APPLICATION OF DIFFERENTIAL METABOLIC CONTROL ANALYSIS TO IDENTIFY NEW TARGETS IN CANCER TREATMENT

A thesis submitted to The University of Manchester for the degree of
Doctor of Philosophy
in the Faculty of Engineering and Physical Sciences

September 2010

Ettore Murabito

School of Chemical Engineering and Analytical Science

LIST OF CONTENTS

LIST OF FIGURES

Chapter 4

Word Count: 45,528

ABSTRACT

In the quest for anti-cancer drugs with high efficacy and low toxicity, cancer metabolism has increasingly been a focus of interest in clinical research. Enhanced glycolysis and robust production of lactate constitute characteristic traits that discriminate many cancerous cells from their normal counterparts. This, in principle, may provide researchers with a general handle on such a complex disease, regardless of the intrinsic genotypic heterogeneity of the single transformed cells. The work carried out during this project and presented in this thesis consists of developing and applying analytical approaches, mainly drawn from the field of metabolic control analysis (MCA), to the study of cancer metabolism. The ultimate goal is to assess whether, and to what extent, the metabolic features of cancer cells may be exploited in the attempt to attack the malignancy more specifically than through traditional clinical approaches. The underlying idea consists of identifying enzymes that represent points of fragility specifically characterising the cancerous metabolic phenotype. These enzymes are such that an alteration in their activity (due for example to the action of an anticancer drug) would elicit the desired response in cancer cells, without affecting their normal counterparts.

The application of MCA relies on a mathematical representation of the system under study. Creating such a model is often hampered by the lack of data about the precise kinetic laws governing the different reaction steps and the value of their corresponding parameters. The most important result reached during this project shows that the metabolic quantities defining the normal and cancer phenotypes (such as fluxes and metabolite concentrations), together with heuristic assumptions about the properties of typical enzyme-catalyzed reactions, already allow for a fast and efficient way to explore the effectiveness of putative drug targets with respect to criteria of high efficacy and low toxicity. The relevance of this result lies in the fact that the quantities defining a

metabolic phenotype are experimentally more accessible than the kinetic parameters of the different enzymatic steps in the system.

DECLARATION

The author declares that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT STATEMENT

i. Copyright in text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made **only** in accordance with instructions given by the Author and lodged in the John Rylands University Library of Manchester. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the Author.

ii. The ownership of any intellectual property rights which may be described in this thesis is vested in The University of Manchester, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement.

iii. Further information on the conditions under which disclosures and exploitation may take place is available from the Head of School.

PREFACE

The goal of this thesis is to apply metabolic control analysis (MCA) to the study of cancer metabolism and to assess whether and to what extent the metabolic features of cancer cells can be exploited to target them more specifically than through traditional approaches, such as chemotherapy or radiotherapy.

Despite focusing only on the metabolic level of a complex phenomenon like cancer, the area to explore is so vast that it provides room for many different routes of investigation. Because the quantification of the control properties of entire metabolic pathways is impractical from an experimental standpoint, the application of MCA often relies on the creation of a dynamic model of the system under investigation. As a first step of this project, then, we aimed to understand where the boundaries of such a model had to be drawn. Although researchers usually focus on the metabolic alterations occurring in the central carbon metabolism, remarkable metabolic differences between normal and cancer cells may lie in pathways not yet deeply studied within the context of cancer research, and a wider portion of the metabolome might need to be considered. At the other extreme, a comprehensive kinetic model of the entire metabolome, besides being extremely challenging to accomplish, would not be necessarily useful from the perspective of understanding the basic mechanisms underlying the biological processes under study. In fact, if the complexity of the model equals the complexity of the system it is meant to reproduce, the former would not provide any particular insight into the causative relationships between different phenomena occurring in the system[1]. Subsequently, the first step of

[1] One could argue, however, that it is the approach adopted in the study of such a genome-scale model, rather than its size or complexity, which determines the limits of our understanding of the system functioning. Indeed, promoters of MCA support the idea that the application of regulation analysis and MCA itself might provide researchers with the means to reach that understanding, despite the complexity of such a comprehensive representation of the system.

this project was intended to set up a criterion to discriminate between sets of reactions and pathways reasonably worthy of being included in the model, and others which could be reasonably neglected. This study was performed through a constraint-based approach ultimately relying on Flux Balance Analysis (FBA). In particular, an algorithm was created to explore the entire space of minimal FBA solutions (*i.e.* all the minimal sets of active reactions consistent with the mass balance requirement) reproducing the metabolic features observed, separately, in cancer and normal cells. In doing so, we aimed to identify two sets of reactions, one for each of the two metabolic phenotypes, which always carry a non-null flux in all the minimal solutions found for the two phenotypes taken separately. These two sets of reactions would constitute a minimal subnetwork able to host the metabolic changes occurring in the emergence of cancer, hence providing us with a minimal scaffold of reactions that could be used to build up the kinetic model necessary to apply MCA.

The methodology proposed, however, turned out to have strong limitations when applied to a genome-scale reconstruction of metabolism, because of the high computational resources requirements. Nevertheless, the algorithm provided a proof of concept for cases in which one wants to explore and identify the minimal subnetworks of a wider system able to capture the essence of the whole network in respect to specific features.

Because of the inapplicability of this algorithm on a genome-scale level, we carried out our study on MCA using an *in silico* representation of the central carbon metabolism covering the pathways that are commonly taken into account in the study of glucose metabolism. From a mechanistic modelling perspective there are other limitations which have to be taken into account, mainly related to the lack of data about detailed dynamic description of various enzymatic steps and uncertainties in the enzyme kinetic parameters. To assess whether the control properties of the system can be determined when detailed knowledge of the system dynamics is missing, we used a probabilistic approach,

where kinetic parameters such as Michaelis-Menten constants of inhibition/activation constants were sampled randomly. This study was based on the assumption that the two metabolic states under comparison (normal and cancer) were completely determined. The rationale of this assumption is that high-throughput metabolomics and fluxomics techniques are now standard procedures in the analysis of cellular metabolism, making fluxes and concentrations experimentally more accessible than kinetic parameters. Using a simplified reconstruction of the central carbon metabolism, and a paradigmatic cancerous phenotype as an example, we showed how our probabilistic approach can provide researchers with some guidance when developing drugs with high efficacy and low toxicity acting at a metabolic level.

The next step consisted of using this sampling approach in a real case study, where we attempted to constrain to a minimum the uncertainties of the system. A more extensive and detailed model of central carbon metabolism was created, where actual rate laws were used to describe the enzymatic kinetics. The two metabolic states under comparison were described relying as much as possible on the experimental data available in literature. For the cancer phenotype, in particular, we aimed to represent specifically the altered metabolic features of breast neoplastic cells. The remaining unknown or uncertain quantities, in terms of metabolic phenotype and kinetic parameters, were finally sampled from intervals representing the range of their possible values, allowing us to make a probabilistic assessment of the suitability of the different enzymes as drug targets.

HOW TO READ THIS THESIS

Because of the different techniques and approaches used during the unfolding of this PhD project, we deemed that the content of this thesis would be better conveyed in the form of independent chapters. Each chapter represents a self-

contained piece of work with its own *Methods* and *Discussion* sections, and is provided in the form of a manuscript.

Chapter 1 – The first chapter is a literature review introducing cancer metabolism and commenting on a variety of modelling approaches in its study. In particular, the historical milestones marking the increasing acknowledgment of cancer metabolism in computational models of cancer are highlighted. The role of mainstream systems biology approaches, such as MCA and FBA, in cancer research is also considered. Their potential role in elucidating some aspect of cancer metabolism, such as its suitability as a new domain of intervention for the development of anticancer drugs, is discussed.

Chapter 2 – This chapter illustrates the algorithm we proposed to explore the set of minimal subnetworks of a larger system, which are consistent with specific requirement of optimality.

Chapter 3 –This chapter focuses on showing how a random sampling approach can help to highlight differences between the control profile of two distinct metabolic phenotypes in a probabilistic manner. This methodology, which combines Monte-Carlo sampling with MCA, is designed to overcome the limitations due to lack of knowledge about the detailed dynamic properties of the system under study, specifically in the form of the exact value of its kinetic parameters. The study is conducted on a simplified representation of the central carbon metabolism. The metabolic states under comparison are representative of a normal phenotype and a paradigmatic cancer phenotype, where the general metabolic features of neoplastic cells are taken into account.

Chapter 4 – A case study is discussed, where the probabilistic approach to MCA is applied to assess the suitability of different enzymes as molecular targets for a drug designed to attack specifically breast cancer. All the data about the system

dynamics and the two metabolic states under comparison (normal and cancer) are retrieved from literature. The results are presented and discussed with regard to different possible clinical strategies, each aimed to hinder or perturb a specific aspect of cancer metabolism.

Chapter 5 – Possible future developments of the study presented in the previous chapters are discussed.

CHAPTER 1

COMPUTATIONAL APPROACHES TO CANCER METABOLISM

INTRODUCTION

The term cancer covers a variety of complicated, multi-stage diseases, which involve the alteration of many interwined cellular processes and mechanisms [1-4]. The advent of high throughput technologies in biological sciences in the last fifteen years has enormously increased the mass of information about the genes involved in the development of cancer as well as the pathways and subcellular circuitry in which they participate [5-9]. This emerging complex picture has not led accelerated progress in treatment options to the extent that some originally envisaged [10]. However it has challenged biologists to find new ways to reason to exploit this apparently rich set of knowledge.

One response to the difficulty of understanding cancer biology based on a purely discursive reasoning has been to adopt computer-assisted approaches to the study of the disease. Non-linear processes dominate the way in which tumour cells interact with their microenvironment. It is clear that the intuitive, verbal reasoning approaches favoured by many oncologists are insufficient to describe the resulting complex system dynamics. Rather, experience from other areas of science has taught us that quantitative methods are needed to develop comprehensive theoretical models for interpretation, organization and integration of this data. In the past decade, computational attempts to describe and understand cancer have been partly reframed by the advent of systems biology, a new paradigm in biological sciences where the functional properties of life are understood as emerging from and studied on the bases of the molecular mechanisms underlying them [11, 12].

It is the argument of this thesis that the role of metabolism in cancer is an area in which a systems biology perspective has a useful contribution to make, particularly through the application of concepts drawn and developed from the

field of metabolic control analysis. Accordingly, this chapter reviews the background biology of cancer, with a particular focus on its metabolic processes. We survey related previous modelling work, and explore the actual and potential role of some of the conceptual and computational systems biology frameworks in the study of cancer metabolism.

CANCER: A BIOLOGICAL PERSPECTIVE

WHAT IS CANCER?

In pluricellular organisms, cells behave in a cooperative way, adjusting their behaviour in accordance to the needs of the tissue, the organ and eventually the whole organism. All the functions that characterise a cell as a living entity are regulated through a tight interaction between the cell and its surroundings. Specific chemical signals are sent and received by each cell in order to coordinate its behaviour with the needs of the "cellular community". Growth, replication and even death occur when they are specifically required. From an ecological perspective, there is no natural selection occurring within a healthy human body. Self-sacrifice rather than survival of the fittest is the general role governing cell populations in pluricellular organisms. In cancer this "social" behaviour is lost and replaced with a behaviour that is characteristic of competition, rather than cooperation. Cells become insensitive to the signals received from their environment and start to behave autonomously. Cancer begins with a single cell mutating in such a way to gain a selective advantage over its neighbours, allowing it to proliferate more quickly and become the founder member of a growing mutant clone, known as a tumour. Successive rounds of mutation, competition and selection lead to progressively less collaborative and more dangerous cells. Thus cancer development can be viewed as somatic (of the body) evolution [13]. In the next stage of disease progression, the uncontrolled cells forming the initial transformed colony (the

tumour) may invade adjacent tissues and spread to other locations of the body via lymphatic or blood channels, forming metastases. This invasive property is precisely what differentiates cancer from a benign tumour, where the original aggregation of abnormal cells is self-limited and does not invade or metastasise.

In the past decades cancer research has collected an enormous amount of information about the differences between cancer cells and their normal counterparts, with the intent of understanding the genetic mechanisms underlying carcinogenesis and ultimately to develop new anticancer drugs [14]. The emerged picture is overwhelmingly complex, and depicts cancer as a disease involving the disruption and alteration of many molecular mechanisms and processes occurring at different hierarchical levels in the cell. Knowledge about gene alterations in different kinds of cancer has been collected and organised by Feutral et al. (2004) in a 'census' which encompasses almost 300 mutated genes, mostly involved in signal transduction processes, that are causally implicated in oncogenesis [15]. The combined knowledge of these oncogenes, the pathways in which they are involved and the interactions of these pathways with each other can enable us to identify differences in the way cancerous and normal cells process biological information [14]. However, the wiring of the altered interaction map alone can be informative to a very limited extent. The non-linear dynamics of many molecular interactions, the presence of feedbacks and the tight cross-talk between different routes make the outcome of the system highly unpredictable. An even more complex picture emerges when also considering the supra-cellular level, in which cancer cells interact directly (by physical contact) or indirectly (through signalling) with each other or adjacent normal cells. Phenomena such as invasion and angiogenesis (the formation of new capillaries induced by cancer cells through cytokine secretion) can be ascribed to the supra-cellular effects of the altered intracellular interaction network.

The molecular processes and genetic alterations involved in the raising of a tumour cell depend strongly on the kind of tumour. Malignant tumour growth is the result of multiple genetic and epigenetic changes, where each one, taken individually, is insufficient by itself to transform the cell, but a number of changes together lead to cancer when summed by accumulation. These changes are quite heterogeneous and no single genetic defect, set of defects or sequence of defects is found in all cells exhibiting a transformed phenotype [16-18]. Nevertheless a set of common traits has been recognised to be universally present in transformed cells, allowing them to escape their normal behaviour and become malignant. These features, described and summarised in Hanahan et al. [19], are the results of genetic alterations involving genes which play an important role in processes and mechanisms such as growth, mitosis, apoptosis. The final outcome of these alterations is a cell with an uncontrolled proliferation and insensitive to apoptotic signals.

THE HALLMARKS OF CANCER

Six essential alterations in cell physiology have been identified and suggested as collectively dictating malignant growth in all types of cancer [19]: self-sufficiency in growth signals, insensitivity to growth-inhibitory (antigrowth) signals, evasion of programmed cell death (apoptosis), limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis. Each of these physiological changes represents the successful breaching of an anticancer defence mechanism hardwired into cells and tissues.

Normal cells require grow signals in order to move from a quiescent into a proliferative state. These signals, received from neighbour cells or through the endocrine systems, bind specific receptors on the cell membrane, and activate signalling pathways that instruct the cell to enter mitosis. In cancer cells, self - sufficiency with respect to growth signals is achieved through a combination of

three mechanisms: endogenous synthesis of growth signals [20, 21], overespression of cell surface receptors of growth signals [20, 22-25] and alterations in the downstream circuitry that receives and processes the signals emitted by ligand-activated growth factors receptors [26-28].

Besides evading normal growth regulation, cancer cells also acquire the ability to evade antiproliferative signals, which constitute another regulatory mechanism through which quiescence and homoeostasis is maintained in normal tissue. In different stages of the cell cycle, normal cells pass through some checkpoints, where, based on the signals received from the external environment, they choose whether to proliferate. Antigrowth signals can block proliferation either by forcing cells to pause in their proliferative cycle and enter into a quiescent state, or by inducing them to permanently relinquish their proliferative potential and remain in a postmitotic state. An example of antigrowth signal is provided by TGF-beta, which blocks the cellular advance in mitosis. In many human cancers, response to this anti-growth signal is negated through a variety of mechanisms, including down-regulation of TGF-beta receptors, or elimination of their downstream targets. The cross-talk occurring between antigrowth signalling pathways and the cell cycle has been investigated extensively [29-35]. Although the description of the interconnectedness of these two cellular mechanisms is still incomplete and subject to further delineation, the necessity to circumvent the antigrowth signalling circuitry by developing cancers has been ascertained [19].

The ability of tumour cell populations to expand in number is determined not only by the deregulation of cell proliferation but also by the evasion of apoptosis, the programmed cell death mechanisms which acts to destroy cells representing a threat to the organism. Surface receptors binds with survival or death factors in the extracellular environment, while intracellular sensors monitor the cell well-being and activate apoptosis in response to detected

abnormalities such as DNA damage, signalling imbalance provoked by oncogene action, survival factor insufficiency or hypoxia [36]. Many apoptotic signals converge to mitochondria, which respond to them by releasing cytochrome c, a potent catalyst of apoptosis [37, 38]. A central role in the apoptosis is played by p53, a tumour suppressor protein that elicits the release of cytochrome c by upregulating the expression of the proapoptotic protein Bax when sensing damages in the DNA. Disruption of the apoptotic machinery can be obtained by impairing its sensoring system (e.g. surface receptors or intracellular "sentinels"), effector components (the proteases that execute the cell disruption) or the transduction of death and survival signals. A most common occurring impairment of the programmed cell death involves mutations in p53, which is found in more than 50% of human cancers.

Self-sufficiency in growth signals, insensitivity to antigrowth signals and evasion of apoptosis, although leading to the uncoupling of the cell growth program from the signals received from the environment, do not ensure the expansion of the initial tumour clone to macroscopic size. A cell-autonomous program is present in virtually all mammalian cells that limits cell multiplication independently of the cell-to-cell signalling pathway. Evasion of this self-limiting program leads to immortalised cells with a limitless replicative potential [39]. Immortalisation of cells involves telomere maintenance [40], *i.e.* in immortalised cells telomeres are maintained at a length above a critical threshold that enables descendant cells to multiply indefinitely [40-48].

The nutrient supply and waste removal provided by blood vessels is crucial for cell survival and proliferation [49-51]. As such, avascular tumours lacking their own network of blood vessels cannot grow beyond a size of 2-3 mm^3. Angiogenesis (the formation of new capillaries) is an important step in the transition from a small, abnormal mass of cells to life-threatening malignant growth. Tumour cells, at some point in their development, begin to synthesise

proteins, such as the vascular endothelial growth factor (VEGF), capable of stimulating angiogenesis. These proteins bind to the receptors of endothelial cells (the building blocks of capillaries), inducing them to penetrate the tumour nodule and begin the process of constructing a network of vessels. As the endothelial cells proliferate, they secrete growth factors that stimulate the growth and motility of tumour cells. Besides in the apoptotic mechanism, p53 also plays an important role in mediating angiogenesis. Here, the role of p53 is to counterbalance the expression of VEGF through upregulating the expression of the anti-angiogenic agent thrombospondin-1 (TSP-1) [52], thus ensuring inappropriate angiogenesis does not occur. In cells bearing mutant p53, hypoxia induced VEGF is not so readily controlled by anti-angiogenic molecules such as TSP-1, hence the neovascularisation may occur [53].

As mentioned above, at the early stage of carcinogenesis, transformed cells remain localised in the original place where the tumour clone formed. Although curable through removal, *in situ* tumours are usually asymptomatic – hence often undetected. In the next stage of tumour progression, altered cells break through the basement membrane that encapsulates the original solid tumour and invade the adjacent tissue. Should a cancer cell successfully enter the circulatory system, either through lymph vessels or breaching of a blood vessel's lining, it will be transported throughout the body and may eventually lodge in the capillaries of another distant organ. Here the cells may begin to multiply, forming a secondary tumour known as a metastasis. Metastases are the cause of around 90% of deaths from cancer [54]. Whilst the primary (original) tumour can be controlled by many available therapies, widespread metastatic disease is very difficult to treat.

METABOLIC FEATURES OF CANCER CELLS

It has long been known that many tumour cells have a metabolic profile that differs from that of their normal counterpart. Otto Warburg, in the late 1920s, was the first to report the metabolic alterations characterising tumour cells. In particular he and his co-workers observed that the cancer metabolic phenotype is characterised by a higher uptake of glucose and a robust production of lactate, even under aerobic conditions (a phenomenon now known as the Warburg effect) [55]. These findings represent a keystone in cancer research, and the metabolic phenotype characterised by an enhanced uptake of glucose and aerobic glycolysis has since been considered as a reliable biomarker for tumours [56]. The glycolytic shift is nowadays exploited in positron emission tomography (PET) to identify malignant and fast-growing tumours and metastasis, emphasising the importance of Warburg's original observations [57, 58].

Based on metabolic data collected from numerous animal and human tumour samples, Warburg hypothesised that the observed cancerous metabolic phenotype arises from impairments in the respiratory chain, and suggested that the very cause of cancer should have been sought in the replacement of respiration of oxygen with fermentation of glucose into lactate [59-61]. These insights, however, met with scepticism, mainly stemming from the predominant concept that the altered metabolism was a non-causative epiphenomenon rather than a mechanistic determinant of carcinogenesis. Indeed, Warburg's theory, seemed inconsistent with evidence of apparently normal respiratory function in some tumour cells [62-70]. Moreover, the additional observation that the metabolic shift featuring in the Warburg effect arises primarily from genomic mutability selected during tumour progression, promoted the idea of cancer as a genetic disease, hence pushing researchers to focus on genetic alterations underlying the emergence of the malignancy. However, the

increasing amount of genomic data collected over several decades showed that a prototypical cancer genotype can not be defined. Malignant cells are remarkably heterogeneous because of the critical role of accumulating random mutations during carcinogenesis [71, 72] and studies of breast and renal cancers have found that every tumour cell exhibited a novel genotype [73, 74].

In recent years, the recognition of the virtual universality of aerobic glycolysis in transformed cells, despite the marked genotypic diversity, has been diverting part of the cancer research focus to the role of such metabolic alterations in the development, sustenance and invasion of cancer [75]. The revitalised interest in the Warburg effect has produced a vast literature addressing the emergence, significance and functional implication of the altered cancer metabolism. Moreover, although cancer is still predominantly considered a genetic disease [76-80], Warburg's deepest hypothesis is also gaining some attention, namely that cancer is caused by the metabolic shift from respiration to lactate production. In particular, Seyfried et al. [81] suggest that cancer is primarily a metabolic disease, as emerging evidence would show [82-90], and propose the hypothesis that damage to cellular respiration precedes and underlies the genome instability that accompanies tumour development.

Whether cancer has genetic or metabolic origins, an altered glycolysis characterised by the features observed by Warburg is considered today not only a good biomarker for tumours but also a possible target for new approaches in cancer therapy [91], as it provides researchers with a clear discriminating factor between normal and altered cells that could be exploited for drug intervention. Indeed, in recent researches a suppression of the phenotypic features of cancer carbon metabolism by either substrate limitation, pharmacological intervention, or genetic manipulation was found to mediate potent tumour-suppressive effects [92].

LINKS BETWEEN HALLMARKS OF CANCER AND CANCER METABOLISM

The metabolic features of neoplastic cells are intricately linked to the principal hallmarks of cancer listed above. The causative relationships between these different aspects of the disease have been partly elucidated, providing arguments for both the ideas of cancer metabolism as an epiphenomenon or as a causative factor in the emergence of the malignancy.

The constitutive activation of growth factor signals causes the metabolic reprogramming of cancer cells through different mechanisms. An example is provided by the cell-autonomous activation of the PI3K/Akt signalling pathway, whose components are frequently altered in human cancers [93]. The activated form of Akt, in particular, stimulates glycolysis by enhancing the activity of glycolytic enzymes such as glucose transporters and phosphofructokinase [94].

A glycolytic enzyme that is often overexpressed in cancer is hexokinase, particularly in its isoform HKII. This enzyme provides one of the links between the cancerous metabolic features and the ability of transformed cells to escape apoptosis. In tumours, HKII is predominantly bound to the outer membrane of mitochondria via the voltage-dependent anionic channel (VDAC) [95]. The VDAC-HKII interaction (possibly enhanced by the constitutive activation of Akt [96]) alters the permeability of the outer membrane and prevents cytochrome c, an important apoptotic signal, to be released into the cytosol, hence suppressing mitochondrial induced cell-death [97].

One of the principal mechanisms promoting the metabolic shift occurring in cancer cells resides in the activation of hypoxia-inducible factor 1 (HIF-1). This transcription factor is activated by the hypoxic stress that cancer cells undergo during the first stages of tumour progression, and is involved in the overexpression of HKII. A recent work indicates that HIF-1 also promotes

angiogenesis by activating the expression of VEGF [98], showing how tumour metabolism and some other hallmarks of cancer can also be elicited in parallel by the same causative event.

HIF-1 also plays an important role in the promotion of invasion and metastasis. In particular, the activation of its α subunit (HIF-1α) is responsible for the loss of E-cadherin [99, 100], a transmembrane protein that ensures cell adhesion. The active form of HIF-1α also promotes the expression of two genes (met and TWIST) that facilitate cell mobility and induce tissue invasion and metastases [101, 102].

Another important link between alterations in signalling pathways and reprogramming of cancer cells metabolism is provided by the loss of p53, which leads to enhanced replicative potential by impairing the apoptotic mechanisms, as mentioned above. Because p53 negatively regulates phosphoglycerate mutase (PGM), the enzyme that converts 3-phosphoglycerate (3PG) to 2-phosphoglycerate (2PG) in glycolysis [103], and transcriptionally activates TIGAR, a glycolysis and apoptosis regulator that inhibits the overall phosphofructokinase activity, loss of p53 result in an enhanced activity of PGM and PFK, contributing to increase the glycolytic flux.

There are many other mechanisms (reviewed in [104]) that link cancer metabolism to the principal hallmarks of cancer previously described. The emerging picture suggests that metabolic alterations occurring in transformed cells can partly be the consequence of non-metabolic oncogenic events, and partly the triggering factors of non-metabolic cancer hallmarks. The metabolic and non-metabolic features of cancer can also be the result of a coevolution caused by the need of simultaneously subverting different tumour suppressor mechanisms.

CANCER: A MODELLING PERSPECTIVE

The vast number of mechanisms involved in carcinogenesis, either directly affected by genetic alterations or indirectly responsive to them, and the highly complex network of interactions in which these mechanisms interplay, have made cancer elude traditional attempts at understanding it. In molecular biology, the focus is on single molecules rather than the networks of interactions in which they are involved. Because this approach fails to capture the systemic nature of cancer (as well as many other complex diseases), the redefinition or recognition of cancer as a Systems Biology disease seems to be most appropriate [14].

The term Systems Biology, first introduced by Zieglgänsberger and Tölle [105], refers to a relatively new trend in biological sciences, where the detailed knowledge of the molecular mechanisms underlying biological functions – obtained through the reductionistic approach of molecular biology – is eventually integrated in a conceptual framework that puts the systemic properties of biological entities at the centre of any investigation. Systems Biology aims to understand how the interactions between the components of a biological system give rise to the properties and functions of that system. This paradigm shift has been applied by an increasing number of researchers, who underline the timeliness of replacing reductionism with a systemic approach in drug discovery, both in general and more specifically in cancer research [14, 106-108]. It is evident, in fact, that the understanding of cancer cannot result from considering single oncogenic events, but can only come from accounting for the combined and in fact integrating action of various extracellular and intracellular triggers. Such a change of perspective also requires the tools and approaches able to address the complexity of the systems under study, which intuition and verbal reasoning alone will fail to grasp. The introduction of computer-assisted analytical tools to help researchers to describe and predict

the behaviour and evolution of cancer is then central to any advance; mathematical models, informed by extant data and continuously revised by new information, become an essential factor for both interpreting experimental data and guiding experimental design [109].

In the following sections we will show how computer-assisted approaches have been historically used in an attempt to describe and understand cancer, and how they have evolved from a simple, mesoscale (multicellular) description of the phenomenon into a more detailed representation that includes some aspects of the system biochemistry. We will also show how this transition, whilst widening the range of modelling approaches, has increasingly acknowledged the role of metabolism in its attempt to describe and predict the disease behaviour, and its response to possible drug intervention strategies. With respect to this point, in the last section we will explore the potential role of some of the most promising mathematical frameworks and computational techniques in addressing the problem of developing anti-cancer drugs acting at a metabolic level with both high specificity and low toxicity.

HISTORICAL APPROACHES

Although the systems biology paradigm is relatively recent and has not been broadly applied yet to drug discovery, mathematical modelling of cancer is not new. Historically, however, mathematical attempts to describe cancerous phenomena have focused on modelling mesoscopic features of tumour growth, with the aim of identifying the main processes governing the dynamics of tumour expansion and its response to different therapies. There exists a substantial literature about the different routes along which tumour growth modelling has developed [110-117]. In this section we highlight the milestones that marked a specific path leading from the first attempts to describe tumour development to more recent approaches, where the metabolic features of

cancer are taken into account and integrated into larger models of tumour growth. A common trait of the models presented in this section is the use of differential equations to describe the mesoscopic characteristics of cancer, but where subcellular molecular processes underlying the development of the tumour physiology are not taken into account.

Many of the early mathematical models found in the literature focus on the growth of multicellular spheroids (MCSs), clusters of cancer cells grown *in vitro* to mimic the early stages of *in vivo* avascular tumour growth and to test the applicability of new cancer treatment strategies [118]. In 1972, Greenspan proposed a model to explain the growth rate of a solid tumour in the earliest stages of development, when, because of the disordered and poor vascularisation, the inner part of the nodule receives nutrients and release waste products primarily by diffusion [119]. The model accounts for the structure of nodular tumours as depicted in Figure 1. In the centre of the nodule the concentration of vital nutrient is below the critical level to sustain life, and a central necrotic core develops. This central core is surrounded by a middle layer of viable, non-proliferating cells and an outer shell where all mitosis occur.

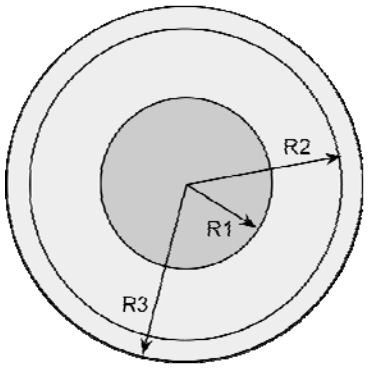

Figure 1 - Internal structure of a nodular carcinoma. *The central core (r<R1) represents the necrotic region which emerges as a consequence of low nutrient viability, unable to sustain cellular life. The central layer (R1<r<R2) is constituted by quiescent cells, while the external shell (R2<r<R3) hosts those cells which undergoes mitosis. Adapted from [119].*

Based on the experimental observation showing that the growth of a solid

tumour by diffusion alone leads asymptotically to a dormant but viable steady-state, the model was used to infer the main processes, such as necrosis or metabolic wastes, affecting the dynamics of solid tumour growth. Existing models of MCS and avascular tumour development are essentially extensions of Greenspan's model, and are effectively used to describe the evolution of the tumour outer boundary in response to vital nutrients (in particular oxygen) and growth factors [120-124].

An important step forward in modelling tumour growth and development was made by adopting population ecology methods, which provide a means for examining tumours, not as an isolated collection of transformed cells, but rather as an invading species in a previously stable multicellular population. The tumour-host interface was modelled by Gatenby [125, 126] as a network of interacting normal and malignant cell populations, using coupled, non-linear differential equations. These interactions were explored to identify the crucial parameters that control tumorigenesis and to demonstrate the limitations of traditional therapeutic strategies. The model, in particular, was used to simulate different stages in tumour progression analogous to the initiation, promotion and invasion stages observed in experimental systems. This was done by changing the parameter values in the differential equations describing the interactions between the normal and neoplastic cell populations. Because the transition from one phase to another could be achieved through different possible changes in the parameter values, the authors proposed the idea of "functional equivalence", in which disparate tumour traits (modelled through the parameters of the differential equations) can play identical roles in tumour growth and invasion. These traits, in turn, were translated into clinical factors, allowing predictions of treatment strategies likely to be successful or unsuccessful at each phase of tumour growth. For example, tumour survival after initiation was shown to be dependent solely on host-generated effects. At this stage, hence, immunological response, modelled through one of the Lotka-

Volterra (competition) factors in the differential equations, emerged as the factor to exploit in order to suppress tumour growth, consistently with the concept of immune surveillance. From a clinical perspective, tumour progression at this early stage could then be stopped by strategies such as reducing the levels of growth factors (e.g. through hormonal manipulation or by blocking their receptors), increasing the immunological response to tumour antigens, or increasing the levels of negative growth factors or receptors.

The first time metabolism was introduced in tumour modelling was in a study interrogating the potential role of the glycolytic phenotype in facilitating tumour invasion. Tumour cell populations, as with any invading population in biology, must directly perturb their environment in such a way as to facilitate their own growth while inhibiting the growth of the original community. The commonality of altered tumour metabolism, in particular the adoption of the glycolytic phenotype in most cancers, led Gatenby and Gawlinski to propose the acid-mediated tumour invasion hypothesis [127] where the tumour cells' increased acid secretion, coupled with their resistance to low extracellular pH, may provide a simple but complete mechanism for cancer invasion. The model consists of a system of three coupled reaction-diffusion equations describing the spatial distribution and temporal evolution of normal tissue density, neoplastic tissue density, and excess concentration of H^+ ions. The equations take into account mesoscale parameters such as the intrinsic growth rate of the normal and neoplastic populations, their carrying capacity, diffusion coefficients and Lotka-Volterra factors. The normal tissue dynamics, in particular, was modelled in order to account for the lethal effect of the acidic environment caused by the excess of H^+ ions, while the cancerous tissue was assumed to be insensitive to the acidic pH. The excess of H^+ ions was assumed to be produced at a rate proportional to the neoplastic cell density and to diffuse away chemically through buffer binding. Early tumour growth was shown to be critically dependent on H^+ production by transformed cells and the level of vasculature,

the latter being represented by the carrying capacity of the cancerous tissue. By varying these factors, a variety of tumour morphologies were reproduced, including tumours growing to large volumes with declining growth rates, highly necrotic growth with the development of tumour chords, and even initial growth followed by a decrease in tumour volume, representing spontaneous regression.

While in Gatenby and and Gawlinski's model the differences in acidity between tumour and normal micro-environments are represented without considering intracellular and extracellular pH separately, Webb et al. developed a mathematical model aiming to elucidate the mechanisms through which cancer cells maintain their internal pH at physiological levels despite an acidic extracellular environment [128]. The model consists of two variables, the intracellular and extracellular H^+ ions concentrations, $[H^+]_i$ and $[H^+]_e$ respectively. The correlation between hypoxia and low external pH [129] is introduced by assuming that the hydrogen ions transported outside the cell are removed from the interstitial space at a rate which is proportional to a parameter V, representing the extent of vascularity. The lower the vascularity, and hence the level of oxygen, the more acidic the interstitial space. The activity of membrane based ion-transport mechanisms is also taken into account by adding in the differential equations for $[H^+]_i$ and $[H^+]_e$ specific terms representing the dependencies of internal and external pH upon these transporters. The activity of the lactate$^-$/H^+ symporter, in particular, is introduced as representative of the tumour glycolytic activity (characterised by production and excretion of lactate) and is assumed to be inhibited by low extracellular pH. The model focuses on the combined activities of these cell membrane ion-transporters in the regulation of the intracellular pH, where the hallmarks of solid tumours, like disorganised vasculature, hypoxia and high glycolytic metabolism, are taken into account. In particular, the model suggests a causative relation between the increased reliance of cancer cells on the

energetically inefficient glycolytic pathway and the fact that their internal pH is less sensitive to external acidity than it is in normal cells.

With Gatenby and Gawinski's model, as well as with that of Webb et al., an interesting trend was initiated in modelling tumour growth, where the mesoscopic features of tumour and cancer, such as vascularisation or tumour colony size, were linked to variables which are more straightforwardly related to the internal metabolic processes of altered cells. Although the dependency of tumour growth on the acidity of the environment was still described in terms of mesoscopic variables, such as the size of the cell-populations and pH, Gatenby's and Webb's work can represent a first step in the attempt to provide an understanding of physiologically relevant tumour properties in terms of the molecular processes underlying them.

The link between the temporal evolution of tumour growth and the features underlying the altered cancerous metabolism was further investigated by Casciari et al. [130]. The authors used a simplified representation of tumour cell glucose metabolism to determine the metabolic profile of altered cells. This metabolic profile was then incorporated into a model of tumour growth, which considers the interaction of tumour cells with oxygen, glucose, lactate, carbon dioxide, hydrogen ions and other ionic species. The oxygen and glucose consumption rates were described through empiric equations, where the functional forms of the rate was based on empirical considerations rather than on biochemistry [131]. The consumption rate of the remaining metabolites was obtained through stoichiometric analysis (conservation of reactants). The model was used to predict and interpret the different phases of EMT6/Ro spheroid growth in terms of the underlying metabolic profile. Many theoretical predictions matched with previously published experimental data. For example, the predicted viable rim thicknesses of the spheroid based on low concentrations of glucose was found to fit data about 1000 microns spheroids

grown at medium glucose concentrations of 5.5 mM or less. The model, however, could not accurately predict other phenomena such as the decrease in oxygen and glucose metabolism seen in spheroids with time, leading to observed growth plateau. In spite of these inaccuracies, to date Casciari et al. is acknowledged to be one of very few experimentally-grounded models of cellular metabolism and tumour growth.

