1–JH2 structure unstable, and possibly contributes to the separation of JH1 and JH2 from the asymmetric mode at an earlier stage, within the first 40 ns simulation (Figure C3d), than that in the wild-type dimer (Figure C3c).

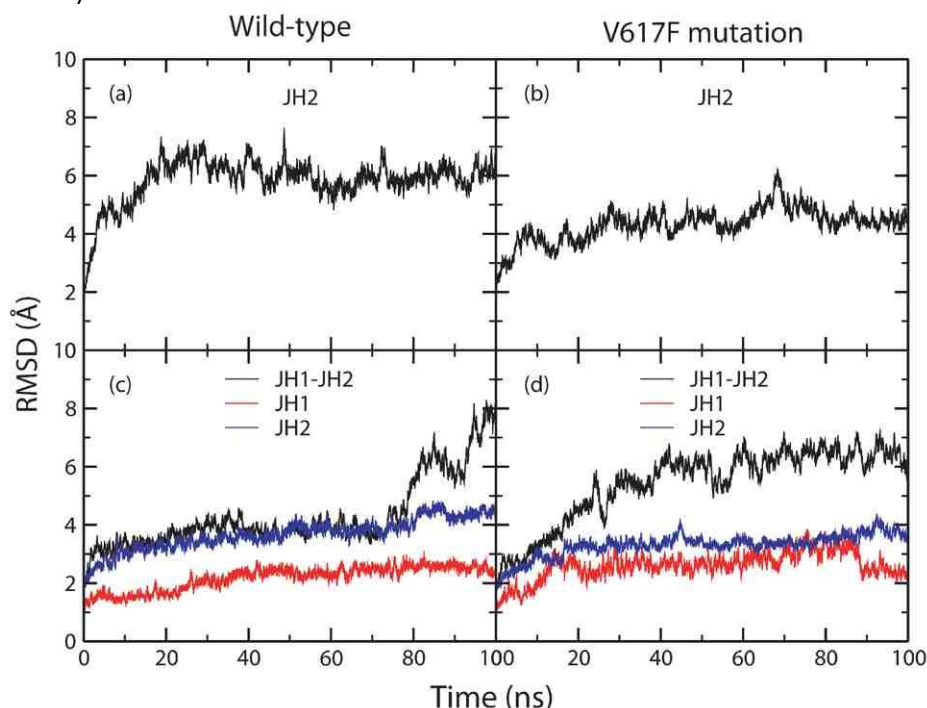


Figure C3. Root-mean-square deviation (RMSD) as function of MD simulation time for the backbone atoms, with respect to the initial structures for (a) wild-type JH2, (b) V617F mutant JH2, (c) wild-type JH1–JH2, and (d) V617F mutant JH1–JH2. The JH2 domain displays a large deviation from the initial structures as shown in (a) and (b). The RMSD increases of the JH1–JH2 dimer in (c) and (d) indicate relative domain movement, possibly evidence of their incipient separation.

Buried surface of JH2 domain. The V617F mutation introduces a large hydrophobic group which is buried at the JH1–JH2 interface (Figure C1b). The interfacial residue F617 has a larger contact surface with the JH1 domain than V617 does (Figure C4). In the modelled asymmetric dimer, the five closest residues of JH1 to residue V617 are N1084, N1085, R1087, N1107 and N1108; four of them are polar

asparagine residues, the last one being a charged arginine. The surface formed by these residues is hydrophilic, and disfavors the bulky hydrophobic sidechain of phenylalanine. Replacement of valine with such a residue could destabilize the association of the JH1–JH2 dimer and make them separate more easily. The separated JH1 and JH2 would thus be able to re-arrange in a form such that JH1 occupies the receiver position and is biologically activated (see further below, and Figure C5). We hypothesize that large hydrophobic residues at position 617 would render the kinase domain constitutively active in a manner similar to that described above for the V617F mutation. The hypothesis is supported by an experimental mutagenesis study which identified that only the substitution of bulky hydrophobic residues at position 617 (V617W, V617M, V617I, V617L and V671F) is able to induce constitutive downstream signalling(C18). The V617I mutation has been reported in patients with hereditary thrombocytosis (C19), while V658L in JAK1, a mutation analogous to V617L in JAK2, been reported in patients with acute leukaemia (C20). V617M and V617W are unlikely to occur as the substitutions require two and three base pair changes in the DNA codon for valine.

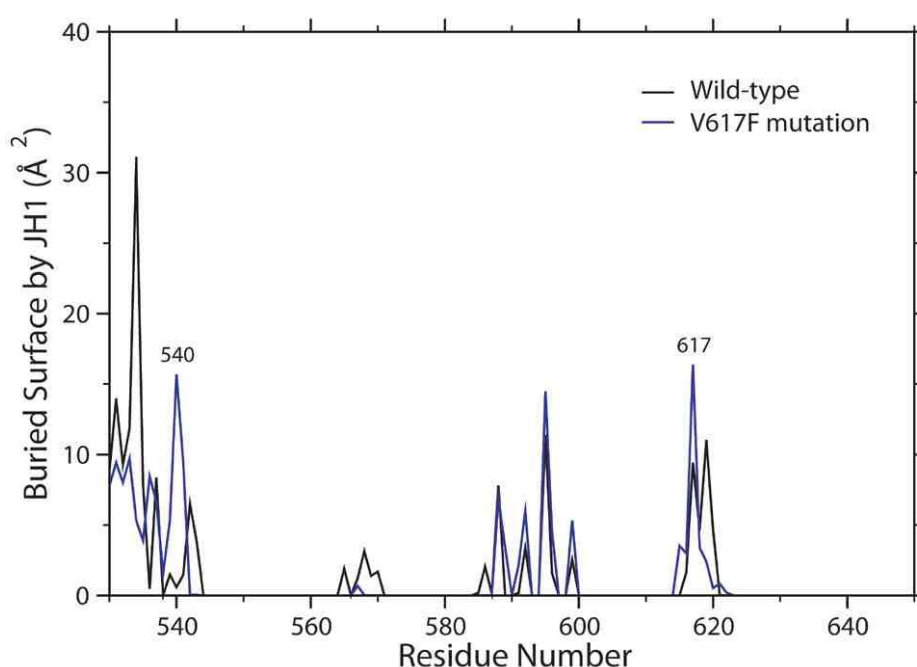


Figure C4 The contact surfaces of JH2 residues buried by the JH1 domain. The substitution of valine by phenylalanine at position 617 introduces a large change in the buried surface area of this residue and also of the link region (residue numbers 530–542) between JH2 and JH3 domains.

Binding free energy of JH1–JH2. We performed MM/PBSA calculations to get quantitative estimates for the binding free energies and their components. The binding free energies averaged over all conformations sampled in the last 80 ns simulations are listed in Table C1 for the wild-type and mutant JH1–JH2 dimers. Our MM/PBSA calculations show that the binding free energy of wild-type JH1–JH2 is 10.4 kcal/mol more favourable than that of the V617F mutant dimer.