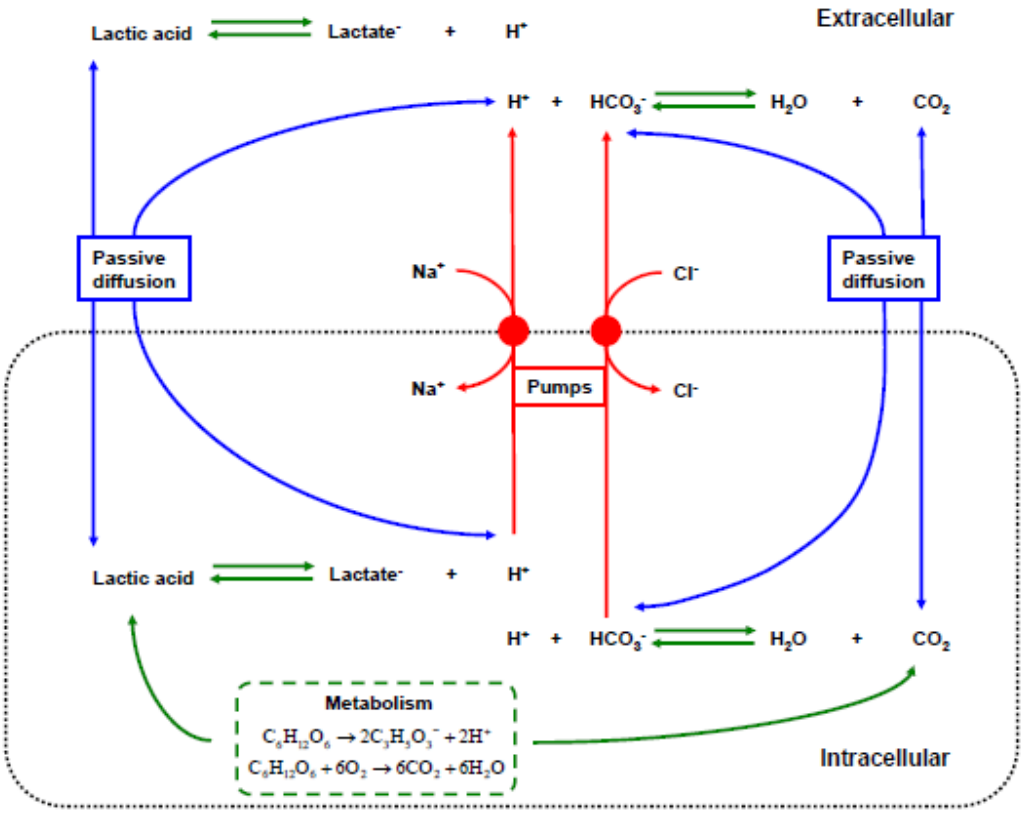

Figure 2 - Simplified model of oxygen, glucose and pH regulation, waste product transport and buffering for a tumour cell and its surroundings. *Reproduced with permission from Smallbone 2005.*

The attempts to relate the observed physiological characteristic of tumour development and invasion with the long observed feature of cancerous metabolism, opens a new possible trend in modelling cancer, where different levels of description are taken into account. However, as pointed out by Rejniak et al. [116], integrating all the different aspects of a complex phenomenon like

cancer into a common description might turn to be as complicated as the disease itself. Hence, although envisaging the potential of an integrated approach [132], current systems biology applications on cancer research still tend to focus on and describe single levels of organisation within the cell.

The stepping stones described above in tumour growth modelling demonstrate how cancer metabolism has gradually gained attention in the cancer modelling community, and how its relevance in the development of the disease has been acknowledged even from a theoretical standpoint. We also mention, however, that a substantial part of the systems biology effort has focused on integrating signal transduction and gene expression data to create *in silico* representations of cancer cells and their functioning [133-135]. A particularly ambitious project is the computer simulation of a human cancer cell by Christopher et al. [136], where the authors have reproduced in silico the connectivity of signal transduction pathways with gene-expression networks through multiple iterations between experiments and model predictions. However, the interconnectedness of signalling pathways with metabolism, and the recognition of the latter as trigger, regulator or end-point of many oncogenic events, makes cancer metabolism a more immediate target of investigation. There also exists a wider spectrum of theoretical frameworks to address the study of metabolic networks, which make metabolism a domain where systems biology could contribute most.

In the next sections we will comment specifically on two mainstream systems biology approaches in the study of metabolism, and how they have been (or might be) used in the context of cancer research.

METABOLIC CONTROL ANALYSIS

In the quest for anti-cancer drugs, cancer metabolism has increasingly been a focus of interest in clinical research. Enhanced glycolysis and a robust production of lactate constitute a characteristic trait that discriminates many cancerous cells from normal cells, albeit to different extents for different cancers. As one of the main problems in defeating cancer is the genetic heterogeneity of the altered cells constituting the malignancy (due to randomly accumulated mutations during tumour progression), the virtual universality of these metabolic features can provide researchers with a general handle on cancer, regardless the specific genotype of the single cells.

In order to develop an anti-cancer drug operating at a metabolic level with requirements of high efficacy and low toxicity, the entire metabolic network must be taken into account. Potent inhibitors of essential enzymes often do not show the expected effect at the cellular level, proving that the ultimate biological effect of targeting specific enzymes is strongly mediated by the whole network in which these enzymes operate [137-139]. Consequently, the traditional assumption that biological processes are controlled by single molecular components or enzymatic steps [140] has to be replaced by the more evidence-supported concept that the control over biological functions is more likely to be spread over much of the entire biochemical network [141, 142]. In the context of clinical research, this change of perspective leads us to consider the network, rather than the single enzyme within a network, as the target of a drug [143]. The specific enzymatic step is then re-thought in terms of "access point" to the network, a site of intervention whereby the action of a drug would elicit the desired response at a system level with respect to some property of interest. In order for a drug to match requirements of high efficacy and low toxicity, it is necessary to identify sites of intervention that represent points of fragility for the altered metabolic phenotype but not for the normal one [14,

143, 144]. Targeting the network through such enzymes would elicit a strong response in the diseased cell (high efficacy) without affecting normal cells (low toxicity). Finding the right balance between efficacy and toxicity is the key to successful drug development.

Metabolic control analysis (MCA) is a quantitative methodology that can be profitably used to identify points of fragility in metabolic networks, by evaluating the importance and relative contribution of individual enzymatic steps in the overall functioning of a given metabolic system. Differential MCA, in particular, has been proposed in clinical research as a tool to understand how the system responds to specific perturbations under different physiological conditions, providing a way to assess both the efficiency and the specificity of a compound designed to target specific enzymes [108, 137, 142-144].
The extent to which any system property is controlled by an enzymatic step is quantified in terms of control coefficients. A control coefficient is defined as the relative change at steady state of the property *P* (the response of the system) divided by the small relative change in the activity of an enzyme *i* (the causative event) [145]. Mathematically:

$$C_i^P = \lim_{\Delta v_i \to 0} \frac{\Delta P / P}{\Delta v_i / v_i} = \frac{d\ln(P)}{d\ln(v_i)} \quad (1)$$

Important examples of control coefficients are the flux control coefficient C_i^J and the concentration control coefficient C_i^S; here the system property *P* of interest is the flux *J* through a specific pathway, and the concentration of a metabolite *S*, respectively. Two important relations, known as summation theorems of MCA, state that the summation of the flux control coefficients and the concentration control coefficients over all the enzymes in the system equals 1 and 0 respectively [146-149].

$$\sum_{i=1}^{n} C_i^J = 1 \tag{2}$$

$$\sum_{i=1}^{n} C_i^S = 0 \tag{3}$$

These relationships allow for the fact that the control over a certain system property need not be exerted by a single enzymatic step, but may well be shared among the different enzymes. When a reaction changes its control over J or S, this is compensated for by changes in the control by all other reactions.

The operational framework of MCA is in part theoretical [147] and in part experimental [150]. From an experimental point of view, infinitesimal changes are neither detectable nor inducible. In practice, quantifiable, non-infinitesimal changes are measured and then used to extrapolate the control for very small changes [142]. The experimental procedures for the determination of the flux-control distribution have been reviewed in [151] and acknowledged for having largely contributed to elucidating how the control of important pathways is spread among the different enzymatic steps in different unicellular organisms. In the context of cancer research, MCA has been undertaken to identify enzymes exerting a major control over important fluxes of the tumour energy metabolism. Boren *et al.* [152] showed that the first reaction in the pentose phosphate pathway, catalysed by G6PDH, has a strong control over the synthesis of ribose (a fundamental component of nucleic acids) in tumour cells. This seems to suggest G6PDH as point of intervention for an anticancer drug aiming to reduce the proliferative capacity of transformed cells by inhibiting nucleotide synthesis. In a similar work, Comin-Anduix et al. proved the highly controlling role of transketolase enzyme, TKL, on the cell proliferation rate in the Ehrlich's ascites tumour model [153]. More recent studies elucidated how the control over the glycolytic flux was distributed among the different glycolytic enzymes. In particular, Marín-Hernández et al. [154] showed that in

AS-30D hepatocarcinoma the main flux control (71%) resides in the first part of the glycolytic pathway (i.e. GLUT and HK), and the rest of the control (29%) is localised in the ALD–LDH segment. In a clinical strategy aimed to kill cancer cells through starvation, the first steps of glycolysis seem then to be the most suitable drug target in terms of efficacy for AS-30D tumour cells.

In all these cases, however, no comparative study with normal cells was made to assess whether the efficacy of a drug targeting those enzyme would also be coupled with low toxicity, *i.e.* a low effect on the non-cancerous cells. Such comparative studies may be difficult to carry out experimentally, as there is currently no simple way to assess *in vivo* the replication rate of non-cancerous cells. The computer assisted approach to MCA might represent a way to overcome this limitation. *In silico* application of MCA is based on a dynamic representation of the system under study, where the stoichiometric, topological and dynamic properties of the metabolic network are taken into account. The system is described as a set of mass-balance equations governing the rate at which the different metabolites are consumed:

$$\frac{d\mathbf{S}}{dt} = \mathbf{N} \cdot \mathbf{v} \tag{4}$$

where $\mathbf{S}$ denotes the set of metabolite concentrations, $\mathbf{N}$ is the stoichiometry matrix, whose generic element $\{N_{ij}\}$ represents the stoichiometric coefficient of metabolite *i* in reaction *j,* and $\mathbf{v}$ is the set of the reaction rates, which depend on the functional form of the rate equations used to describe the kinetics of each enzymatic step. By expressing the rates $\mathbf{v}$ as explicit functions of the metabolite concentrations, Eq.(4) becomes a set of differential equations that can be solved numerically if initial conditions and parameter values are known. The model can then be used to find (if it exists) a set of fluxes and metabolite concentrations satisfying the steady-state condition (where, by definition, $d\mathbf{S}/dt = 0$). By taking the derivatives of these steady-state fluxes or

concentrations with respect to the enzyme activities, the control coefficients can be readily calculated through software packages of numerical analysis such as Copasi [155]. The repository of live models JWS-online also provides a facility to evaluate the control coefficients of any of the models present in the database [156]. This procedure provides researchers with a virtual benchwork for a rapid and simultaneous assessment of the control exerted by all the enzymes in the network over a variety of systemic properties (often referable to specific fluxes). Because cancer and normal cells have different metabolic phenotypes, it is reasonable to expect they also have different control profiles. The predicted differential response of cancerous versus normal cells to specific drug interventions may be assessed by comparing the control coefficients of two distinct dynamic models, one representing the metabolic network in the normal physiological state and one reproducing the cancerous metabolic features. Good putative drug targets would be those enzymes showing a major control on a property of interest in the former model and and a minor control (over the same property) in the latter model. Theoretical comparative studies have not been performed at present (with the exception of Bakker et al. in their work on *T.brucei* drug design [157]). However, the potentiality of the *in silico* approach of MCA in cancer research has been shown in a recent work by Reuss [158], where the author predicted a high control of G6PDH over the replication rate of HEPG2 liver cancer cells, confirming the previous experimental result obtained by Boren et al. [152]. Similarly, kinetic modelling of glycolysis in AS-30D and HeLa tumour cells has revealed that indeed GLUT and HK together with HPI are the main flux-controlling steps in both tumours (A. Marín-Hernández, R. Moreno-Sánchez and E. Saavedra, manuscript in preparation) in accordance to previous experimental results [154].

Because the specific features of cancer metabolism provide neoplastic cells with selective advantages over their normal counterparts, targeting cancer metabolism in order to elicit specific metabolic responses is one on the

branches that cancer research is currently trying to explore [152, 159-163]. From this perspective, metabolism becomes the target of the treatment as well as the domain of the functional properties one aims to alter. This purely metabolic perspective allows for a simplified description of the cancer phenomenon, where, in a first attempt, the metabolic network is considered as a separate subsystem from the more complex cellular context. However, metabolism is widely interconnected with other cellular processes. Signalling events lead to related metabolic reactions which, in turn modify other metabolic functions or gene expression. It seems then reasonable and desirable to widen the focus of MCA from purely metabolic events to other cellular processes. An MCA extension which includes gene-expression and signalling transduction has been already formalised and proposed with the name of hierarchical regulation analysis [164, 165].

FLUX BALANCE ANALYSIS

Drugs are designed to affect one or more specific properties of the cells needing treatment. These properties usually represent what differentiates diseased cells from their normal counterparts, or a pathogenic organism from its host. The property one chooses to affect can vary depending on the specific clinical strategy pursued. Because neoplastic cells grow and replicate at a considerably faster rate than their normal counterparts, the rationale behind many of the possible choices in cancer treatment consists of halting the proliferative potential of the malignant tissue. Indeed, traditional clinical approaches such as chemotherapy and radiotherapy aim to kill cancer cells by disrupting their replication machinery. Similarly, in drug intervention at the metabolic level, the preliminary step consists of identifying a property which characterises the altered phenotype and which is therefore sensible to target. In this respect, constraint-based modelling approaches, and particularly Flux Balance Analysis (FBA), can provide us with a way to identify these properties.

FBA is a constraint-based approach widely used to study the metabolic capabilities of a cell or a subcellular system. In the classical formulation of FBA only the stoichiometric information is considered, while the dynamic aspects of the system are neglected. The underlying assumption of FBA is that any metabolic system is aimed to fulfil a specific biological task at the best of its possibilities. The goal of FBA is then to explore the range of all the possible metabolic phenotypes (in terms of flux distributions at steady states) that allow the system to optimise that task. The mathematical representation of the biological task of interest is called objective function and in the existing FBA approaches takes the form of a linear combination of the (unknown) flux-variables. The general formulation of FBA can be written as follows:

$$\begin{aligned} &\text{maximise} && Z = \mathbf{f}^T \cdot \mathbf{v} \\ &\text{subject to} && \begin{matrix} \mathbf{N} \cdot \mathbf{v} = \mathbf{0} \\ \mathbf{v}^L \le \mathbf{v} \le \mathbf{v}^U \end{matrix} \end{aligned} \tag{5}$$

where $\mathbf{N}$ is the stoichiometric matrix of the system, $\mathbf{v}$ is the flux vector whose elements represent the rate at which the different reactions occur in the network, $\mathbf{v}^L$ and $\mathbf{v}^U$ are the lower and upper bound of the fluxes, and Z is the objective function.

If the system under study is known to optimise a certain biological requirement, then that requirement might be considered as the property one may want to target in order to disrupt the metabolic phenotype of the cell. However, the identification of the biological task that the objective function should represent is not always easy. For studies involving *E.coli* metabolism, Z is usually defined to represent the yield of biomass [166], assuming that bacteria aim to grow as fast as possible (although this assumption does not reflect a generally valid

principle in microbiology [167]). By contrast, for human cells, things are not so straightforward.

Since cancer cells grow at a much higher rate than their normal counterparts, it would seem reasonable to adopt the same approach as for *E.coli* by choosing the maximisation of biomass production rate as the optimisation criterion. Although this intuitive choice may seem sensible, the resulting FBA solution highlights a flux-pattern which does not match with the observed characteristic of cancer metabolism discussed above [59]. Because of the high demand of ATP in the production of biomass, the flux pattern corresponding to the maximal yield shows the glucose uptake flux entirely entering the TCA cycle, with no lactate production. To retrieve a flux pattern highlighting the cancer metabolic features previously discussed (a constitutive activation of the branch leading to lactate production and, possibly, the reduction of the flux entering the TCA cycle), the FBA problem has then to be formulated differently. A possible way to do so consists of replacing the maximal yield of biomass with a different criterion of optimality. In a recent work, Simeonidis et al. showed how an appropriate reformulation of FBA can be used to reproduce the Crabtree effect, an experimentally-observed behaviour whereby *S.cerevisiae* produces ethanol aerobically in the presence of high external glucose concentrations rather than producing biomass through the TCA cycle [168]. The authors hypothesised that (one of) the "driving forces" behind yeast metabolism is resource preservation (see also [169]). By minimising the number of active reactions (and hence the number of enzymes) needed to produce a required amount of biomass, the flux patterns obtained as solutions of the FBA problem showed the characteristic switch from respiration to fermentation that occurs when the concentration of glucose in the growing medium is increased above a certain threshold. Because of the commonalities in the metabolic features of fermenting yeast and cancerous cells [170], a similar argument might be applied to reproduce the constitutive metabolic changes occurring in carcinogenesis. From an FBA

perspective, higher concentration of glucose in the growth medium and higher rate of glucose uptake due to over-expression of glycolytic enzymes are both implemented by increasing the upper limit of the glucose uptake flux. In both cases, the requirement of resource preservation would force the system to switch from respiration to fermentation/lactate production as soon as the glycolytic flux becomes high enough to provide the cell with the amount of ATP needed for the required production of biomass.

A related issue that FBA could address is whether cancer cells are committed to optimise different biological functions concurrently. Indeed, the enhanced replication rate of neoplastic cells, combined with a predilection for fermentation (which is not the most efficient way to produce ATP) would seem to support a multi-functional optimisation hypothesis, whereby different criteria of optimisation have to be satisfied simultaneously. As initially hypothesised by Gatenby and Gawlinski [127], the production and excretion of lactic acid constitutes a way for cancer cells to compete with their normal counterpart by creating an hostile environment for normal cells. However, the fact that sometimes the TCA cycle is nevertheless active (although to a smaller extent than it could) makes evident that competing through excretion of lactate is not the only task that cancer cells try to optimise. Using a specular argument, one could say that, despite the enhanced replication rate of cancer cells, the fact that the TCA cycle is somehow hampered shows that replicating at the highest possible rate is not the (only) objective that drives cancer cells, or, in other words, that there are multiple goals pushing the system toward a different metabolic flux pattern. The relevance of different possible optimisation criteria in the functioning of the system and their relative weights could also help to elucidate why the phenotypic traits of cancer metabolism are present to different extents in different types of cancer.

There are other points that an FBA approach might help to elucidate. Knowledge of the metabolic shift occurring in tumorigenesis predominantly involves central carbon metabolism. However, the shift may extend beyond central metabolism, and remarkable metabolic differences between normal and cancer cells may lie in pathways not yet studied within the context of cancer research. A further application of FBA could highlight particularly active metabolic pathways in cancer on a genome-scale level, and identify the regions where the flux pattern differs most between cancer and normal cells. Shlomi et al. [171] have recently used an FBA approach to describe the tissue specificity of human metabolism, where tissue-specific gene and protein expression data were integrated with a genome-scale reconstruction of the human metabolic network. Different integer values were assigned to different gene-expression states, so to distinguish among highly (1), lowly (-1) and moderately (0) expressed genes. The objective function of the FBA problem was then set to account for the differences between the activity of each reaction in the predicted pattern of fluxes and the integer representation of the corresponding experimental gene-expression level. By minimising such an objective function, the authors were able to retrieve stoichiometrically and topologically consistent flux patterns on a genome-scale level with the maximum number of reactions whose activity was in accordance with their expression state. This study may establish a FBA-based computational approach for the genome-wide study of normal and cancer human metabolism in a tissue-specific manner.

Another interesting point FBA might address is the following. Given the selective pressure that biological systems undergo when functioning under mutual competition, it seems reasonable to assume that cancer cells fulfil their specific biological tasks in the most economic way. In other words, given the available external substrates and given a set of functionally important targets to accomplish, the cell would employ its resources most 'effortlessly'. From a metabolic perspective, this would translate into the employment of a minimal

number of active reactions, or, more generally, a minimal employment of resources. In *E.coli*, for example, experimental results have shown that fitness increases while unused catabolic functions decrease, this reduction being beneficial and therefore favoured by selection [172]. In the context of FBA, this 'principle of minimal effort' has been used in different forms to identify the pattern of fluxes that best portraits the system functioning with respect to specific criteria of optimality [169, 173]. On the other hand, there exist different flux patterns that are equally optimal with respect to a certain criterion or set of criteria. Extension of FBA to find alternate optimal solutions [174] or alternate optimal patterns of fluxes [175] have been developed. In particular, as part of this thesis work I have already proposed an algorithm able to find all the minimal and equally optimal flux patterns of a metabolic network with respect to a given functional task [175]. The superposition of all minimal optimal flux patterns allows us to identify those pathways or sets of reactions that *must* be active in order for the system to optimally fulfil a given function, and other sets of reactions that can be alternatively active. The application of such an approach in the context of cancer research might help to identify and predict the narrowest region of human metabolism necessary to observe the carcinogenic metabolic shift. From the perspective of developing a kinetic model of cancer metabolism, these results might also provide modellers with a concise set of reactions that can be used as a backbone for a mechanistic representation of the system under study, as well as an idea about which pathways and reactions can be reasonably neglected.

CONCLUSIONS

Despite the genotypic heterogeneity of cancer cells, the common metabolic features that have been observed in a variety of different kind of cancers make cancer metabolism an increasing focus of interest from a clinical perspective [176]. Moreover, metabolism is tightly related to many aspects of cellular

functioning, and the metabolic alterations occurring during carcinogenesis have been proven to play a major role in the emergence and progression of the disease [177]. Hence, cancer metabolism seems to be an anchor-point that might be exploited to directly or indirectly affect many aspects of the disease.
The application of systems biology approaches in the study of cancer metabolism is deemed to deepen our understanding of an important aspect of the disease and to strongly enhance our ability to develop drugs with higher efficacy and lower toxicity.

There is a need to further develop models of cellular metabolic processes in healthy and cancerous tissue, and to analyse them with new tools and approaches derived from the systems biology perspective. By admitting that cancer is truly a systemic disease, one my may begin to understand its emergent behaviour as more than the sum of its constituent parts. Moreover, through providing a theoretical framework to which the vast array of available metabolic data may be fused, one may begin to uncover the non-linear interactions that govern this complex disease.

GLOSSARY

- **adenoma**: Benign epithelial tumour in which the cells form recognisable glandular structures.
- **angiogenesis**: The formation of new capillary blood vessels.
- **anthracycline therapy**: Type of chemotherapy that acts to prevent cell division by disrupting the structure of the DNA.
- **apoptosis**: Programmed cell death, as signalled by the nuclei in normally functioning cells.
- **basement membrane**: Extracellular matrix characteristically found under epithelial cells. There are two distinct layers: the basal lamina, immediately adjacent to the cells, and the reticular lamina.
- **benign**: Not malignant. Benign tumours do not invade or metastasise, having lost growth control but not positional control. They are usually surrounded by a fibrous capsule of compressed tissue, and treatment or removal is curative.
- **carcinogen**: Chemical, virus or radiation that can induce cancer.
- **carcinoma**: Malignant epithelial tumour.
- **clastogenic**: Altering the structure of chromosomes.
- **differentiation**: Process undergone by cells as they mature into normal cells. Differentiated cells have distinctive characteristics, perform specific functions and are less likely to divide.
- **epigenetic**: Differentiation due to selective gene activation and expression. Not due to changes in the genome.
- **epithelium**: Covering of internal and external surfaces of the body, including the lining of vessels and other small cavities. It consists of cells joined by small amounts of 'cementing' substances. Epithelium is classified

into types on the basis of its depth (in terms of cell number) and the shape of the superficial cells.

- **extracellular matrix (ECM)**: Any material produced by cells and secreted into the surrounding medium, usually applied to the noncellular portion of tissues. Although produced by cells, the ECM can influence the behaviour of cells quite markedly, an important factor to consider when growing cells in vitro.
- **heritable**: Capable of being transmitted from parent to child.
- **hyperplasia**: Abnormal increase in the number of normal cells in a tissue.
- **hypoxia**: Diminished oxygen supply.
- ***in situ***: Localised. A carcinoma that has not breached the basement membrane.
- **invasion**: Movement of cells into adjacent tissue normally occupied by a different cell type.
- ***in vitro***: Cell manipulation outside the body.
- ***in vivo***: In the living body. An experimental procedure using an intact live animal.
- **malignant**: Tending to become progressively worse and to result in death. Malignant tumours are invasive and have the capacity to metastasise. Compare benign.
- **mesoscale**: Description of tumours that does not rely on the detailed biochemical mechanisms underlying the emergence of the malignancy. The variable and parameters taken into account are referable to multicellular properties and/or are averaged to take into account the global effect of a population of cells.
- **metastasis**: Transfer of cells from one organ or part to another, not directly connected with it. This usually occurs through the blood vessels, lymph channels or spinal fluid.

- **oncogene**: Overexpressed version of a normal gene (the proto-oncogene) that promotes excessive growth. Compare tumour suppressor gene.
- **necrosis**: Non-programmed cell death.
- **quiescence**: The state of not dividing.
- **somatic**: Characteristic of the body.
- **stroma**: (syn: interstitium) Connective tissue framework of an organ, gland or other structure, in contrast to the functional cells. Rich in extracellular matrix.
- **tumour**: (syn: neoplasm) Abnormal mass of tissue serving no useful function to the host, resulting from excessive cell division that is uncontrolled and progressive. May be either benign or malignant.
- **tumour suppressor gene (TSG)**: (syn: anti-oncogene) A gene negatively regulating cell division that, when inactivated (through mutation for example), allows escape from normal growth constraints. Compare oncogene.

REFERENCES

1. Weinstein, B., *The Origins of Human Cancer: Molecular Mechanisms of Carcinogenesis and Their Implications for Cancer Prevention and Treatment -- Twenty-seventh G. H. A. Clowes Memorial Award Lecture.* Cancer Research, 1988. **48**: p. 4135-4143.
2. Weber, G.F., *Molecular Mechanisms of Cancer*. 2007: Springer.
3. Frank, S.A., *Dynamic of Cancer - Incidence, Inheritance and Evolution*. Princeton Series in Evolutionary Biology, ed. H.A. Orr. 2007, Princeton: Princeton University Press.
4. Armitage, P., *Multistage Models of Carcinogenesis.* Environmental Health Perspectives, 1985. **63**: p. 195-201.
5. Engle, L.J., C.L. Simpson, and J.E. Landers, *Using high-throughput SNP technologies to study cancer.* Oncogene, 2006. **25**(11): p. 1594-1601.
6. Kallioniemi, O.-P., et al., *Tissue microarray technology for high-throughput molecular profiling of cancer.* Hum. Mol. Genet., 2001. **10**(7): p. 657-662.
7. Sanchez-Carbayo, M., *Use of High-Throughput DNA Microarrays to Identify Biomarkers for Bladder Cancer.* Clin Chem, 2003. **49**(1): p. 23-31.
8. Vogelstein, B. and K. Kinzler, *Cancer genes and the pathways they control.* Nature medicine, 2004. **10**(8): p. 789-799.
9. Wood, L., et al., *The Genomic Landscapes of Human Breast and Colorectal Cancers.* Science, 2007. **318**(5853): p. 1108-1113.
10. Venter, C., et al., *The Sequence of the Human Genome.* Science, 2001. **291**(5507): p. 1304-1351.
11. Kitano, H., *Systems Biology: A Brief Overview.* Science, 2002. **295**(5560): p. 1662-1664.
12. Westerhoff, H. and B. Palsson, *The evolution of molecular biology into systems biology.* Nature Biotechnology, 2004. **22**(10): p. 1249-1252.
13. Boland, C. and A. Goel, *Somatic evolution of cancer cells.* Seminars in Cancer Biology, 2005. **15**(6): p. 436-450.
14. Hornberg, J., et al., *Cancer: a Systems Biology disease.* Bio Systems, 2006. **83**(2-3): p. 81-90.
15. Futreal, P.-A., et al., *A census of human cancer genes.* Nature Reviews Cancer, 2004. **4**: p. 177-183.
16. Bishop, J.M., *The molecular genetics of cancer.* Science, 1987. **235**(4786): p. 305-311.

17. Weinberg, R.A., *Oncogenes, antioncogenes, and the molecular bases of multistep carcinogenesis.* Cancer research, 1989. **49**(14): p. 3713-3721.

18. Cho, K.R. and B. Vogelstein, *Genetic alterations in the adenoma--carcinoma sequence.* Cancer, 1992. **70**(6 Suppl): p. 1727-1731.

19. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. **100**(1): p. 57-70.

20. Fedi, P., et al., *Growth factors*, in *Cancer Medicine*. 1997, Baltimore: Williams & Wilkins. p. 41-64.

21. Laurent, A., et al., *Controlling Tumor Growth by Modulating Endogenous Production of Reactive Oxygen Species.* Cancer Research, 2005. **65**(3): p. 948-956.

22. Slamon, D.J., et al., *Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene.* Science (New York, N.Y.), 1987. **235**(4785): p. 177-182.

23. Yarden, Y. and A. Ullrich, *Growth Factor Receptor Tyrosine Kinases.* Annual Review of Biochemistry, 1988. **57**: p. 443-478.

24. Lukashev, M.E. and Z. Werb, *ECM signalling: orchestrating cell behaviour and misbehaviour.* Trends in cell biology, 1998. **8**(11): p. 437-441.

25. Giancotti, F.G. and E. Ruoslahti, *Integrin signaling.* Science (New York, N.Y.), 1999. **285**(5430): p. 1028-1032.

26. Medema, R.H. and J.L. Bos, *The role of p21ras in receptor tyrosine kinase signaling.* Critical reviews in oncogenesis, 1993. **4**(6): p. 615-661.

27. Hunter, T., *Oncoprotein networks.* Cell, 1997. **88**(3): p. 333-346.

28. Kinzler, K.W. and B. Vogelstein, *Lessons from hereditary colorectal cancer.* Cell, 1996. **87**(2): p. 159-170.

29. Weinberg, R.A., *The retinoblastoma protein and cell cycle control.* Cell, 1995. **81**(3): p. 323-330.

30. Fynan, T.M. and M. Reiss, *Resistance to inhibition of cell growth by transforming growth factor-beta and its role in oncogenesis.* Critical reviews in oncogenesis, 1993. **4**(5): p. 493-540.

31. Markowitz, S., et al., *Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability.* Science, 1995. **268**(5215): p. 1336-1338.

32. Schutte, M., et al., *DPC4 gene in various tumor types.* Cancer Research, 1996. **56**(11): p. 2527-2530.

33. Chin, L., J. Pomerantz, and R.A. DePinho, *The INK4a/ARF tumor suppressor: one gene--two products--two pathways.* Trends in biochemical sciences, 1998. **23**(8): p. 291-296.

34. Zuo, L., et al., *Germline mutations in the p16INK4a binding domain of CDK4 in familial melanoma.* Nature Genetics, 1996. **12**(1): p. 97-99.

35. Dyson, N., et al., *The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product.* Science (New York, N.Y.), 1989. **243**(4893): p. 934-937.

36. Evan, G. and T. Littlewood, *A matter of life and cell death.* Science (New York, N.Y.), 1998. **281**(5381): p. 1317-1322.

37. Green, D.R. and J.C. Reed, *Mitochondria and apoptosis.* Science (New York, N.Y.), 1998. **281**(5381): p. 1309-1312.

38. Guo, L., et al., *Effects of cytochrome c on the mitochondrial apoptosis-induced channel MAC.* Am J Physiol Cell Physiol, 2004. **286**(5): p. C1109-1117.

39. Hayflick, L., *Mortality and immortality at the cellular level. A review.* Biochemistry. BiokhimiiÍ¡a, 1997. **62**(11): p. 1180-1190.

40. Shay, J.W. and S. Bacchetti, *A survey of telomerase activity in human cancer.* European journal of cancer (Oxford, England : 1990), 1997. **33**(5): p. 787-791.

41. Bryan, T.M. and T.R. Cech, *Telomerase and the maintenance of chromosome ends.* Current opinion in cell biology, 1999. **11**(3): p. 318-324.

42. Bryan, T.M., et al., *Telomere elongation in immortal human cells without detectable telomerase activity.* The EMBO journal, 1995. **14**(17): p. 4240-4248.

43. Bodnar, A., et al., *Extension of Life-Span by Introduction of Telomerase into Normal Human Cells.* Science, 1998. **279**(5349): p. 349-352.

44. Vaziri, H. and S. Benchimol, *Reconstitution of telomerase activity in normal human cells leads to elongation of telomeres and extended replicative life span.* Current biology : CB, 1998. **8**(5): p. 279-282.

45. Counter, C., et al., *Dissociation among in vitro telomerase activity, telomere maintenance, and cellular immortalization.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(25): p. 14723-14728.

46. Halvorsen, T., G. Leibowitz, and F. Levine, *Telomerase Activity Is Sufficient To Allow Transformed Cells To Escape from Crisis.* Mol. Cell. Biol., 1999. **19**(3): p. 1864-1870.

47. Zhu, J., et al., *Telomerase extends the lifespan of virus-transformed human cells without net telomere lengthening.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(7): p. 3723-3728.

48. Greenberg, R., et al., *Short dysfunctional telomeres impair tumorigenesis in the INK4a(delta2/3) cancer-prone mouse.* Cell, 1999. **97**(4): p. 515-525.

49. Bouck, N., V. Stellmach, and S.C. Hsu, *How tumors become angiogenic.* Advances in cancer research, 1996. **69**: p. 135-174.

50. Hanahan, D. and J. Folkman, *Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis.* Cell, 1996. **86**(3): p. 353-364.

51. Folkman, J., et al., *Tumour angiogenesis*, in *Cancer Medicine*. 1997, Baltimore: Williams & Wilkins. p. 181-204.

52. Dameron, K.M., et al., *Control of angiogenesis in fibroblasts by p53 regulation of thrombospondin-1.* Science, 1994. **265**(5178): p. 1582-1584.

53. Ravi, R., et al., *Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1α.* Genes and Development, 2000. **14**: p. 34-44.

54. Sporn, M.B., *The war on cancer.* Lancet, 1996. **347**(9012): p. 1377-1381.

55. Warburg, O., F. Wind, and E. Negelein, *The Metabolism of Tumors in the Body.* The Journal of general physiology, 1927. **8**(6): p. 519-530.

56. Gottlieb, E., *Cancer: The fat and the furious.* Nature, 2009. **461**(7260): p. 44-45.

57. Delbeke, D., *Oncological Applications of FDG PET Imaging: Brain Tumors, Colorectal Cancer Lymphoma and Melanoma.* J Nucl Med, 1999. **40**(4): p. 591-603.

58. Czernin, J., *Oncological Applications of FDG-PET*, in *PET: molecular imaging and its biological applications*. 2004, Springer-Verlag: New York.

59. Warburg, O. and F. Dickens, *The Metabolism of Tumors.* The American Journal of the Medical Sciences, 1931. **182**(1): p. 123.

60. Warburg, O., *On the origin of cancer cells.* Science (New York, N.Y.), 1956. **123**(3191): p. 309-314.

61. Warburg, O. *The prime cause of cancer and prevention*. in *Annual meeting of Nobelists*. 1969. Lindau, Germany.

62. Moreno-Sánchez, R., et al., *Energy metabolism in tumor cells.* The FEBS journal, 2007. **274**(6): p. 1393-1418.

63. Moreno-Sánchez, R., et al., *The bioenergetics of cancer: Is glycolysis the main ATP supplier in all tumor cells?* BioFactors, 2009. **35**(2): p. 209-225.

64. Bonnet, S.b., et al., *A mitochondria-K+ channel axis is suppressed in cancer and its normalization promotes apoptosis and inhibits cancer growth.* Cancer cell, 2007. **11**(1): p. 37-51.

65. Semenza, G., *HIF-1 mediates the Warburg effect in clear cell renal carcinoma.* Journal of bioenergetics and biomembranes, 2007. **39**(3): p. 231-234.

66. Aisenberg, A.C., *The Glycolysis and Respiration of Tumors.* 1961.

67. Fantin, V.R. and P. Leder, *Mitochondriotoxic compounds for cancer therapy.* Oncogene, 2006. **25**(34): p. 4787-4797.

68. Hervouet, E., et al., *A new role for the von Hippel-Lindau tumor suppressor protein: stimulation of mitochondrial oxidative phosphorylation complex biogenesis.* Carcinogenesis, 2005. **26**(3): p. 531-539.

69. Weinhouse, S., *The Warburg hypothesis fifty years later.* Journal of cancer research and clinical oncology, 1976. **87**(2): p. 115-126.

70. WeinHouse, S., et al., *On Respiratory Impairment in Cancer Cells.* Science, 1956. **124**(3215): p. 267-272.

71. Loeb, L.A., *Mutator phenotype may be required for multistage carcinogenesis.* Cancer research, 1991. **51**(12): p. 3075-3079.

72. Rabinovitch, P.S., et al., *Progression to cancer in Barrett's esophagus is associated with genomic instability.* Laboratory investigation; a journal of technical methods and pathology, 1989. **60**(1): p. 65-71.

73. Kerangueven, F., et al., *Genome-wide Search for Loss of Heterozygosity Shows Extensive Genetic Diversity of Human Breast Carcinomas.* Cancer Research, 1997. **57**(24): p. 5469-5474.

74. Jiang, F., et al., *Construction of evolutionary tree models for renal cell carcinoma from comparative genomic hybridization data.* Cancer research, 2000. **60**(22): p. 6503-6509.

75. Gatenby, R. and E. Gawlinski, *The glycolytic phenotype in carcinogenesis and tumor invasion: insights through mathematical models.* Cancer research, 2003. **63**(14): p. 3847-3854.