The JH1 domain has very similar free energy in the two simulations (Table C1) as the C-lobe of the JH1 domains is relatively stable and thus hardly affected by mutations in JH2 on the sampled timescale. Direct comparison of the energy components between the two protein systems is not straightforward, because competing effects ($\Delta\Delta G_{PB}$ and $\Delta\Delta E_{elec}$ for example) can be quite different but tend to cancel each other out. Nevertheless, it becomes clear from the energy decomposition (Table C1) that the direct inter-domain electrostatic interactions are favourable to the binding for

both wild-type and mutant systems, although the energies could not compensate the large desolvation penalties associated with the binding event, thereby leading to an unfavourable total electrostatic contribution. The V617F mutation makes the electrostatic interaction between two domains much weaker ($\Delta\Delta E_{\text{elec}} = 96.5$ kcal/mol), which cannot be fully compensated by the smaller solvation penalty ($\Delta\Delta G_{\text{PB}} = -82.3$ kcal/mol). The more favourable binding free energy of the wild-type JH1–JH2 is due to its less unfavourable total electrostatic interactions compared with the V617F mutant. This finding is not unexpected as the residue 617 interacts with a polar/charged region of JH1; the intrusion of a large hydrophobic group should indeed be energetically unfavourable.

Table C.I JH1–JH2 interactions (kcal/mol) calculated using the MM/PBSA method.

Type ^a	Wild-type				V617F				$\Delta\Delta^c$
	Complex	JH1	JH2	Δ^b	Complex	JH1	JH2	Δ^b	
E_{vdW}	-4857.5	-2466.7	-2291.7	-99.1	-4817.4	-2430.1	-2284.6	-102.6	-3.5
E_{elec}	-40544.6	-20531.4	-19826.4	-186.9	-40171.8	-20427.8	-19653.6	-90.4	96.5
G_{PB}	-6801.0	-3310.7	-3730.7	240.4	-7144.6	-3456.8	-3845.9	158.1	-82.3
G_{SA}	150.7	79.1	82.3	-10.8	158.3	84.1	85.2	-11.0	-0.2
E_{MM}	-45402.1	-22998.1	-22118.1	-286.0	-44989.2	-22857.9	-21938.2	-193.0	93.0
G_{solv}	-6650.2	-3231.5	-3648.4	229.6	-6986.3	-3372.7	-3760.8	147.2	-82.4
Total	-52052.4	-26229.6	-25766.4	-56.3	-51975.5	-26230.6	-25699.0	-45.9	10.4

^aSee Equation (C5); ^b $\Delta = E_{\text{Complex}} - E_{\text{JH1}} - E_{\text{JH2}}$; ^c $\Delta\Delta = \Delta_{\text{V617F}} - \Delta_{\text{Wild-type}}$. The uncertainties were estimated from the block average method (C21), with block size of 1ns.

Asymmetric vs anti-symmetric models. Two different molecular mechanisms are behind the proposed dimeric forms of the kinase(-like) domains of JAK2 for its activation and regulation. In the anti-symmetric conformation, the A-loop and α C-helix of kinase domain JH1 directly interact with the pseudokinase JH2 domain (Figure C1a,c). Regulation is achieved by restricting the A-loop in a closed conformation, while maintaining the JH1 domain in an inactive conformation. Mutations change the inter-domain interactions, causing the A-loop of JH1 to become extended and hence the active site is highly accessible (C22).

In the asymmetric model, one domain acts as a receiver while the other acts as an activator(C23). The receiver can adopt a structurally active conformation with specific A-loop conformation and α C-helix orientation (Figure C1d), although JAK2 may still remain in a catalytically inactive state when the JH2 domain acts as the receiver. In this model, the α C-helix but not the A-loop of the receiver interacts with the other domain directly. The inter-domain interaction rotates the α C-helix to the active state which then promotes the conformational switch of the A-loop from an inactive packed position to an active extended more open form (Figure C1d). The activation and regulation of JAK2 are determined by how well the JH1 domain occupies the receiver position. Although both JH1 and JH2 can act as the receiver, under most physiological conditions the JH2 domain is believed to adopt this position, putting JAK2 into a catalytically inactive state. Mutations in the JH2 domain weaken the JH1-JH2 interaction, making JH1 more likely to adopt the receiver position.

C3.5 Conclusion

The kinase regulation of JAK2 has been proven to rely on interaction between the two JAK homology domains JH1 and JH2 (C24), which maintains JH1 in an inactive conformation. In this study, a molecular mechanism has been investigated for this regulation caused by an asymmetric dimerization of JH1 and JH2 domains, and the mutational effects analyzed with respect to the association of the JH1–JH2 dimer. We have built an asymmetric dimer of JH1 and JH2 domains, and

used molecular dynamics simulations to study the JH1–JH2 interactions responsible for the negative regulation of JH2 on JAK2 activation.

This asymmetric dimer exhibits interactions reminiscent of those found in members of the ErbB family and other protein kinases. The EGFR/ErbB2 heterodimer consists of a specific arrangement after EGF stimulation, in which EGFR adopts the receiver position in the asymmetric dimer and is activated first (C23). EGFR activation is followed by separation of the EGFR/ErbB2 dimer, and reformation of the dimer with ErbB2 now in the receiver position. This mechanism, and the corresponding structural arrangement of EGFR/ErbB2, are likely to be utilized by other ErbB dimers (C23), and possibly by other kinases. We hypothesize here that, under normal physiological conditions, JAK2s are brought together in an unproductive configuration. Their phosphate-transferring domains form an asymmetric homodimer in which the pseudokinase domain JH2 occupies the position of “receiver”, while the kinase domain is the “activator” (Figure C5a). Upon cytokine receptor activation, the inhibitory JH1–JH2 conformation is displaced and the domains are rearranged into an active state (C25). The JH1–JH2 model proposed in this study could be one dimer in a tetramer which is formed by four phosphate-transferring domains from two JAK2s. The existence of tetramers and even higher order oligomers has been demonstrated for EGFR (C26) and other protein kinases (C27).

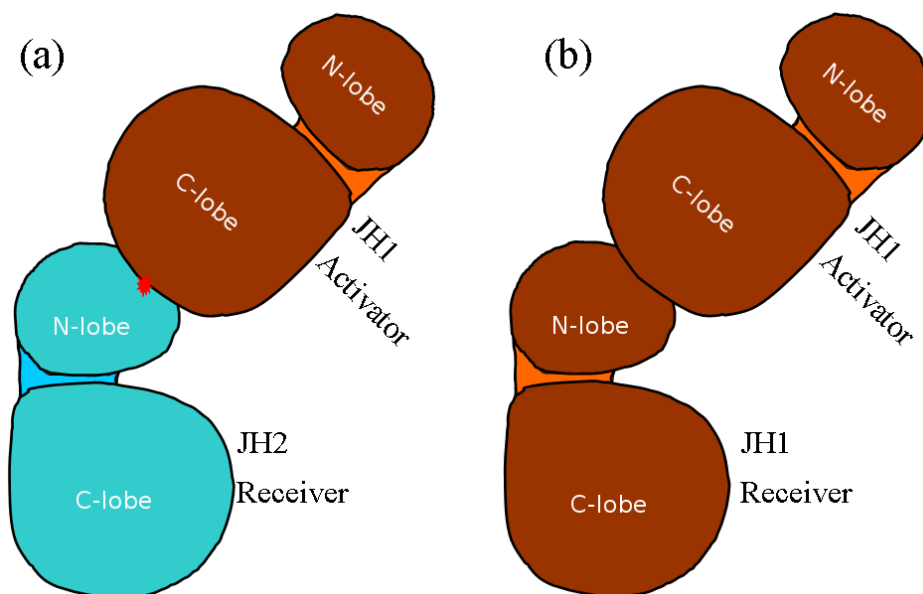


Figure C5 A model for the function of the JH2 domain in activating JAK2. Under normal physiological conditions, JAK2 is autoinhibited through an asymmetric dimer in which JH2 (blue) acts as the receiver and JH1 the activator (orange) (a). Activating mutation (red star) at the dimer interface (a) results in the separation and rearrangement of the kinase domains, possibly forming a JH1 homodimer (b) and leading to the activation of the JH1 domain.