76. Kim, J.-w. and C. Dang, *Cancer's Molecular Sweet Tooth and the Warburg Effect.* Cancer Research, 2006. **66**(18): p. 8927-8930.

77. Hsu, P. and D. Sabatini, *Cancer Cell Metabolism: Warburg and Beyond.* 2008. **134**(5): p. 703-707.

78. Shaw, R., *Glucose metabolism and cancer.* Cell division, growth and death / Cell differentiation, 2006. **18**(6): p. 598-608.

79. Jones, R. and C. Thompson, *Tumor suppressors and cell metabolism: a recipe for cancer growth.* Genes & Development, 2009. **23**(5): p. 537-548.

80. Vizán, P., S. Mazurek, and M. Cascante, *Robust metabolic adaptation underlying tumor progression.* Metabolomics, 2008. **4**(1): p. 1-12.

81. Seyfried, T. and L. Shelton, *Cancer as a metabolic disease.* Nutrition & metabolism, 2010. **7**(1): p. 7.

82. Seyfried, T. and P. Mukherjee, *Targeting energy metabolism in brain cancer: review and hypothesis.* Nutrition & metabolism, 2005. **2**.

83. Chen, Y., et al., *Oxygen consumption can regulate the growth of tumors, a new perspective on the Warburg effect.* PloS one, 2009. **4**(9).

84. Ramanathan, A., C. Wang, and S. Schreiber, *Perturbational profiling of a cell-line model of tumorigenesis by using metabolic measurements.*

Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(17): p. 5992-5997.

85. John, A.P., *Dysfunctional mitochondria, not oxygen insufficiency, cause cancer cells to produce inordinate amounts of lactic acid: the impact of this on the treatment of cancer.* Medical hypotheses, 2001. **57**(4): p. 429-431.

86. Booyens, J., *Cancer: A simple metabolic disease?* Medical Hypotheses, 1983. **12**(3): p. 195-201.

87. Galluzzi, L., et al., *Mitochondrial gateways to cancer.* Molecular aspects of medicine, 2009. **31**(1): p. 1-20.

88. Cuezva, J., et al., *The Bioenergetic Signature of Cancer.* Cancer Research, 2002. **62**(22): p. 6674-6681.

89. Kiebish, M., et al., *Cardiolipin and electron transport chain abnormalities in mouse brain tumor mitochondria: lipidomic evidence supporting the Warburg theory of cancer.* Journal of lipid research, 2008. **49**(12): p. 2545-2556.

90. Arismendi-Morillo, G. and A. Castellano-Ramirez, *Ultrastructural mitochondrial pathology in human astrocytic tumors: potentials implications pro-therapeutics strategies.* J Electron Microsc (Tokyo), 2008. **57**(1): p. 33-39.

91. *Targeting tumour metabolism.* Nature Reviews Drug Discovery, 2010. **9**: p. 503 - 504.

92. Wittig, R. and J. Coy, *The role of glucose metabolism and glucose-associated signalling in cancer.* Perspectives in medicinal chemistry, 2008. **1**: p. 64-82.

93. Vara, J.Á.F., et al., *PI3K/Akt signalling pathway and cancer.* Cancer Treatment Reviews, 2004. **30**(2): p. 193-204.

94. Manning, B.D. and L.C. Cantley, *AKT/PKB Signaling: Navigating Downstream.* Cell, 2007. **129**(7): p. 1261-1274.

95. Rose, I.A. and J.V.B. Warms, *Mitochondrial Hexokinase. Release, rebinding and location.* The Journal of Biological Chemistry, 1967. **242**(7): p. 1635-1645.

96. Miyamoto, S., A.N. Murphy, and J.H. Brown, *Akt mediates mitochondrial protection in cardiomyocytes through phosphorylation of mitochondrial hexokinase-II.* Cell Death and Differentiation, 2008. **15**: p. 521–529.

97. Pastorino, J.G., N. Shulga, and J.B. Hoek, *Mitochondrial Binding of Hexokinase II Inhibits Bax-induced Cytochrome c Release and Apoptosis.* The Journal of Biological Chemistry, 2002. **277**: p. 7610-7618.

98. Kim, J.-w., et al., *Hypoxia-inducible factor 1 and dysregulated c-Myc cooperatively induce vascular endothelial growth factor and metabolic switches hexokinase 2 and pyruvate dehydrogenase kinase 1.* Molecular and Cellular Biology, 2007. **27**(21): p. 7381-7393.

99. Esteban, M.A., et al., *Regulation of E-cadherin Expression by VHL and Hypoxia-Inducible Factor.* Cancer Research, 2006. **66**(7): p. 3567-3575.

100. Pouysségur, J., F. Dayan, and N.M. Mazure, *Hypoxia signalling in cancer and approaches to enforce tumour regression.* Nature, 2006. **441**(25): p. 437-443.

101. Pennacchietti, S., et al., *Hypoxia promotes invasive growth by transcriptional activation of the met protooncogene.* Cancer Cell, 2003. **3**(4): p. 347-361.

102. Yang, M.-H., et al., *Direct regulation of TWIST by HIF-1α promotes metastasis.* Nature Cell Biology, 2008. **10**(3): p. 295 - 305.

103. Kondoh, H., et al., *Glycolytic enzymes can modulate cellular life span.* Cancer Research, 2005. **65**(1): p. 177-185.

104. Kroemer, G. and J. Pouyssegur, *Tumor Cell Metabolism: Cancer's Achilles' Heel.* Cancer Cell, 2008. **13**(6): p. 472-482.

105. Zieglgänsberger, W. and T.R. Tölle, *The pharmacology of pain signalling.* Current opinion in neurobiology, 1993. **3**(4): p. 611-618.

106. Ahn, A., et al., *The clinical applications of a systems approach.* PLoS medicine, 2006. **3**(7).

107. Laubenbacher, R., et al., *A systems biology view of cancer.* Biochimica et biophysica acta, 2009. **1796**(2): p. 129-139.

108. Westerhoff, H., et al., *Integrating systems approaches into pharmaceutical sciences.* European journal of pharmaceutical sciences, 2008. **35**(1-2): p. 1-4.

109. Gatenby, R. and P. Maini, *Mathematical oncology: Cancer summed up.* Nature, 2003. **421**(6921): p. 321-321.

110. Preziosi, L., *Cancer Modelling and Simulation*. 2003: Chapman & Hall / CRC Press.

111. Araujo, R.P. and D.L. McElwain, *A history of the study of solid tumour growth: the contribution of mathematical modelling.* Bulletin of mathematical biology, 2004. **66**(5): p. 1039-1091.

112. Byrne, H.M., et al., *Modelling aspects of cancer dynamics: a review.* Philosophical transactions. Series A, Mathematical, physical, and engineering sciences, 2006. **364**(1843): p. 1563-1578.

113. Sanga, S., et al., *Predictive oncology: A review of multidisciplinary, multiscale in silico modeling linking phenotype, morphology and growth.* NeuroImage, 2007. **37**: p. S120-S134.

114. Bellomo, N., M. Chaplain, and E. De Angelis, *Selected Topics in Cancer Modeling: Genesis, Evolution, Immune Competition and Therapy*. 2008: BirkhÃ¤user Boston.

115. Byrne, H. and D. Drasdo, *Individual-based and continuum models of growing cell populations: a comparison.* Journal of mathematical biology, 2009. **58**(4-5): p. 657-687.

116. Rejniak, K. and L. McCawley, *Current trends in mathematical modeling of tumor-microenvironment interactions: a survey of tools and applications.* Experimental biology and medicine (Maywood, N.J.), 2010. **235**(4): p. 411-423.

117. Anderson, A., M.A.J. Chaplain, and K. Rejniak, *Single-Cell-Based Models in Biology and Medicine*. Mathematics and Biosciences in Interaction. 2007: Birkhäuser Basel.

118. Sutherland, R.M., *Cell and environment interactions in tumor microregions: the multicell spheroid model.* Science, 1988. **240**(4849): p. 177-184.

119. Greenspan, H.P., *Models for the Growth of a Solid Tumor by Diffusion.* Studies in Applied Mathematics, 1972. **LI**(4).

120. McElwain, D., *Apoptosis as a volume loss mechanism in mathematical models of solid tumor growth.* Mathematical Biosciences, 1978. **39**(1-2): p. 147-157.

121. McElwain, D.L.S. and P.J. Ponzo, *A model for the growth of a solid tumor with non-uniform oxygen consumption.* Mathematical Biosciences, 1977. **35**(3-4): p. 267-279.

122. Greenspan, H.P., *On the growth and stability of cell cultures and solid tumors.* Journal of theoretical biology, 1976. **56**(1): p. 229-242.

123. Byrne, H.M. and M.A.J. Chaplain, *Free boundary value problems associated with the growth and development of multicellular spheroids.* European Journal of Applied Mathematics, 1997. **8**(06): p. 639-658.

124. Byrne, H.M., *A weakly nonlinear analysis of a model of avascular solid tumour growth.* Journal of mathematical biology, 1999. **39**(1): p. 59-89.

125. Gatenby, R.A., *Models of tumor-host interaction as competing populations: implications for tumor biology and treatment.* Journal of theoretical biology, 1995. **176**(4): p. 447-455.

126. Gatenby, R.A., *Application of competition theory to tumour growth: implications for tumour biology and treatment.* European journal of cancer (Oxford, England : 1990), 1996. **32A**(4): p. 722-726.

127. Gatenby, R. and E. Gawlinski, *A Reaction-Diffusion Model of Cancer Invasion.* Cancer Research, 1996. **56**(24): p. 5745-5753.

128. Webb, S.D., J.A. Sherratt, and R.G. Fish, *Mathematical modelling of tumour acidity: regulation of intracellular pH.* Journal of theoretical biology, 1999. **196**(2): p. 237-250.

129. Helmlinger, G., et al., *Interstitial pH and pO2 gradients in solid tumors in vivo: High-resolution measurements reveal a lack of correlation.* Nature Medicine, 1997. **3**(2): p. 177-182.

130. Casciari, J.J., S.V. Sotirchos, and R.M. Sutherland, *Mathematical modelling of microenvironment and growth in EMT6/Ro multicellular tumour spheroids.* Cell proliferation, 1992. **25**(1): p. 1-22.

131. Casciari, J.J., S.V. Sotirchos, and R.M. Sutherland, *Variations in tumor cell growth rates and metabolism with oxygen concentration, glucose concentration, and extracellular pH.* Journal of cellular physiology, 1992. **151**(2): p. 386-394.

132. Lengeler, J.W., *Metabolic networks: a signal-oriented approach to cellular models.* Biological Chemistry, 2000. **381**(9-10): p. 911-920.

133. Rajasethupathy, P., S.J. Vayttaden, and U.S. Bhalla, *Systems modeling: a pathway to drug discovery.* Current Opinion in Chemical Biology, 2005. **9**(4): p. 400-406.

134. Schwab, E.D. and K.J. Pienta, *Modeling signal transduction in normal and cancer cells using complex adaptive systems.* Medical Hypotheses, 1997. **48**(2): p. 111-123.

135. Lamb, J., *The Connectivity Map: a new tool for biomedical research.* Nature Reviews Cancer, 2007. **7**: p. 54-60.

136. Christopher, R., et al., *Data-Driven Computer Simulation of Human Cancer Cell.* Annals of the New York Academy of Sciences, 2004. **1020**: p. 132-153.

137. Bakker, B., et al., *Network-based selectivity of antiparasitic inhibitors.* Molecular biology reports, 2002. **29**(1-2): p. 1-5.

138. Gerber, S., et al., *Drug-efficacy depends on the inhibitor type and the target position in a metabolic network--a systematic study.* In Memory of Reinhart Heinrich, 2008. **252**(3): p. 442-455.

139. Westerhoff, H.V., *Systems biology: new paradigms for cell biology and drug design.* Ernst Schering Research Foundation workshop, 2007(61): p. 45-67.

140. Krebs, H.A., *Rate control of the tricarboxylic acid cycle.* Advances in enzyme regulation, 1970. **8**: p. 335-353.

141. Groen, A.K., et al., *Control of Gluconeogenesis in Rat Liver Cells.* The Journal of Biological Chemistry, 1983. **258**(23): p. 14346-14353.

142. Moreno-Sánchez, R., et al., *Metabolic control analysis indicates a change of strategy in the treatment of cancer.* Mitochondrion, 2010. **In press**.

143. Hornberg, J., et al., *Metabolic control analysis to identify optimal drug targets*, in *Systems Biological Approaches In Infectious Diseases. Progress in drug research.*, H.I. Boshoff and C.E. Barry III, Editors. 2007. p. 171-189.

144. Cascante, M., et al., *Metabolic control analysis in drug discovery and disease.* Nature Biotechnology, 2002. **20**(3): p. 243-249.

145. Burns, J.A., et al., *Control of metabolic systems.* 1985. **10**(16).

146. Heinrich, R. and T.A. Rapoport, *A Linear Steady-State Treatment of Enzymatic Chains. General Properties, Control and Effector Strength.* The European Journal of Biochemistry / FEBS, 1974. **42**(1): p. 89-95.

147. Kacser, H. and J.A. Burns, *The control of flux.* Symposia of the Society for Experimental Biology, 1973. **27**: p. 65-104.

148. Giersch, C., *Control analysis of metabolic networks. 1. Homogeneous functions and the summation theorems for control coefficients.* European journal of biochemistry / FEBS, 1988. **174**(3): p. 509-513.

149. Westerhoff, H.V. and K. Van Dam, *Thermodynamics and Control of Biological Free Energy Transduction*. 1987: Elsevier.

150. Groen, A.K. and H.V. Westerhoff, *Modern Control Theories: A Consumer's Test*, in *Control of Metabolic Processes*, A. Cornish-Bowden and M.L. Cárdenas, Editors. 1990, Springer. p. 101-118.

151. Moreno-Sánchez, R., et al., *Metabolic Control Analysis: A Tool for Designing Strategies to Manipulate Metabolic Pathways.* Journal of Biomedicine and Biotechnoloy, 2008. **2008**(597913).

152. Boren, J., et al., *Metabolic control analysis aimed at the ribose synthesis pathways of tumor cells: a new strategy for antitumor drug development.* Molecular biology reports, 2002. **29**(1-2): p. 7-12.

153. Comín-Anduix, B., et al., *The effect of thiamine supplementation on tumour proliferation. A metabolic control analysis study.* European journal of biochemistry / FEBS, 2001. **268**(15): p. 4177-4182.

154. Marín-Hernández, A., et al., *Determining and understanding the control of glycolysis in fast-growth tumor cells. Flux control by an over-expressed but strongly product-inhibited hexokinase.* The FEBS journal, 2006. **273**(9): p. 1975-1988.

155. Hoops, S., et al., *COPASI--a COmplex PAthway SImulator.* Bioinformatics, 2006. **22**(24): p. 3067-3074.

156. Snoep, J., *The Silicon Cell initiative: working towards a detailed kinetic description at the cellular level.* Current opinion in biotechnology, 2005. **16**(3): p. 336-343.

157. Bakker, B.M., et al., *Metabolic control analysis of glycolysis in trypanosomes as an approach to improve selectivity and effectiveness of drugs.* Molecular and Biochemical Parasitology, 2000. **106**: p. 1-10.

158. Reuss, M. *Multiscale Modeling in Cancer Therapy and Synergistic Interplay of Target Identification and Drug Delivery*. in *Metabolic Engineering VIII Conference*. 2010. Jeju Island, Korea.

159. Fischer, K., et al., *Inhibitory effect of tumor cell–derived lactic acid on human T cells.* Blood, 2007. **109**(9): p. 3812-3819.

160. Sahra, I.B., et al., *Targeting Cancer Cell Metabolism: The Combination of Metformin and 2-Deoxyglucose Induces p53-Dependent Apoptosis in Prostate Cancer Cells.* Cancer Research, 2010. **70**(6): p. 2465-2475.

161. *Targeting tumour metabolism.* Nature Reviews Drug Discovery. **9**(7): p. 503-504.

162. Cao, X., et al., *Glucose uptake inhibitor sensitizes cancer cells to daunorubicin and overcomes drug resistance in hypoxia.* Cancer Chemotherapy and Pharmacology, 2007. **59**(4): p. 495-505.

163. Harmon, A.W. and Y.M. Patel, *Naringenin inhibits glucose uptake in MCF-7 breast cancer cells: a mechanism for impaired cellular proliferation.* Breast Cancer Research and Treatment, 2004. **85**(2): p. 103-110.

164. Snoep, J., et al., *DNA supercoiling in Escherichia coli is under tight and subtle homeostatic control, involving gene-expression and metabolic regulation of both topoisomerase I and DNA gyrase.* European journal of biochemistry / FEBS, 2002. **269**(6): p. 1662-1669.

165. Kahn, D. and H. Westerhoff, *Control theory of regulatory cascades*.* Journal of Theoretical Biology, 1991. **153**(2): p. 255-285.

166. Reed, J. and B. Palsson, *Genome-scale in silico models of E. coli have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states.* Genome research, 2004. **14**(9): p. 1797-1805.

167. Schuster, S., T. Pfeiffer, and D. Fell, *Is maximization of molar yield in metabolic networks favoured by evolution?* Journal of Theoretical Biology, 2008. **252**(3): p. 497-504.

168. Simeonidis, E., et al., *Why does yeast ferment? A Flux Balance Analysis study.* Biochemical Society Transactions, 2010. **38**(5).

169. León, M.P.d., H. Cancela, and L. Acerenza, *A Strategy to Calculate the Patterns of Nutrient Consumption by Microorganisms Applying a Two-Level Optimisation Principle to Reconstructed Metabolic Networks.* Journal of Biological Physics, 2008. **34**(1-2): p. 73–90.

170. Diaz-Ruiz, R., et al., *Tumor cell energy metabolism and its common features with yeast metabolism.* Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 2009. **1796**(2): p. 252-265.

171. Shlomi, T., et al., *Network-based prediction of human tissue-specific metabolism.* Nat Biotech, 2008. **26**(9): p. 1003-1010.

172. Cooper, V.S. and R.E. Lenski, *The population genetics of ecological specialization in evolving Escherichia coli populations.* Nature, 2000. **407**(6805): p. 736-739.

173. Holzhütter, H.-G., *The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks.* European journal of biochemistry / FEBS, 2004. **271**(14): p. 2905–2922.

174. Lee, S., et al., *Recursive MILP model for finding all the alternate optima in LP models for metabolic networks.* Computers & Chemical Engineering, 2000. **24**(2-7): p. 711-716.

175. Murabito, E., et al., *Capturing the essence of a metabolic network: A flux balance analysis approach.* Journal of Theoretical Biology, 2009. **260**(3): p. 445-452.

176. Annibaldi, A. and C. Widmann, *Glucose metabolism in cancer cells.* Current Opinion in Clinical Nutrition and Metabolic Care, 2010. **13**(4): p. 466-470.

177. Hsu, P.P. and D.M. Sabatini, *Cancer Cell Metabolism: Warburg and Beyond.* Cell, 2008. **134**(5): p. 703-707.

CHAPTER 2

CAPTURING THE ESSENCE OF A METABOLIC NETWORK: A FLUX BALANCE ANALYSIS APPROACH*

* The content and the organization of this chapter is given in the same format in which the manuscript was submitted to the Journal of Theoretical Biology.

ABSTRACT

As genome-scale metabolic reconstructions emerge, tools to manage their size and complexity will be increasingly important. Flux Balance Analysis (FBA) is a constraint-based approach widely used to study the metabolic capabilities of cellular or subcellular systems. FBA problems are highly underdetermined and many different phenotypes can satisfy any set of constraints through which the metabolic system is represented.

Two of the main concerns in FBA are exploring the space of solutions for a given metabolic network and finding a specific phenotype which is representative for a given task such as maximal growth rate. Here we introduce a recursive algorithm suitable for overcoming both of these concerns. The method proposed is able to find the alternate optimal patterns of active reactions of a FBA problem and identify the minimal subnetwork able to perform a specific task as optimally as the whole.

Our method represents an alternative to and an extension of other approaches conceived for exploring the space of solutions of an FBA problem. It may also be particularly helpful in defining a scaffold of reactions upon which to build up a dynamic model, when the important pathways of the system have not yet been well-defined.

INTRODUCTION

FBA is a constraint-based approach suitable for studying the range of possible phenotypes of a metabolic system [1-5]. Despite this declared intent, it is common practice to rely on one single optimal solution (usually the first one found) for further studies on the system under investigation [6-8], even though it is well-known that FBA problems generally have many more than one equivalent solutions that all satisfy the optimality criteria – a behaviour of the system that is known in optimisation as *degeneracy*. Available software packages and optimisation solvers do not typically find alternate optima, but stop after the first solution is found. This solution is then treated as representative of the pattern of flux in the real system, but no computational or biological evidence actually supports this arbitrary choice. In order to make the simulated phenotype more representative of reality, a common approach is to reduce the feasible space of solutions by enriching the system with additional information (*i.e.* constraints) retrieved from experimental measurements [9-11]. These may refer, for example, to the maximum activity of the enzymes catalysing the different reactions in the network or even to the current value of certain fluxes under the conditions considered. Moreover, in order to guarantee the feasibility of the solutions, reaction thermodynamics is often included in FBA by imposing additional non-linear constraints describing energy balance with the chemical potential [12, 13].

In the original formulation of FBA, only topological and stoichiometric information is considered, while the dynamic aspects of the system are neglected. The topology and stoichiometry of a metabolic network can be represented through the *stoichiometric matrix* **S**, whose generic element s_{ij} is the stoichiometric coefficient of metabolite *i* in reaction *j* [14]. The mass balance

equations at steady-state for all the metabolites can then be expressed through the following expression

$$\mathbf{S} \cdot \mathbf{v} = \mathbf{0} \tag{1}$$

where **v** is the *flux vector* whose elements v_j represent the rate of each reaction j in the network. Additional constraints can also be added in order to bind the fluxes within a certain range of values

$$v_j^L \leq v_j \leq v_j^U \qquad j = 1,\ 2,\ ...,n \tag{2}$$

where v_j^L and v_j^U are lower and upper bounds respectively for fluxes v_j. These constraints are generally used to determine the reversibility of the reactions as well as to establish a maximum value for their flux when the corresponding $V_{\max}$ has been measured experimentally, or to specify the availability of specific nutrients [3, 11, 15]. The polytope defined by the set of constraints (1) and (2) is called the *space of feasible solutions* or, more simply, the *feasible space*.

The main concern in FBA is finding, within the feasible space of solutions, the flux distributions that allow the system to optimally perform a specific task (*e.g.* maximization of biomass production). The general formulation of a FBA problem may then be expressed as follows

$$\begin{gathered} \max f(\mathbf{v}) \\ f : \mathbf{R}^n \rightarrow \mathbf{R} \\ \textit{subject to} \\ (1),\ (2) \end{gathered} \tag{G.F.}$$

where f is a mathematical function, called *objective function*, representing the response of the system in terms of its ability to perform the task we are interested in. The aim is then to identify feasible sets of steady-state fluxes optimizing the stated cellular function represented by f, for a metabolic network subject to a set of mass conservation constraints such as (1) [16].

The objective function f is usually represented by a linear combination of flux variables v_j

$$f = \sum_{j=1}^{n} \alpha_j v_j \tag{3}$$

This, together with the linearity of constraints (1) and (2), casts the general formulation of FBA into a linear programming (LP) problem.

Alternate optimal solutions and alternate optimal patterns

FBA problems are in general strongly underdetermined. The underlying assumption of the use of a linear objective function, in fact, implies that its level sets are hyperplanes and that the level set corresponding to the optima may be parallel to one of the edges or facets of the convex polytope defined by constraints (1) and (2). Although this leads to an infinite number of optimal solutions, we are often interested in a finite subset of them, where each solution differs from the others in the set of its active reactions (or non null vectorial components). More precisely, we are interested in finding all the different patterns of active reactions that are compatible with the criterion of optimality of an FBA problem. This set of optimal patterns is in part identified by the so called *alternate optimal solutions*, represented by the extreme points that lie at the intersection between the hyperplane of the objective function and the polytope of the feasible space [17], but is not limited to them. The optimal patterns, in fact, can also be partly derived through overpositions of different alternate optima solutions. Given the similarity of these two concepts, from now on we will refer to the set of optimal patterns as *alternate optimal patterns*.

The principle of minimal effort

It is not a trivial task to find all the alternate optimal patterns, nor to assess which one is best at representing the real pattern of fluxes in the cell under the

specific conditions considered. Attempts in addressing these two issues have been proposed separately. The literature reports a number of algorithms for finding the extreme points of the convex polytope defined by equations (1-2) [18, 19]. Lee *et al*., in particular, proposed a recursive method for finding the alternate optimal *solutions* in linear programming (LP) models for metabolic networks [17]. On the other hand, different ways have been proposed to find solutions representative of the system under examination [20-22]. Holzhütter (2004), in particular, suggested an approach based on the principle of "minimal effort" to infer the solution most likely to represent the true pattern of fluxes.

The algorithm presented in this paper is suitable for both these purposes, although the principle of minimal effort has been formulated differently. Here the set of active reactions is minimized rather than the total flux through the network (as in Holzhütter's work), hence the problem is turned into finding the minimal subnetwork with the same optimal capabilities as the whole network. The biological rationale behind this approach can be explained as follows: given a specific cellular function to accomplish and given a set of available external substrates, the cell should minimize the number of enzymes necessary to optimally perform this function. This formulation of the principle of minimal effort turns to be particularly suitable when the cellular function to optimize is the production of biomass. There is experimental evidence that decrease in the number of active reaction steps and increase in growth rate can occur in parallel. More specifically, turning off reactions that do not contribute to growth, allows nutrients and proteins that were used in these processes to be "reallocated" and used to produce biomass [23-25]. A more practical reason for applying the principle of minimal effort as stated above is that the minimal optimal solution provides a concise set of reactions that can be used as a backbone for the development of kinetic models which capture the essence of the whole system, being at the same time as small as possible. This would allow

researchers to build large-scale kinetic models of organisms, by providing an idea about which pathways can be reasonably neglected.

To illustrate the algorithm, two examples are presented. First, a "toy model" is used as a workbench to assess the suitability of the algorithm. Then, a simplified reconstruction of the central carbon metabolism of *E. Coli* is used to compare the results between our algorithm and the one presented by Lee *et al*. [17]. The detailed formulation of the algorithm is given in the *Methods* section.

RESULTS

A TOY NETWORK

The "toy model" used in this first example (Fig. 1) has been conceived to easily assess the suitability of the algorithm for the aims it addresses. The network consists of 16 reactions and 10 compounds. All fluxes are bound between -10 and 10 (arbitrary units), except for reaction 1 where the lower bound is set to 0 and the upper bound to 1.

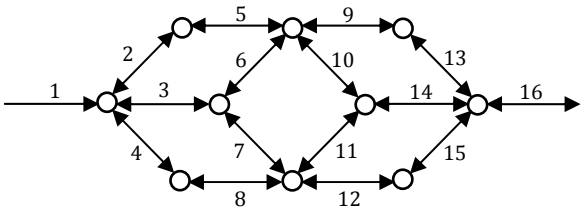

Figure 1 – A toy model consisting of 16 reactions and 10 compounds. *The number next to each reaction is the identifier of that reaction. The double pointed arrows represent reversible reactions. All the reactions in the model are reversible except reaction 1.*

The criterion of optimality is the maximization of the flux through reaction 16, which may be seen as the output of the system. By applying our algorithm, 109 alternate optimal patterns were found, with an optimum value of $v_{16} = 1$.

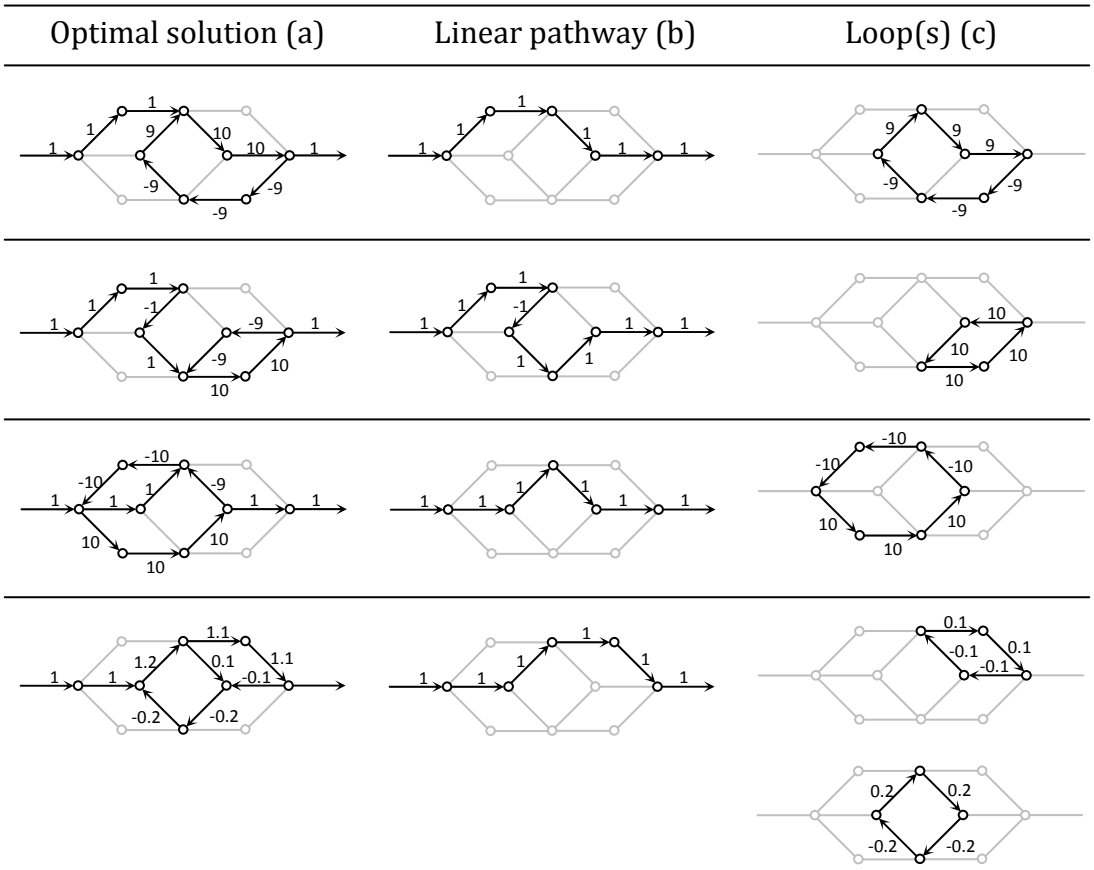

Figure 2 – Some of the alternate optimal patterns found for the toy model of Fig. 1 *when maximising the flux through reaction 16. Each of the solutions shown in column* ***a*** *is the superposition of a linear pathway, connecting the input and the output of the system (column* ***b****), and one or more loops of reactions (column* ***c****). The numbers on the active reactions represent their corresponding flux. It must be noted that the flux through an active reaction might be decomposed in two or more different fluxes (differing sometimes even in their directions) in accordance to the decomposition of the optimal solution into the corresponding linear pathway and loops.*

Some of these patterns are shown graphically in Fig. 2. It is easy to see how they can be decomposed into a linear pathway, connecting the input (v_1) and the

output (v_{16}) of the system, and one or more loops of reactions. These loops are referred as infeasible cycles because they can carry a non-null flux without any communication between the system and the external environment in terms of exchange of matter [26].

The presence of infeasible cycles is one of the major causes of degeneracy in an FBA problem. In order to prevent the optimal solutions containing any infeasible cycles, reaction thermodynamics is sometimes included in FBA by imposing additional non-linear constraints describing energy balance with the chemical potential [12, 13]. The infeasibility of these reaction loops makes them not only biophysically unachievable but also unnecessary for the optimality of the system. Consequently, in *minimal optimal solutions* (*i.e.* solutions satisfying the previous criterion of optimality *and* having the minimum possible number of active reactions) no infeasible cycles can be present. In order to find these solutions, and hence their corresponding pattern of active reactions, we fixed the flux of reaction 16 at the optimal value identified previously and we run the algorithm with a new criterion of optimality: minimizing the number of non-zero fluxes. The algorithm found 8 minimal alternate optimal patterns consisting of 6 active reactions each (Fig. 3). Thanks to the simple topology and the symmetry of this network it is easy to see that a minimal solution cannot have less than 6 active reactions and that there are no minimal patterns other than the eight identified, providing evidence of the suitability of the algorithm for finding all the alternate optimal patterns of a linear optimization problem and its minimal solutions.

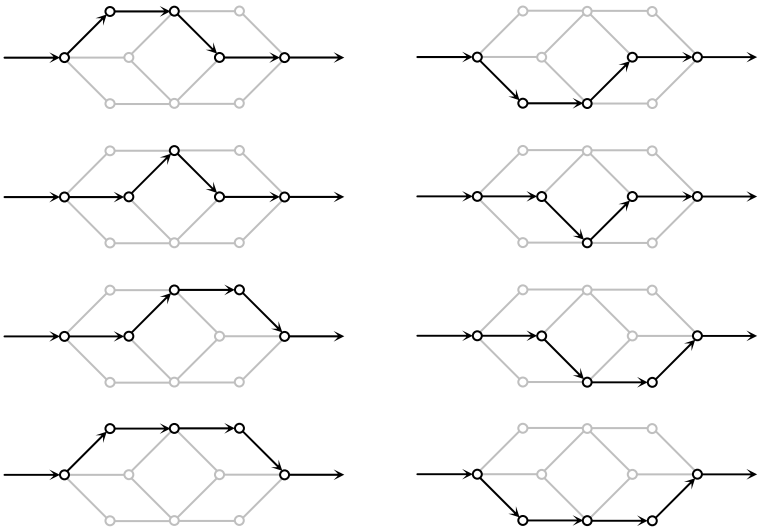

Figure 3 – Complete set of minimal optimal patterns for the model in Fig. 1. *These solutions don't present any reactions loop and are not decomposable into simpler pathways.*

A MODEL OF THE CENTRAL CARBON METABOLISM IN E. COLI.

After testing the algorithm on the toy model, we applied it on a simplified reconstruction of the central carbon metabolism of *E. Coli* (Fig. 4). This model is taken from Lee *et al* [17] in a study aimed to identify the alternatives in the carbon trafficking in mutants lacking pyruvate kinase. The model consists of 30 metabolites and 34 reactions or, in terms of the stoichiometry matrix, 30 equations (mass balances) and 34 variables. The complete set of constraints used in Lee *et al*. (balance equations, lower and upper bounds) is reported in the Appendix. Since some of the reactions are supposed to have a constant flux (equations a.17-a.27 of Appendix) it is possible to remove them from the set of variables and deal with a reduced system consisting of 18 equations and 23 variables.

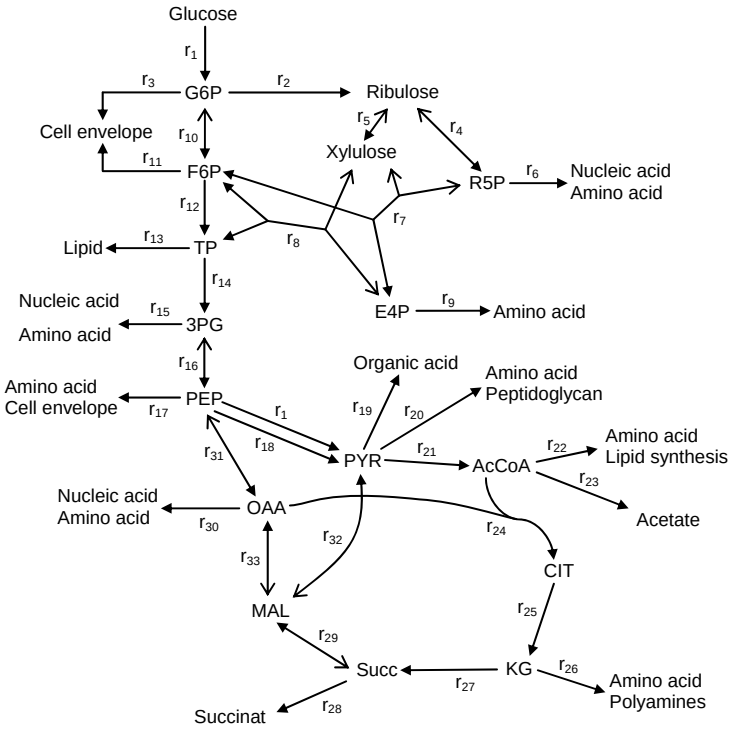

Figure 4 – Simplified metabolic map of E.coli central carbon metabolism *as proposed in Lee et al [17].*

Our algorithm was run to find the alternate optimal patterns while minimizing the flux through reaction 18 (pyruvate kinase). The algorithm found 28 alternate optimal *patterns* – versus the 9 alternate optimal *solutions* obtained by Lee *et al*. Although Lee's approach has a slightly different scope than ours (rather than finding *alternate optimal patterns* it finds *alternate optimal solutions*, which are a subset of the former), these results show an improvement in respect to similar approaches previously proposed to explore the feasible space. The complete set of optimal patterns is presented in Table 1.