Our work explains the abolition of the JH2-mediated regulation of JAK2 by the activating mutation V617F. The interactions between the two domains in the asymmetric dimer (Figure C5a) preclude the JH1 domain from adopting an active conformation. The V617F mutation in JH2 relieves negative regulation by preventing JH1–JH2 association from the asymmetric form as shown in Figure C5a. After their separation, the JH1 and JH2 domains may re-arrange to interact in other conformations, possibly including a JH1–JH1 homodimer (Figure C5b)(25) because of the favourable interaction between two JH1 domains (C24). The two domains could still be arranged in an asymmetric form, but one of the kinase domains JH1 takes the receiver position and adopts a conformation that fully activates JAK2 kinase (Figure C5b). Based on the JH1–JH2 model proposed in this study, the dual

(both negative and positive) regulation of the JH2 domain (C17) could be explained by the possible weakened/strengthened JH1-JH2 interaction induced by JH2 mutations, which makes the domains easier/harder to separate for conformational re-arrangement. It will be important to obtain x-ray crystallography structures of JH1-JH2 domains as these may help to establish whether JAK2 is regulated through the asymmetric dimer, as suggested by our study. Verification of the mechanism proposed here must await detailed structural and biochemical analyses.

C4 Other applications in progress

C4.1 Further analysis of EGFR simulations

In Deliverable D12.2, we reported a long MD study of EGFR using Anton resources (C28). Close inspection of the conformations and interactions within the binding pocket reveals, converse to the wild type, that the mutant EGFR prefers to bind gefitinib, a targeted anticancer drug, rather than ATP, offering an explanation for why gefitinib is more effective in patients with EGFR mutations than those without. Further analyses of the trajectories show that part of the EGFR undergoes a conformational transition into a disorder state, which was confirmed by recent long-timescale molecular dynamics simulations and biophysical experiments (C29). Many proteins display functions requiring intrinsic disorder. The disorder in the kinase domain of EGFR locates at the interface of asymmetric dimer, indicating that the dimerization process may utilize the folding upon binding mechanisms. The involvement of disorder-to-order transformation is useful for enabling high specificity coupled with low affinity, which makes it easier for two EGFRs to form a dimer. The mutation makes the kinase domain more disordered, and hence easier to form a dimer than wild-type. Further analyses, including quantifying the degrees of disorder, the preference of forming a dimer, are in progress.

C4.2 KRAS simulation and analyses

To study the activation mechanism, we have performed targeted MD (TMD) simulations for both wild-type and mutant KRAS. The conformations are guided from an active/inactive state towards an inactive/active state by means of steering forces, respectively. TMD is a very useful method for the simulation of large conformational transitions. Analyses in progress include: conformational changes during the simulations, principal component analysis, energy profiles, etc. This application aims at unravelling the dynamic and energetic information encoded in the mutant KRAS, and revealing its activation mechanism. The study could explain the inherent activation induced by activation mutations, which makes its activity independent to upstream EGFR activation.

C5 References

- C1. Karplus, M.; McCammon, J. A. *Nature Structural Biology* 2002, 9, 646-652.
- C2. Wan, S.; Coveney, P. V. *J.R.Soc.Interface* 2011, 8, 1114-1127.
- C3. Matthews, D. J.; Gerritsen, M. E. *Targeting Protein Kinases for Cancer Therapy*; John Wiley & Sons, Inc.: New Jersey, 2010.
- C4. Fiser, A.; Sali, A. *Macromolecular Crystallography*, Pt D 2003, 374, 461-491.
- C5. Stamos, J.; Sliwowski, M. X.; Eigenbrot, C. *Journal of Biological Chemistry* 2002, 277, 46265-46272.
- C6. Antonyamy, S.; Hirst, G.; Park, F.; Sprengeler, P.; Stappenbeck, F.; Steensma, R.; Wilson, M.; Wong, M. *Bioorganic & Medicinal Chemistry Letters* 2009, 19, 279-282.
- C7. Sadiq, S. K.; Wright, D.; Watson, S. J.; Zasada, S. J.; Stoica, I.; Coveney, P. V. *J.Chem.Inf.Model.* 2008, 48, 1909-1919.
- C8. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J.Chem.Phys.* 1983, 79, 926-935.
- C9. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; Kollman, P. J. *Comput.Chem.* 2003, 24, 1999-2012.
- C10. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J.Comput.Chem.* 2005, 26, 1781-1802.
- C11. Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. *J.Am.Chem.Soc.* 1998, 120, 9401-9409.
- C12. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J.Comput.Chem.* 2005, 26, 1668-1688.
- C13. Dixit, A.; Verkhivker, G. M. *PLoS Comput.Biol.* 2009, 5, e1000487.
- C14. Mi, L. Z.; Lu, C. F.; Li, Z. L.; Nishida, N.; Walz, T.; Springer, T. A. *Nat.Struct.Mol.Biol.* 2011, 18, 984-989.
- C15. Lindauer, K.; Loerting, T.; Liedl, K. R.; Kroemer, R. T. *Protein Engineering* 2001, 14, 27-37.
- C16. Krieger, E.; Joo, K.; Lee, J.; Lee, J.; Raman, S.; Thompson, J.; Tyka, M.; Baker, D.; Karplus, K. *Proteins: Struct.Funct.Bioinf.* 2009, 77, 114-122.
- C17. Bandaranayake, R. M.; Ungureanu, D.; Shan, Y.; Shaw, D. E.; Silvennoinen, O.; Hubbard, S. R. *Nature Structural & Molecular Biology* 2012, 19, 754-759.
- C18. Dusa, A.; Staerk, J.; Elliott, J.; Pecquet, C.; Poirel, H. A.; Johnston, J. A.; Constantinescu, S. N. *Journal of Biological Chemistry* 2008, 283, 12941-12948.
- C19. Mead, A. J.; Rugless, M. J.; Jacobsen, S. E. W.; Schuh, A. *New England Journal of Medicine* 2012, 366, 967-969.
- C20. Gordon, G. M.; Lambert, Q. T.; Daniel, K. G.; Reuther, G. W. *Biochemical Journal* 2010, 432, 255-265.
- C21. Wan, S. Z.; Stote, R. H.; Karplus, M. *J.Chem.Phys.* 2004, 121, 9539-9548.
- C22. Lee, T. S.; Ma, W. L.; Zhang, X.; Kantarjian, H.; Albitar, M. *Bmc Structural Biology* 2009, 9.
- C23. Macdonald-Obermann, J. L.; Piwnica-Worms, D.; Pike, L. J. *Proceedings of the National Academy of Sciences of the United States of America* 2012, 109, 137-142.
- C24. Saharinen, P.; Takaluoma, K.; Silvennoinen, O. *Molecular and Cellular Biology* 2000, 20, 3387-3395.
- C25. Saharinen, P.; Vihinen, M.; Silvennoinen, I. *Molecular Biology of the Cell* 2003, 14, 1448-1459.
- C26. Clayton, A. H. A.; Walker, F.; Orchard, S. G.; Henderson, C.; Fuchs, D.; Rothacker, J.; Nice, E. C.; Burgess, A. W. *Journal of Biological Chemistry* 2005, 280, 30392-30399.
- C27. Lemmon, M. A.; Schlessinger, J. *Cell* 2010, 141, 1117-1134.
- C28. Wan, S. Z.; Wright, D. W.; Coveney, P. V. *Molecular Cancer Therapeutics* 2012, 11, 2394-2400.