We also applied our algorithm to find the minimal optimal solution(s) of the system. To do so, we fixed the flux of reaction 18 to the optimal (minimal) value found previously (this ensures the optimality of the minimal solutions we wanted to find) and we minimized, as objective function, the number of active

reactions. The algorithm found one minimal solution consisting of 27 active reactions (including the reactions with fixed flux). The solution is represented in Table 1 by the highlighted column.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
r_1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_8	1	1	1	1	0	0	0	0	0	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1
r_{10}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_{12}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_{14}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_{16}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_{18}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
r_{19}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_{21}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_{23}	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0
r_{24}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_{25}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_{27}	1	1	1	1	1	1	1	1	1	0	1	0	1	1	0	0	1	1	0	1	0	1	1	1	1	1	1	1
r_{28}	1	1	1	0	1	0	1	0	1	0	1	0	1	1	0	0	1	0	0	0	0	1	1	1	1	0	0	0
r_{29}	1	0	1	1	1	1	0	1	1	0	1	0	1	0	0	0	1	1	0	1	0	0	0	1	0	1	1	1
r_{31}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
r_{32}	0	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	0	1	1	0	1	1	1	0	0	1	1
r_{33}	1	1	1	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	1	0	0	1	1	0	0	1	1	0
r_{34}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 1 *– The columns of the matrix represent the optimal patterns of active reactions found in E.Coli central carbon metabolism. A reaction is labelled 1 or 0 when active or inactive respectively. Reactions which were always active or inactive in all the solutions found have been omitted (except for reaction 18, whose value was minimized according to the criterion of optimality considered). The highlighted column is the solution with the minimum set of active reactions.*

DISCUSSION

Even though FBA is meant in principle to provide insights on the whole range of possibilities for a metabolic network rather than identifying a specific

phenotype, it is common practice to rely only on one specific solution for further studies and analysis [6, 7]. Sometimes different solutions are found by changing the constraints on the flux variables in order to simulate, for example, different growth conditions [21, 27]. This leads *de facto* to a different FBA problem and doesn't answer the question about how many alternative ways the system has to optimally achieve a specific task in a given condition specified by a fixed set of constraints.

Finding a way to explore more extensively the range of possibilities of a metabolic network and identifying a pattern of fluxes which can reasonably represent the real trafficking of matter in the system are two of the main directions toward which the original formulation of FBA has been developed in this paper. The algorithm presented here has been conceived to address both these important issues. It represents an extension in respect to algorithms previously proposed to explore the feasible space of solutions and it can also be used to identify the minimal subnetwork(s) able to capture the essence of the whole system in terms of its ability to optimally perform a specific task.

By applying this algorithm on a toy model, we assessed its suitability to achieve these two goals. The algorithm was used firstly to enumerate the alternate optimal patterns of the system, regardless of the number of their active reactions, and secondly to identify its minimal optimal solutions (and hence the subnetworks highlighted by their patterns of active reactions). We note that this second task was performed independently from the first one and that the identification of the minimal optimal solutions can be done directly, without passing through the previous enumeration of the whole set of alternate optimal patterns. This would prove particularly useful in the case of a large and highly interconnected network, where finding all the alternate optimal patterns could be extremely time-consuming.

The requirement for an optimal (or suboptimal) solution to have the minimal set of active reactions has previously been adopted by León *et al.* in a recent study meant to calculate bacterial responses to different growth media [21]. For a given set of bounds on the exchange fluxes (defining a specific growth medium) only one solution was found. The approach we propose here has the advantage of combining the minimal effort principle, as stated in the *Introduction*, with an algorithm able to enumerate the alternate optimal patterns that optimises a given cellular function under a specified set of constraints. Merging these two features in the same algorithm provides a straightforward way to extend the results of previous studies where the focus on identifying the most "simple" solution in terms of number of active reactions was uncoupled from the problem of enumerating the different patterns in which such a solution can occur.

A second, larger network was then used to compare the results obtained through our algorithm with those presented by Lee *et al*. [17]. We found 28 alternate optimal *patterns* versus the 9 alternate optimal *solutions* identified in Lee's original work, suggesting our algorithm represents an improvement upon previous approaches proposed to explore the alternatives in the trafficking of matter in metabolic networks. The reason for this higher number of solutions (patterns) lies in the slightly different scope of the two approaches. Here we are looking for different flux distributions in terms of active reactions, regardless of the specific value of the non-zero fluxes. In Lee's work the focus was on finding solutions that were identified by both a specific pattern of active reactions and unique values of their fluxes, thus limiting the solutions found to a subset of all the possible alternate optimal patterns.

We also mention that the set of the minimal optimal patterns constitutes a subset of the *elementary flux modes* of a metabolic network (see [28] for their definition). Recently, Figueiredo et al. [29] proposed an algorithm to compute

the shortest of these elementary flux modes, of which the minimal optimal patterns are still a subset. Figueiredo's algorithm can indirectly provide with an efficient way to identify also the minimal optimal patters, as we defined them. Once the shortest elementary modes have been found, one has just to check whether they can carry a pattern of fluxes satisfying the given criterion of optimality. We underline, however, that the benefit of our algorithm consists of combining in the same formulation a means to explore the different ways in which a given biological function can be performed in optimal conditions with a way to identify the minimal metabolic resources that have to be put in place to achieve that same degree of optimality.

Further improvements in our approach may be achieved by taking into account the thermodynamics of the system. Currently, the only aspect related to thermodynamic properties is introduced through the upper and lower bounds of the flux variables. Setting one of these bounds to zero, in fact, provides the corresponding reaction with a preferential directionality. Our approach could benefit from considering the thermodynamics of the whole optimal solutions. Following a method proposed by Holzhütter [22], a possible way to do that is by using the equilibrium constants of all the enzymatic steps to weight the fluxes through their corresponding reactions and hence assign an energy cost to the whole solution. This cost would then be used to rank the different solutions by assessing to what extent they are thermodynamically favourable. Although our algorithm is conceived to find alternate optimal patters, which are not necessarily identified by unique sets of fluxes (unless they corresponds to alternate optimal *solutions*), the implementation of this method does not require any major change to our original algorithm and would provide it with a second criterion to assess the suitability of the solutions in representing the real metabolic phenotype.

METHODS

The central idea of the algorithm proposed in this paper consists in introducing for each flux v_j a binary variable w_j^0 which flags the corresponding reaction as active or inactive:

$$w_j^0 = 1 \Leftrightarrow v_j = 0 \tag{4}$$

In order to make this flagging mechanism work, the general FBA formulation set out in the *Introduction* (model G.F.) was enriched with a new set of constraints. In particular:

- Every flux variable v_j was split into three different variables: v_j^0, v_j^- and v_j^+.

$$v_j = v_j^0 + v_j^- + v_j^+ \tag{5.a}$$

  These new variables are meant to carry respectively the "almost null", negative and positive part of the flux. If flux v_j is positive, for example, then $v_j^- = v_j^0 = 0$ and $v_j = v_j^+ > 0$ (similarly when the flux is negative). If the flux is "reasonably" close to zero to be considered null (*i.e.* its value is smaller than a given threshold ε), then we have $v_j^- = v_j^+ = 0$ and $v_j = v_j^0 \simeq 0$.

- For each flux v_j three binary variables, w_j^0, w_j^- and w_j^+, were further introduced in the model. These variables are subject to the following constraint

$$w_j^0 + w_j^- + w_j^+ = 1 \tag{5.b}$$

so that only one of them can be non zero (i.e. 1) at the same time. The introduction of these binary variables recast the FBA problem into a mixed integer linear programming (MILP) problem.

- Binary variables w_j^0, w_j^- and w_j^+ were coupled respectively with the previously introduced continuous variables v_j^0, v_j^- and v_j^+ through the following constraints:

$$v_j^L(1-w_j^0) \le v_j \le v_j^U(1-w_j^0) \quad (5.c)$$

$$-\varepsilon\, w_j^0 < v_j^0 < \varepsilon\, w_j^0 \quad (5.d)$$

$$v_j^L w_j^- \le v_j^- \le \varepsilon\, w_j^- \quad (5.e)$$

$$\varepsilon\, w_j^+ \le v_j^+ \le w_j^+ v_j^U \quad (5.f)$$

where ε, as mentioned above, is a very small number used as the *tolerance* for considering a flux as zero or non-zero (a flux is considered zero when $|v_j| < \varepsilon$).

Constraint (5.c) implements the forward implication of equation (4), *i.e.* when w_j^0 equals 1 the corresponding flux v_j is forced to zero. On the other hand, the backward implication is assured by the constraints (5.d-f) and (5.b) It's easy to prove this statement using a *reductio ad absurdum* argument. If the flux v_j was zero (i.e. $|v_j| < \varepsilon$) with $w_j^0 = 0$, then either $w_j^+ = 1$ or $w_j^- = 1$ (equation 5.b). Let's assume $w_j^+ = 1$ (this won't make our argument lose in generality). In this case, $w_j^0 = w_j^- = 0$ (equation 5.b), and consequently $v_j^0 = v_j^- = 0$ (equations 5.d and 5.e). This implies that $v_j = v_j^+ \ge \varepsilon$ (equation 5.a and 5.f) contradicting the assumption that v_j equals zero (Q.E.D).

The new formulation of the FBA problem can then be written as follows:

$$\begin{array}{c} \max f(\mathbf{v}) \\ f:\mathbf{R}^n \to \mathbf{R} \\ s.t. \\ (1),\ (5.a\text{-}f) \end{array} \qquad \text{(N.F.1)}$$

The tolerance value ε used as a threshold to discriminate between zero and non-zero fluxes was set to 10^{-6}. Model N.F.1 was formulated in GAMS (General Algebraic Modelling System) and solved using the CPLEX solver.

Alternate optimal patterns

In order to find alternate optimal patterns, the model N.F.1 was run iteratively. At each iteration a new constraint (an *integer cut*) was added to "cut away" from the feasible space solutions already found in previous iterations. Specifically, at iteration K the following set of integer cuts have been added to the system:

$$\sum_{j\in Z^k} w_j^0 - \sum_{j\in \overline{Z}^k} w_j^0 \le \left|Z^k\right| - 1, \qquad k = 1,\ 2,\ ...,\ K-1 \tag{6}$$

where Z^k is the set of indexes corresponding to the null fluxes (*i.e.* inactive reactions) at iteration *k*, $\left|Z^k\right|$ denotes its cardinality and $\overline{Z}^k$ is the complement of Z^k over the whole index set.

More specifically, the algorithm used to find alternate optimal patterns can be described as follows:

Step 1

a) In the first iteration ($K = 1$), model G.F. is solved. The objective function is chosen according to the optimal criterion we want to use.

b) Set $Z^k\left(k = 1\right)$ is defined.

Step K **, for** $K \geq 2$.

a) Model N.F.1 is solved.

b) Set Z^k is defined.

c) If the optimal value of the objective function is not smaller than the value found in the previous iteration the algorithm repeats step *K*, otherwise it stops.

Finding the minimal subnetworks

To find the minimal subnetwork(s) with the same optimal capability of the whole system, an optimisation was performed using the following formulation:

$$\max \sum_j w_j^0$$

(N.F.2) subject to

(1), (5.a-f)

The only difference between N.F.2 and N.F.1 consists in the objective function, here defined as $f := \sum_j w_j^0$, *i.e.* the objective is to maximise the number of reactions with a zero flux. Maximising *f* forces the solution(s) to have the minimal number of active fluxes.

To assure the optimality of the minimal solution(s) found through N.F.2, the general FBA formulation (model G.F.) was first run, in order to find the optimal value of the flux we wanted to optimise; subsequently the flux optimized through G.F. was fixed to the optimal value identified and model N.F.2 was then solved to find the minimal solution(s). As before, model N.F.2 was formulated in GAMS and solved using the CPLEX solver.

NOMENCLATURE

We present below the complete list of variables and parameters used in the algorithm and their correspondent description

Indices	i	Index of metabolites
	j	Index of reactions
	k	Index of iterations
Sets	Z^k	Set of indexes j corresponding to the null fluxes at iteration k.
	$\overline{Z}^k$	Complement of Z^k over the whole index set of the reactions.
Continuous variables	v_j	Variable representing the flux through reaction j.
	v_j^0	Variable carrying the almost zero part of flux v_j.
	v_j^-	Variable carrying the negative part of flux v_j.
	v_j^+	Variable carrying the positive part of flux v_j.
Binary variables	w_j^0	Equals 1 if v_j is zero, equals 0 otherwise.
	w_j^-	Equals 1 if v_j is negative, equals 0 otherwise.

	w_j^+	Equals 1 if v_j is positive, equals 0 otherwise.
Parameters	**S**	Stoichiometric matrix
	v_j^L	Lower bound of flux v_j
	v_j^U	Upper bound of flux v_j
	ε	Tolerance for considering a flux as zero or non-zero (flux v_j is considered zero when $\left\|v_j\right\| < \varepsilon$).

APPENDIX

Set of balance equations defining the system represented in Fig. 4.

$$r_1 - r_2 - r_3 - r_{10} = 0 \quad \text{(a.1)}$$

$$r_{10} - r_{12} - r_{11} + r_7 + r_8 = 0 \quad \text{(a.2)}$$

$$r_{14} - r_{15} - r_{16} = 0 \quad \text{(a.3)}$$

$$-r_1 + r_{16} - r_{18} - r_{17} - r_{31} = 0 \quad \text{(a.4)}$$

$$r_1 + r_{18} - r_{21} - r_{19} - r_{20} + r_{32} = 0 \quad \text{(a.5)}$$

$$-r_{31} - +r_{24} + r_{30} - r_{33} = 0 \quad \text{(a.6)}$$

$$r_{21} - r_{24} - r_{23} - r_{22} = 0 \quad \text{(a.7)}$$

$$r_{24} - r_{25} = 0 \quad \text{(a.8)}$$

$$r_{25} - r_{27} - r_{26} = 0 \quad \text{(a.9)}$$

$$-r_{29} + r_{27} - r_{28} = 0 \quad \text{(a.10)}$$

$$r_2 - r_4 - r_5 = 0 \quad \text{(a.11)}$$

$$-r_7 + r_4 - r_6 = 0 \quad \text{(a.12)}$$

$$-r_7 - r_8 + r_5 = 0 \quad \text{(a.13)}$$

$$r_7 - r_8 - r_9 = 0 \quad \text{(a.14)}$$

$$2r_{12} + r_8 - r_{14} - r_{13} = 0 \quad \text{(a.15)}$$

$$r_{29} - r_{33} - r_{32} = 0 \quad \text{(a.16)}$$

$$2.5r_3 = 0.205 \quad \text{(a.17)}$$

$$2.5r_{11} = 0.0709 \quad \text{(a.18)}$$

$$2.5r_{13} = 0.129 \quad \text{(a.19)}$$

$$2.5r_{15} = 1.493 \quad \text{(a.20)}$$

$$2.5r_{17} = 0.7191 \quad \text{(a.21)}$$

$$2.5r_6 = 0.897 \quad \text{(a.22)}$$

$$2.5r_9 = 0.361 \quad \text{(a.23)}$$

$$2.5r_{20} = 2.833 \quad \text{(a.24)}$$

$$2.5r_{22} = 2.928 \quad \text{(a.25)}$$

$$2.5r_{26} = 1.078 \quad \text{(a.26)}$$

$$2.5r_{30} = 1.786 \quad \text{(a.27)}$$

$$2r_2 + r_{25} + r_{32} = 7.2 \quad \text{(a.28)}$$

$$\begin{aligned} &r_{\mathrm{ATP}} + r_{12} - 3r_{14} - r_{18} - r_{23} \\ &-3r_{27} - 2r_{21} - 2r_{33} - r_{29} \\ &-r_{31} + 2r_{19} = 0 \end{aligned} \quad \text{(a.29)}$$

ACKNOWLEDGMENT

We acknowledge the support of the BBSRC/EPSRC Grant BB/C008219/1 “The Manchester Centre for Integrative Systems Biology” (MCISB). Particular thanks to Lazaros Papageorgiu for the useful discussion.

REFERENCES

1. Edwards, J.S. and B.O. Palsson, *The Escherichia coli MG1655 in silico metabolic genotype: Its definition, characteristics, and capabilities.* PNAS, 2002. **97**(10): p. 5528-5533.
2. Kauffman, K.J., P. Prakash, and J.S. Edwards, *Advances in flux balance analysis.* Current Opinion in Biotechnology, 2003. **14**: p. 491-496.
3. Förster, J., et al., *Genome-Scale Reconstruction of the Saccharomyces cerevisiae Metabolic Network.* Genome Research, 2003. **13**: p. 244-253.
4. Varma, A., B.W. Boesch, and B.O. Palsson, *Biochemical Production Capabilities of Escherichia Coli.* Biotechnology and Bioengineering, 1993. **42**: p. 59-73.
5. Edwards, J.S. and B.O. Palsson, *Systems Properties of the Haemophilus influenzae Rd Metabolic Genotype.* Journal of Biological Chemistry, 1999. **274**(25): p. 17410-17416.
6. Savinell, J.M. and B.O. Palsson, *Network analysis of intermediary metabolism using linear optimization. II. Interpretation of hybridoma cell metabolism.* Journal of theoretical biology, 1992. **154**(4): p. 455-473.
7. Ramakrishna, R., et al., *Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints.* Am J Physiol Regul Integr Comp Physiol, 2001. **280**(3): p. 695-704.
8. Fell, D.A. and R.J. Small, *Fat synthesis in adipose tissue: an examination of stoichiometric constraints.* Biochemical Journal, 1986. **238**: p. 781-786.
9. Bonarius, H.P.J., et al., *Metabolic Flux Analysis of Hybridoma Cells in Different Culture Media Using Mass Balances.* Biotechnology and Bioengineering, 1996. **50**: p. 299-318.
10. Llaneras, F. and J. Picó, *Stoichiometric Modelling of Cell Metabolism.* Journal of Bioscience and Bioengineering, 2008. **105**(1): p. 1-11.
11. Varma, A. and B. Palsson, *Stoichiometric Flux Balance Models Quantitatively Predict Growth and Metabolic By-Product Secretion in Wild-Type Escherichia coli W3110.* Applied and environmental Microbiology, 1994. **60**(10): p. 3724-3731.
12. Beard, D.A., et al., *Thermodynamic constraints for biochemical networks.* Journal of Theoretical Biology, 2004. **228**: p. 327-333.
13. Beard, D.A., S.-d. Liang, and H. Qian, *Energy Balance for Analysis of Complex Metabolic Networks.* Biophysical Journal, 2002. **83**(1): p. 79-86.
14. Lee, J.M., E. P.Gianchandani, and J.A. Papin, *Flux balance analysis in the era of metabolomics.* Briefings in Bioinformatics, 2006. **7**(2): p. 140-150.

15. Schilling, C.H., et al., *Genome-Scale Metabolic Model of Helicobacter pylori 26695.* Journal of Bacteriology, 2002. **184**(16): p. 4582-4593.

16. Price, N.D., J.L. Reed, and B.Ø. Palsson, *Genome-scale models of microbial cells: evaliating the consequences of constraints.* Nature, 2004. **2**: p. 886-897.

17. Lee, S., et al., *Recursive MILP model for finding all the alternate optima in LP models for metabolic networks.* Computers and Chemical Engineering, 2000. **24**(2-7): p. 711-716.

18. Matheiss, T.H. and D.S. Rubin, *A Survey and Comparison of Methods for Finding All Vertices of Convex Polyhedral Sets.* Mathematics of Operations Research, 1980. **5**(2): p. 167-185.

19. Swart, G., *Finding the convex hull facet by facet.* Journal on Algorithms, 1985. **6**: p. 17-48.

20. Smallbone, K. and E. Simeonidis, *Flux balance analysis: A geometric perspective.* Journal of Theoretical Biology, 2009. **258**: p. 311-315.

21. León, M.P.d., H. Cancela, and L. Acerenza, *A Strategy to Calculate the Patterns of Nutrient Consumption by Microorganisms Applying a Two-Level Optimisation Principle to Reconstructed Metabolic Networks.* Journal of Biological Physics, 2008. **34**(1-2): p. 73-90.

22. Holzhütter, H.-G., *The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks.* European Journal of Biochemistry, 2004. **271**(14): p. 2905–2922.

23. Cooper, V.S. and R.E. Lenski, *The population genetics of ecological specialization in evolving Escherichia coli populations.* Nature, 2000. **407**: p. 736-739.

24. Acerenza, L. and M. Graña, *On the Origins of a Crowded Cytoplasm.* Journal of Molecular Evolution, 2006. **63**: p. 583-590.

25. Graña, M. and L. Acerenza, *A model combining cell physiology and population genetics to explain Escherichia coli laboratory evolution.* BMC Evolutionary Biology, 2001. **1**(12).

26. Schilling, C.H., D. Letscher, and B.O. Palsson, *Theory for the Systemic Definition of Metabolic Pathways and their use in Interpreting Metabolic Function from a Pathway-Oriented Perspective.* Journal of Theoretical Biology, 2000. **203**: p. 229-248.

27. Edwards, J.S. and B.O. Palsson, *Metabolic flux balance analysis and in silico analysis of Escherichia coli K-12 gene deletions.* BMC Bioinformatics, 2000. **1**(1).

28. Schuster, S. and C. Hilgetag, *On elementary flux modes in biochemical reaction systems at steady state.* Journal of Biological Systems, 1994. **2**(2): p. 165-182.

29. de Figueiredo, L.F., et al., *Computing the shortest elementary flux modes in genome-scale metabolic networks.* Bioinformatics, 2009. **25**(23): p. 3158–3165.

CHAPTER 3

A PROBABILISTIC APPROACH TO IDENTIFY PUTATIVE DRUG TARGETS IN BIOCHEMICAL NETWORKS†

† The content and the organization of this chapter is given in the same format in which the manuscript was submitted to the Journal of the Royal Society Interface.

ABSTRACT

Network-based drug design holds great promise in clinical research as a way to overcome the limitations of traditional approaches in the development of drugs with high efficacy and low toxicity. This novel strategy aims to study how a biochemical network as a whole, rather than its individual components, responds to specific perturbations in different physiological conditions. Proteins exerting little control over normal cells and larger control over altered cells may be considered as good candidates for drug targets. The application of network-based drug design would greatly benefit from using an explicit computational model describing the dynamics of the system under investigation. However, creating a fully characterized kinetic model is not an easy task, even for relatively small networks, as it is still significantly hampered by the lack of data about kinetic mechanisms and parameters values.

Here we propose a Monte Carlo approach to identify the differences between flux control profiles of a metabolic network in different physiological states, when information about the kinetics of the system is partially or totally missing. Based on experimentally-accessible information on metabolic phenotypes, we develop a novel method to determine probabilistic differences in the flux control coefficients between the two observable phenotypes. Knowledge of how differences in flux control are distributed among the different enzymatic steps is exploited to identify points of fragility in one of the phenotypes. Using a prototypical cancerous phenotype as an example, we demonstrate how our approach can assist researchers in developing compounds with high efficacy and low toxicity.

INTRODUCTION

The main challenge in drug discovery consists of developing drugs which are both effective and selective, hence non-toxic. In this respect, drug development approaches have often assumed that the enzyme chosen as target plays the role of "rate limiting step" for the biological function of interest [1]. However, single and completely rate-limiting steps barely exist [2] and it is not easy to identify them if they do. Potent inhibitors of essential enzymes often do not show the expected effect at the cellular level [3-5]. Often, the network of biochemical interactions in living cells buffers changes introduced in a single enzymatic step. Likewise, local changes may induce unforeseen adverse side-effects on the whole system [4]. This interconnectedness of biological function usually results in poor predictive power with respect to the requirement of high efficiency and low toxicity for a drug.

With the advent of Systems Biology, drug discovery has been shifting its focus from a single-molecule to a system-level perspective, and the concept of network-based drug design has been introduced [3-6]. Within this new paradigm, the aim is to study a biochemical network as a whole and identify points of fragility specifically characterizing an altered phenotype (as in a cancer cell) [7, 8]. By targeting these points of fragility, a significant response in altered cells is induced. Differential network-based drug design then maximizes the difference in response of target cell versus normal cell [9]. A possible implementation of network-based drug design is based on Metabolic Control Analysis (MCA), where the concept of fragility is expressed in terms of control coefficients [10]. *Differential* MCA, in particular, has been proposed as a tool to understand how the system responds to specific perturbations under different physiological conditions or in different host cells, hence providing a way to

assess both the efficiency and the specificity of a compound designed to target specific enzymes [3, 9, 11].

However, one fundamental requirement for applying MCA is the availability of a fully-characterized kinetic model of the system under study. Creating such a model is not an easy task, even for relatively small metabolic networks [12, 13]. In most cases, the detailed enzymatic mechanisms governing the dynamics of the different metabolic steps are unknown and precise knowledge of kinetic parameters under the relevant *in vivo* conditions is usually not available. The resulting uncertainty in predicting the dynamic behaviour is significant and increases drastically with the size of the system [12]. Randomized sampling of the parameter space represents a way in which such uncertainty can be quantified and probabilistic insights about the dynamical behaviour of the system can be obtained. Such probabilistic approaches have been recently used in a number of different studies, spanning from applications on MCA [14, 15], on metabolic engineering [16], to a description of the dynamics in metabolic networks [17-19]. Here we build upon these ideas and present a Monte Carlo approach to identify differences between the control profiles of a metabolic network in two different settings, one representative of a tumour cell and the other of a corresponding normal cell. Our aim is to show how putative targets for drugs operating at the metabolic level may be identified in a probabilistic manner, when only partial knowledge is available with respect to the kinetic properties of the system. In particular, our analysis is based on experimentally accessible information on metabolic phenotypes, such as observable concentrations and fluxes, rather than on detailed knowledge of kinetic parameters. It is demonstrated that a combination of such phenotypic data, together with heuristic assumptions about the properties of typical enzyme-catalyzed reactions, already allows for a fast and efficient way to explore the effectiveness of putative drug targets. Our method makes use of biophysical constraints on the metabolic network, as provided by mass-conservation and

thermodynamics, and implements an efficient numerical scheme to allow scanning of a large parameter-space, making it applicable to networks of large size. As a proof-of-concept, we apply our methodology to identify the points of fragility characterizing a paradigmatic cancer metabolic phenotype. We demonstrate that our method allows us to identify those enzymes that exert a high differential control upon a given relevant system property, and thus represent suitable sites to specifically target the cancerous phenotype.

METHODS

METABOLIC CONTROL ANALYSIS

The dynamic behaviour of a metabolic system, consisting of *m* metabolites and *r* reactions, can be described by a set of ordinary differential equations of the form

$$\frac{d\mathbf{S}}{dt} = \mathbf{N} \cdot \mathbf{v} \tag{1}$$

where $\mathbf{S}$ is an *m*-dimensional vector denoting the set of metabolite concentrations and $\mathbf{N}$ is the $m \times r$ stoichiometry matrix. The *r*-dimensional vector $\mathbf{v} = \mathbf{v}(\mathbf{S}, \mathbf{K})$ describes the functional form of the reaction rates, which depend on the metabolite concentrations $\mathbf{S}$ and the set of kinetic parameters $\mathbf{K}$. The presence of mass-conservation relationships and conserved chemical moieties in the network results in linear dependencies among the rows of $\mathbf{N}$. In this case, one distinguishes between a set of independently variable metabolite concentrations $\mathbf{S}^{\text{ind}}$ and a set of dependent metabolite concentrations $\mathbf{S}^{\text{dep}}$. To characterize the dynamics of the system it is sufficient to consider the set of independent variables, making use of a reduced stoichiometry matrix $\mathbf{N'}$ consisting of linearly independent rows only. The link between $\mathbf{N'}$ and $\mathbf{N}$ is provided by the expression $\mathbf{N} = \mathbf{L} \cdot \mathbf{N'}$, where $\mathbf{L}$ denotes the link matrix [20].

To evaluate the system, each reaction has to be assigned particular rate equation and each kinetic parameter needs to be assigned a specific quantitative value. Given a set of initial conditions, Eq. (1) can then be solved computationally, using standard methods of numerical integration. Usually, the model gives rise to one or more steady states $\left(\mathbf{v}^0,\mathbf{S}^0\right)$ that can be compared to experimentally observed values. The response of the system to a perturbation of a kinetic parameter, representing for example the action of a drug, can be quantified in terms of *control coefficients*, as introduced by Metabolic Control Analysis (MCA) [21-24]. In particular, given an effector p_i which acts directly on the enzymatic step *i*, the (*scaled*) *concentration control coefficient* C_i^S is defined as [20, 25]

$$C_i^S := \left(\frac{dS}{dp_i}\frac{p_i}{S}\right) \Bigg/ \left(\frac{dv_i}{dp_i}\frac{p_i}{v_i}\right) = \left(\frac{d\ln S}{d\ln v_i}\right) \tag{2}$$

and quantifies the ratio between the relative change of the steady-state concentration of metabolite *S* and the relative change in the catalytic action of enzyme *i* (induced by an infinitesimal change in the parameter p_i). Analogously, the *flux control coefficient* C_i^J is defined as

$$C_i^J := \left(\frac{dJ}{dp_i}\frac{p_i}{J}\right) \Bigg/ \left(\frac{dv_i}{dp_i}\frac{p_i}{v_i}\right) = \left(\frac{d\ln J}{d\ln v_i}\right), \tag{3}$$

and quantifies the response of the system in terms of the relative change in a steady-state flux *J*. Reder [20] showed that the matrix $\mathbf{C}^S$ of the concentration control coefficients can be expressed as

$$\mathbf{C}^S = -\left(\mathbf{D}_{\mathbf{S}^0}\right)^{-1} \cdot \mathbf{L} \cdot \left(\mathbf{J}'\right)^{-1} \cdot \mathbf{N}' \cdot \mathbf{D}_{\mathbf{v}^0} \tag{4}$$

where **J'** denotes the Jacobian of the system with respect to the independent variables, while $\mathbf{D}_{\mathbf{v}^0}$ and $\mathbf{D}_{\mathbf{S}^0}$ denote square matrices with elements $\mathbf{S}^0$ and $\mathbf{v}^0$ on the diagonal, respectively, and zero elsewhere.

For the matrix of flux control coefficients $\mathbf{C}^J$ the following expression holds:

$$\mathbf{C}^J = \mathbf{1} + \mathbf{D}_{\mathbf{v}^0}^{-1} \cdot \left.\frac{\partial \mathbf{v}}{\partial \mathbf{S}}\right|_0 \cdot \mathbf{D}_{\mathbf{S}^0} \cdot \mathbf{C}^S . \tag{5}$$

Because **J'** can be written in terms of the link matrix and the reduced stoichiometry matrix as

$$\mathbf{J'} = \mathbf{N'} \cdot \frac{\partial \mathbf{v}}{\partial \mathbf{S}} \cdot \mathbf{L}, \tag{6}$$

it follows from Eqs. (4)-(5)-(6) that, once the network topology (as represented by **L** and **N'**) and the network "phenotype" (as represented by $\mathbf{D}_{\mathbf{v}^0}$ and $\mathbf{D}_{\mathbf{S}^0}$) are known, the only quantities to be evaluated in order to retrieve the control coefficients are the partial derivatives $\partial \mathbf{v}/\partial \mathbf{S}$. These, in turn, depend of the functional form of the rate equations governing the dynamics of each enzymatic step, the kinetic parameters, and the metabolic state where the derivatives are evaluated. When full kinetic information is available, the control coefficients can be readily calculated using available software packages, such as Copasi [26], or online simulation tools for biochemical models that reside in repositories, such as JWS Online [27].

DIFFERENTIAL METABOLIC CONTROL ANALYSIS

To locate suitable drug targets, one aims to identify those enzymes where a perturbation – usually a decrease in enzyme activity induced by an applied inhibitor – elicits a large response in at least one vital aspect of the functioning

of the diseased cell, whereas a similar decrease in the activity of the same enzyme is less detrimental for the functioning, or "phenotype", of normal cells. We note that the diseased metabolic phenotype often corresponds to changes in either enzyme concentrations (because of mutations in regulatory elements or signal transduction genes), or availability of external substrates (influx), whereas most Michaelis-Menten constants remain unchanged. Given detailed kinetic models of both phenotypes, the putative action of a drug is applied to both models and the difference in the response is assessed using differential MCA or direct simulations.

The vital aspect of the diseased cell that is targeted may be the production flux of ATP or it could be the growth rate itself. Because it is often a flux, we shall denote it by J_{target}. Our aim is then to locate suitable sites for the action of a drug which result in the maximal differential response in this desired flux between diseased phenotype and normal phenotype. Suitable drug targets are chosen according to the following, alternative or simultaneous, criteria:

- **Maximal selectivity**. We assume that the effect we aim for is the decrease of J_{target} in the diseased cells. In this case we are interested in enzymes which exert a positive control upon J_{target} in the altered phenotype: $C_{i\,(\text{disease})}^{J} > 0$. Independent of whether the effect induced in normal cells consists of a decrease or increase of J_{target}, we require this effect to be lower, in magnitude, than in altered cells. In terms of flux control coefficients of an enzyme *i*, this criterion can be expressed as follow

$$S_i^{J_{\text{target}}} := C_{i\,(\text{disease})}^{J_{\text{target}}} - \left| C_{i\,(\text{normal})}^{J_{\text{target}}} \right| > 0 \qquad (7)$$

  The higher the value of $S_i^{J_{\text{target}}}$, the stronger the differential response between normal and altered cells. We note that Eq. (7) does not impose any

restriction on the sign nor the value of $C_{i\,(\text{normal})}^{J_{\text{target}}}$, provided that its magnitude is smaller than the magnitude of $C_{i\,(\text{disease})}^{J_{\text{target}}}$. Indeed, enzymes with non-negligible negative value of $C_{i\,(\text{normal})}^{J_{\text{target}}}$ may be also considered suitable targets as long as the difference in magnitude with $C_{i\,(\text{disease})}^{J_{\text{target}}}$ allows to reach an acceptable specificity through an appropriate dosage of an applied inhibitor.

- **Minimal toxicity**. To avoid other system properties (not captured by the flux of interest J_{target}) undergoing important changes in normal cells, we require the normal phenotype to be *wholly* robust against perturbation in the drug target activity. In this respect we define the *toxicity coefficient*, which quantifies how far the normal phenotype deviates from its original metabolic state after an inhibition of enzyme *i*:

$$T_i = \frac{1}{N}\sum_j \left| C_{i\,(\text{normal})}^{J_j} \right| \quad j \in \{\text{all exchange fluxes}\}, \tag{8}$$

where the summation is over all the *N* exchange fluxes of the system. This definition of fragility reflects a black-box perspective, where the behaviour of the system is assessed through its inputs and outputs. However, other definitions of fragility are possible and are examined later in *Discussion*.

A MONTE CARLO STRATEGY

A straightforward implementation of the strategy described above is only rarely applicable, as detailed kinetic models, describing the normal as well as the diseased phenotype, are usually not available. To overcome this limitation, we implement a Monte Carlo strategy which allows us to deal with incomplete knowledge of kinetic parameters and explicitly takes into account available

phenotypic data. In particular, we assume that for both states of the system the measured metabolic phenotype, characterized by a set of concentrations and fluxes, is known. In this respect, our approach is based on the fact that high-throughput metabolomics and fluxomics studies are now standard techniques in the analysis of cellular metabolism [28-34]. We proceed according to the following rationales:

(i) We require that the map of the metabolic network of interest is known and is the same for both the diseased and the normal phenotype. In addition, we assume that the topology of regulatory interactions is, as far as possible, available.

(ii) For some of the enzymatic steps, information about the detailed functional form of the rate laws may be available. However, in the absence of such information, we employ heuristic assumptions about generic reaction characteristics to describe the dependencies of flux rates with respect to substrates, products and allosteric effectors.

(iii) To obtain a probabilistic understanding on how the control is distributed, the kinetic parameters are sampled randomly from pre-assigned intervals. Each sampled set of parameters is made compliant with the given metabolic phenotypes and additional thermodynamic and biophysical constraints, by rescaling the maximal activity of every reaction step.

(iv) For each sampled set of parameters the differential response in the normal and the disease phenotype is evaluated.

Our workflow is illustrated in Fig. 1. In the following, each step is described more thoroughly and the details of its implementation are given.

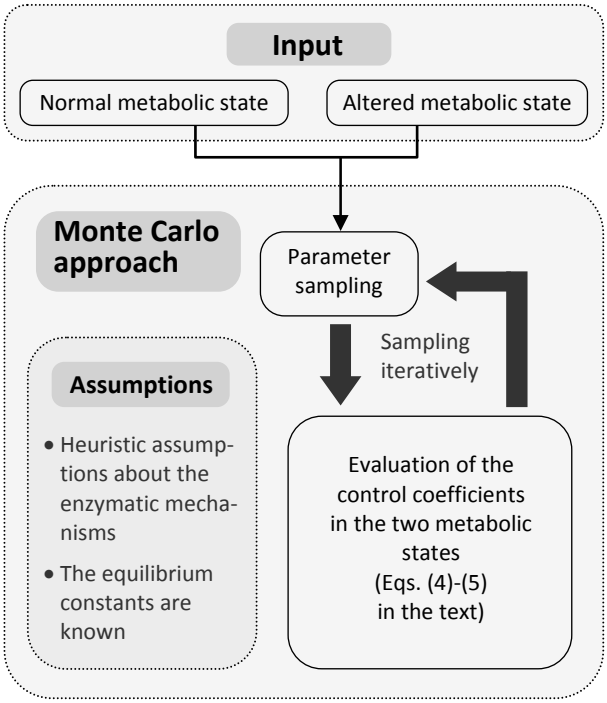

Figure 1 – Workflow of our Monte Carlo approach. *The Monte Carlo approach described in the text receives as input the two metabolic states under comparison (each defined in terms of fluxes and metabolite concentrations at stationary condition). The computation of the control coefficients, following the sampling of the parameter values, is done assuming that the rate equations and the equilibrium constants are known. Where the detailed enzyme mechanism is unknown, a heuristic approximate rate equation can be used.*

DEFINING GENERIC REACTION CHARACTERISTICS

To evaluate the differential response, each rate equation has to be assigned a specific mathematical form. To describe the kinetics of enzyme catalyzed reactions for which the true kinetics is unknown, several possible heuristic approximate rate equations have been proposed in the literature [35-37]. In general, good results are obtained for functional forms that follow generalized Michaelis-Menten equations. Given a reaction converting a set of substrates A into a set of products B

$$\alpha_1 A_1 + \alpha_2 A_2 + ... \rightarrow \beta_1 B_1 + \beta_2 B_2 + ...$$

such rate equation can be written as follows

$$v = f_{reg} \cdot \frac{V_{\max} \prod \tilde{a}_i^{\alpha_i}}{F(\tilde{a}, \tilde{b})} \cdot \left(1 - \frac{\Gamma}{K_{eq}} \right) \tag{9}$$

where

- f_{reg} is a dimensionless prefactor describing the interactions of the enzyme with allosteric regulators. Following the definition given in [37], we write

$$f_{reg} = \left(\prod \frac{K_{P_k}^I}{K_{P_k}^I + P_k} \right) \cdot \left(\prod \frac{Q_l}{K_{Q_l}^A + Q_l} \right) \tag{10}$$

  with P_k and Q_l denoting respectively a generic (allosteric) inhibitor and activator, and $K_{P_k}^I$ and $K_{Q_l}^A$ denoting their corresponding binding constant. This equation assumes that the activators and inhibitors are non competitive.

- $V_{\max}$ is the forward maximal rate of the enzyme and implicitly depends on the enzyme concentration.

- $\tilde{a}_i := \frac{A_i}{K_{A_i}^M}$ is the concentration of reactant A_i divided by its Michaelis constant $K_{A_i}^M$

- $\tilde{b}_j := \frac{B_j}{K_{B_j}^M}$ is the concentration of reactant B_j divided by its Michaelis constant $K_{B_j}^M$

- $F(\tilde{a},\tilde{b})$ is a positive polynomial of the scaled reactant concentrations. Following [37], we choose

$$F\left(\tilde{a},\tilde{b}\right)=\prod_i\left(1+\tilde{a}_i+...+\tilde{a}_i^{\alpha_i}\right)+\prod_j\left(1+\tilde{b}_j+...+\tilde{b}_j^{\beta_j}\right)-1 \qquad (11)$$

- Γ is the mass-action ratio, defined as

$$\Gamma=\frac{\Pi_l B_l^{\beta_l}}{\Pi_k A_k^{\alpha_k}},$$

- K_{eq} denotes the equilibrium constant of the reaction.

The functional form described above captures the generic characteristics of enzyme catalyzed reactions, such as reversibility, product inhibition, saturation and a reaction direction that only depends upon the mass-action ratio relative to the equilibrium constant. In the case study presented in this paper, we used equations in the form proposed in [37] for every reaction.

DEFINING THE KINETIC PARAMETERS

Once the functional form of each rate equation is specified, the equations have to be populated with their respective kinetic constants. Again, we may assume that a (usually small) number of kinetic parameters is available, either by direct experimentation or from data repositories [38, 39]. To account for the remaining unknown or uncertain values, we implement Monte Carlo approach to systematically evaluate the behaviour of the network in a probabilistic manner. The idea is to sample the parameters so that their values comply with known phenotypic data and satisfy basic biophysical and thermodynamic constraints.

Our starting point consists of a metabolic state as characterized by a flux distribution $\mathbf{v}^0$ and a set of metabolite concentrations $\mathbf{S}^0$. The flux distribution $\mathbf{v}^0$ satisfies the stationarity constraint $d\mathbf{S}^0/dt \equiv \mathbf{N} \cdot \mathbf{v}^0 = \mathbf{0}$, and the direction of each flux must be consistent with the set of concentrations, relative to its equilibrium constant. Reaction parameters are then assigned or sampled according to the following criteria:

- **Equilibrium constants**. The equilibrium constants are physicochemical quantities which reflect the change in standard Gibbs free energy occurring in a reaction. They do not depend on the specific organism or cell type, but may depend on intracellular parameters, such as temperature. While as yet a large-scale detailed experimental quantification of the change in free energy is not available, a number of algorithms exist that allow for a reasonable computational approximation [40-44]. In the case study presented here, the values assigned to the equilibrium constants are obtained from [45].

- **Michaelis-Menten constants**. The Michaelis-Menten constants are genuine enzyme kinetic parameters and are sampled randomly in intervals $[K_M^{\min}, K_M^{\max}]$. Different options are available to specify the interval borders. Here, the intervals boundaries are chosen according to the observed metabolite concentrations. We sample the values of the parameters from intervals covering at least two orders of magnitudes around the steady-state concentrations of the corresponding metabolite. In particular, if $[S]_1$ and $[S]_2$ denote the concentration of a metabolite *S* in the two metabolic states, the affinity, inhibition and activation constants of any enzyme in respect to *S* are sampled between $\min\{[S]_1, [S]_2\} \times 10^{-\alpha}$ and $\max\{[S]_1, [S]_2\} \times 10^{\beta}$, with α and β representing adjustable factors. To ensure that the sampled values are evenly spread among the different

order of magnitudes around $[S]_1$ and $[S]_2$, the sampling is performed using a logarithmic distribution. In the sampling process no distinction is made between parameters relating to regulatory interaction and substrate and product affinities. The results shown in the next section refer to sampling conditions where $\alpha = \beta = 1$. The robustness of our results was subsequently tested for different choices of α and β, and different sampling distributions.

- **Maximum rates**. Given the metabolic state and the parameters defined above, it follows from Eq. (9) that the maximum reaction velocity $V_{\max}$ is unambiguously determined. In other terms, to make the sampled values of Michaelis-Menten constants compliant with the metabolic state $\left(\mathbf{v}^0, \mathbf{S}^0\right)$ and the thermodynamics of the system (determined by the equilibrium constants K_{eq}), the maximum reaction velocity is computed by reversing Eq. (9) with respect to $V_{\max}$.

EVALUATING THE CONTROL COEFFICIENTS.

The sampling is performed iteratively. For each sampled set of parameters, the partial derivatives $\partial \mathbf{v} / \partial \mathbf{S}$ are evaluated based on the generalized Michaelis-Menten rate equation described above. Both metabolic states were tested for stability by evaluating their corresponding Jacobian. Parameter sets resulting in Jacobians with positive real part among their spectrum of eigenvalues were discarded, otherwise the control coefficients were evaluated using Eqs (4)-(5). We note that the computation only employs basic matrix inversion and multiplication – a procedure which is orders of magnitude faster than numerical integration, making our approach applicable to systems of large size. The process is repeated iteratively until $2{,}5 \times 10^4$ samples is obtained.

ASSESSING THE SUITABILITY OF DRUG TARGET CANDIDATES

The differences in the control profiles between diseased and normal phenotype are evaluated according to the criteria of maximal selectivity and minimal toxicity specified above. Because of the probabilistic nature of our approach, these criteria have to be reformulated such that they describe the *average*, over all the saxmpling iterations, of the two quantities defined in Eqs (7)–(8). As an additional criterion, we evaluate the reliability of the estimated average selectivity. As a quantitative measure of the quality of our prediction, we define the *reliability coefficient* as the ratio of the average selectivity to the standard deviation of the sampled selectivities:

$$R_i^{J_{\text{target}}} \equiv \frac{\overline{S_i^{J_{\text{target}}}}}{\sigma\left(S_i^{J_{\text{target}}}\right)}. \tag{12}$$

DEFINING THE DISEASED PHENOTYPE

To show the applicability of our method on a system of reasonable complexity, we consider a case study to identify suitable sites for drug intervention specifically targeting a cancerous phenotype. As to the best of our knowledge no detailed characterization of a cancer metabolic phenotype (in terms of fluxes and metabolite concentrations) exists, we employ a modelling approach to obtain a set of fluxes and concentrations representative of a generic cancer phenotype.

Our starting point consists of a fully defined metabolic map of human erythrocyte metabolism [45], modified for our purpose (Fig. 2). Two different

sets of maximal activities ($V_{\max}$) are used to reproduce the flux patterns characteristic of a normal cell and a paradigmatic cancer cell.

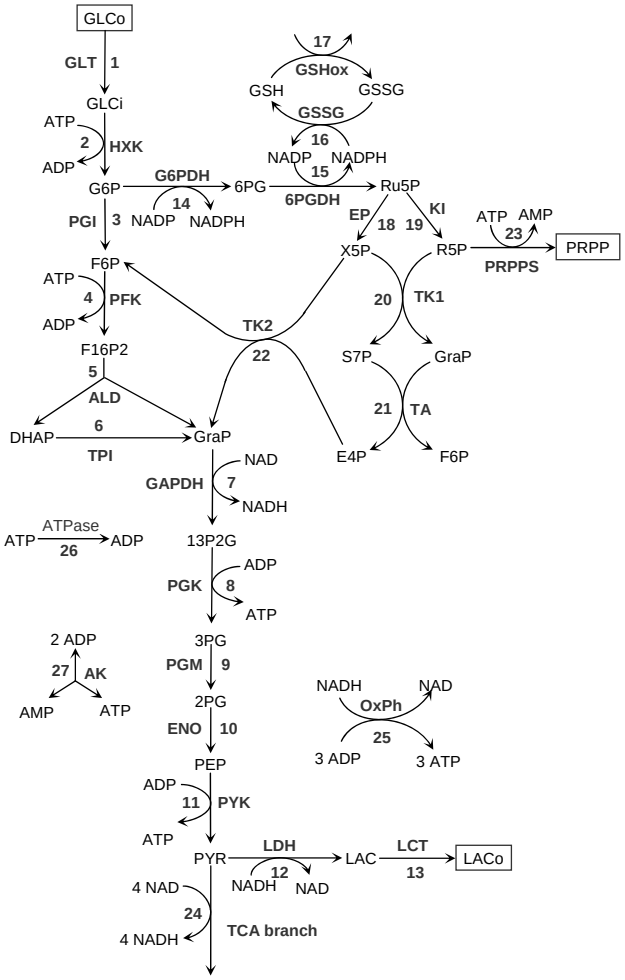

Figure 2 – Metabolic map of central carbon metabolism. *The metabolic map was derived, in its main features, from Holzhütter's model of erythrocyte metabolism and subsequently enriched with a reaction representing the TCA cycle and a reaction representing the oxidative phosphorylation process.*

In particular, the cancerous metabolic phenotype is obtained by increasing the $V_{\max}$ of those enzymes which are often overexpressed in cancer, namely the glucose transporter [46, 47], hexokinase (HXK) [48-51] and phosphofructo-kinase (PFK) [48, 51, 52]. The activity of a fourth enzyme, the glutathione

oxidase (GSHox), is also increased in order to achieve a higher flux through the pentose phosphate pathway, as observed by Richardson *et al*. [53].

We emphasize that the model itself is not used in further analysis and its purpose is solely to obtain a kinetically and thermodynamically consistent set of concentrations and fluxes that represents the common features shared amongst a wide panel of different cancerous cell lines, in particular a higher uptake of glucose and an enhanced production of lactate [54-56]. Details of model construction and analysis are given in the *Supplementary Material*.

RESULTS

IDENTIFYING THE CONTROL PROFILE OF THE SYSTEM BASED ON THE TOPOLOGY AND THE METABOLIC STATE.

The conjecture underlying our probabilistic approach is that some of the control properties of a metabolic system are independent from the precise magnitude of the enzyme parameters. The rationale behind this conjecture is that the signs and magnitudes of the control coefficients are determined to an appreciable extent by the topology of the system as well as the metabolic state under consideration [20, 57]. To investigate our conjecture, we distinguish between two extreme scenarios: if the control coefficients are entirely determined by the metabolic map and phenotype, then their value should be independent from the specific choice of the kinetic parameters; on the other hand, if stoichiometry and phenotype had no bearing on control properties at all, then the control coefficients calculated for randomized values of the kinetic parameters should have infinitely wide and flat distributions.

Fig. 3 shows the distributions obtained for the control exerted by selected enzymes upon the glucose import flux in the normal phenotype. In some of the

cases, the control coefficients are distributed fairly narrowly with respect to the entire likely range of value. For example, the control coefficient of the glucose transporter (GLT, Fig. 3a) is distributed almost entirely between 0.01 and 0.05.

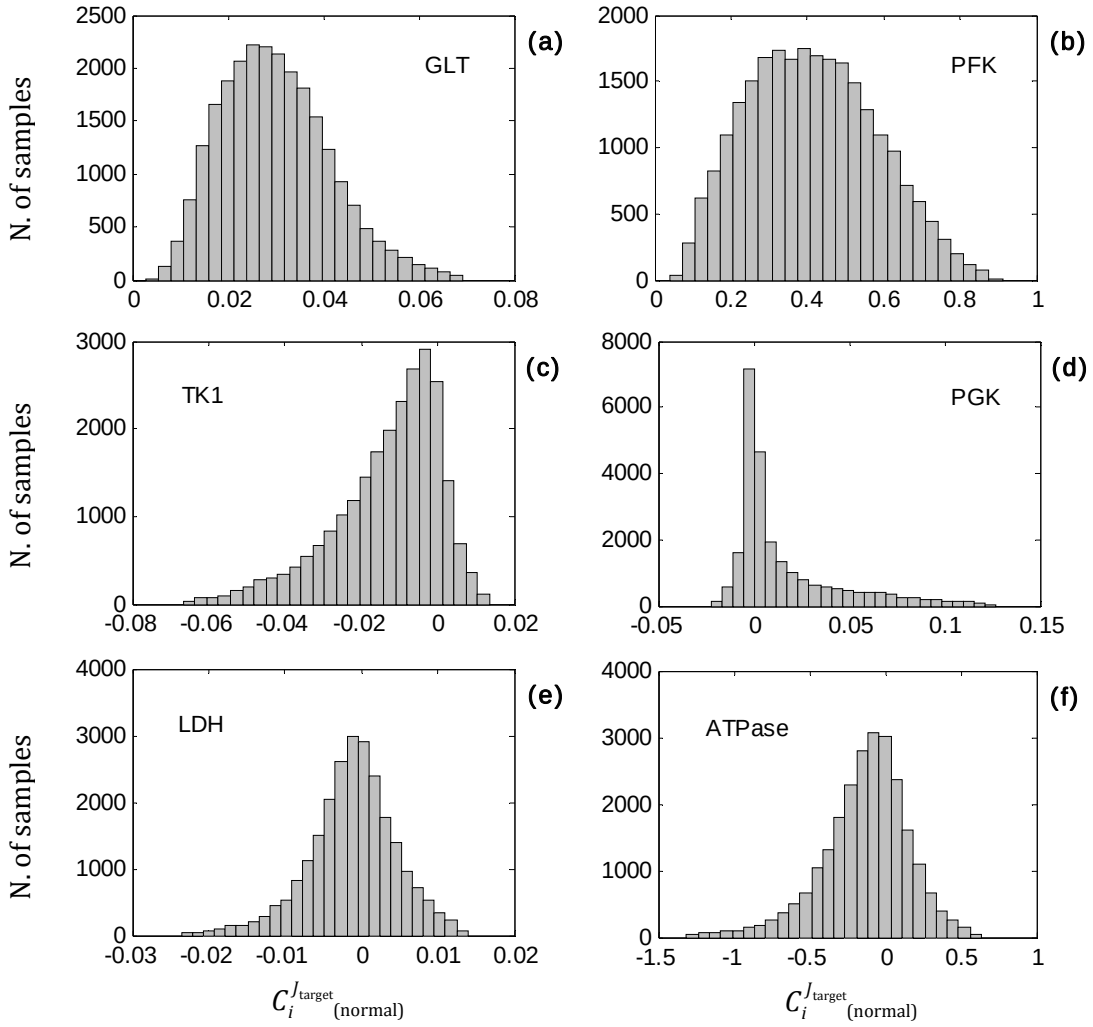

Figure 3 – Calculated distributions of the control exerted by some enzymes over the glucose uptake flux in the normal phenotype. *(a) Glucose transporter (GLT) (b) Phosphofructokinase (PFK), (c) Transketolase 1 (TK1),* ***(d)*** *Phosphoglyceratekinase (PGK), (e) Lactate dehydrogenase (LDH), (f) ATPase.*

This reinforces an earlier conclusion from a precise calculation that this control coefficient is small and that inhibitors of the glucose transporter are unlikely to be toxic to human erythrocytes [58]. More importantly, it confirms the strength of our conjecture for this case; *i.e.* within the limits of experimental accuracy

[59], the value of this flux control coefficient can be estimated on the basis of the metabolic map, the metabolic phenotype and very limited kinetic information. The other panels of Fig. 3 show that this phenomenon is not unique. It is also valid for the control exerted on this same flux by other enzymes in the system, also for ones with higher flux control. Fig. 3 also shows that the accuracy by which metabolic map and metabolic phenotype determine the flux control coefficient is not always the same. For the phosphofructokinase, for example, the estimated flux control coefficients on glucose import exhibit a broad distribution covering almost the entire interval between zero and one (Fig 3b). The other panels in Fig. 3 are chosen as representative cases where the distribution of the control coefficients over the uptake of glucose is either mainly confined in the negative semiaxes (as opposed to glucose transporter and phosphofructokinase) or is spread evenly around zero, allowing no best guess on the sign of the actual control coefficient.

Fig. 4 summarizes the distributions of the estimated control coefficients for the entire network. Each position in the matrix corresponds to the control exerted by an enzyme (columns) upon a flux within the network (rows) and shows the (colour-coded) percentage of calculated control coefficients that is positive. If the box is white then the flux control coefficients is positive independent, to a good extent, of the values of the kinetic constants. For example, the figure shows that the control exerted by the first 6 enzymes on all the first 13 reactions is positive, as indeed might be expected from the network topology shown in Fig. 2.

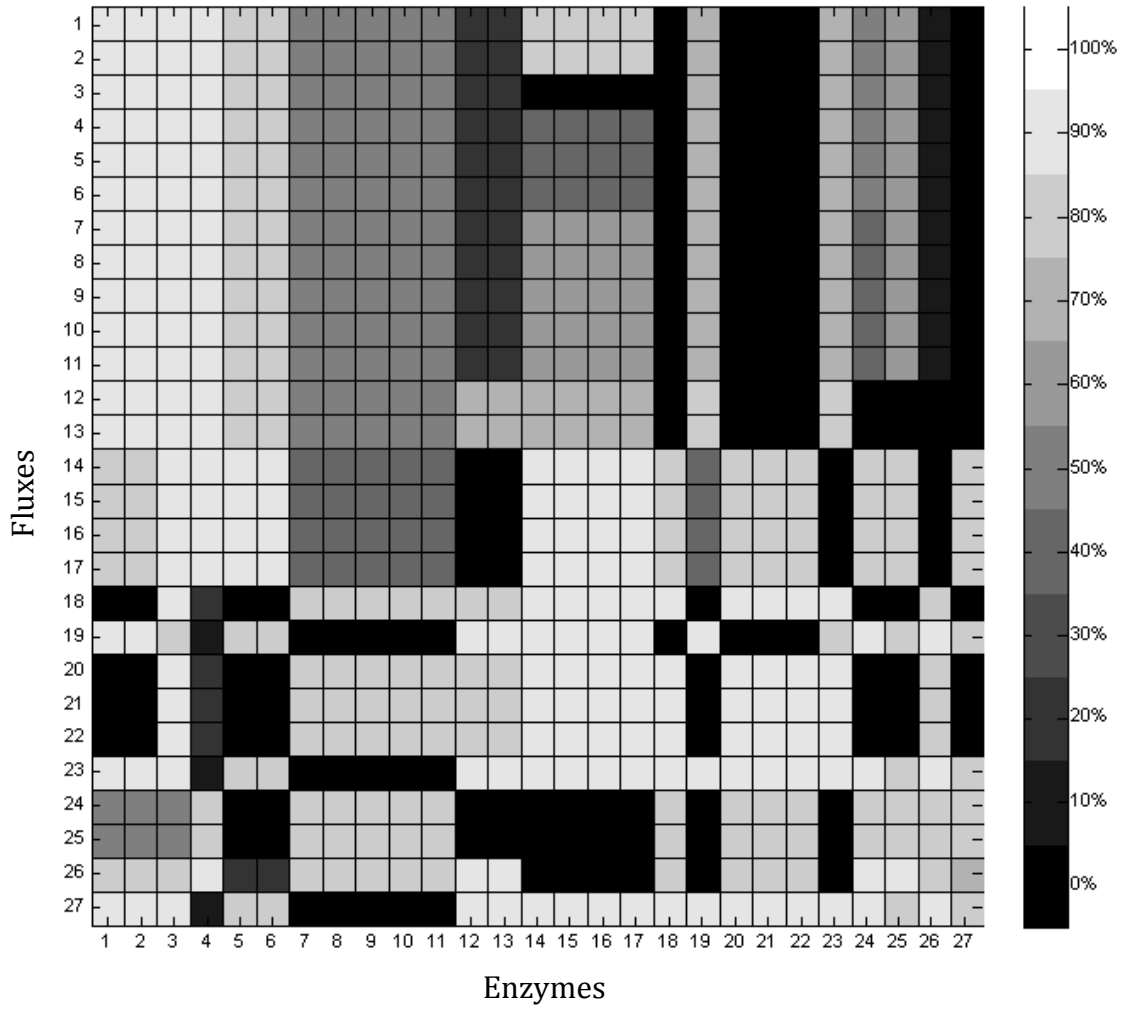

Figure 4 – grey-scale representation of the matrix of the flux control coefficient in the normal metabolic state. *The matrix of the flux control coefficients in the normal metabolic phenotype is represented as a grey-scale matrix. The entry (j, i) is associated to the statistical control exerted by enzyme i (column index) upon flux j (row index). More in particular, the shade of the entry represents the percentage of calculated control coefficients that is positive. The ends of the colours scale represent the extreme situations in which the distribution lies entirely over positive (white) or negative (black) values. The numbers at the left and the bottom of the matrix refers to the different reaction steps in the system as depicted in Fig.2.*

Likewise, the enzymes 14-18 (corresponding to G6PDH, 6PGD, GSSG, GSHox and EP) mostly exert a negative control on the flux through reaction 3 (PGI), again consistently with our expectations from the network topology. However, other results are less straightforward to interpret, showing the utility of our

calculations for complex reaction network. Overall, Fig. 4 confirms our conjecture that often at least the sign of the control coefficient is, to a good extend, independent from the precise values of kinetic parameters -- enabling us to obtain a best guess for putative drug targets in the face of incomplete information.

IDENTIFYING CANDIDATE TARGETS FOR DRUG INTERVENTION

From the perspective of limiting the survival or proliferation of cells that function as parasites in the human body, a possible strategy consists of inhibiting glycolysis, provided the pathway is phenotypically different in the parasitic versus the host cells [3, 5, 60, 61]. From a MCA perspective, this translates into inhibiting the activity of an enzyme which exerts a major control over (for example) the uptake of glucose in the diseased/parasitic phenotype and a minor control over the same flux in the normal phenotype. The rationale of such an approach is to starve the diseased/parasitic cells without significantly affecting the host. From here on we will refer to this specific strategy to present and comment on our results.

When identifying potential drug molecular targets, several criteria m**a**y be important and taken simultaneously into account. One criterion considered here is the effect that a certain fractional inactivation of the molecular target has on the target flux J_{target}. We shall call this the *criterion of maximal selectivity*. In particular we used a combination of effectiveness and selectivity by introducing the *selectivity coefficient* $S_i^{J_{\text{target}}}$ defined in Eq (7). As evident from the previous section, we can calculate the selectivity coefficient for each specific diseased cell compared to each specific normal cell with the same kinetic parameters but $V_{\max}$'s adjusted to reflect the differences in their phenotype. This enables us to calculate a distribution of the selectivities between all diseased and normal cells that we consider statistically feasible in the sense of the parameter sampling we do. Fig. 5 shows the

distributions we found for these selectivity coefficients for the same enzymes of Fig. 3.

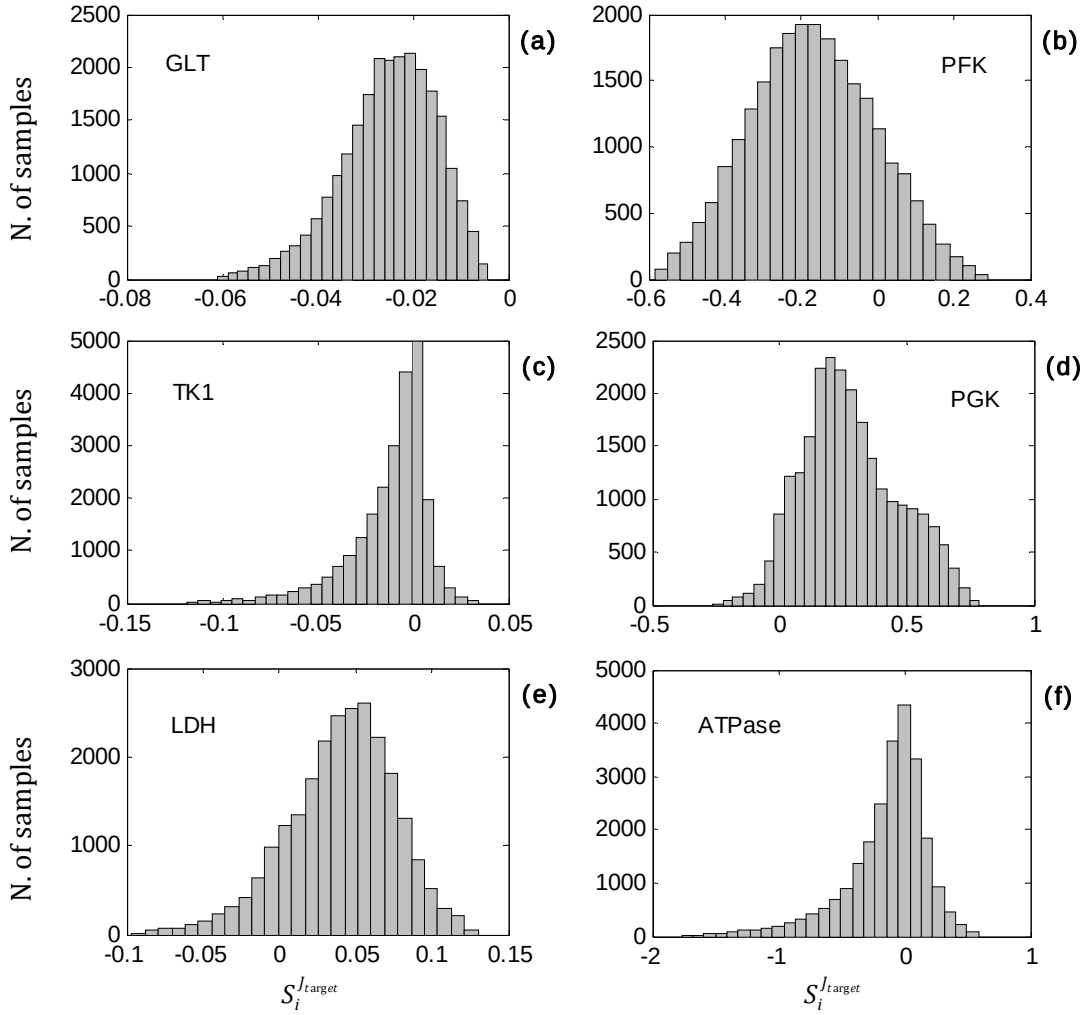

Figure 5 – Calculated distributions of the selectivity coefficient. *The distribution of the selectivity coefficients with respect to the uptake of glucose are shown for the same enzymes as in Fig.3.*

One may expect at least some flux control coefficients to differ between diseased and healthy cells. However, because the diseased phenotype was taken to be due to overexpression of four enzymes (see *Methods*), those differences would be such that the flux control by those enzymes in the diseased cells would be smaller than in the normal cells. As shown by Fig. 5a-b for the glucose

transporter and phosphofructokinase, this is indeed what we observed for these four enzymes, which, consequently, cannot be considered good candidate drug targets (cf. [62]). The issue therefore became whether the overexpression of these four enzymes would increase the control coefficients of *other* enzymes.

Fig. 5.d-e shows that for PGK and LDH this was indeed the case; for these enzymes there was the desired selectivity between diseased and normal cells in terms of their control on the target flux. We will further use a high magnitude of the average selectivity $\overline{S_i^{J_{\text{target}}}}$ (averaged over all the sampling iterations) as our first criterion in deciding about the best molecular drug targets. Importantly, for PGK and LDH this average selectivity was stronger than the uncertainty of its magnitude. This brings us to the *criterion of reliability* of the calculated average selectivity. For the statements about selectivity to be meaningful we should like them to be relatively insensitive to the uncertainty in the values of the kinetic parameters. As a quantitative measure of this criterion we used the *reliability coefficient* defined as in Eq. (12).

A further criterion we wish to consider is the toxicity of the drug. As a measure of toxicity, here we used the average relative effect of the inhibition of the target enzyme on all cellular functional fluxes. In quantitative terms, this was done through Eq. (8), where the toxicity coefficient T_i is defined as the average magnitude of the flux control coefficients with respect to the target enzyme, where the average is computed over all the exchange fluxes. In this paper we consider the average $\overline{T_i}$ of this toxicity coefficient, computed over all the sampling iterations.

We now have three criteria by which to compare potential molecular drug targets. In any actual situation we will need to look at all three criteria

simultaneously. To make the weighing of the three criteria as transparent as possible, we introduce the following scaled quantities:

$$\sigma_i \equiv \frac{\overline{S_i^{J_{\text{target}}}}}{\max\limits_{\text{all } i}\left\{S_i^{J_{\text{target}}}\right\}} \tag{13}$$

$$\rho_i \equiv \frac{\overline{R_i^{J_{\text{target}}}}}{\max\limits_{\text{all } i}\left\{R_i^{J_{\text{target}}}\right\}} \tag{14}$$

$$\frac{1}{\tau_i} \equiv \frac{\max\left\{\overline{T_i}\right\} - \overline{T_i}}{\max\left\{\overline{T_i}\right\} - \min\left\{\overline{T_i}\right\}}, \tag{15}$$

where σ_i, ρ_i and $1/\tau_i$ are the normalized selectivity, the normalized reliability and the normalized safety, respectively. We note that the latter increases with decreasing toxicity. By restricting our search for putative targets to only those enzymes with positive average selectivity $\overline{S_i^{J_{\text{target}}}}$, *i.e.* enzymes which tend to produce the wanted inhibiting effect on J_{target} in the diseased phenotype, the three normalized criteria are bound between 0 and 1.

Fig. 6 depicts the values of the normalized selectivity, reliability and safety for all the enzymes with positive average selectivity $\overline{S_i^{J_{\text{target}}}}$. Because it is hard to interpret the 3-D figure, Fig.7 gives the projections onto the three planes: safety vs. reliability, selectivity vs. reliability and selectivity vs. safety. A most interesting result was that one enzyme target was both most selective and most reliable, and was also hardly toxic. This target was phosphoglyceratekinase (PGK). The fact that all three criteria come out with a single enzyme target suggests that that target could indeed be exceptionally valuable. It also suggests that the methodology we have introduced may be useful; a lesser result would have been if the target that was most selective would have been least reliable

and most toxic, or if the least toxic target would also have been least selective. However, in general, we expect that no single target is best according to all three criteria introduced above. The more important result therefore is that our methodology leads to a clear separation of the potential targets at least in the dimensions of selectivity and reliability. In this respect the criterion of toxicity appeared to be less discriminatory for the enzymes shown.

We may wish to classify drug targets with a single score, which takes into account all three criteria (selectivity, reliability and lack of toxicity), perhaps with different weights. To do so, we define the following quantity

$$Z_i = \frac{w_\sigma \cdot \sigma_i + w_\rho \cdot \rho_i + \dfrac{w_\tau}{\tau_i}}{w_\sigma + w_\rho + w_\tau} \qquad (16)$$

as the score to be assigned to enzyme *i*. Eq. (16) represent a plane in the three-dimensional space defined by σ_i, ρ_i and $1/\tau_i$.

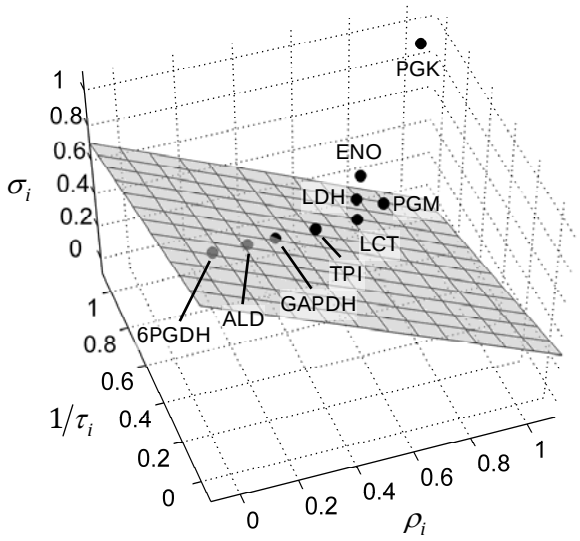

Figure 6 *– The normalized selectivity (σ_i), safety ($1/\tau_i$) and reliability (ρ_i), defined by Eqs. (13), (14) and (15) respectively, are plotted versus each other. Enzymes with negative values of* $\overline{S_i^{J_{\text{target}}}}$ *(see text) are not represented in the graph. The grey plane represents the*

set of points in the three-dimensional space sharing the same score, as defined by Eq. (16). Different scores are represented by different planes, parallel to each other.

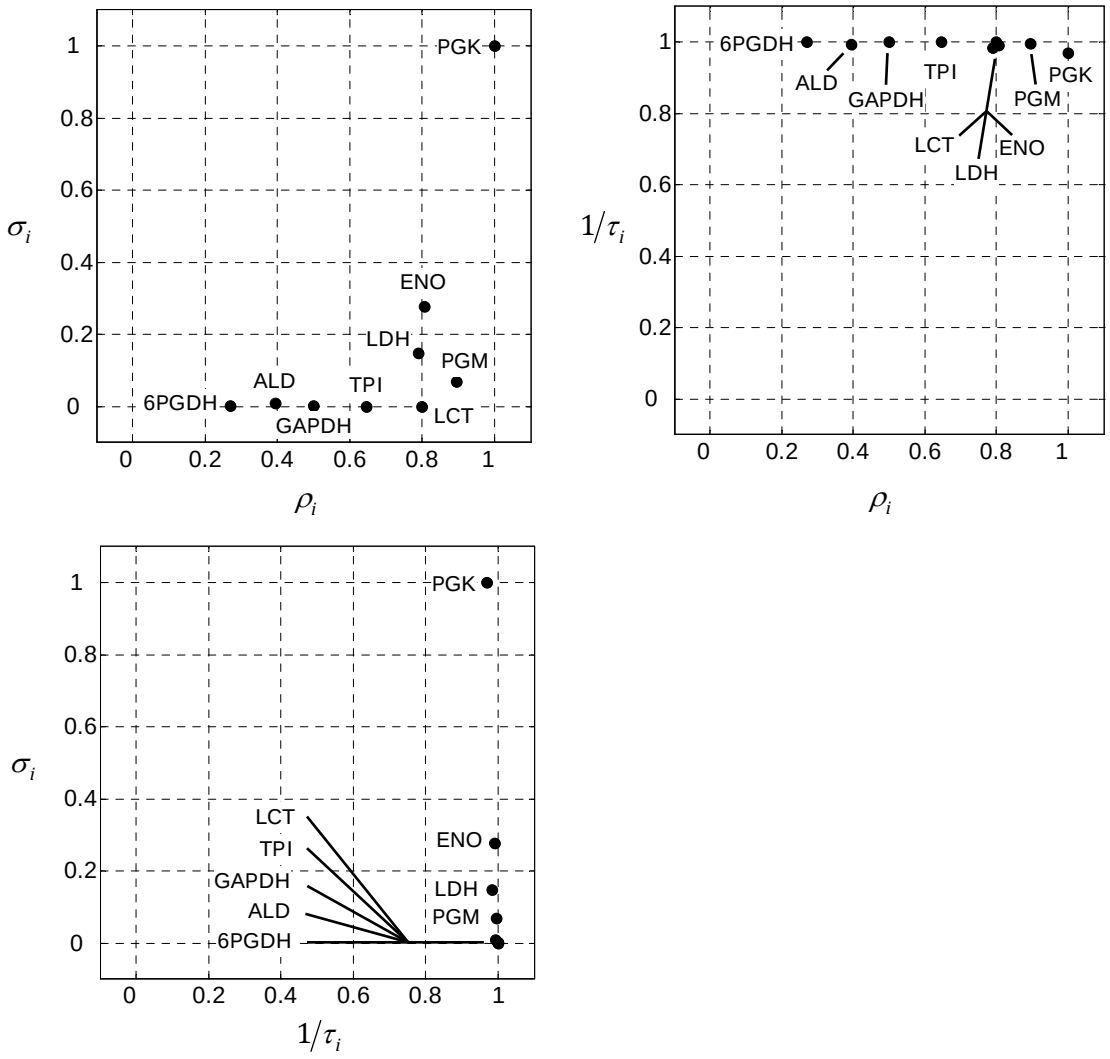

Figure 7 *– Orthogonal projections of the three-dimensional scatter-plot shown in Fig.6.*

For a minimally required score Z^*, one may draw the corresponding plane in Fig. 6 and require all enzyme targets one wishes to consider for further development to lie above that plane. In Fig. 6, we have drawn the plane corresponding to $Z^* = 1/3$ and weight factors $w_\sigma = 4$, $w_\tau = 2$ and $w_\rho = 1$. This specific choice for the weight factors value was made to prioritize the maximal selectivity over the minimal toxicity, and the latter over the requirement of

minimal uncertainty. Table 1 shows the list of enzyme with positive average selectivity, sorted by decreasing value of *Z*.