C29. Shan, Y. B.; Eastwood, M. P.; Zhang, X. W.; Kim, E. T.; Arkhipov, A.; Dror, R. O.; Jumper, J.; Kuriyan, J.; Shaw, D. E. Cell 2012, 149, 860-870.

Chapter 6 [M]

TUMOUR BIOMECHANISM FOCUSED MODELS: VARIOUS MOLECULAR MODELS

M1 Spatiotemporal and population-based modeling of P53 and its effects on proliferation and differentiation of tumor stem cells (SCs)

Background Tumours are a wide family of diseases, exhibiting a wide span of phenotypic characteristics. Two common features are shared by all kinds of tumors: the derangement of both cell proliferation and differentiation. The p53 gene is mutated in half of all existing tumours. This renders p53 the main oncosuppressor gene deserving the title of the guardian of the genome. Not surprisingly, p53 is heavily involved in both the control of proliferation and cellular differentiation. As far as proliferation is concerned, it is well known that p53 triggers the blockage of cell cycle, and also apoptotic pathways in response to moderate and strong stress signals respectively. As far as cellular differentiation is concerned, p53 tightly controls the polarity of division of stem cells (SCs) (Cicalese, Bonizzi et al., Cell, 2009). Suppression of p53 induces a strong increase of the probability of symmetric differentiation of breast SCs (Cicalese, Bonizzi et al., Cell, 2009). Moreover, drug-driven activation of p53 induces rapid differentiation of human embryonic SCs, as has been shown in (Maimets et al., Oncogene, 2008).

It is important to stress that p53 under-expression is often observed in tumours carrying wild-type p53, which is particularly frequent in children (Sonnemann et al., Eur J Canc, 2011). This phenomenon is caused in those tumours by the over-expression of the main competitor of p53: the Mdm2 protein (Vogelstein et al., Nature, 2000). For example, this may occur when gene p14, which inhibits Mdm2 by sequestering it into the nucleolus, is deleted. Deletion of p14 has been observed in breast, brain and lung cancers (Vogelstein et al., Nature, 2000). Another avenue leading to the underexpression of wild type p53 is given by Mdm2 amplification (Vogelstein et al., 2000; Tovar et al, PNAS 2006). Finally, binding to viral proteins in infected cells is the third major cause of underexpression of p53 (Vogelstein et al., 2000), and it is typical of chronic-infections- related cancers, such as liver cancers (HCV-related), cervix cancer (HPV related) and lymphomas. The relevance of mutation-independent dysfunctions of p53 has been summarized by Vogelstein, Lane and Levine by the following phrase: "the amount of p53 protein in cells is determined mainly by the rate at which it is degraded, rather than the rate at which it is produced". Due to the above outlined characteristics, underexpressed wild-type p53 is seen as a primary candidate target for antitumor therapies (Vogelstein et al., Nature, 2000).

Among the p53-targeting drugs, a family of small molecules play a prominent role: nutlins (Vassilev et al., Science, 2004). Nutlins are able to bind to Mdm2 exactly in the "binding pockets" where p53 binds, thus impeding the formation of p53-Mdm2 complexes, first described in (Vassilev et al., Science, 2004), thus inducing a rapid increase of levels of p53. Since the activation of p53 causes both a reactivation of the apoptotic pathways and the trigger of differentiation of tumor SCs, nutlins are regarded as a potentially important antitumoral agent. In their study, Vassilev and coworkers clearly showed that Nutlins exhibit a potent antitumor action on wt-p53 tumor cell lines, where the cell viability is halved at low concentration (2 μ M). They also showed that nutlins have only a marginal

effect on mutant p53 cell lines (such as SW 480, MDA-MB-435, PC3). From a dynamical and macroscopic point of view, Vassilev and colleagues also showed that the pattern of action of Nutlin-3 on a human osteosarcoma (SJSA-1) xenograft on nude mice is similar to the one of doxorubicin. In this cellular line, Vassilev and coworkers showed that the p53 is under-expressed due to gene amplification of Mdm2, which is amplified of roughly 25-folds, although very likely not all the extra copies are fully functional (Tovar et al, PNAS, 2006). The same research group (Carvajal et al., Cancer Research, 2005) showed as different subtypes of Nutlins may have a differential action on tumour cell lines. Namely, both Nutlin3-a and Nutlin3-b were very effective on HCT116 colon cancer cell, whereas on colon cancer RKO cells the Nutlin 3-a was considerably less active.

A number of other preclinical studies showed that Nutlin is a potent antitumor drug for cellular lines from important types of cancers carrying dysfunctional wild type p53, such as: B-cells chronic lymphocytic leukemia, prostate cancer, Hodgkin lymphoma, Ewing's sarcoma, glioblastoma multiforme. It has been observed that the Nutlins-mediated restoring of p53 levels does not automatically guarantee beneficial effects if other modules of p53-related pathway are dysfunctional. Finally, although in all the above mentioned tumours the effect of Nutlin is believed to lie mainly in the activation of apoptotic pathways, it is important to remark that in the above mentioned studies concerning the role of p53 in the differentiation of SCs - embryonic (Maimets et al., Oncogene, 2008) and breast cancer (Cicalese et al, 2011) - the drug used for activating p53 is Nutlin. This evidence suggests that Nutlin might induce synergies between increased apoptosis and reduction of symmetric division of cancer SCs.

One may roughly summarize the above experimental findings on the effect of Nutlin on wt-p53 tumours as follows. The binding of Nutlin to Mdm2 by eliminating a competitor to p53 implies an increase of the latter which positively influences tumor response. This is achieved in part due to the onset of massive apoptosis and in part - for SCs based tumors – due to the establishment of a more physiological pathway of cellular differentiation.

Recent experimental results. Very recently, the role of p53 in determining the cell fate and the role of nutlins in modulating the senescence-related function of p53 have been explored by means of computation-driven “wet” experiments by Lahav and coworkers (Purvis et al., Science, 2012). In particular they showed that the pulsed dynamics of p53 molecule numbers may be transformed in a sustained more or less constant level. This changed dynamics modifies the cell fate and induces an acceleration of senescence-related factors.

Our work The aim of the present work is a preliminary attempt to build quantitative models in which the main currently known biochemical interplay between p53 and mdm2 is explicitly included and to compare (to some extent) the models with experimental data.

- For the effects of p53 on cellular proliferation and apoptosis, we have finalized a plausible stochastic model of the p53/Mdm2 interplay and of the Nutlin pharmacodynamics on the single-cell level.
- We have developed an individual based model of cellular proliferation of a stem cell based cellular population, both in the absence and in presence of therapies
- In order to study the effects of p53 on SC differentiation, we have executed a series of simulations of a stochastic mean field plausible model of symmetric/asymmetric pattern formation in proliferating cells. This model is based on the recently introduced paradigm of wave-pinning for the pattern formation (Mori et al, Biop J, 2008; Walther et al, Bull Math Biol, 2012).