Enzyme	Score
PGK	0.99
ENO	0.56
LDH	0.48
PGM	0.45
LCT	0.40
TPI	0.38
GAPDH	0.36
ALD	0.35
6PGDH	0.32

Table 1 – Suitability of the different enzymes as drug targets. *The suitability of each enzyme as drug target is evaluated according to the three criteria of selection described in the text. The score is computed through Eq. (16) with* $w_\sigma = 4$, $w_\tau = 2$ *and* $w_\rho = 1$.

ROBUSTNESS OF THE RESULTS

As a probabilistic method, our approach crucially relies on the robustness of the results with respect to different ways of sampling the parameter values. To this end, we repeated the data generation process and the subsequent analysis using different sampling conditions. In particular we changed the range of values from which the parameters were sampled and the sampling distribution. For each different set of sampling conditions, we ranked the suitability of the different enzyme as drug targets as explained above.

Table 2 summarizes the results obtained for the different sampling conditions, showing PGK as the highest scoring target for all of them. As expected, enzymes with lower scores than PGK did not always keep their position in the ranking list. The lower the score the higher the chance for the enzyme to enter the list in a different position, when the sampling condition was altered. In most of the

cases, however, the top 6 entries of the list remained the same. In particular PGK always emerged as the best guess as putative target, while Enolase (ENO) and Phosphoglycerate mutase (PGM) were always found within the top 4 entries of the list, and Lactate dehydrogenase (LDH) was almost always found between the 3rd and the 6th position (see *Supplementary Materials* for more details).

α	β	Sampling function	Best guess
1	1	Logarithmic	PGK
1	2	Logarithmic	PGK
2	1	Logarithmic	PGK
2	2	Logarithmic	PGK
1	1	Linear	PGK
1*	1*	Logarithmic	PGK

Table 2 – Best drug target guesses for different sampling conditions. *The first two columns refer to the parameters α and β used to define the intervals from which the parameter values were sampled. The first 4 rows refer to sampling conditions where the intervals were defined separately for each parameter, as discussed in Methods. The last row (*) refers to a sampling performed on the same interval for all the parameters. The borders of this interval were defined as* $\min\{[\mathbf{S}]_1,[\mathbf{S}]_2\}\times 10^{-\alpha}$ *and* $\max\{[\mathbf{S}]_1,[\mathbf{S}]_2\}\times 10^{\beta}$*, where* $\{[\mathbf{S}]_1,[\mathbf{S}]_2\}$ *denotes the set of concentrations of all the metabolites in the two metabolic states.*

The use of a linear sampling distribution caused the most different results in respect to the scoring list obtained with a logarithmic distribution, even for the high-scoring enzymes (except for PGK, which was always at the top of the list). This fact can be ascribed to the highly asymmetric sampling of the parameter values in respect to the metabolite concentrations. When using a linear distribution, there is a much higher probability that the sampled value of the Michaelis-Menten constants exceed the concentration of the corresponding metabolites. In our case, where $\alpha = \beta = 1$ (see *Methods* for their definition), for

each reaction step statistically at least 90% of the sampled values were larger than the concentration of their corresponding metabolites. This implies a scenario where the saturation level of all the enzymes in respect to all their reactants and modifiers was almost always greater that 50%. On the other hand, when using a logarithmic distribution, the value of the Michaelis-Menten constants tended to be evenly sampled around the metabolite concentrations, or, more precisely, among the orders of magnitude spanned by the sampling intervals.

COMPARISON WITH THE DYNAMIC MODEL

To assess the significance of our result, we compared the insights gained through our statistical approach with the results obtainable directly through a dynamic simulation. The dynamic model used to retrieve the cancer metabolic state (see *Methods* and *Supplementary Material*) was used as workbench for this assessment.

Metabolic branch/process	Normal state	Cancer state
Glucose uptake	-7%	-44%
Lactate production	-61%	-82%
TCA cycle	-1%	-8%
Phoshoribosylpyrophosphate synthetase (PRPPS)	+9%	+13.0%
ATPase	-2%	-12%
Oxidative phosphorilation	-1%	-8%

Table 3 – Relative changes in the principal fluxes of the system after decreasing the activity of PGK. *The fluxes considered are the same as in Figure 8. The relative changes reported are the result of decreasing the activity of PGK from 12.9 to 6.5 mM/s.*

The two sets of parameters characterising the two metabolic states (normal and cancer) were both modified by decreasing the activity ($V_{\max}$) of PGK from 12.9 to 6.5 mM/h, simulating the addition of a noncompetitve inhibitor. Fig. 8 shows how some of the fluxes of the metabolic system changed in the two phenotypes in response to this perturbation. Table 3 shows the relative changes recorded once the system reached the new steady state after the perturbation.

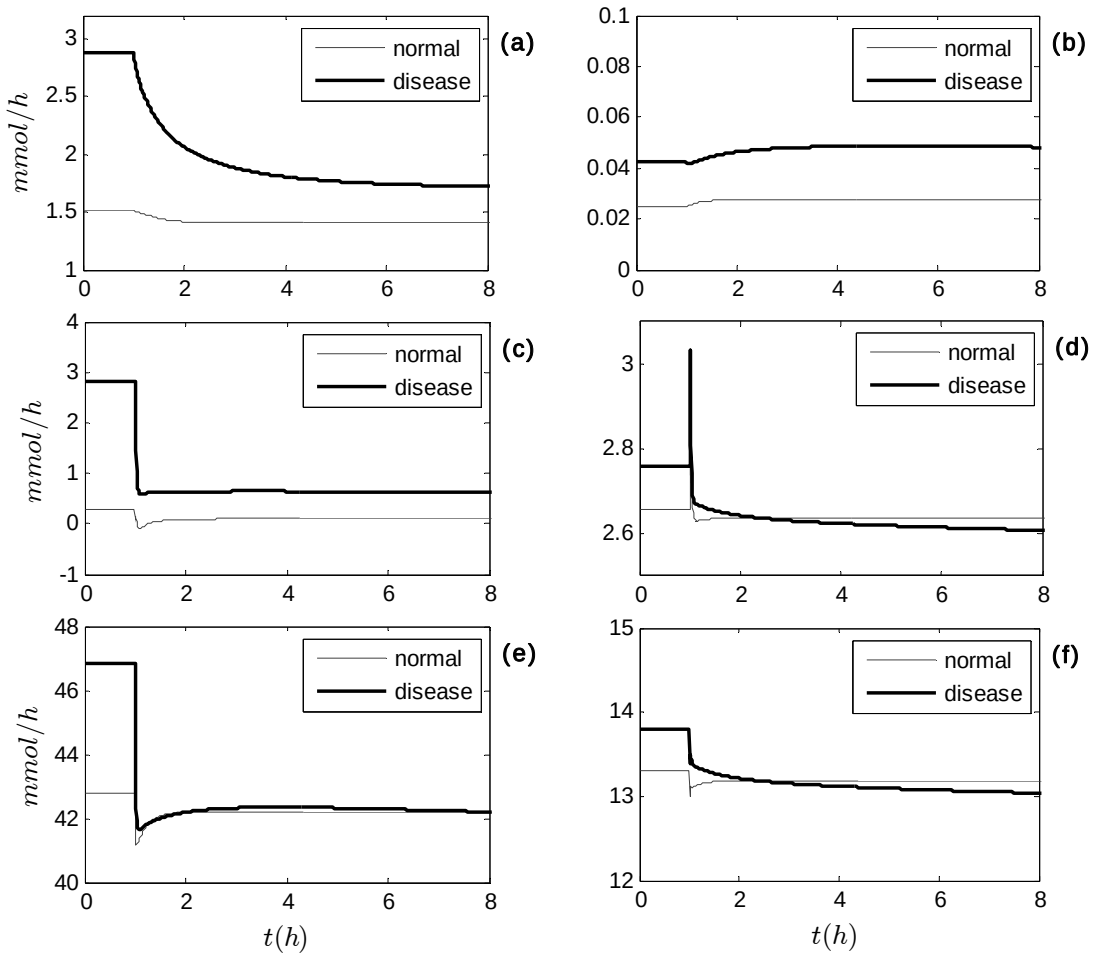

Figure 8 – Effects induced on the principal fluxes of the system by decreasing the activity of PGK. *The plots show the effect of decreasing the activity of PGK from 12.9 to 6.5 mM/h in the two phenotypes (normal and disease). Six main fluxes are considered: (a) glucose uptake, (b) PRPPS, (c) lactate production, (d) flux entering the TCA cycle, (e) ATPase and (f) flux through the oxidative phosphorylation process.*

DISCUSSION

Despite a surge in the perspectives of using mechanistic mathematical modelling in clinical development [63-66], parametric uncertainty remains a challenge to mechanistic approaches in medicine [13]. However, contrary to the scarcity of kinetic data, the comprehensive quantification of all concentrations and fluxes within a metabolic system is, at least in principle, experimentally feasible. The question we addressed in this paper is whether the knowledge of a metabolic phenotype only – expressed in terms of fluxes and metabolite concentrations at steady-state – allows for a *probabilistic* understanding of how the control properties of a biochemical system are distributed among the different enzymatic steps and metabolic processes. We developed a Monte Carlo approach which aims to provide researchers with a probabilistic description of how the control properties of a metabolic network differ between two fully characterized metabolic phenotypes, when only minimal knowledge of the system is available with respect to its kinetic parameters. Our goal was to locate points of fragility in a diseased/pathogen phenotype which can be considered for drug interventions with maximal effectiveness and minimal toxicity. In particular, enzymes which exert a major control over a certain property of interest in a diseased/pathogen phenotype, and a minor control over the same property in the normal/host phenotype, represent good putative targets. In our method, the control profiles characterizing the two phenotypes under comparison are determined for sampled values of the unknown affinity, inhibition and activation constants. Our Monte Carlo approach provides us with a statistical understanding of how the control is differentially distributed between the two metabolic states, where the word 'statistical' should be interpreted in the sense of uncertainty rather than in the sense of population dispersion.

The system under study was a simplified reconstruction of the central carbon metabolism, with two metabolic states that were representative of a normal and a paradigmatic cancerous phenotypes. The results presented in this work refer to a clinical strategy aimed to starve specifically cancer cells, *i.e.* to locate enzymes which exhibit a high differential control upon the uptake of glucose, denoted as J_{target}, between the two phenotypes.

In the development of new drugs, pharmaceutical companies tend to give priority to the maximal effect that a compound can induce in the disease cells. Among the palette of compounds which pass this filter, a second screening is performed to look for compounds which leave the normal phenotype unaltered as much as possible. Our approach is different. To assess the suitability of the enzymes as possible targets for an anti-cancer drug, three different criteria were considered in this paper: maximal selectivity, maximal safety (or minimal toxicity) and maximal reliability. The first criterion was formulated in such a way as to encompass both the requirements of high effectiveness and high selectivity with respect to the specific property one wants to affect in the diseased cells. High values of the selectivity coefficient correspond to a high (positive) control over J_{target} in the diseased cell, and a relatively small control over the same flux in the normal cells. In particular, we introduced the selectivity coefficient, as defined in Eq.(7), to quantify the differential response of the system to inhibition of enzyme *i*. Alternative definitions than Eq.(7) may also be considered. For example, Bakker et al. defined it as the ratio between the control over J_{target} in the parasitic/diseased cell over the same control in the host/normal cell [3]. Such a definition, however, may result in over-ranking enzymes with small control in the diseased phenotype as putative targets. An enzyme A with $C_{A\ (\text{normal})}^{J_{\text{target}}} = 0.01$ and $C_{A\ (\text{diseased})}^{J_{\text{target}}} = 0.1$, for instance, would be considered preferable to an enzyme B with $C_{\text{B}\ (\text{normal})}^{J_{\text{target}}} = 0.3$ and $C_{\text{B}\ (\text{diseased})}^{J_{\text{target}}} = 0.9$, as

the ratio $C_{A\ (\text{diseased})}^{J_{\text{target}}} / C_{A\ (\text{normal})}^{J_{\text{target}}}$ would be around three times the ratio $C_{\text{B}\ (\text{diseased})}^{J_{\text{target}}} / C_{\text{B}\ (\text{normal})}^{J_{\text{target}}}$. However, using an appropriate dosage of a drug that inhibits enzyme B, one could reduce the effect of the enzyme inhibition in the normal cell to a negligible extent while maintaining in the diseased cell a much higher effect than would be feasible by inhibition of enzyme A.

The criterion of maximal safety (or minimal toxicity) was introduced to assure that the normal phenotype was robust to a perturbation of the drug target. The predicted toxic effect of inhibiting an enzyme *i* was quantified through the toxicity coefficient defined in Eq. (8). We defined this coefficient as the average magnitude of the flux control coefficient of enzyme *i* with respect to all the exchange fluxes. Our definition reflects a black-box perspective, where the behaviour of the system is assessed solely through its input and output fluxes. However, the definition of toxicity is far from unique and its suitability may depend on the specific situation. For example, an alternative choice would be to define the toxicity as the maximal control coefficient over all the fluxes within the network. Also, there may be reactions or pathways for which a change in the flux does not induce any major effect on the vital functions of the cell, while even small alterations in other fluxes might entail significant stress in the normal cellular physiology. In this case, the criterion of minimal toxicity can be augmented by assigning different weights to the respective control coefficients. Another possibility in defining the toxicity coefficient would be to take into account the concentration of key metabolites that are known to have toxic effects in the normal phenotype. In this case, the toxicity is assessed in terms of the concentration control coefficients of the enzyme targets with respect to these metabolites.

Finally, the third *criterion of maximal reliability* was introduced in order to prefer enzymes whereby the computed average selectivity was affected by the

least possible uncertainty.

To evaluate the suitability of an enzymes as a drug target, different weights can be assigned to these criteria, depending on the relevance they have for the researcher. In the present work we chose to prioritize the maximal selectivity over the minimal toxicity, and the latter over the requirement of minimal uncertainty. By doing so, PGK was identified as the best guess for a suitable drug target. Although in this case PGK was both the most selective and most reliable enzyme target, besides being hardly toxic, we note that in general the outcome of our analysis can sensibly depend on the priority ascribed to the three criteria mentioned above. From a clinical perspective, performing our analysis with different weight factors may then provide a palette of best drug targets referring to different thresholds amongst the required properties of selectivity, safety and reliability.

The robustness of our results was tested against different choices of the sampling conditions. The fact that PGK always emerged as the best guess as putative target represents an important result, as it suggests that the sampling approach proposed in this paper can provide a relevant insight about the control profile of a metabolic system that is reasonably robust with respect to alterations in the numerical methods. If the top entries of the scoring list were to change completely based on the sampling condition, the method would have proven to be of less utility. We note, however, that in general different choices of the sampling distribution may lead to different enzymes as best putative drug target. The use of different sampling distributions may then be thought and adopted as a way to provide a more restrictive assessment of how the emergence of a specific enzyme as best target is due to topological constraints rather than specific parameter values (or range of values).

The conclusions achieved through our probabilistic approach, leading to the identification of PGK as the best putative target, were compared with the outcome of a mechanistic approach by decreasing the activity ($V_{\max}$) of PGK in the dynamic model introduced in *Methods* and described in *Supplementary Material*. The response of the system to this perturbation showed that the normal phenotype was indeed less sensitive to an inhibition of PGK than the diseased phenotype. This lower sensitivity has been observed not only in respect to the glucose uptake flux, but in the general behaviour of the system. The exchange fluxes and the main metabolic processes were in fact less affected in the normal phenotype (Fig. 8), showing good agreement with our statistical result. Regarding specifically the production of lactate, we registered a significant decrease in the flux through LDH in both normal and diseased phenotype (-61% and -82%, respectively). Since this strong decrease occurred on a branch which is virtually unused in normal cells (except for muscle cells under anaerobic conditions and erythrocytes), this result does not invalidate PGK as putative drug target candidate.

In conclusion, the statistical approach proposed in this paper provides us with a useful strategy for assessing how the control profile is differently distributed in two distinct metabolic phenotypes. It also highlights which enzyme can best represent a putative target with respect to requirements such as high effectiveness and low toxicity. The significance of the results obtained through this kind of analysis, however, may be improved at different levels, not necessarily only related to the probabilistic nature of our approach. For example, a higher degree of detail in the representation of the metabolic map would reduce the approximation introduced by considering lumped reactions representing more complex biochemical pathways or processes. In the example provided in this paper, two of such reactions were used to represent the TCA cycle and oxidative phosphorylation. Expanding a lumped reaction into the

entire set of enzymatic steps it represents, could lead to new and important insights about how the control properties are distributed in the system. This is mostly the case when metabolic intermediates that are only involved in a lumped pathway have regulatory effects on enzymatic steps outside that pathway. A more detailed description of the lumped pathway would then result not only in a "higher resolution" of how the control properties are locally distributed amongst the different steps which were lumped together, but also in a different overall control profile. Just as in the case of conventional kinetic modelling, the level of detail in which the metabolic map is represented can determine the level of accuracy of the regulatory map, and consequently have a non negligible effect on the results.

A related aspect is the choice of rate equations of the enzymatic steps. While generic rate equations are commonly used to capture generic aspects of metabolic networks, the actual, experimentally determined, rate equations may result in a slightly modified control properties. In particular, the experimental determination of reaction functions may also identify further unknown regulatory interactions, as well as possible cooperativity between metabolic compounds. Regarding the use of the generic rate equation, we also note that different choices are possible. The specific instance of generalized Michaelis-Menten equation proposed by Liebermeister *et al.* [37], and used in this paper, takes into account generic characteristics of enzyme catalyzed reactions – such as reversibility, product inhibition, saturation, reaction direction that only depends upon the mass-action ratio relative to the equilibrium constant. However, alternative choices have been proposed by Rohwer et al. [67] that also account for competition between substrates and products and take possible cooperativity into account. We note though, in contrast to explicit kinetic models, we are only interested in the derivatives of the rate equations, such that minor differences in the precise functional form often have no major effect. In

most applications, we also expect that at least partial information on some kinetic parameters is available. Such partial information allows to further constraint the sampling intervals and to obtain results that are specific for the system under study. In this sense, our approach can be straightforwardly incorporated into an interative scheme that allows to quantify uncertainty in the control profile, and hence allows to pinpoint further experiments to increase the specificity and reliability of the results.

ACKNOWLEDGMENT

We acknowledge the support of the Doctoral Training Centre ISBML by EPSRC and BBSRC (EP/F500009/1), of the Manchester Centre of Integrative Systems Biology by BBSRC, EPSRC (BB/C0082191) and additional BBSRC support (BBD0190791), of SysMO (BBF003535281, 35361, 35521) and of the EC Network of Excellence BioSim.

SUPPLEMENTARY MATERIAL

THE MECHANISTIC MODEL

A full detailed characterization of a cancer metabolic phenotype in terms of fluxes and metabolites concentrations has not been performed. To obtain a set of fluxes and concentrations representative of a generic cancer phenotype we decided to use a modelling approach. First of all we implemented a kinetic model reproducing the system shown in Fig.2 of the main text. The model was created using Copasi [26]. The dynamic of each reaction step was described in terms of a generalized Michaelis-Menten equation, in the form of the convenience kinetics [37]. The list of allosteric regulation considered in the model is listed in Table S1.

Enzymatic step	Activators	Inhibitors
HXK		G6P
PFK	AMP	ATP
PK	F16P2	
G6PDH	NADP	
TCA cycle	ADP/ATP	

Table S1 – List of the allosteric effectors considered in the kinetic model.

Where possible, affinity, inhibition and activation constants were retrieved from BRENDA database [38], otherwise their value was chosen within the same order of magnitude of the corresponding substrate or modifier concentration. For the equilibrium constants we used the value listed in Holzhütter 2004 [45]. Finally, the maximum activity (V_{max}) of each enzyme was computed in order to fit the set of fluxes and concentrations defining the normal metabolic state.

The cancerous metabolic phenotype was retrieved by increasing by a factor of four the V_{max} of GLT, HXK, PFK and GSHox. The list of fluxes and metabolite concentrations defining the two metabolite states are provided in Tables S2-S3.

Metabolite	Normal state	Cancer state
GLC	4.57E+00	4.81E+00
G6P	3.95E-02	8.46E-02
F6P	1.56E-02	3.28E-02
F16P2	9.68E-03	4.09E-01
DHAP	1.49E-01	1.02E+00
GraP	6.06E-03	4.16E-02
13P2G	4.81E-04	2.60E-03
3PG	6.58E-02	1.43E-01
2PG	8.44E-03	1.82E-02
PEP	1.09E-02	2.45E-02
PYR	8.40E-02	1.01E-01
LAC	1.68E+00	1.68E+00
ATP	1.60E+00	1.73E+00
ADP	3.24E-01	2.78E-01
AMP	7.47E-02	1.79E-01
NAD	6.53E-02	6.53E-02
NADH	1.56E-04	1.97E-04
6PG	2.51E-02	4.37E-02
Ru5P	4.72E-03	1.29E-02
X5P	1.27E-02	3.48E-02
R5P	1.40E-02	3.83E-02
S7P	1.54E-02	2.52E-02
E4P	6.27E-03	3.29E-02
GSH	3.11E+00	3.03E-02
GSSG	1.87E-04	6.22E-04
NADP	2.34E-05	5.47E-05
NADPH	4.89E-03	4.86E-03

Table S2 – Set of metabolite concentrations characterizing the normal and cancer states. *All the concentrations are expressed in mM.*

Reaction	Normal state	Cancer state
GLT	1.51E+00	2.61E+00
HXK	1.51E+00	2.61E+00
PGI	1.42E+00	2.34E+00
PFK	1.47E+00	2.48E+00
ALD	1.47E+00	2.48E+00
TPI	1.47E+00	2.48E+00
GAPDH	2.95E+00	5.04E+00
PGK	2.95E+00	5.04E+00
PGM	2.95E+00	5.04E+00
ENO	2.95E+00	5.04E+00
PK	2.95E+00	5.04E+00
LDH	0.30E+00	2.10E+00
LAC_Tran	0.30E+00	2.10E+00
G6PDH	9.70E-02	2.71E-01
6PGDH	9.70E-02	2.71E-01
GSSGR	1.94E-01	5.43E-01
GSHox	1.94E-01	5.43E-01
EP	4.80E-02	1.46E-01
KI	4.90E-02	1.25E-01
TK1	2.40E-02	7.29E-02
TA	2.40E-02	7.29E-02
PRPPS	2.50E-02	5.25E-02
TK2	2.40E-02	7.29E-02
TCA	2.66E+00	2.94E+00
OxPh	1.33E+01	1.47E+01
ATPase	4.28E+01	4.90E+01
AK	0.00E+00	0.00E+00

Table S3 – Patterns of fluxes characterising the normal and cancer metabolic states. *The fluxes are expressed in mM/s.*

List of the rate equations and parameter values used in the model

The values of the kinetic parameters used in the mechanistic model have been chosen according to the following criteria:

- The equilibrium constants (K_{eq}) have been taken from Holzhütter 2004 [45]. For the equilibrium constant of the TCA pseudo-reaction we used the product of the K_{eq} of the enzymatic steps taking part in the Krebs cycle. The equilibrium constant for the oxidative phosphorylation process was set to an arbitrary high value.

- The affinity, inhibition and activation constants ($K_{S_i}^{M}, K_{S_i}^{I}, K_{S_i}^{A}$) have been retrieved, where possible, from Brenda database [38]. When more than one value was available, factors such as taxonomy as well as temperature and pH at which the parameter was measured have been taken into account. When no value was found for a given parameter, we used as reference point the concentration of the metabolite to which that parameter refers. In this case the value of the parameter was set to 10^m where *m* is the order of magnitude of the corresponding metabolite concentration expressed in mM.

- The value of the maximal activities ($V_{\max}$) have been computed in order to make the set of concentration at steady-state and the parameters values compliant with the set of fluxes representative of the normal metabolic state (see main text).

The chemical reactions reported below are written following the direction of their positive flux in the normal physiological metabolic state. The values of the equilibrium constants refers to this same direction.

1. Glucose uptake (GLT) – $Glc(e) \rightarrow Glc$

$$v_{GLT} = \frac{V_{\max} \cdot \frac{[GLC](e)}{K_{Glc}^{M}} \cdot \left(1 - \frac{[GLC]/[GLC](e)}{K_{eq}}\right)}{1 + \frac{[GLC](e)}{K_{Glc}^{M}} + \frac{[GLC]}{K_{Glc}^{M}}}$$

$V_{\max} = 3.69\text{E+01};\ K_{eq} = 1.00\text{E+00};\ K_{Glc}^{M} = 1.00\text{E+00};$

2. Hexokinase (HXK) – $Glc + ATP \rightarrow G6P + ADP$

$$v_{HXK} = \left(\frac{K_{G6P}^{I}}{[G6P] + K_{G6P}^{I}}\right) \cdot \frac{V_{\max} \cdot \frac{[Glc] \cdot [ATP]}{K_{Glc}^{M} \cdot K_{ATP}^{M}} \cdot \left(1 - \frac{\frac{[G6P] \cdot [ADP]}{[Glc] \cdot [ATP]}}{K_{eq}}\right)}{\left(1 + \frac{[Glc]}{K_{Glc}^{M}}\right) \cdot \left(1 + \frac{[ATP]}{K_{ATP}^{M}}\right) + \left(1 + \frac{[G6P]}{K_{G6P}^{M}}\right) \cdot \left(1 + \frac{[ADP]}{K_{ADP}^{M}}\right)}$$

$V_{\max} = 6.88\text{E+01};\ K_{eq} = 3.90\text{E+03};\ K_{Glc}^{M} = 3.80\text{E+00};\ K_{ATP}^{M} = 1.51\text{E+00};\ K_{G6P}^{M} = 1.00\text{E-02};$
$K_{ADP}^{M} = 1.00\text{E-01};\ K_{G6P}^{I} = 2.9\text{E-02};$

3. Phosphoglucose isomerase (PGI) – $G6P \rightarrow F6P$

$$v_{PGI} = \frac{V_{\max} \cdot \frac{[G6P]}{K_{G6P}^{M}} \cdot \left(1 - \frac{[F6P]/[G6P]}{K_{eq}}\right)}{1 + \frac{[G6P]}{K_{G6P}^{M}} + \frac{[F6P]}{K_{F6P}^{M}}}$$

$V_{\max} = 2.79\text{E+03};\ K_{eq} = 3.92\text{E-01};\ K_{G6P}^{M} = 3.51\text{E10-01};\ K_{F6P}^{M} = 1.86\text{E-02};$

4. Phosphofructo-kinase (PFK) – $F6P + ATP \rightarrow F16P2 + ADP$

$$v_{PFK} = \left(\frac{K_{ATP}^{I}}{[ATP] + K_{ATP}^{I}}\right) \cdot \left(\frac{[AMP]}{[AMP] + K_{AMP}^{I}}\right) \cdot \frac{V_{\max} \cdot \frac{[F6P] \cdot [ATP]}{K_{F6P}^{M} \cdot K_{ATP}^{M}} \cdot \left(1 - \frac{\frac{[F16P2] \cdot [ADP]}{[F6P] \cdot [ATP]}}{K_{eq}}\right)}{\left(1 + \frac{[F6P]}{K_{F6P}^{M}}\right) \cdot \left(1 + \frac{[ATP]}{K_{ATP}^{M}}\right) + \left(1 + \frac{[F16P2]}{K_{F16P2}^{M}}\right) \cdot \left(1 + \frac{[ADP]}{K_{ADP}^{M}}\right) - 1}$$

$V_{\max} = 1.13\text{E+02};\ K_{eq} = 1.00\text{E-05};\ K_{F6P}^{M} = 1.00\text{E-01};\ K_{ATP}^{M} = 1.20\text{E-01};\ K_{F16P2}^{M} = 1.00\text{E-02};$
$K_{ADP}^{M} = 4.90\text{E-01};\ K_{ATP}^{I} = 2.5\text{E-01};\ K_{AMP}^{A} = 1.00\text{E-01};$

5. Aldolase (ALD) – $F16P2 \rightarrow DHAP + GAP$

$$v_{ALD} = \frac{V_{\max} \cdot \frac{[F16P2]}{K^M_{F16P2}} \cdot \left(1 - \frac{[DHAP] \cdot [GAP]}{[F16P2] \over K_{eq}}\right)}{\left(1 + \frac{[F16P2]}{K^M_{F16P2}}\right) + \left(1 + \frac{[DHAP]}{K^M_{DHAP}}\right) \cdot \left(1 + \frac{[GAP]}{K^M_{GAP}}\right) - 1}$$

$V_{\max} = 1.41\text{E+}02;\ K_{eq} = 1.14\text{E-}01;\ K^M_{F16P2} = 9.00\text{E-}03;\ K^M_{DHAP} = 1.00\text{E-}02;\ K^M_{GAP} = 5.70\text{E-}02;$

6. Triosephosphate isomerase (TPI) – $DHAP \rightarrow GAP$

$$v_{TPI} = \frac{V_{\max} \cdot \frac{[DHAP]}{K^M_{DHAP}} \cdot \left(1 - \frac{[GAP]/[DHAP]}{K_{eq}}\right)}{1 + \frac{[DHAP]}{K^M_{DHAP}} + \frac{[GAP]}{K^M_{GAP}}}$$

$V_{\max} = 3.61\text{E+}03;\ K_{eq} = 4.07\text{E-}02;\ K^M_{DHAP} = 5.00\text{E-}01;\ K^M_{GAP} = 2.60\text{E-}01;$

7. Glyceraldehyde 3-Phosphate Dehydrogenase (GAPDH) –

$GraP + NAD \rightarrow 13P2G + NADH$

$$v_{GAPDH} = \frac{V_{\max} \cdot \frac{[GraP] \cdot [NAD]}{K^M_{GraP} \cdot K^M_{NAD}} \cdot \left(1 - \frac{[13P2G] \cdot [NADH]}{[GraP] \cdot [NAD]} \over K_{eq}\right)}{\left(1 + \frac{[GraP]}{K^M_{GraP}}\right) \cdot \left(1 + \frac{[NAD]}{K^M_{NAD}}\right) + \left(1 + \frac{[13P2G]}{K^M_{13P2G}}\right) \cdot \left(1 + \frac{[NADH]}{K^M_{NADH}}\right) - 1}$$

$V_{\max} = 8.18\text{E+}03;\ K_{eq} = 1.92\text{E-}04;\ K^M_{GraP} = 7.00\text{E-}02;\ K^M_{NAD} = 1.00\text{E-}01;\ K^M_{13P2G} = 1.00\text{E-}02;$
$K^M_{NADH} = 1.00\text{E-}02;$

8. Phosphoglycerate kinase (PGK) – $13P2G + ADP \rightarrow 3PG + ATP$

$$v_{PGK} = \frac{V_{\max} \cdot \frac{[13P2G] \cdot [ADP]}{K^M_{13P2G} \cdot K^M_{ADP}} \cdot \left(1 - \frac{[3PG] \cdot [ATP]}{[13P2G] \cdot [ADP]} \over K_{eq}\right)}{\left(1 + \frac{[13P2G]}{K^M_{13P2G}}\right) \cdot \left(1 + \frac{[ADP]}{K^M_{ADP}}\right) + \left(1 + \frac{[3PG]}{K^M_{3PG}}\right) \cdot \left(1 + \frac{[ADP]}{K^M_{ADP}}\right) - 1}$$

$V_{\max} = 1.29\text{E+}01;\ K_{eq} = 1.46\text{E+}03;\ K^M_{13P2G} = 1.00\text{E-}04;\ K^M_{ADP} = 2.30\text{E-}01;\ K^M_{3PG} = 5.90\text{E-}01;$
$K^M_{ATP} = 1.00\text{E+}00;$

9. Phosphoglycerate mutase (PGM) – $3PG \rightarrow 2PG$

$$v_{PGM} = \frac{V_{\max} \cdot \frac{[3PG]}{K^M_{3PG}} \cdot \left(1 - \frac{[2PG]/[3PG]}{K_{eq}}\right)}{1 + \frac{[2PG]}{K^M_{2PG}} + \frac{[3PG]}{K^M_{3GP}}}$$

$V_{\max} = 1.30\text{E+}02;\ K_{eq} = 1.45\text{E-}01;\ K^M_{3PG} = 2.20\text{E-}01;\ K^M_{2PG} = 4.10\text{E-}02;$

10. Enolase (ENO) – $2PG \rightarrow PEP$

$$v_{ENO} = \frac{V_{\max} \cdot \frac{[2PG]}{K^M_{2PG}} \cdot \left(1 - \frac{[PEP]/[2PG]}{K_{eq}}\right)}{1 + \frac{[2PG]}{K^M_{2PG}} + \frac{[PEP]}{K^M_{PEP}}}$$

$V_{\max} = 3.54\text{E+}02;\ K_{eq} = 1.70\text{E+}00;\ K^M_{2PG} = 2.20\text{E-}01;\ K^M_{PEP} = 2.00\text{E-}01;$

11. Pyruvate kinase (PYK) – $PEP + ADP \rightarrow PYR + ATP$

$$v_{PYK} = \left(\frac{[F16P2]}{K^A_{F16P2} + [F16P2]}\right) \cdot \frac{V_{\max} \cdot \frac{[PEP] \cdot [ADP]}{K^M_{PEP} \cdot K^M_{ADP}} \cdot \left(1 - \frac{\frac{[PYR] \cdot [ATP]}{[PEP] \cdot [ADP]}}{K_{eq}}\right)}{\left(1 + \frac{[PEP]}{K^M_{PEP}}\right) \cdot \left(1 + \frac{[ADP]}{K^M_{ADP}}\right) + \left(1 + \frac{[PYR]}{K^M_{PYR}}\right) \cdot \left(1 + \frac{[ADP]}{K^M_{ADP}}\right) - 1}$$

$V_{\max} = 2.86\text{E+}04;\ K_{eq} = 1.38\text{E+}04;\ K^M_{PEP} = 1.70\text{E-}01;\ K^M_{ADP} = 2.40\text{E-}01;\ K^M_{PYR} = 4.80\text{E-}01;$
$K^M_{ATP} = 3.50\text{E-}01;\ K^A_{F16P2} = 1.00\text{E+}00;$

12. Lactate dehydrogenase (LDH) – $PYR + NADH \rightarrow LAC + NAD$

$$v_{LDH} = \frac{V_{\max} \cdot \frac{[PYR] \cdot [NADH]}{K^M_{PYR} \cdot K^M_{NADH}} \cdot \left(1 - \frac{\frac{[LAC] \cdot [NAD]}{[PYR] \cdot [NADH]}}{K_{eq}}\right)}{\left(1 + \frac{[PYR]}{K^M_{PYR}}\right) \cdot \left(1 + \frac{[NADH]}{K^M_{NADH}}\right) + \left(1 + \frac{[LAC]}{K^M_{LAC}}\right) \cdot \left(1 + \frac{[NAD]}{K^M_{NAD}}\right) - 1}$$

$V_{\max} = 1.36\text{E+}03;\ K_{eq} = 9.09\text{E+}03;\ K^M_{PYR} = 3.00\text{E-}02;\ K^M_{NADH} = 7.00\text{E-}03;\ K^M_{LAC} = 1.00\text{E+}00;$
$K^M_{NAD} = 1.00\text{E-}02;$

13. Lactate efflux (LCT) – $LAC \rightarrow LAC(e)$

$$v_{LCT} = \frac{V_{\max} \cdot \frac{[LAC]}{K^M_{LAC}} \cdot \left(1 - \frac{[LAC](e)/[LAC]}{K_{eq}}\right)}{1 + \frac{[LAC]}{K^M_{LAC}} + \frac{[LAC](e)}{K^M_{LAC}}}$$

$V_{\max} = 4.21\text{E+03};\ K_{eq} = 1.00\text{E+00};\ K^M_{LAC} = 1.00\text{E+00};$

14. Glucose -6-phosphate dehydrogenase (G6PDH) –

$G6P + NADP \rightarrow GA6P + NADPH$

$$v_{G6PDH} = \left(\frac{[NADP]}{K^A_{NADP} + [NADP]}\right) \cdot \frac{V_{\max} \cdot \frac{[G6P] \cdot [NADP]}{K^M_{G6P} \cdot K^M_{NADP}} \cdot \left(1 - \frac{\frac{[GA6P] \cdot [NADPH]}{[G6P] \cdot [NADP]}}{K_{eq}}\right)}{\left(1 + \frac{[G6P]}{K^M_{G6P}}\right) \cdot \left(1 + \frac{[NADP]}{K^M_{NADP}}\right) + \left(1 + \frac{[GA6P]}{K^M_{GA6P}}\right) \cdot \left(1 + \frac{[NADPH]}{K^M_{NADPH}}\right) - 1}$$

$V_{\max} = 6.15\text{E+03};\ K_{eq} = 2.00\text{E+03};\ K^M_{G6P} = 1.00\text{E-02};\ K^M_{NADP} = 2.38\text{E-03};\ K^M_{GA6P} = 8.64\text{E-01};$
$K^M_{NADPH} = 1.40\text{E-02};\ K^A_{NADP} = 1.00\text{E-02};$