- In order to describe the stem cell symmetric/asymmetric division, we have defined a second model that significantly extends the Gamba et al model (PNAS 2005) on the role of PTEN in determining cellular polarization
- All the above models have to take into account stochastic effects, both intrinsic and extrinsic. Modeling of symmetric/asymmetric cell division has also to take into account spatial noise.

Although in the above summarized biological literature on p53 and the preclinical literature on Nutlins the experiments focus on Mdm2, another major player shaping the dynamics of p53 network is PTEN. Indeed p53 is a transcription factor with positive feedback for PTEN which in turn (through PIP and AKT) induces cytoplasmatic Mdm2. Since these processes are enacted within timescales not substantially different from those typical for p53- and Mdm2 related processes, one cannot eliminate PTEN via a quasi-steady state approach. On the contrary, since we are interested in analyzing the dynamic response of the cell to time-varying Nutlin concentration and the symmetric/asymmetric pattern formation, PTEN is a primary actor whose subnetwork has to be integrally represented.

M2 Modeling the pharmacodynamics of Nutlin

This part of the work has been done in collaboration with A. Gandolfi (IASI CNR, Rome, Italy) and K. Puszinski (Silesian Technical University, Gliwice, Poland). Concerning the pharmacodynamics of Nutlin, we have finalized the work (described in the previous deliverable D12.2) on the modeling of the key mechanism of saturable cellular import of extracellular Nutlin with Mdm2-related activations. We have taken into account not only the Mdm2-p53 interplay but also a detailed description of the major factors interacting with them, mainly PTEN.

With respect to the molecular biomechanism part of the previous deliverable (D12.3) which focused on only one route of the p53 action i.e. the induction of apoptosis, we have put special emphasis here on the role of p53 on the cell cycle and in particular on the suspension of cellular proliferation. Some experimental data have been also fitted.

M3 Modeling the proliferation of Stem-Cell based cellular populations and the effect of Nutlin on their dynamics

This research is done in collaboration with P. Ubezio ("M. Negri" Pharmacological Institute, Biophysics Unit, Milan). We have computationally investigated the role of a Nutlin-based treatment on the dynamics and phenotypic distribution of a physiologic or neoplastic cellular population (e.g. of a mammosphere) driven by SCs, in particular with reference to the above-mentioned model of Nutlin pharmacodynamics and to the experiments described by (Cicalese, Bonizzi et al 2009).

A classical biological description is the one representing the structure of such a cellular population as composed of various kinds of cells: Active SCs, Quiescent SCs (QSCs), Progenitor Cells (PCs) (i.e. semi-differentiated cells that have no or very limited renewal ability), mature cells and, in some cases, Quiescent Progenitors. It is noted that with the term Quiescent Cell one can describe in reality three kinds of cells:

1. Quiescent cells that have stopped proliferating (i.e. that entered in the so called G0 phase) due to a program aimed at having a reservoir of SCs in case of “emergency scenarios. After some time (very likely quite long) they restart cycling (i.e. that exit from the G0 phase)
2. Quiescent cells that have stopped proliferating because of an anti-tumor therapy. They should restart cycling when the level of drug decreases under a given threshold
3. Quiescent Cells that stopped cycling due to poor levels of oxygen or nutrients. They restart cycling after the level of nutrients is again adequate. Otherwise, if the nutrients further decrease they become necrotic.

It is noted that in cases 2 and 3 quiescence is environment-driven. Initially we shall only consider the first two classes of quiescent cells (QCs), since in order to be properly taken into the account, the third class requires a spatial description of the dynamics of the system. The above view has been recently challenged by some biomathematics-driven experimental studies performed by Trumpp and coworkers (Wilson et al, Cell, 2008) and by Mantz and collaborators (Takizawa et al, J Exp Med, 2011), where it is hypothesized that QSCs are unnecessary for explaining cell population dynamics. The latter can be described as composed by multiple and heterogeneous SCs sub-phenotypes with remarkably different cell cycle durations.

Namely:

1. In (Takizawa et al, J Exp Med, 2011) it is assumed that in a cellular population multiple classes of SCs coexist without differentiation links between them and that they show a wide range of cell cycle lengths
2. In (Wilson et al, Cell, 2008) and in its follow-up computational paper (van der Wath et al PlosOne 2009) it is suggested that a hierarchy of two types of SCs exists, where the first class is very slowly dividing, and these cells can semi-differentiate into more fast dividing SCs. This second class, in turn, generates the remaining population by differentiating into progenitor cells etc.

We, however, believe that there is a wide corpus of experimental work indicating that the presence of quiescent cells is fundamental. In other words, we believe that the above-postulated presence of very-slowly dividing SCs (in parallel/series with more fast dividing SCs) is a bona-fide mathematical artifact due to the absence of QSCs in the above two models. However, the presence of multiple classes of hierarchically organized SCs by (Wilson et al, Cell, 2008) is very intriguing and should be explored, since it would allow a fine-tuning of the cellular multipotency. This might also contribute to a better understanding of tumor heterogeneity.

Within this framework we have designed and implemented (in C++ programming language) an individual-based model where each single cell belonging to the various cellular sub-types is explicitly represented as a distinct “object”. This kind of models allows for a realistic description of the population dynamics. Indeed, for example, the lifetime of each cell is represented as a random variable with a generic bell-shaped distribution, instead of an exponentially distributed one, which is the standard underlying assumption of mesoscale Gillespie Dynamics (Gillespie, J Phys Chem, 1977). Indeed, the assumption of exponential distribution for lifetimes is perfectly realistic for chemical dynamics, but is only a rough adaptation for population dynamics. Of course, the price to pay is that the system is no more Markovian and each individual agent has to be taken into the account. This also implies that in mean field dynamics partial integro-differential equations are needed, where the time and also the “cell ages” (since the start of their cell cycle) have to appear.

As far as the cell types are concerned, the structure of the population is as follows:

1. We consider n classes of stem cells (SCs: Sc_1, Sc_2, \dots, Sc_n), and p classes of progenitors (P_1, \dots, P_p), and a class of mature cells: M .
2. A SC of class j may become a SC of Class $j+1$ if $j < n$. If $j = n$, then the newborn cells may become of type P_1 .
3. Similarly, a progenitor cell of class j may become a PC of Class $j+1$ if $j < p$. If $j = p$ then both the newborn cells become of class M

Roughly speaking, the above structure can be described serially, since the cell types can be ordered on a segment. However, without remarkable changes, one might consider a finer and more realistic structure where the various classes can be organized along a tree.

In the absence of therapy the proposed algorithm is as follows:

1. When a SC is born, its cell cycle length T_c is extracted from a given bell-shaped distribution. It is also determined if it has to enter quiescence, with a given pre-established probability. Its apoptosis probability and death time (within its cycle length) are also determined.
2. If the newborn SC has to enter quiescence the time T_q when it will get quiescent is extracted (from a given probability density). Of course, $T_q < T_c$. Here we have considered two sub-cases: A and B. In Case A, the cell enters quiescence at the end of G1. In other words, T_q is the duration of its G1 phase. In Case B we consider that a cell enters quiescence just after it is born (Ubezio et al, Phys Rev E, 1998). In such a case $T_q = 0$, and the cell becomes quiescent immediately.
3. When a SC becomes quiescent the time T_b when it will restart cycling is stochastically determined. It should be (and we expect that) $T_b \gg T_c$, but this is not a constraint we set.
4. When a SC just after exiting from quiescence restarts cycling, its residual time of cycling is $T_{res} = T_c - T_q$. This means that the object representing a QSC has memory of its past as an active cell. In other words, here quiescence is transient.
5. When a SC divides, the differentiation fate of the two daughter cells has to be decided. Indeed, based on a given probability of symmetric division, it is decided whether one or both the daughter cells are SCs or if none of them remains SC (see below for the case of therapy).
6. In the case of asymmetric division and of one class of SCs, one of the daughter cells belongs to a class of first order progenitor, the other remains SC. In case of multiple classes of SCs, if the division is asymmetric, one of the two daughters of a SC of class j becomes a SC of class $j+1$, unless n is the last class of SCs. In such a case, one of the daughter cells enters in the first class of progenitor cells.
7. As far as progenitor cells (PCs) are concerned, from a mathematical point of view their dynamics is formally identical to the one of the SCs. However, there is a quantitative fundamental difference: the probability of symmetric division for PCs is null or small, depending on whether or not we allow that PCs have a residual renewal ability.