15. 6-phosphogluconolactonase (6PGDH) – $GA6P + NADP \rightarrow Ru5P + NADPH$

$$v_{6PGDH} = \frac{V_{\max} \cdot \frac{[GA6P] \cdot [NADP]}{K^M_{GA6P} \cdot K^M_{NADP}} \cdot \left(1 - \frac{\frac{[Ru5P] \cdot [NADPH]}{[GA6P] \cdot [NADP]}}{K_{eq}}\right)}{\left(1 + \frac{[GA6P]}{K^M_{GA6P}}\right) \cdot \left(1 + \frac{[NADP]}{K^M_{NADP}}\right) + \left(1 + \frac{[Ru5P]}{K^M_{Ru5P}}\right) \cdot \left(1 + \frac{[NADPH]}{K^M_{NADPH}}\right) - 1}$$

$V_{\max} = 1.66\text{E+02};\ K_{eq} = 1.42\text{E+02};\ K^M_{GA6P} = 1.40\text{E-02};\ K^M_{NADP} = 1.30\text{E-02};\ K^M_{GA6P} = 1.00\text{E-02};$
$K^M_{NADPH} = 1.00\text{E-02};$

16. Glutathione reductase (GSSG) – $GSSG + NADPH \rightarrow 2\,GSH + NADP$

$$v_{GSSG} = \frac{V_{\max} \cdot \frac{[GSSG] \cdot [NADPH]}{K^M_{GSSG} \cdot K^M_{NADPH}} \cdot \left(1 - \frac{\frac{[GSH]^2 \cdot [NADP]}{[GSSG] \cdot [NADPH]}}{K_{eq}}\right)}{\left(1 + \frac{[GSSG]}{K^M_{GSSG}}\right) \cdot \left(1 + \frac{[NADPH]}{K^M_{NADPH}}\right) + \left[1 + \frac{[GSH]}{K^M_{GSH}} + \left(\frac{[GSH]}{K^M_{GSH}}\right)^2\right] \cdot \left(1 + \frac{[NADP]}{K^M_{NADP}}\right) - 1}$$

$V_{\max} = 9.05\text{E+01};\ K_{eq} = 1.04\text{E+00};\ K_{GSSG}^{M} = 1.90\text{E-02};\ K_{NADPH}^{M} = 8.00\text{E-03};\ K_{GSH}^{M} = 6.50\text{E-02};$
$K_{NADP}^{M} = 1.00\text{E-04};$

---

17. Glutathione oxidation (GSHox) – $2\ GSH \rightarrow GSSG$

$$v_{GSHox} = \frac{V_{\max} \cdot \left(\frac{[GSH]}{K_{GSH}^{M}} \right)^{2} \cdot \left(1 - \frac{[GSSG]/[GSH]^{2}}{K_{eq}} \right)}{1 + \frac{[GSH]}{K_{GSH}^{M}} + \left(\frac{[GSH]}{K_{GSH}^{M}} \right)^{2} + \frac{[GSSG]}{K_{GSSG}^{M}}}$$

$V_{\max} = 1.18\text{E+02};\ K_{eq} = 1.00\text{E+05};\ K_{GSSG}^{M} = 1.00\text{E-03};\ K_{GSH}^{M} = 6.90\text{E-01};$

---

18. Phosphoribulose epimerase (EP) – $Ru5P \rightarrow X5P$

$$v_{EP} = \frac{V_{\max} \cdot \frac{[Ru5P]}{K_{Ru5P}^{M}} \cdot \left(1 - \frac{[X5P]/[Ru5P]}{K_{eq}} \right)}{1 + \frac{[Ru5P]}{K_{Ru5P}^{M}} + \frac{[X5P]}{K_{X5P}^{M}}}$$

$V_{\max} = 1.02\text{E+04};\ K_{eq} = 2.70\text{E+00};\ K_{Ru5P}^{M} = 1.90\text{E-01};\ K_{X5P}^{M} = 1.00\text{E-02};$

---

19. Ribose phosphate isomerise (KI) – $Ru5P \rightarrow R5P$

$$v_{KI} = \frac{V_{\max} \cdot \frac{[Ru5P]}{K_{Ru5P}^{M}} \cdot \left(1 - \frac{[R5P]/[Ru5P]}{K_{eq}} \right)}{1 + \frac{[Ru5P]}{K_{Ru5P}^{M}} + \frac{[R5P]}{K_{R5P}^{M}}}$$

$V_{\max} = 6.13\text{E+02};\ K_{eq} = 3.00\text{E+00};\ K_{Ru5P}^{M} = 6.60\text{E-01};\ K_{R5P}^{M} = 2.20\text{E+00};$

---

20. Transketolase I (TK1) – $X5P\ +\ R5P \rightarrow GraP + S7P$

$$v_{TK1} = \frac{V_{\max} \cdot \frac{[X5P] \cdot [R5P]}{K_{X5P}^{M} \cdot K_{R5P}^{M}} \cdot \left(1 - \frac{\frac{[GraP] \cdot [S7P]}{[X5P] \cdot [R5P]}}{K_{eq}} \right)}{\left(1 + \frac{[X5P]}{K_{X5P}^{M}} \right) \cdot \left(1 + \frac{[R5P]}{K_{R5P}^{M}} \right) + \left(1 + \frac{[GraP]}{K_{GraP}^{M}} \right) \cdot \left(1 + \frac{[S7P]}{K_{S7P}^{M}} \right) - 1}$$

$V_{\max} = 1.38\text{E+01};\ K_{eq} = 1.05\text{E+00};\ K_{X5P}^{M} = 1.50\text{E-01};\ K_{R5P}^{M} = 3.00\text{E-01};\ K_{GraP}^{M} = 4.90\text{E+00};$
$K_{S7P}^{M} = 4.00\text{E+00};$

---

21. Transaldolase (TA) – $GraP + S7P \rightarrow E4P + F6P$

$$v_{TA} = \frac{V_{\max} \cdot \frac{[GraP] \cdot [S7P]}{K^M_{GraP} \cdot K^M_{S7P}} \cdot \left(1 - \frac{\frac{[E4P] \cdot [F6P]}{[GraP] \cdot [S7P]}}{K_{eq}}\right)}{\left(1 + \frac{[GraP]}{K^M_{GraP}}\right) \cdot \left(1 + \frac{[S7P]}{K^M_{S7P}}\right) + \left(1 + \frac{[E4P]}{K^M_{E4P}}\right) \cdot \left(1 + \frac{[F6P]}{K^M_{F6P}}\right) - 1}$$

$V_{\max} = 2.05\text{E+02}$; $K_{eq} = 1.05\text{E+00}$; $K^M_{GraP} = 3.80\text{E-02}$; $K^M_{S7P} = 1.80\text{E-01}$; $K^M_{E4P} = 7.00\text{E-03}$; $K^M_{S7P} = 3.43\text{E-01}$;

---

22. Transketolase 2 (TK2) – $E4P + X5P \rightarrow GraP + F6P$

$$v_{TK2} = \frac{V_{\max} \cdot \frac{[X5P] \cdot [E4P]}{K^M_{X5P} \cdot K^M_{E4P}} \cdot \left(1 - \frac{\frac{[GraP] \cdot [F6P]}{[X5P] \cdot [E4P]}}{K_{eq}}\right)}{\left(1 + \frac{[X5P]}{K^M_{X5P}}\right) \cdot \left(1 + \frac{[E4P]}{K^M_{E4P}}\right) + \left(1 + \frac{[GraP]}{K^M_{GraP}}\right) \cdot \left(1 + \frac{[F6P]}{K^M_{F6P}}\right) - 1}$$

$V_{\max} = 5.98\text{E+02}$; $K_{eq} = 1.20\text{E+00}$; $K^M_{X5P} = 1.50\text{E-01}$; $K^M_{E4P} = 3.60\text{E-01}$; $K^M_{GraP} = 4.90\text{E+00}$; $K^M_{F6P} = 7.00\text{E+00}$;

---

23. Phosphoribosylpyrophosphate synthetase (PRPPS) –

$R5P + ATP \rightarrow PRPP + AMP$

$$v_{PRPPS} = \frac{V_{\max} \cdot \frac{[R5P] \cdot [ATP]}{K^M_{R5P} \cdot K^M_{ATP}} \cdot \left(1 - \frac{\frac{[\mathrm{P}RPP] \cdot [AMP]}{[R5P] \cdot [ATP]}}{K_{eq}}\right)}{\left(1 + \frac{[R5P]}{K^M_{R5P}}\right) \cdot \left(1 + \frac{[ATP]}{K^M_{ATP}}\right) + \left(1 + \frac{[PRPP]}{K^M_{\mathrm{P}RPP}}\right) \cdot \left(1 + \frac{[AMP]}{K^M_{AMP}}\right) - 1}$$

$V_{\max} = 1.60\text{E-01}$; $K_{eq} = 1.00\text{E+05}$; $K^M_{R5P} = 3.30\text{E-02}$; $K^M_{ATP} = 1.40\text{E-01}$; $K^M_{PRPP} = 1.00\text{E+00}$; $K^M_{AMP} = 1.22\text{E-01}$;

---

24. TCA pseudo-reaction – $PYR + 4\ NAD \rightarrow 4\ NADH$

$$v_{TCA} = \left(\frac{\frac{ADP}{ATP}}{K^A_{ADP/ATP} + \frac{ADP}{ATP}}\right) \cdot \frac{V_{\max} \cdot \frac{[PYR] \cdot [NAD]}{K^M_{PYR} \cdot K^M_{NAD}} \cdot \left(1 - \frac{\frac{[NADH]}{[PYR] \cdot [NAD]}}{K_{eq}}\right)}{\left(1 + \frac{[PYR]}{K^M_{PYR}}\right) \cdot \left(1 + \frac{[NAD]}{K^M_{NAD}}\right) + \left(1 + \frac{[NADH]}{K^M_{NADH}}\right) - 1}$$

$V_{\max} = 5.63\text{E+01};\ K_{eq} = 5.71\text{E+15};\ K^{M}_{PYR} = 1.00\text{E-01};\ K^{M}_{NAD} = 1.00\text{E-01};\ K^{M}_{NADH} = 1.00\text{E-03};$
$K^{A}_{ADP/ATP} = 3.61\text{E-01};$

---

25. Oxidative phosphorylation (OxPh) – $3\,ADP + NADH \rightarrow 3\,ATP + NAD$

$$v_{OxPh} = \frac{V_{\max} \cdot \frac{[ADP]\cdot[NADH]}{K^{M}_{ADP}\cdot K^{M}_{NADH}} \cdot \left(1 - \frac{\frac{[ATP]\cdot[NAD]}{[ADP]\cdot[NADH]}}{K_{eq}}\right)}{\left(1+\frac{[ADP]}{K^{M}_{ADP}}\right)\cdot\left(1+\frac{[NADH]}{K^{M}_{NADH}}\right)+\left(1+\frac{[ATP]}{K^{M}_{ATP}}\right)\cdot\left(1+\frac{[NAD]}{K^{M}_{NAD}}\right)-1}$$

$V_{\max} = 2.15\text{E+02};\ K_{eq} = 1.00\text{E+09};\ K^{M}_{ADP} = 1.00\text{E-01};\ K^{M}_{NADH} = 1.00\text{E-03};\ K^{M}_{ATP} = 1.00\text{E+00};$
$K^{M}_{NAD} = 1.00\text{E-01};$

---

26. ATPase – $ATP \rightarrow ADP$

$$v_{ATPase} = \frac{V_{\max} \cdot \frac{[ATP]}{K^{M}_{ATP}} \cdot \left(1 - \frac{[ADP]/[ATP]}{K_{eq}}\right)}{1+\frac{[ATP]}{K^{M}_{ATP}}+\frac{[ADP]}{K^{M}_{ADP}}}$$

$V_{\max} = 1.56\text{E+02};\ K_{eq} = 1.00\text{E+05};\ K^{M}_{ADP} = 1.00\text{E-01};\ K^{M}_{ATP} = 1.00\text{E+00};$

---

27. Adenylate kinase (AK) $AMP + ATP \rightarrow 2\,ADP$

$$v_{AK} = \frac{V_{\max} \cdot \frac{[ATP]\cdot[AMP]}{K^{M}_{ATP}\cdot K^{M}_{AMP}} \cdot \left(1 - \frac{[ADP]^2}{[ADP]\cdot[AMP]}{K_{eq}}\right)}{\left(1+\frac{[ATP]}{K^{M}_{ATP}}\right)\cdot\left(1+\frac{[AMP]}{K^{M}_{AMP}}\right)+\left[1+\frac{[ADP]}{K^{M}_{ADP}}+\left(\frac{[ADP]}{K^{M}_{ADP}}\right)^2\right]-1}$$

$V_{\max} = 8.47\text{E-02};\ K_{eq} = 2.50\text{E+01};\ K^{M}_{AMP} = 3.80\text{E-01};\ K^{M}_{ATP} = 1.30\text{E+01};\ K^{M}_{ADP} = 1.00\text{E-01};$

ROBUSTNESS OF THE RESULTS

The parameter values were sampled according to the corresponding metabolite concentrations. In particular, the affinity, inhibition or activation constant for any enzyme in respect to metabolite S (acting as a substrate or as a modifier) was sampled between $\min\{[S]_1,[S]_2\}\times 10^{-\alpha}$ and $\max\{[S]_1,[S]_2\}\times 10^{\beta}$, where $[S]_1$ and $[S]_2$ are the concentration of S in the two metabolic states under comparison, and α and β are two adjustable factors.

The parameter sampling and the subsequent data generation process described in the main text under *Methods* was repeated for different sampling conditions. In particular, different sampling distributions and different values for α and β were used. Tables S4-7 show the list of the enzymes with positive average selectivity $\overline{S_i^{J_{target}}}$ (see main text for its definition), sorted in descending scoring order, for each sampling condition.

Enzyme	Score
PGK	0.99
ENO	0.56
LDH	0.48
PGM	0.45
LCT	0.40
TPI	0.38
GAPDH	0.36
ALD	0.35
6PGD	0.32

Table S4 – *Sampling function: logarithmic. Sampling intervals:* $\alpha=1$, $\beta=1$

Enzyme	Score
PGK	0.99
ENO	0.51
PGM	0.44
TPI	0.40
LDH	0.40
GAPDH	0.38
ALD	0.37
LCT	0.34
EP	0.31
6PGD	0.30
TA	0.29
AK	0.25

Table S5 *– Sampling function: logarithmic Sampling intervals:* $\alpha = 1,\ \beta = 2$

Enzyme	Score
PGK	0.99
ENO	0.64
LDH	0.51
PGM	0.45
LCT	0.42
6PGD	0.35
TPI	0.33
GSSG	0.29
GAPDH	0.29

Tables S6 *– Sampling function: logarithmic. Sampling intervals:* $\alpha = 2,\ \beta = 1$

Enzyme	Score
PGK	0.99
ENO	0.58
PGM	0.44
LDH	0.41
LCT	0.35
TPI	0.38
6PGD	0.34
GAPDH	0.30
ALD	0.30
GSSG	0.29

Tables S7 *– Sampling function: logarithmic. Sampling intervals:* $\alpha = 2,\ \beta = 2$

Enzyme	Score
PGK	0.94
PRPP	0.93
ENO	0.74
PGM	0.43
GSSG	0.43
LDH	0.41
6PGD	0.39
PGI	0.38
KI	0.36
G6PD	0.35
LCT	0.35
TPI	0.33

Tables S8 *– Sampling function: logarithmic. Sampling intervals:* $\alpha = 1,\ \beta = 1$ ***

** In this case, the sampling is performed on the same interval for all the parameters. The borders of this interval are defined as* $\min\{[\mathbf{S}]_1,[\mathbf{S}]_2\}\times 10^{-\alpha}$ *and* $\max\{[\mathbf{S}]_1,[\mathbf{S}]_2\}\times 10^{\beta}$*, where* $[\mathbf{S}]_1$ *and* $[\mathbf{S}]_2$ *denote the complete sets of metabolite concentrations in the two metabolic states.*

Enzyme	Score
PGK	0.99
AK	0.75
PFK	0.70
ENO	0.47
GAPDH	0.43
PGM	0.43
ALD	0.42
TPI	0.41
EP	0.39
TA	0.38
PGI	0.37
TK1	0.33
PK	0.29

Table S9 *– Sampling function: linear. Sampling intervals:* $\alpha = 1,\ \beta = 1$*.*

REFERENCES

1. Majumder, H.K., *Drug targets in kinetoplastid parasites*. Advances in Experimental Medicine and Biology. Vol. 625. 2008: Springer.
2. Groen, A.K., et al., *Quantification of the contribution of various steps to the control of mitochondrial respiration.* The Journal of Biological Chemistry, 1982. **257**(6): p. 2754-2757.
3. Bakker, B.M., et al., *Network-based selectivity of antiparasitic inhibitors.* Molecular Biology Reports, 2002. **29**: p. 1-5.
4. Gerber, S., et al., *Drug-efficacy depends on the inhibitor type and the target position in a metabolic network — A systematic study.* Journal of Theoretical Biology, 2008. **252**: p. 442-455.
5. Westerhoff, H.V., *Systems Biology: New Paradigms for Cell Biology and Drug Design*, in *Systems Biology*. 2007, Springer Berlin Heidelberg. p. 45-67.
6. Guimera, R., M. Sales-Pardo, and L.A.N. Amaral, *A network-based method for target selection in metabolic networks.* Bioinformatics, 2007. **23**(13): p. 1616-1622.
7. Kitano, H., *Cancer as a robust system: implications for anticancer therapy.* Nature Reviews. Cancer, 2004. **4**(3): p. 227-235.
8. Kitano, H., *The Theory of Biological Robustness and Its Implication in Cancer*, in *Systems Biology. Applications and Perspectives*. 2007, Springer Berlin Heidelberg. p. 69-88.
9. Hornberg, J.J., et al., *Metabolic control analysis to identify optimal drug targets*, in *Systems Biological Approaches in Infectious Diseases*. 2007, Birkhäuser Basel. p. 171-189.
10. Westerhoff, H.V. *Emerging Principles: a theory for robustness*. in *ICSB2006*. 2006. Yokohama.
11. Cascante, M., et al., *Metabolic control analysis in drug discovery and disease.* Nature Biotechnology, 2002. **20**: p. 243-249.
12. Schellenberger, J. and B.Ø. Palsson, *The use of randomized sampling for analysis of metabolic networks.* Journal of Biological Chemistry, 2008. **284**(9): p. 5457–5461.
13. Luan, D., M. Zai, and J.D. Varner, *Computationally derived points of fragility of a human cascade are consistent with therapeutic strategies.* Computational Biology, 2007. **3**(7): p. 1347-1359.
14. Wang, L. and V. Hatzimanikatis, *Metabolic engineering under uncertainty. I: Framework development.* Metabolic Engineering, 2006. **8**: p. 133-141.
15. Wang, L. and V. Hatzimanikatis, *Metabolic engineering under uncertainty—II: Analysis of yeast metabolism.* Metabolic Engineering, 2006. **8**: p. 142-159.
16. Tran, L.M., M.L. Rizk, and J.C. Liao, *Ensemble Modeling of Metabolic Networks.* Biophysical Journal, 2008. **95**(12): p. 5606–5617.

17. Steuer, R., et al., *Structural kinetic modeling of metabolic networks.* Proceedings of the National Academy of Sciences, 206. **103**(32): p. 11868–11873.

18. Steuer, R., *Computational approaches to the topology, stability and dynamics of metabolic networks.* Phytochemistry, 2007. **68**(16-18): p. 2139-2151.

19. Grimbs, S., et al., *The stability and robustness of metabolic states: identifying stabilizing sites in metabolic networks.* Molecular Systems Biology, 2007. **3**(146).

20. Reder, C., *Metabolic control theory: a structural approach.* Journal of Theoretical Biology, 1988. **135**(2): p. 175-201.

21. Burns, J.A., et al., *Control analysis of metabolic systems.* Trends in Biochemical Sciences, 1985. **10**(16).

22. Kacser, H. and J.A. Burns, *The control of flux.* Symposia of the society for Experimental Biology, 1973 **27**: p. 65-105.

23. Heinrich, R. and T.A. Rapoport, *A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. .* European Journal of Biochemistry, 1974. **42**(1): p. 89-95.

24. Groen, A.K., et al., *Quantification of the Contribution of Various Steps to the Control of Mitochondrial Respiration.* The Journal of Biological Chemistry. **257**(6): p. 2754-2757.

25. Westerhoff, H.V. and K. VanDam, *Thermodynamics and Control of Biological Free-energy Transduction*. 1987, Amsterdam: Elsevier.

26. Hoops, S., et al., *COPASI - a COmplex PAthway SImulator.* Bioinformatics, 2006. **22**(24): p. 3067-3074.

27. Snoep, J.L., *The Silicon Cell initiative: working towards a detailed kinetic description at the cellular level.* Current Opinion in Biotechnology, 2005. **16**(3): p. 336-343.

28. Fiehn, O., *Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks.* Comparative and Functional Genomics, 2001. **2**(3): p. 155-168.

29. Kell, D.B., *Metabolomics and systems biology: making sense.* Current Opinion in Microbiology, 2004. **7**(3): p. 296-307.

30. Goodacre, R., et al., *Metabolomics by numbers: acquiring and understanding global metabolite data.* Trends in Biotechnology, 2004. **22**(5): p. 254-252.

31. Steuer, R., et al., *Structural kinetic modeling of metabolic networks.* Proceedings of the National Academy of Sciences, 2006. **103**(32): p. 11868–11873.

32. Sauer, U., *Metabolic networks in motion: 13C-based flux analysis.* Molecular Systems Biology, 2006. **2**.

33. Zamboni, N., et al., *13C-based metabolic flux analysis.* Nature Protocols, 2009. **4**(6): p. 878-892.

34. Kruger, N.J. and R.G. Ratcliffe, *Insights into plant metabolic networks from steady-state metabolic flux analysis.* Biochimie, 2009. **91**(6): p. 697-702.

35. Bulik, S., et al., *Kinetic hybrid models composed of mechanistic and simplified enzymatic rate laws – a promising method for speeding up the kinetic modelling of complex metabolic networks.* The FEBS Journal, 2009. **276**(2): p. 410-424.

36. Dräger, A., et al., *Modeling metabolic networks in C. glutamicum: a comparison of rate laws in combination with various parameter optimization strategies.* BMC Systems Biology, 2009. **3**(5).

37. Liebermeister, W. and E. Klipp, *Bringing metabolic networks to life: convenience rate law and thermodynamic constraints.* Theoretical Biology and Medical Modelling, 2006. **3**(41).

38. Pharkya, P., E.V. Nikolaev, and C.D. Maranas, *Review of the BRENDA Database.* Metabolic Engineering, 2003. **5**: p. 71-73.

39. Rojas, I., et al., *SABIO-RK: a database for biochemical reactions and their kinetics*, in *BioSysBio 2007: Systems Biology, Bioinformatics and Synthetic Biology*. 2007, BMC Systems Biology 2007: Manchester, UK.

40. Alberty, R.A., *The role of water in the thermodynamics of dilute aqueous solutions.* Biophysical Chemistry, 2003. **100**(1-3): p. 183-192.

41. Goldberg, R.N., Y.B. Tewari, and T.N. Bhat, *Thermodynamics of enzyme-catalyzed reactions—a database for quantitative biochemistry.* Bioinformatics, 2004. **2**(16): p. 2874–2877.

42. Jankowski, M.D., et al., *Group contribution method for thermodynamic analysis of complex metabolic networks.* Biophysical Journal, 2008. **95**(3): p. 1487-1499.

43. Henry, C.S., et al., *Genome scale thermodynamic analysis of Escherichia coli metabolism.* Biophysical Journal, 2006. **90**(4): p. 1453-1461.

44. Feist, A.M., et al., *A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information.* Molecular Systems Biology, 2007. **3**.

45. Holzhütter, H.-G., *The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks.* European Journal of Biochemistry, 2004. **271**: p. 2905–2922.

46. Wood, S. and P. Trayhurn, *Glucose transporters (GLUT and SGLT): expanded families of sugar transport proteins.* British Journal of Nutrition, 2003. **89**: p. 3-9.

47. Macheda, M.L., S. Rogers, and J.D. Best, *Molecular and Cellular Regulation of Glucose Transporter (GLUT) Proteins in Cancer.* Journal of Cellular Physiology, 2005. **202**: p. 654–662.

48. Marín-Hernández, A., et al., *Determining and understanding the control of glycolysis in fast-growth tumor cells. Flux control by an over-expressed but strongly product-inhibited hexokinase.* FEBS Journal, 2006. **273**: p. 1975–1988.

49. Wilson, J.E., *Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function.* The Journal of Experimental Biology, 2003. **206**: p. 2049-2057.

50. Parry, D.M. and P.L. Pedersen, *Intracellular localization and properties of particulate hexokinase in the Novikoff ascites tumor. Evidence for an outer mitochondrial membrane location.* Journal of Biological Chemistry, 1983. **258**(18): p. 10904-10912.

51. Nakashima, R.A., et al., *Purification and Characterization of a Bindable Form of Mitochondrial Bound Hexokinase from the Highly Glycolytic AS-30D Rat Hepatoma Cell Line.* Cancer Research, 1988. **48**: p. 913-919.

52. Vora, S., J.P. Halper, and D.M. Knowles, *Alterations in the Activity and Isozymic Profile of Human Phosphofructokinase during Malignant Transformation in Vivo and in Vitro: Transformation- and Progression-linked Discriminants of Malignancy.* Cancer Research, 1985. **45**: p. 2993-3001.

53. Richardson, A.D., et al., *Central carbon metabolism in the progression of mammary carcinoma.* Breast Cancer Research and Treatment, 2008. **110**: p. 297–307.

54. Warburg, O., F. Wind, and E. Negelein, *The Metabolism of Tumors in the Body.* Journal of General Physiology, 1927. **8**(6): p. 519-530.

55. Weinhouse, S., *The Warburg hypothesis fifty years later.* Journal of Cancer Research and Clinical Oncology, 1976. **87**(2): p. 115-126.

56. DeBerardinis, R.J., et al., *The Biology of Cancer: Metabolic Reprogramming Fuels Cell Growth and Proliferation.* Cell Metabolism, 2008. **7**(1): p. 11-20.

57. Westerhoff, H.V. and D.B. Kell, *Matrix Method for Determining Steps Most Rate-Limiting to Metabolic Fluxes in BiotechnologicaI Processes.* Biotechnology and Bioengineering, 1987. **30**: p. 101-107.

58. Bakker, B.M., et al., *Metabolic control analysis of glycolysis in trypanosomes as an approach to improve selectivity and effectiveness of drugs.* Molecular and Biochemical Parasitology, 2000. **106**: p. 1-10.

59. Groen, A.K., et al., *Quantification of the contribution of various steps to the control of mitochondrial respiration.* Journal of Biological Chemistry, 1982. **257**(6): p. 2754 -2757.

60. Ljvorzhinski, D., et al., *A novel proteomic coculture model of prostate cancer cell growth.* Proteomics, 2004. **4**(10): p. 3268-3275.

61. DiPaola, R.S., et al., *Therapeutic Starvation and Autophagy in Prostate Cancer :A New Paradigm for Targeting Metabolism in CancerTherapy.* The Prostate, 2008. **68**(16): p. 1743-1752.

62. Hornberg, J.J. and H.V. Westerhoff, *Oncogenes are to lose control on signaling following mutation. Should we aim off target?* Molecular Biotechnology, 2006. **34**(2): p. 109-116.

63. Kitano, H., *Systems Biology: A Brief Overview.* Science, 2002. **295**(5560): p. 1662-1664.

64. Kitano, H., *Computational systems biology.* Nature, 2002. **420**: p. 206-210.

65. Assmus, H.E., et al., *Dynamics of biological systems: role of systems biology in medical research.* Expert Review of Molecular Diagnostics, 2006. **6**(6): p. 891-902.

66. Arnaud, C.H., *Systems Biology's Clinical Future.* Chemical & Engineering News, 2006. **84**(31): p. 17-26.

67. Rohwer, J.M., et al., *Evaluation of a simplified generic bi-substrate rate equation for computational systems biology.* Systems Biololy (Stevenage), 2006. **153**(5): p. 338-341.

CHAPTER 4

TARGETING BREAST CANCER METABOLISM

ABSTRACT

In many diseases, such as cancer, cells show a specific metabolic shift from their normal physiological state. The differences between the normal and the altered metabolic phenotypes may be exploited to identify points of fragility characterising the disease, and hence to specifically target altered cells. The application of Metabolic Control Analysis (MCA) has been proposed as a possible way to identify such points of fragility at the metabolic level. Here we use an MCA approach to assess the suitability of different enzymes as molecular targets for drugs designed to attack breast cancer. We base our study on experimental data characterising the metabolic features of breast cancer, and make use, where possible, of actual kinetic equations in the attempt to provide the most realistic description of the system under study. Unknown metabolic and kinetic quantities are sampled randomly, providing us with a probabilistic assessment of the control profile of the system in the two metabolic phenotypes. The suitability of the different enzymes as molecular targets is subsequently assessed with respect to criteria of both high efficacy and low toxicity.

INTRODUCTION

For almost a century, it has been known that the emergence of cancer is accompanied by specific metabolic alterations. In particular, most cancer cells are characterized by an increased glucose consumption and an aerobic glycolytic activity (known as the Warburg effect) [1]. This metabolic shift, observed almost universally during carcinogenesis, has always been considered a reliable biomarker for tumours [2], and today researchers are assessing the possibility to exploit it in order to target cancer cells more specifically than through traditional approaches [3]. There is currently a quest to find anticancer drugs operating at the metabolic level with both high efficacy and low toxicity. The underlying idea consists of identifying enzymes that represent points of fragility that specifically characterise the cancerous metabolic phenotype [4-6]. These enzymes are such that an alteration in their activity (due for example to the action of an anticancer drug) would elicit the desired response in cancer cells, without affecting their normal counterparts. Metabolic Control Analysis (MCA) is a conceptual framework that can profitably be used to identify such points of fragility [7-9]. The aim of MCA is to understand how the control upon a system's property is distributed among the different enzymatic steps of a metabolic network [7]. Enzymes which exert a strong control over a property of interest in the cancer metabolic phenotype and a low control in the normal phenotype can be considered good candidate targets for a drug aimed to elicit a high differential response between neoplastic and normal cells.

One way to apply MCA and gain insights on the suitability of the different enzymes as putative molecular targets is to generate a fully characterized computational representation of the system under study [10]. Unfortunately, a complete dynamic characterization of a metabolic network is often hindered by the lack of data about the kinetic mechanism of the different enzymatic steps

and the value of many parameters to relevant *in vivo* conditions. A possible way to circumvent this limitation consists of sampling the uncertain or unknown quantities and predicting the control properties of the system on a probabilistic basis [11-17]. In a previous work [18] we showed how putative targets for drugs operating at the metabolic level may be identified through MCA in a probabilistic manner, when minimal knowledge is available about the dynamic properties of the system. In particular we showed that the complete set of fluxes and concentrations defining the two metabolic states under comparison, combined with heuristic assumptions on the properties of typical enzyme-catalysed reactions, already allows for a fast and efficient way to explore the effectiveness of putative drug targets in the abundant cases where detailed kinetic models are unavailable or incomplete. As a proof-of-concept, we applied our methodology to identify points of fragility characterizing a paradigmatic cancer metabolic phenotype, while using only generalised Michaelis-Menten equations to describe the kinetics of the different enzymatic steps.

Using the same conceptual framework, here we address a similar issue but with a more specific clinical implication. In particular, we aim to assess the suitability of different metabolic enzymes as putative molecular targets for a drug specifically designed to attack breast cancer. In order to do so, we based our study on the experimental data currently available in literature about the metabolic features of breast cancer, and made use, where possible, of actual kinetic equations in the attempt to minimize the uncertainty introduced in the description of the system dynamics. The reason for choosing this type of cancer lies on the fact that breast cancer is, to our knowledge, the most extensively characterized in terms of the pattern of the metabolic fluxes acquired by the cells during carcinogenesis and some of the metabolite concentrations [19, 20]. The uncertainties of the system, such as unknown parameter values or unquantified metabolite concentrations, are randomly sampled allowing for a probabilistic assessment of how the control profile of the system differs

between the two metabolic states: cancer and normal. These differences are then used to identify the best putative enzyme targets with respect to specific clinical strategies aimed to attack breast cancer cells in an effective and non-toxic way.

METHODS

MCA TO LOCATE POINTS OF FRAGILITY OF A METABOLIC SYSTEM

The concept of *control coefficient* is central in MCA and provides a way to evaluate – at steady-state – the extent to which a property of interest changes in response to a perturbation in the activity of an enzyme [7, 21, 22]. Important examples of control coefficients are the *flux control coefficient* C_i^J and the *concentration control coefficient* C_i^S, defined as

$$C_i^J := \lim_{\Delta v_i \to 0} \frac{\Delta J/J}{\Delta v_i/v_i} = \frac{dJ/J}{dv_i/v_i} = \frac{d\ln(J)}{d\ln(v_i)} \tag{1}$$

$$C_i^S := \lim_{\Delta v_i \to 0} \frac{\Delta S/S}{\Delta v_i/v_i} = \frac{dS/S}{dv_i/v_i} = \frac{d\ln(S)}{d\ln(v_i)} \tag{2}$$

where J is the steady-state flux of a given pathway, S the steady-state concentration of a given metabolite and v_i is the catalytic activity of enzyme *i*. In Eqs.(1)-(2) it is assumed that the change in the enzymatic activity v_i (which elicits changes in steady-state fluxes and concentrations) is caused by the action of an effector (for example a drug) acting directly and selectively on the enzymatic step *i*. Given a metabolic state $\left(\mathbf{v}^0, \mathbf{S}^0\right)$, described in terms of the complete set of fluxes and metabolite concentrations at steady-state, Reder [23] showed that the matrix of the flux control coefficients $\mathbf{C}^J$ and the matrix of the concentration control coefficients $\mathbf{C}^S$ may be expressed as

$$\mathbf{C}^S = -\left(\mathbf{D}_{\mathbf{S}^0}\right)^{-1} \cdot \mathbf{L} \cdot \left(\mathbf{N}' \cdot \left.\frac{\partial \mathbf{v}}{\partial \mathbf{S}}\right|_0 \cdot \mathbf{L} \right)^{-1} \cdot \mathbf{N}' \cdot \mathbf{D}_{\mathbf{v}^0} \tag{3}$$

$$\mathbf{C}^J = \mathbf{1} + \mathbf{D}_{\mathbf{v}^0}^{-1} \cdot \left.\frac{\partial \mathbf{v}}{\partial \mathbf{S}}\right|_0 \cdot \mathbf{D}_{\mathbf{S}^0} \cdot \mathbf{C}^S \tag{4}$$

where $\mathbf{D}_{\mathbf{v}^0}$ and $\mathbf{D}_{\mathbf{S}^0}$ denote diagonal matrices of elements $\mathbf{S}^0$ and $\mathbf{v}^0$, while $\mathbf{N}'$ and $\mathbf{L}$ denote respectively the reduced stoichiometric matrix and the link matrix (see [23] for definitions), both completely determined by the stoichiometry of the system.

In this paper we use Eqs. (1)-(2) to evaluate the control properties of the system under two different metabolic states, one representing the cancerous phenotype and the other representing the normal phenotype. The information needed to evaluate the control coefficients cover the stoichiometry of the system (given by $\mathbf{N}'$ and $\mathbf{L}$), its dynamic properties (reflected in the partial derivatives $\partial\mathbf{S}/\partial\mathbf{v}$) and the two sets of fluxes and metabolite concentrations defining the two metabolic states (as described by $\mathbf{D}_{\mathbf{v}^0}$ and $\mathbf{D}_{\mathbf{S}^0}$). We base this information as much as possible on literature data in an attempt to minimize uncertainties in the definition of the system properties and the metabolic states under comparison. As we will see later in this section, unknown or uncertain quantities such as kinetic parameters or some of the metabolite concentrations are sampled from reasonable ranges of values.

DEFINITION OF THE METABOLIC MAP

The system under study is a representation of the central carbon metabolism as depicted in Fig. 1, and consists of glycolysis, the pentose phosphate pathway, the TCA cycle and a simplified representation of the respiratory chain.