In the presence of therapy:

1. Each cell embeds its stochastic model of p53-Mdm2-PTEN (etc..) network and their interplay with Nutlin
2. The events symmetric/asymmetric division are decided based on the number of p53 molecules at the division time
3. The death of the cell is determined also by the level of p53.
4. Going into and coming back from the quiescent state is no more uniquely determined based on pre-assigned probabilities, but it also depends on the levels of p53. This is a quite complex point of which we have to find a satisfactory solution

5. Extrinsic stochasticity (e.g. bounded noise in drug pharmacokinetics, and in the rate constants of the bio-molecular species) will be considered, in line with (d’Onofrio and Gandolfi, Phys Rev E, 2010)

It is noted that the inclusion of the therapy makes the model multi-scale, since four time-scales have to be considered: the dynamics of the cellular population, the dynamics of the p53 network, and the dynamics of the drug delivery and the drug pharmacokinetics.

The simulation of the above stochastic model requires many repetitions in order to infer the heuristic time varying probability densities of the number of various classes of cells. However, in absence of therapy, for given experimental data numerous simulations are needed in order to fit the parameter values to the data, such as the quiescence and the symmetric division probabilities, and the average and variance of the probability densities of T_c , T_q , T_b etc... This means that a very large parameter space has to be explored.

Thus it was deemed necessary to consider a mean field model where the dynamics of the heterogeneous cellular population is modeled by appropriate differential equations with delays (Bertuzzi et al, Math Biosci, 1997; Ubezio et al., Phys Rev E, 1998).

Initially, we will fit available data by using this simplified model. This will allow inferring a guessing the average values of the various parameters including variances. A Monte-Carlo fitting procedure will be thus employed by using as starting points the above-mentioned guesses.

In the third stage, we shall also include some important features:

1. The inclusion of space. This will require at least to include: i) the description of random movements of single cells; ii) the description of their mechanical interplay; iii) the description of the diffusion of nutrients and of the drug. For points i and ii we shall take into the account the previous literature, for examples the works by D. Drasdo and coworkers (Ramis-Conde et al, Biop J, 2008)
2. The inclusion of intercellular signaling modulating, for example the exit from quiescent state of QSCs.

It is noted that, whereas the basic model is linear in the sense that the cells do not interact, the addition of nutrients and intercellular signaling will induce a strong degree of nonlinearity due to competition of nutrient and space, and to the feedback related to the intercellular signaling. In parallel we shall consider how to adapt this approach so as to potentially link it to the clinical Oncosimulator framework.

M4 The p53-regulated SCs Polarity in Breast Cancer: the wave-pinning model

Background As stated by (Onsum and Rao, Curr Op Cell Biology, 2009) cells are dynamically reorganized responding to both external inductions and internal spontaneous cues including random signals. In particular, spontaneous forming of specialized internal domains of regulatory elements is crucial for cellular differentiation. As we have already mentioned, p53 seems involved also in forming these domains in SCs (Cicalese, Bonizzi et al., Cell, 2009). This new and fascinating role for p53 might

be of paramount relevance for anti-tumor therapies, but its study has some an additional layer of difficulty with respect to classical systems biology modeling, due to the inherent spatial nature of the problem. A further layer of difficulty is due to the fact that properly taking into the account realistic spatiotemporal extrinsic stochastic phenomena is quite complex. The extrinsic stochastic fluctuations perturbing kinetic rate constants have to be bounded in order to avoid artifacts as stressed in other contexts (d’Onofrio, Phys Rev E, 2010). Thus one must model the perturbations as spatiotemporally bounded stochastic processes, as those described in the previous report (d’Onofrio and de Franciscis, Phys Review E, 2012; submitted, 2012b) (both works are on www.arxiv.org).

A model that has been advocated for manifold biological pattern formation is the classical Gierer-Meinhardt paradigm of the small range activator- large range inhibitor, based on the so-called Turing bifurcation, and the similar model of Substrate-Depletion (Murray, Mathematical Biology, Springer, 2002). Recently, in (Mori et al, Biop J, 2008) some doubts have been cast on the adequacy of this mechanism to explain sufficiently rapid intracellular domain formation. As an alternative and more rapid mechanism in (Mori et al, Biop J, 2008) the so-called wave-pinning it has been proposed consisting of the formation of advancing fronts of large protein concentrations; fronts that decelerate and then freeze. From the biochemical point of view, this mechanism may arise when a chemical is present in two forms: an active form A that is bounded to the cellular membrane, and an inactive form I that freely diffuses into the cytosol. A interconverts into I and vice-versa. Due to the difference in diffusion coefficients and to the constant total amount of the chemical it has been shown (Mori et al, Biop J, 2008) that in one dimension this biochemical mechanism produces the wave-pinning phenomenon, and that some key proteins linked to cell polarization such as rho-GTPases fit well into this framework.

Results. The simulations (d’Onofrio and de Franciscis, 2012c, submitted, available on www.arxiv.org) show that the wave pinning mechanism is relatively robust in case of cued polarization, which is fundamental in case of chemotaxis, where the cell polarization is the response to an external gradient. In case of spontaneous cell polarization, in absence of extrinsic noise it seems less adequate since in this case part of cells do not polarize at all, but exhibit a central pattern (i.e. the factors concentrate in the middle of cells, thus forming a spatial pattern but not a polarization).

However, quite interestingly, if a moderate amount of external noise is added, the model suggests that no more cells exhibits a central non-polarized pattern. The remaining cells either are polarized or they have a uniform concentration of polarization-related factors.

This is shown in Fig. M1, where each distribution of the concentration of factor A (e.g. a rho-GTPase) is represented by a point. In this plot the X axes indicates the direction of polarization. Thus if $X = 0.5$ then there is no polarization. The Y axis indicates the variance of the concentration of A. Thus if $Y = 0$ then the distribution of A is flat. B indicates the amplitude of noise. We see that for a low noise level, in many cases one may observe a central pattern corresponding to no-polarization. In no cases the distribution of chemical is uniform. For a moderate level of noise instead, the anomalous central pattern disappears and the cells are either polarized or non-polarized with uniform distribution of A.

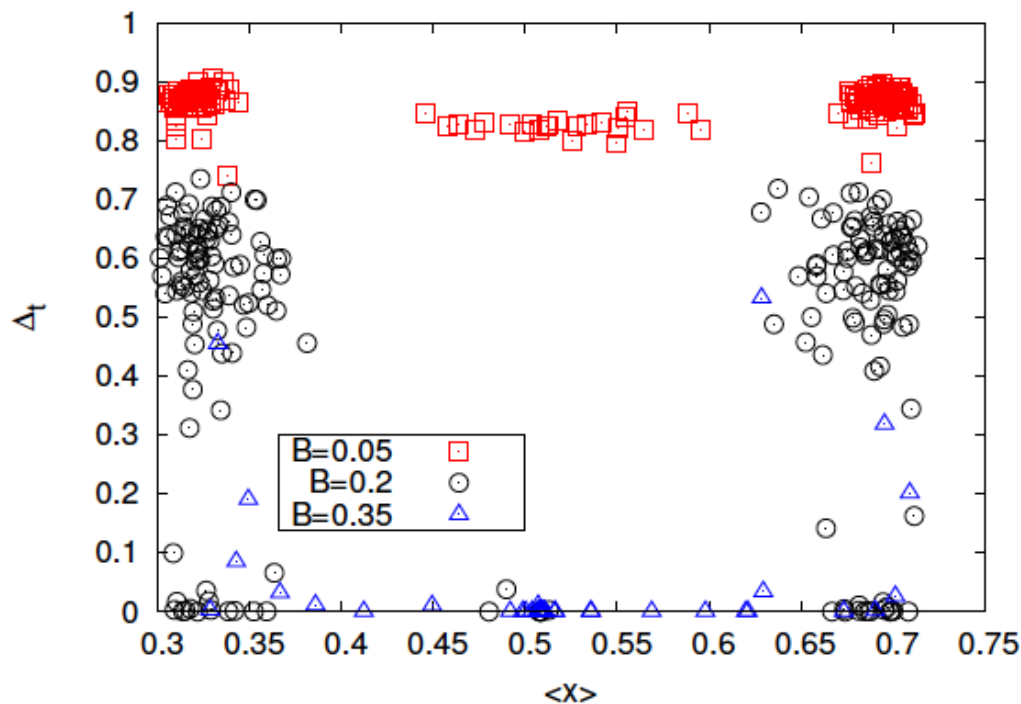


Fig. M1 Effect of spatiotemporal noise on the degradation rate d of the chemical A. B denotes the noise amplitude (express as fraction of d).

M5 The p53-regulated SC Polarity in Breast Cancer: the phase model

This research component is being done in collaboration with A. Gamba (Turin Polytechnical University, Turin, and “Candiolo Cancer Research Centre”, Candiolo). The above results obtained by considering a toy-model of wave-pinning stimulated us in finding a more realistic model to describe the role of p53 in determining the polarization of cancer SCs, and the role of spatiotemporal extrinsic noises. We found that in 2005, Gamba and colleagues (Gamba et al, PNAS 2005; Gamba et al. J Stat Mech, 2009; Semplice et al, PLoSOne, 2012) proposed an interesting stochastic model of chemotaxis-induced spatial symmetry breaking in a single cell. The central core of this model was the interplay between two membrane-bound molecules and two cytosolic molecules. The presence of some feedbacks (with related bistability) and the difference of diffusion coefficients induced a mechanism of phase separation that is different from the ones based on Turing's bifurcation. Indeed, this new model is based on the Lifshitz-Slyozov patch coalescence mechanism (Lifshitz and Slyozov, J Phys Chem sol, 1961) underlying the dynamics of many phase transitions

Quite interestingly:

1. In the supplementary materials of (Gamba et al, PNAS 2005) Gamba et al. proposed a simplified and mean-field version of their model that exhibits, in one spatial dimension, a traveling wave whose velocity decreases until the front stops. In other words, the wave-

pinning mechanism described in the previous section is a particular case of the phase-separation model by Gamba et al

2. In (Gamba et al J Stat Mech, 2009; Semplice et al PloSOne, 2012) it is suggested that chemicals involved in the polarization of epithelial cells might be PIP2, PIP3 (in the membrane) and PTEN, PIK3

Thus, since PTEN is a fundamental part of the p53-mdm2 network, we believe that the experimental results obtained in (Cicalese, Bonizzi et al, Cell, 2009) might be explained by means of a model that we have built, which integrates our model of p53-mdm2 interplay, and the Gamba spatiotemporal model.

M6 References

(Cicalese, Bonizzi et al., Cell, 2009) =

Cicalese A, Bonizzi G, Pasi CE, Faretta M, Ronzoni S, Giulini B, Briskin C, Minucci S, Di Fiore PP, Pelicci PG. "The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells." *Cell*. 2009 Sep 18;138(6):1083-95. doi: 10.1016/j.cell.2009.06.048.

(Maimets et al., Oncogene, 2008) =

Maimets T, Neganova I, Armstrong L, Lako M. "Activation of p53 by nutlin leads to rapid differentiation of human embryonic stem cells." *Oncogene*. 2008 Sep 11;27(40):5277-87. doi: 10.1038/onc.2008.166. Epub 2008 Jun 2.

(Sonnemann et al., Eur J Canc, 2011)=

Sonnemann J, Palani CD, Wittig S, Becker S, Eichhorn F, Voigt A, Beck JF. "Anticancer effects of the p53 activator nutlin-3 in Ewing's sarcoma cells." *Eur J Cancer*. 2011 Jun;47(9):1432-41

(Vogelstein et al., Nature, 2000) =

Vogelstein B, Lane D, Levine AJ. "Surfing the p53 network." *Nature*. 2000 Nov 16;408(6810):307-311

(Tovar et al, PNAS 2006)=

Tovar C, Rosinski J, Filipovic Z, Higgins B, Kolinsky K, Hilton H, Zhao X, Vu BT, Qing W, Packman K, Myklebost O, Heimbrook DC, Vassilev LT. "Small-molecule MDM2 antagonists reveal aberrant p53 signaling in cancer: implications for therapy." *Proc Natl Acad Sci U S A*. 2006 Feb 7;103(6):1888-9

(Vassilev et al., Science, 2004)=

Vassilev LT, Vu BT, Graves B, Carvajal D, Podlaski F, Filipovic Z, Kong N, Kammlott U, Lukacs C, Klein C, Fotouhi N, Liu EA. "In vivo activation of the p53 pathway by small-molecule antagonists of MDM2." *Science*. 2004 Feb 6;303(5659):844-8

(Carvajal et al., Cancer Research, 2005) =

Carvajal D, Tovar C, Yang H, Vu BT, Heimbrook DC, Vassilev LT. "Activation of p53 by MDM2 antagonists can protect proliferating cells from mitotic inhibitors." *Cancer Res*. 2005 Mar 1;65(5):1918-24.

(Purvis et al., Science, 2012)=

Purvis JE, Karhohs KW, Mock C, Batchelor E, Loewer A, Lahav G. "p53 dynamics control cell fate." *Science*. 2012 Jun 15;336(6087):1440-4.

(Mori et al, Biop J, 2008) =

Mori Y, Jilkine A, Edelstein-Keshet L. "Wave-pinning and cell polarity from a bistable reaction-diffusion system." *Biophys J*. 2008 May 1;94(9):3684-97. doi: 10.1529/biophysj.107.120824. Epub 2008 Jan 22.

(Walther et al, Bull Math Biol, 2012)=

Walther GR, Marée AF, Edelstein-Keshet L, Grieneisen VA. "Deterministic versus stochastic cell polarisation through wave-pinning." *Bull Math Biol*. 2012 Nov;74(11):2570-99. doi: 10.1007/s11538-012-9766-5. Epub 2012 Sep 7.

(Wilson et al, Cell, 2008)=

Wilson A, Laurenti E, Oser G, van der Wath RC, Blanco-Bose W, Jaworski M, Offner S, Dunant CF, Eshkind L, Bockamp E, Lió P, Macdonald HR, Trumpp A. "Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair." *Cell*. 2008 Dec 12;135(6):1118-29.

(Takizawa et al, J Exp Med, 2011)=

H. Takizawa, RR Regoes, CS Boddupalli, S Bonhoeffer, and MG Manz, "Dynamic variation in cycling of hematopoietic stem cells in steady state and inflammation," *J Exp Med* vol. 208 no. 2 273-284 (2012)

(van der Wath et al PlosOne 2009) =

van der Wath RC, Wilson A, Laurenti E, Trumpp A, Lió P. "Estimating dormant and active hematopoietic stem cell kinetics through extensive modeling of bromodeoxyuridine label-retaining cell dynamics." *PLoS One*. 2009 Sep 22;4(9):e6972. doi: 10.1371/journal.pone.0006972.

(Gillespie, J Phys Chem, 1977)=

Gillespie, DT “Exact Stochastic Simulation of Coupled Chemical Reactions” The Journal of Physical Chemistry 81 (25): 2340–2361 (1977).

(Ubezio et al, Phys Rev E, 1998)=

F. Montalenti, G. Sena, P. Cappella, and P. Ubezio “Simulating cancer-cell kinetics after drug treatment: Application to cisplatin on ovarian carcinoma,” Phys. Rev. E 57, 5877 (1998)

(d’Onofrio and Gandolfi, Phys Rev E, 2010)=

Alberto d’Onofrio and Alberto Gandolfi “Resistance to antitumor chemotherapy due to bounded-noise-induced transitions,” Phys. Rev. E 82, 061901 (2010)

(Bertuzzi et al, Math Biosc, 1997)=

A Bertuzzi, A Gandolfi, C Sinisgalli, D Iacoviello, “Steel’s potential doubling time and its estimation in cell populations affected by nonuniform cell loss , “ Mathematical Biosciences, Volume 143, Issue 2, (1997), Pages 61-89

(Ramis-Conde et al, Biop J, 2008)=

Ramis-Conde I, Drasdo D, Anderson AR, Chaplain MA. “Modeling the influence of the E-cadherin-beta-catenin pathway in cancer cell invasion: a multiscale approach.” Biophys J. 2008 Jul;95(1):155-65. doi: 10.1529/biophysj.107.114678. Epub 2008 Mar 13.

(Onsum and Rao, Curr Op Cell Biology, 2009)=

Onsum MD, Rao CV, “Calling heads from tails: the role of mathematical modeling in understanding cell polarization.” Curr Opin Cell Biol. 2009 Feb;21(1):74-81. doi: 10.1016/j.ceb.2009.01.001. Epub 2009 Jan 23.

(Cicalese et al, 2009)=

Cicalese A, Bonizzi G, Pasi CE, Faretta M, Ronzoni S, Giulini B, Briskin C, Minucci S, Di Fiore PP, Pelicci PG. “The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells.” Cell. 2009 Sep 18;138(6):1083-95. doi: 10.1016/j.cell.2009.06.048.

(d’Onofrio and de Franciscis, Phys Review E, 2012)=

Sebastiano de Franciscis and Alberto d’Onofrio, “Spatiotemporal bounded noises and transitions induced by them in solutions of the real Ginzburg-Landau model,” Phys. Rev. E 86, 021118 (2012)

(d’Onofrio and de Franciscis, submitted, 2012b)=

S de Franciscis, A d’Onofrio, “Spatiotemporal sine-Wiener Bounded Noise and its effect on Ginzburg-Landau model,” Submitted to Nonlinear Dynamics Preprint available at <http://arxiv.org/abs/1206.6020>

(Murray, Mathematical Biology, Springer, 2002)=

JD Murray, Mathematical Biology, Springer, 2002

(d’Onofrio ad de Franciscis, 2012c, submitted, available on www.arxiv.org)=

Sebastiano de Franciscis, Alberto d’Onofrio, “Cellular polarization: interaction between extrinsic bounded noises and wave-pinning mechanism, “Submitted to phys rev e Preprint: <http://arxiv.org/abs/1212.4996>

(Gamba et al PNAS 2005)=

Gamba A, De Candia A, Di Talia S, Coniglio A, Bussolino F, et al., “Diffusion limited phase separation in eukaryotic chemotaxis. “Proc Natl Acad Sci U S A 102: 16927–16932 (2005)

(Gamba et al J Stat Mech, 2009)=

Gamba A, Kolokolov I, Lebedev V, Ortenzi G, “Universal features of cell polarization processes.” J Stat Mech 2009: P02019

(Semplice et al PloSOne, 2012)=

Semplice M, Veglio A, Naldi G, Serini G, Gamba A, “A Bistable Model of Cell Polarity.” PLoS ONE 7(2): e30977. (2012)

(Lifshitz and Slyozov, J Phys Chem sol, 1961)=

Lifshitz IM, Slyozov VV , “The kinetics of precipitation from supersaturated solid solutions. J Phys Chem Solids 19: 35–50 (1961)

(Cicalese, Bonizzi et al, Cell, 2008)=

Cicalese A, Bonizzi G, Pasi CE, Faretta M, Ronzoni S, Giulini B, Briskin C, Minucci S, Di Fiore PP, Pelicci PG. "The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells." Cell. 2009 Sep 18;138(6):1083-95. doi: 10.1016/j.cell.2009.06.048.

Chapter 7 [CO] CONCLUSIONS

In this document the major aspects of the current developmental status of the p-medicine Oncosimulator and the special biomechanism models have been delineated. The particular biomechanism models, except for their own independent significance might also be used for the refinement of specific submodules of the Oncosimulator. This refers inter alia to the probability of a tumour stem cell to undergo symmetric division and the probability of a tumour cell to undergo apoptosis following chemotherapeutic treatment. It should be noted that the nephroblastoma branch of the Oncosimulator, being an extension of a system developed in previous research projects has reached the most advanced stage. This is why the focus of the nephroblastoma branch in this document has been on its code execution acceleration rather on its pure modeling aspects. The models of both the antiangiogenic agent based targeted treatment of breast cancer and the acute lymphoblastic leukemia (ALL) treatment are still under development and refinement. However, important results have already been achieved for both branches. Important advances in the special biochemical and molecular biomechanisms have also been made.

In conclusion the development of both the p-medicine Oncosimulator and the special biomechanism models is more or less within the bounds of the original plan. Therefore, the implementation of the VPH modeling workpackage of the p-medicine project can be claimed to be successful up to now.

Appendix A - Abbreviations and Acronyms

ALL	Acute Lymphoblastic Leukemia
CAU	Christian-Albrechts-Universität zu Kiel
CTL	Cytotoxic T Lymphocytes
FDR	False Discovery Rate
HR	High Risk
ICCS-NTUA	Institute of Communication and Computer Systems – National Technical University of Athens
IEO	Istituto Europeo di Oncologia
IS	Immune System
MD	Molecular Dynamics
MRD	Minimal Residual Disease
PDB	Protein Data Bank
PSNC	Poznan Supercomputing and Networking Center
SB	Systems Biology
SR	Small Risk
TC	Tumour Cell
UCL	University College London
USAAR	University of Saarland
WP	workpackage
WT	Wilms Tumour (nephroblastoma)