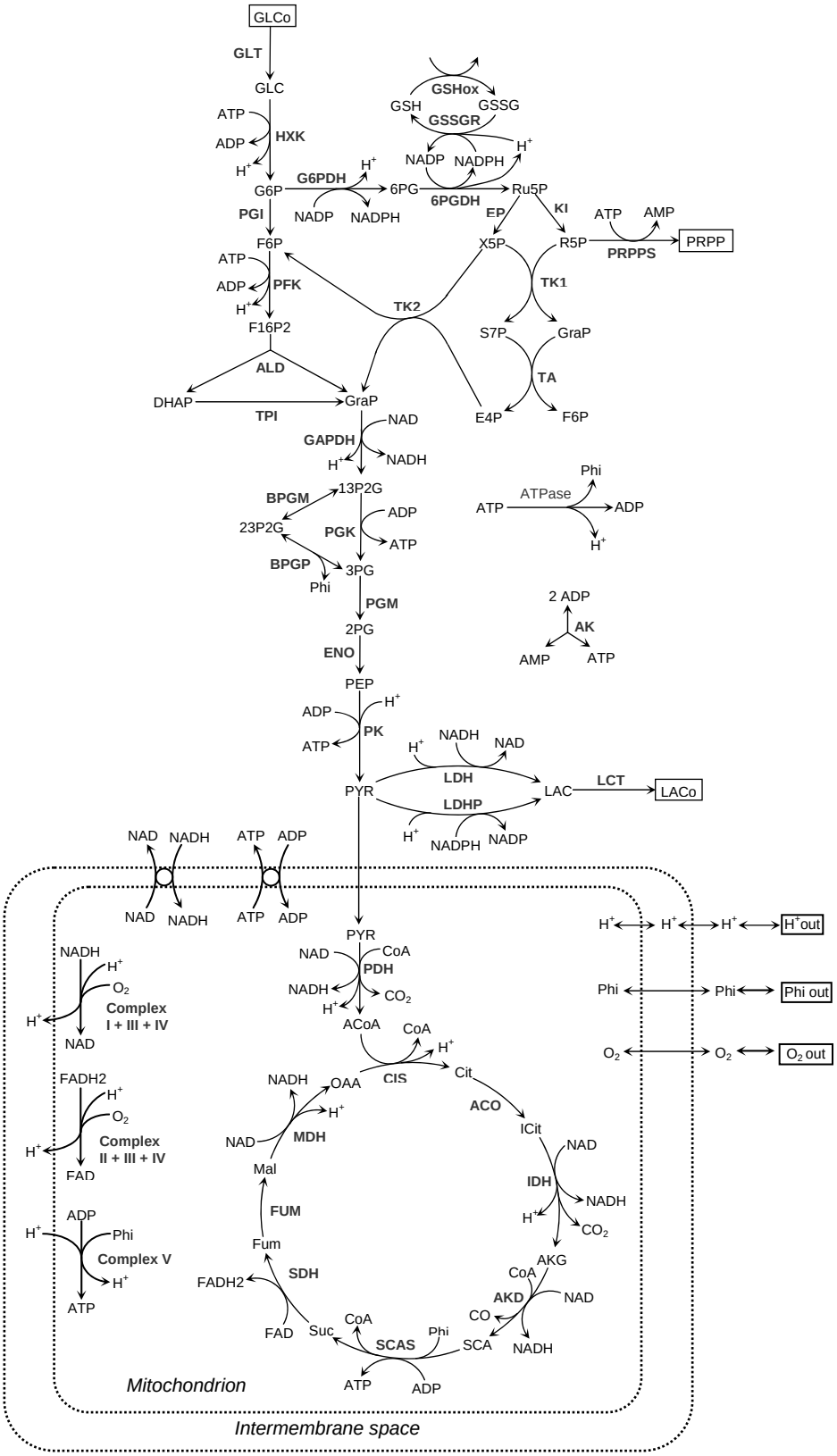

Figure 1 – Metabolic map of central carbon metabolism. *The pathways and metabolic processes taken into account are glycolysis, the pentose phosphate pathway, the TCA cycle and a simplified representation of the oxidative phosphorylation.*

Three different compartments are taken into account: the cytosol, hosting glycolysis and the pentose phosphate pathway; the mitochondria, where the TCA takes place; and the intermembrane space, where the protons produced in mitochondria are pumped in and subsequently released from for the mitochondrial synthesis of ATP. We note that considering the intermembrane space adds an unnecessary degree of details to our representation of central carbon metabolism, as protons equilibrate immediately between intermembrane space and cytosol. However, we chose to consider explicitly all the three compartments to provide a better schematic representation of the system under study.

Our model represents an extension of the reconstruction of erythrocyte central carbon metabolism by Schuster and Holzhütter [24]. The choice to adopt the latter model as our starting point was motivated by the fact that erythrocyte metabolism is the most extensively studied and characterized. Schuster and Holzhütter's model, in particular, provides us with a detailed description of the kinetics of each enzymatic step and a comprehensive regulatory map of the metabolic regulations occurring in glycolysis as well as in the pentose phosphate pathway. Additional reactions were added to this initial model in order to take into account the TCA cycle and the oxidative phosphorylation, which are absent in human erythrocytes. The ANT transporter and a phenomenological translocation step accounting for the shuttling activity of NAD/NADH were also introduced to connect the cytoplasm to the mitochondria.

Following Li et al. [25], the electron transport chain and oxidative phosphorylation are described through three lumped reactions (see S*upplementary Material* for details).

RATE LAWS OF THE REACTION STEPS

The enzyme kinetics of glycolysis and the pentose phosphate pathway is described through the same rate equations used in Schuster's reconstruction of erythrocyte central carbon metabolism [24]. With respect to a generic biochemical reaction $\alpha_1 A_1 + \alpha_2 A_2 + ... \rightarrow \beta_1 B_1 + \beta_2 B_2 + ...$, all these equations can be written in the following general form [24]:

$$v = V_{\max} \cdot \left(\Pi_k A_k^{\alpha_k} - \frac{\Pi_l B_l^{\beta_l}}{K_{eq}} \right) \cdot f(\mathbf{A}, \mathbf{B}, \mathbf{E}, \mathbf{K}) \qquad (4)$$

where K_{eq} denotes the equilibrium constant, $V_{\max}$ the maximal forward rate, and *f* is a function containing all the non-linearities due to saturation, allostery, etc. This function depends on the concentrations of reactants (**A**), products (**B**) and effectors (**E**), and on the kinetic parameters (**K**) such as Michaelis constants or activation/inhibition constants. The benefit of using reaction rates that can be expressed through the general form of Eq.(4) is that the thermodynamic properties, described by the factor $\left(\Pi_k A_k^{\alpha_k} - \Pi_l B_l^{\beta_l} / K_{eq} \right)$, are maintained separated from the specific enzymatic mechanism governing the reaction dynamics. In expanding Schuster and Holzhütter's model to encompass the respiration pathway, we tried to describe the kinetics of the reactions in the TCA cycle following this same principle. The rate equations, in particular, were taken from Wu et al. [26] (for PDH, CIS, ACO, IDH and AKD) and Mogilevskaya et al. [27] (for SDH and FUM). The kinetics of SCAS and MDH were described in terms of generalized Michaelis-Menten rate laws, in the specific implementation proposed by Liebermeister et al. [28].

The kinetics of the three lumped reactions representing the electron chain and oxidative phosphorylation were described through the same rate laws used in Li

et al. [25], where the general form of Eq.(4) is modified in order to take into account the dependency of the reaction rates on the proton motive force (see *Supplementary Material* for details).

The intercompartment translocation of metabolites depends on the action of specific carriers. The kinetics of carrier mediated transport was described through the general rate equation proposed by Li et al. [25]. For the facilitated translocation of an uncharged metabolite S from compartment $c1$ to compartment $c2$, this rate equation can be written as follows:

$$v = T_{\max}\left(\frac{S_{c1}}{K_M + S_{c1}} - \frac{S_{c2}}{K_M + S_{c2}}\right) \tag{5}$$

where S_{c1} and S_{c2} denote the concentrations of S in compartment $c1$ and $c2$ respectively, $T_{\max}$ is the maximal transport rate from compartment $c1$ to compartment $c2$, and K_M is the Michaelis-Menten constant. A modified version of Eq.(5) was used to describe the translocation of ATP/ADP via ANT and the apparent transport of NAD/NADH between cytosol and mitochondria in order to take into account the effect of the mitochondrial membrane potential on these charged cofactors (details given in *Supplementary Material*).

The non facilitated transport was described through a passive diffusion rate equation:

$$v = \lambda\left(S_{c1} - S_{c2}\right) \tag{6}$$

where λ is the permeability coefficient for diffusion from $c1$ to $c2$.

DEFINING THE METABOLIC STATES UNDER COMPARISON

The normal phenotype

Schuster and Holzhütter's model of erythrocyte metabolism was used as starting point not only for the creation of the metabolic map depicted in Fig.1, but also in the definition of the normal metabolic state. In particular, its metabolite concentrations were used as representative of typical physiological values. The pattern of fluxes through the different branches was also maintained as in Schuster and Holzhütter's original model, with the exception of the lactate dehydrogenase flux, which was mainly diverted toward the TCA cycle in order to represent the functioning of a normal cell, where glucose is processed through the respiration pathway. To evaluate the fraction of glycolytic flux entering the TCA cycle, the rates of glucose uptake and lactic acid secretion are set to the physiological value measured in skeletal muscle cells at normal resting condition, respectively 0.195 and 0.09 mM/min [25]. To accomplish this, we rescale the fluxes through glycolysis and the pentose phosphate pathway in order to preserve the original flux distribution pattern, (*i.e.* the fluxes maintain the same relative value with respect to each other). The flux entering the TCA cycle has been obtained by subtracting the conversion rate of pyruvate into lactate from the pyruvate kinase flux, thus ensuring the mass balance of pyruvate at steady-state.

By allowing the flux from pyruvate to enter the TCA cycle, the oxidation of NADH into NAD through lactate dehydrogenase can only occur at a lower rate than in the complete fermentative regime of Shuster and Holzhütter's model, thus breaking the original redox balance in the cytosol. The rate of NAD-NADH translocation between cytosol and mitochondria was set to restore this balance, while maintaining the original cytosolic concentration of the two cofactors. This translocation step was introduced to represent the transport of NAD and NADH

via the glutamate-aspartate cycle and the glycerol phosphate cycle. The net production and transportation rate of NADH in and toward mitochondria was in turn counterbalanced by the reduction of NADH into NAD in the electron transport chain and oxidative phosphorylation, where mitochondrial ADP is phosphorylated to ATP. The balance between mitochondrial ADP and ATP was maintained by allowing the ATP produced in the oxidative phosphorylation to be released in the cytosol, while importing an equal amount of cytosolic ADP into mitochondria.

Disease phenotype

To our knowledge, no study has been performed yet that characterizes the metabolic cancer phenotype in terms of both fluxes and metabolite concentrations. For a partial characterization of the cancerous metabolic phenotype we used the pattern of fluxes measured by Richardson et al. in the most advanced stage of tumour progression in breast cancer [19]. This pattern of fluxes is characterized by a different distribution of the glucose uptake flux amongst the different pathways of the central carbon metabolism. In particular, the flux entering the pentose phosphate pathway now accounts for ~26% of the pyruvate production versus the 2% of the normal metabolic state.

The analysis of Richardson et al., based on ^{13}C labelling techniques, implies the quantification of specific metabolites as a prerequisite to assess the downstream trafficking of carbon influx. However, the set of metabolites for which the concentration is provided in their study do not overlap with the metabolic intermediates of glycolysis, pentose phosphate pathway and TCA cycle, as depicted in Fig. 1 (the only exceptions are lactate and succinate). Moreover, the concentrations are given in terms of relative changes between the least and most advanced stages of cancer progression considered in their study, making them unsuitable for retrieving absolute concentrations starting from a normal metabolic phenotype.

Some of the concentrations of glycolytic intermediates, however, can be retrieved from earlier studies. Cohen et al. [20], for example, quantified the differences in the phosphate metabolite levels in a number of breast cancer cell lines, thus providing us with a range of values that can be used to constrain some of the glycolytic intermediate concentrations. Some general features of cancer physiology can also help us to reduce the uncertainty in the characterization of the cancerous metabolic profile. For example, lactate is generally found to be present in tumours at levels much higher than in the corresponding normal tissues [29-32]. On the other hand, despite the increased acid production, different studies have consistently demonstrated that the intracellular pH of tumours is the same or slightly alkaline compared with that of normal cells [33], as tumour cells excrete protons through up-regulation of the Na^+/H^+ antiport and other membrane transporters. Consequently, extracellular pH is substantially lower (usually by ~0.5 *pH* unit) than normal [34-36]. Some other constraints may be inferred through some general considerations. For example, in normal cells, as a consequence of the activity of the respiration chain and the coupled ATP synthesis, pH in mitochondria is higher than in cytosol. On the other hand, in the absence of any respiration activity the proton concentration in mitochondria cannot exceed the proton concentration in cytosol. This provides us with a lower and upper bound for the mitochondrial pH in cancer cells, where the respiratory activity, relative to the glucose uptake, is lower than it would be in a full respiratory regime. Table 1 lists the range of values we set for some of the metabolic intermediates in the cancerous phenotype and the corresponding bibliographic references.

Metabolite	Range of values	Reference
Glucose (Glc)	0.05 – 0.95	[37]
Glucose-6-phosphate (G6P)	0.0153 – 0.90 (a)	[20]
Fructose-6-phosphate (F6P)	0.0765 – 1.131 (a)	[20]
GSH	0.4 – 2.3	[37]
Lactic acid (Lac)	5.02 – 10.7	[29-32, 37]
Succinate (Succ)	0.9 – 3.2 (b)	[37]
pH (cytosol)	7.1	[33]
pH (mitochondria)	< 7.6 ; > 7.1	

Table 1 – List of constraints for metabolite concentrations in the cancerous phenotype.

(a) The concentration of G6P and F6P were originally expressed in $\mu mol/10^8$ cells. To convert these values to mM, we evaluated the cellular volume assuming spherical cells of typical diameter 50 μm.

(b) The original values refer to the concentration of succinate with respect to the total cell volume. To convert those values to mitochondrial concentrations, we multiply them by a factor of 10.

SAMPLING METABOLITE CONCENTRATIONS

The uncertainty in the definition of the cancerous phenotype due to the lack of unique values for most of the metabolite concentrations has an effect on the outcome of the system in terms of its control profile. This effect was assessed through a random sampling approach, where the uncertain metabolite concentrations were sampled and the control coefficients subsequently evaluated.

Sampling metabolite concentrations in a sensible manner is not trivial. These concentrations have to satisfy the thermodynamic constraints imposed by the reactions in which the corresponding metabolites are involved and the set of steady-state fluxes characterizing the metabolic state under consideration. The rate equations used in our model are such that the logarithmic form of the corresponding thermodynamic constraints can be expressed as linear inequalities

$$\sum_{i=1}^{m} N_{ij} X_i \lessgtr b_j \quad \text{with } j = 1,2,...,n \tag{7}$$

where X_i is the logarithmic concentration of metabolite S_i and N_{ij} is the stoichiometric coefficient of S_i in reaction *j*. The direction of the inequalities in Eq.(7) depends on the specific reaction *j*, and is determined in particular by the sign of the flux at steady-state and the specific rate equation from which the constraint is derived. For the metabolite concentrations to be thermodynamically meaningful, their logarithmic values have to satisfy simultaneously all the *n* linear inequalities of Eq. (7). To sample values which are compliant with this requirement, we used the known property according to which, given a set of solutions $\left\{\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, ..., \mathbf{X}^{(K)}\right\}$ of Eq.(7), any linear combination of the form

$$\frac{\sum_{k=1}^{K} \varphi_k \mathbf{X}^{(k)}}{\sum_{k=1}^{K} \varphi_k} \tag{8}$$

is also a solution of the same set of inequalities. Thanks to this property, once an initial representative set of solutions is found, a thermodynamically compliant way to sample the metabolite concentrations consists of combining these solutions linearly with random coefficients φ_k. To find a first representative set of solutions $\left\{\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, ..., \mathbf{X}^{(K)}\right\}$ we used a linear programming approach, in the form of the algorithm proposed by Lee et al. [38]. More details are provided in *Supplementary Materials*.

SAMPLING KINETIC PARAMETERS

Another source of uncertainty is represented by the value of the kinetic parameters. As a reference point we adopted the parameter values originally used in the publications from which the kinetic equations were taken. However one needs to be aware that the values of kinetic parameters are often retrieved by fitting a model to the experimentally observed behaviour of the system under study. Partial knowledge of the enzymatic mechanisms or simplifying assumptions on the topology of the network or its regulatory map may cause the fitted kinetic parameters to differ appreciably from their *in vivo* value. Because of these reasons we also considered the kinetic parameters among the quantities to be sampled with the aim to assess to what extent their precise value is relevant in determining the control profile of the system. From an operational standpoint, we distinguish between different kinds of parameters:

1. The equilibrium constants K_{eq} were taken from Holzhütter [39] for all the reactions and assumed to be known.

2. Affinity, inhibition or activation constants were sampled randomly. The sampling was performed logarithmically and covered, for each parameter, two orders of magnitude around its original value (*i.e.* the value provided in the work from which the corresponding kinetic law was taken).

3. Maximal activities such as $V_{\max}$ and $T_{\max}$ were adjusted at each sampling iteration in order to make the metabolic state (*i.e.* metabolite concentrations and fluxes) compliant with the steady-state condition.

IDENTIFYING PUTATIVE DRUG TARGETS

The suitability of the different enzymes as putative drug targets was assessed with respect to the same three criteria used in our previous work [18]. We briefly summarise them here. In doing so we assume that the aimed action of a drug hitting an enzyme *i* is meant to induce a decrease in a given flux J_{target}.

1. Maximal selectivity. The *selectivity coefficient*

$$S_i^{J_{target}} \equiv C_{i\ (\text{cancer})}^{J_{target}} - \left| C_{i\ (\text{normal})}^{J_{target}} \right| \qquad (12)$$

was used to quantify the differential response of the system in the two metabolic states under comparison, cancer and disease. Here $C_{i\ (\text{cancer})}^{J_{target}}$ and $C_{i\ (\text{normal})}^{J_{target}}$ denote the control coefficients of enzyme *i* with respect to J_{target} in the cancer and normal metabolic state respectively. The higher the (positive) value of $S_i^{J_{target}}$, the higher the differential response elicited by a drug hitting enzyme *i*. We are interested in particular in enzymes with the highest possible value of $\overline{S_i^{J_{target}}}$, where the average is computed over all the sampling iterations.

2. Minimal toxicity. The *toxicity coefficient* T_i

$$T_i \equiv \frac{1}{N} \sum_j \left| C_{i\ (\text{normal})}^{J_j} \right| \qquad j \in \{\text{all the } N \text{ exchange fluxes}\} \qquad (13)$$

was used to assess to which extent perturbations in the activity of enzyme *i* affect the overall behaviour of the system in the normal metabolic state. The propensity of the system to move away from its original status due to the action of a drug hitting enzyme *i* is measured through the average sensitivity of the

input and output of the system in respect to catalytic activity of that enzyme. The requirement of low toxicity translates into low values of $\overline{T_i}$.

3. Maximal reliability. The reliability of the calculated average selectivity $\overline{S_i^{J_{target}}}$ was evaluated through the *reliability coefficient*, defined as:

$$R_i^{J_{target}} \equiv \frac{\overline{S_i^{J_{target}}}}{\sigma\left(S_i^{J_{target}}\right)} \tag{14}$$

where $\sigma\left(S_i^{J_{target}}\right)$ denotes the standard deviation of the calculated selectivities over all the sampling iterations.

To make the three criteria above quantitatively comparable, we normalized the three coefficients defined in Eq. (12)-(13)-(14) as follows:

$$\sigma_i \equiv \frac{\overline{S_i^{J_{target}}}}{\max\limits_{\text{all } i}\left\{\overline{S_i^{J_{target}}}\right\}} \tag{15}$$

$$\rho_i \equiv \frac{\overline{R_i^{J_{target}}}}{\max\limits_{\text{all } i}\left\{\overline{R_i^{J_{target}}}\right\}} \tag{16}$$

$$\frac{1}{\tau_i} \equiv \frac{\max\left\{\overline{T_i}\right\} - \overline{T_i}}{\max\left\{\overline{T_i}\right\} - \min\left\{\overline{T_i}\right\}} \tag{17}$$

When considering only enzymes with positive $\overline{S_i^{J_{target}}}$, all the three normalized coefficients have values spanning from 0 to 1. Here 0 represents the worst scenario in terms of either selectivity (normal and cancer cells are equally responsive in magnitude), toxicity (highest alteration of the normal metabolic

phenotype) and reliability (maximal uncertainty of the average selectivity). On the other end, 1 represents the best possible result for all the three criteria. We used the normalized coefficients σ_i, $1/\tau_i$ and ρ_i to compute, for each enzyme *i*, a unique score Z_i accounting simultaneously for all the three criteria mentioned above:

$$Z_i = \frac{w_\sigma \cdot \sigma_i + w_\rho \cdot \rho_i + \frac{w_\tau}{\tau_i}}{w_\sigma + w_\rho + w_\tau} \quad (18)$$

where w_σ, w_ρ and w_τ are weight factors that can be chosen according to the specific relation of priority one wants to give to the three criteria.

RESULTS

DIFFERENT STRATEGIES FOR TARGETING CANCER METABOLISM

The results reported in this section refer to three possible strategies of intervention to target cancer metabolism.

- **Strategy 1** – The first strategy consists of blocking glucose uptake in an attempt to starve cancer cells specifically. Because of the high dependence of cancer cells on glucose as source of energy [40], decreasing the glucose uptake is likely to weaken their proliferative potential [41].

- **Strategy 2** – The second strategy consists of decreasing the pentose phosphate pathway flux in order to hinder the synthesis of ribose, a fundamental component of nucleic acids. The rationale behind this strategy is that the most (around 75~90%) of ribose recovered from nucleic acids of certain tumour cells arrives directly or indirectly through the PPP [42, 43].

Hence, inhibiting the PPP flux would result in hindering cancer cell replication [44].

- **Strategy 3** – A third clinical approach may consist of inhibiting the excretion of lactic acid. Cancerous cells, in fact, compete with adjacent normal cells by creating an acidic extra-cellular environment which is harmful to the non-neoplastic tissue [45]. By inhibiting lactic acid excretion (which is the main cause of the low extracellular pH in the cancer microenvironment), one would hinder one of the means by which cancerous cells invade in the pre-existing normal cell population [46, 47]. At the same time, a decreased lactic acid efflux would induce the self-poisoning of cancer cells through the excess of endogenous lactic acid production [48-50].

STRATEGY 1: STARVING CANCER CELLS

Because we already had a reference value for the kinetic parameters (see *Methods*), we started to evaluate the control coefficients from sampled values only of the (unknown/uncertain) metabolite concentrations of the cancerous phenotype, in order to assess to what extent the uncertainties of the cancerous metabolic state alone were affecting the control profile of the system. By limiting the sampling process to the metabolite concentrations, we obtained one single value for each of the control coefficients in the normal metabolic state, as no quantity defining the non-cancerous metabolic phenotype was involved in the sampling process. Conversely, in the cancerous metabolic state a distribution of values was obtained for each control coefficient, reflecting the effect that the uncertainty of the input data has on the response of the system.

Fig. 2 shows the distributions of the calculated selectivity coefficients of some of the enzymes in the system with respect to the uptake of glucose. The selectivity coefficients of hexokinase (Fig. 2.a) are entirely distributed around negative values. According to the definition of the selectivity coefficient provided in

Eq.(12), this means that the magnitude of the control exerted by this enzyme on the uptake of glucose is larger in the normal phenotype than it is in the cancerous. This result seems in accordance with the fact that hexokinase (HXK) is one of the most overexpressed enzymes in many tumours [51], including breast cancer [52, 53], especially in its isoform HXKII. Indeed, when the concentration (hence the activity) of a specific enzyme increases, its control is theoretically expected to decrease as a consequence of the lower degree of saturation of the enzyme with respect to its substrate [54].

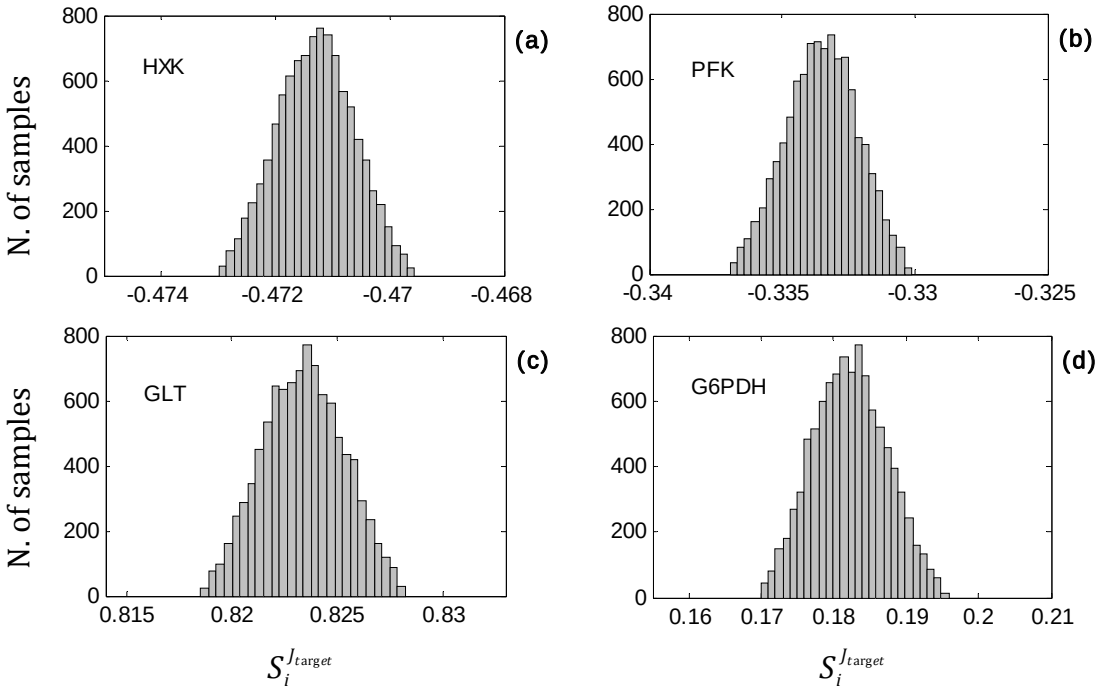

Figure 2 – Selectivity coefficient distributions – Strategy 1. *The selectivity coefficients are shown for HXK (a), PFK (b), GLT (c) and G6PDH (d). The results refer to a clinical strategy aimed to starve cancer cells.*

Similarly, phosphofructokinase (PFK) also shows negative values of its selectivity coefficient (Fig. 2.b), accordingly with what one would expect on the bases of the same reasoning exposed above. Interestingly, however, the selectivity coefficient of the glucose transporter (Fig. 2.c) is highly positive, despite this enzyme being known to be strongly overexpressed in breast cancer

[55-59]. This result, seemingly in contrast with known experimental data about enzyme expression levels, might not be in contradiction with the behaviour that the system would have *in vivo*. The overexpression of a particular enzyme, in fact, must be seen in relation with the (differential) levels of expression of all the other enzymes, as they also contribute in determining the control profile of the system. Hence overexpression of a particular enzyme does not *necessarily* result in a decreased of its control. The high control that the glucose transporter shows over the uptake of glucose in the cancerous phenotype might find some substantiation in a study on AS-30D and HeLa tumour cells revealing that GLUT is amongst the main flux-controlling steps in both tumours (A. Marín-Hernández, R. Moreno-Sánchez and E. Saavedra, personal communication).

Another interesting result is the positive values of the selectivity coefficient of G6PDH (Fig. 2.d), the first enzymatic step of the pentose phosphate pathway. These positive values might well reflect the fact that the glucose uptake flux is diverted into the PPP to a much greater extent in the cancer phenotype than in the normal one. By hindering the flux through the PPP one should then elicit a greater inhibitory effect on the glucose uptake in cancer cells. This result might also suggest that G6PDH has a higher control over the PPP flux in cancer cells than it has in normal cells. We will verify this assertion later, when presenting the results obtained for the second clinical strategy (decreasing the PPP flux in order to hinder the synthesis of ribose).

A more comprehensive picture of the suitability of the different enzymes as molecular targets can be achieved by considering not only the maximal selectivity as discriminator factor, but also the criteria of minimal toxicity (or maximal safety) and maximal reliability introduced in the previous section. In order to provide a general picture of how the three criteria are met by the different enzymes, Fig. 3 shows the normalized selectivity σ_i, safety $1/\tau_i$ and

reliability ρ_i plotted against each other. Only enzymes with a positive average selectivity coefficient are shown, as the others represent bad candidate drug targets. Interestingly, GLT appears to be the best candidate in respect to maximal selectivity as well as minimal toxicity (maximal safety), and it is also among the best candidates in terms of reliability.

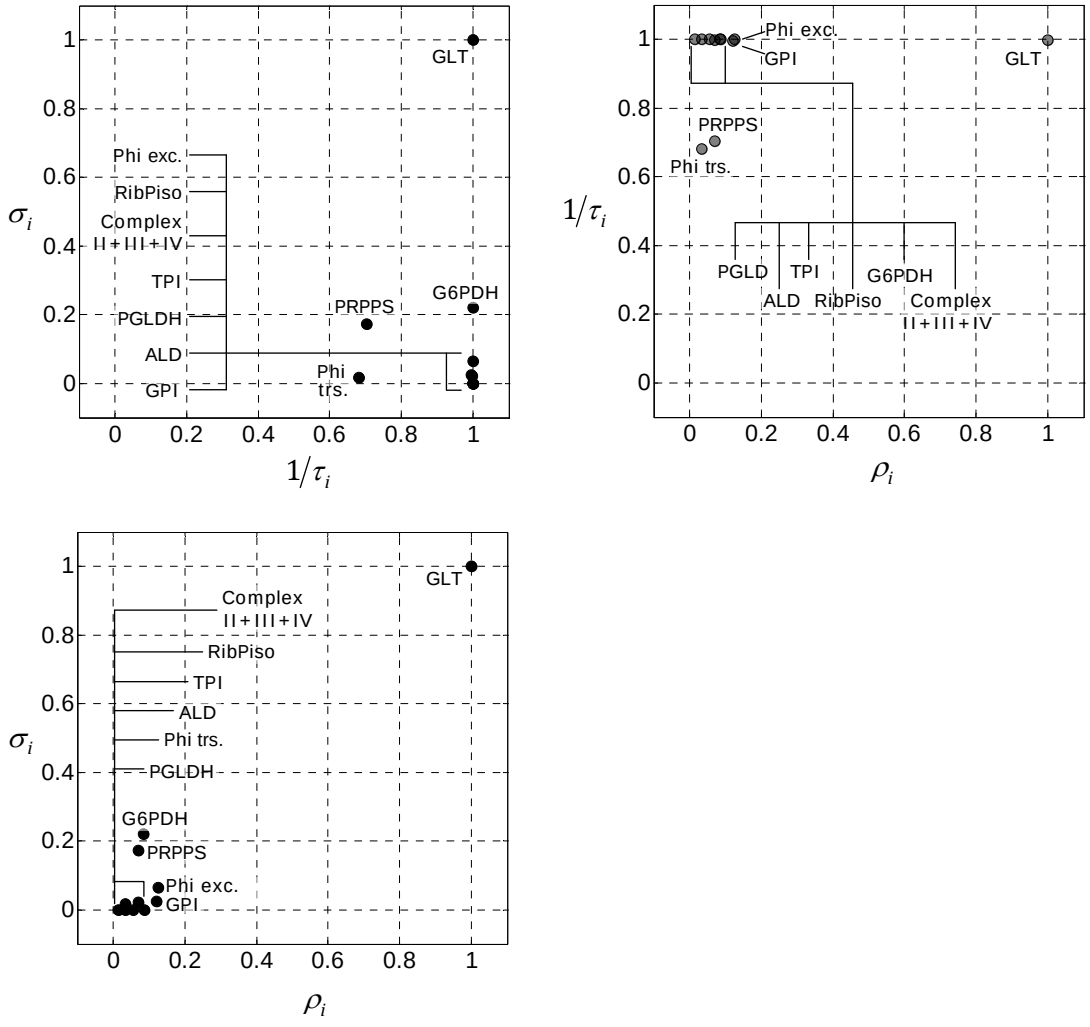

Figure 3 – Normalized selectivity, safety and reliability coefficients plotted versus each other – Strategy 1. *The value of the coefficients was evaluated with regard of the clinical strategy aimed to strarve cancer cells.*

For each enzyme we used Eq.(18) to calculate a unique score representing the overall suitability of the enzyme as putative target, where the three criteria are

simultaneously taken into account. The weight coefficients in Eq.(18) were set as follows: $w_\sigma = 4$, $w_\tau = 2$ and $w_\rho = 1$ in order to prioritize the maximal selectivity over the minimal toxicity (maximal safety) and the latter over the maximal reliability. The first two columns of Table 2 list the enzymes shown in Fig. 3 and their corresponding score.

Strategy 1		Strategy 2		Strategy 3	
Reaction step	**Score**	**Reaction step**	**Score**	**Reaction step**	**Score**
GLT	1.00	G6PDH	1.00	GLT	1.00
G6PDH	0.46	GLT	0.60	G6PDH	0.46
Phi exc.	0.39	AK	0.48	RibPiso	0.36
GPI	0.36	TrKet1	0.41	GPI	0.35
RibPiso	0.36	GAPDH	0.37	TPI	0.34
Comp. II+III+IV	0.35	GSSGRD	0.36	Phiexch	0.34
TPI	0.34	RibPepi	0.36	PGLDH	0.34
ALD	0.34	TrKet2	0.35	PPRPPS	0.33
PGLDH	0.34	TrAld	0.35		
PRPPS	0.33	PGK	0.35		
Phi trs.	0.24	BPGP	0.32		
		ATPase	0.10		

Table 2 – Suitability of the enzymes as drug target when only metabolite concentrations are sampled. *The suitability of each enzyme as drug target is assessed with regard to the three clinical strategies described in the text. The score associated to each enzyme is computed through Eq. (18) with weight factors* $w_\sigma = 4$, $w_\tau = 2$ *and* $w_\rho = 1$. *Only enzymes with positive average selectivity are shown.*

STRATEGY 2: HINDERING THE PRODUCTION OF RIBOSE

The same type of analysis was repeated for a drug designed to hinder the production of ribose by inhibiting the flux through the pentose phosphate pathway. Fig. 4 shows the distribution of the selectivity coefficient for some of the reaction steps in the system. The high positive values of the selectivity coefficient of G6PHD (Fig. 4.a) reflect the fact that the first step of PPP has a stronger control in the cancerous phenotype than in the normal, hence

confirming our previous supposition. Although no comparative study has been performed yet, this result is partly corroborated by experimental studies [44] and theoretical studies [60] showing that G6PDH exerts a higher control in cancer cells over the flux of the oxidative part of PPP.

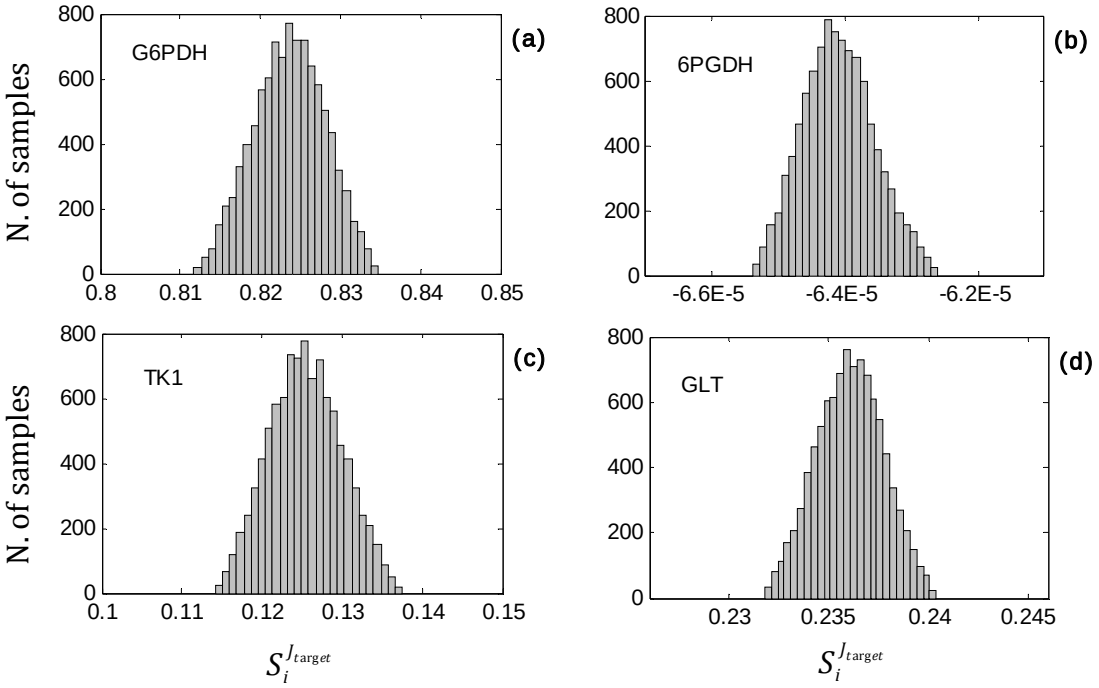

Figure 4 – Selectivity coefficient distributions – Strategy 2. *The selectivity coefficients are shown for G6PDH (a), 6PGDH (b), TK1 (c) and GLT (d). The results refer to a clinical strategy aimed to hinder the production of ribose.*

On the other hand, the selectivity coefficient of 6PGDH, the second step of pentose phosphate pathway, is distributed around very low (and negative) values (Fig. 4.b), meaning that the control exerted by that enzymatic step would have a very similar amplitude in both the normal and the cancerous phenotype. The glucose transporter, GLT, also shows appreciable positive values of its selectivity coefficient (Fig. 4.d). This is not a surprising result considering that the control exerted by GLT on the glucose uptake is higher in the cancer phenotype than in the normal. Because the glucose influx through GLT is split between PPP and the downstream steps of glycolysis, it is reasonable to expect

that the positive differential control of GLT on the uptake of glucose is "echoed" in the control over the flux entering the pentose phosphate pathway. This fact makes GLT an interesting candidate target as its inhibition would elicit the desired response in the system with respect to two possible strategies of intervention, one aiming to starve cancer cells, and one aiming to hinder their replication potential by decreasing the production of ribose.

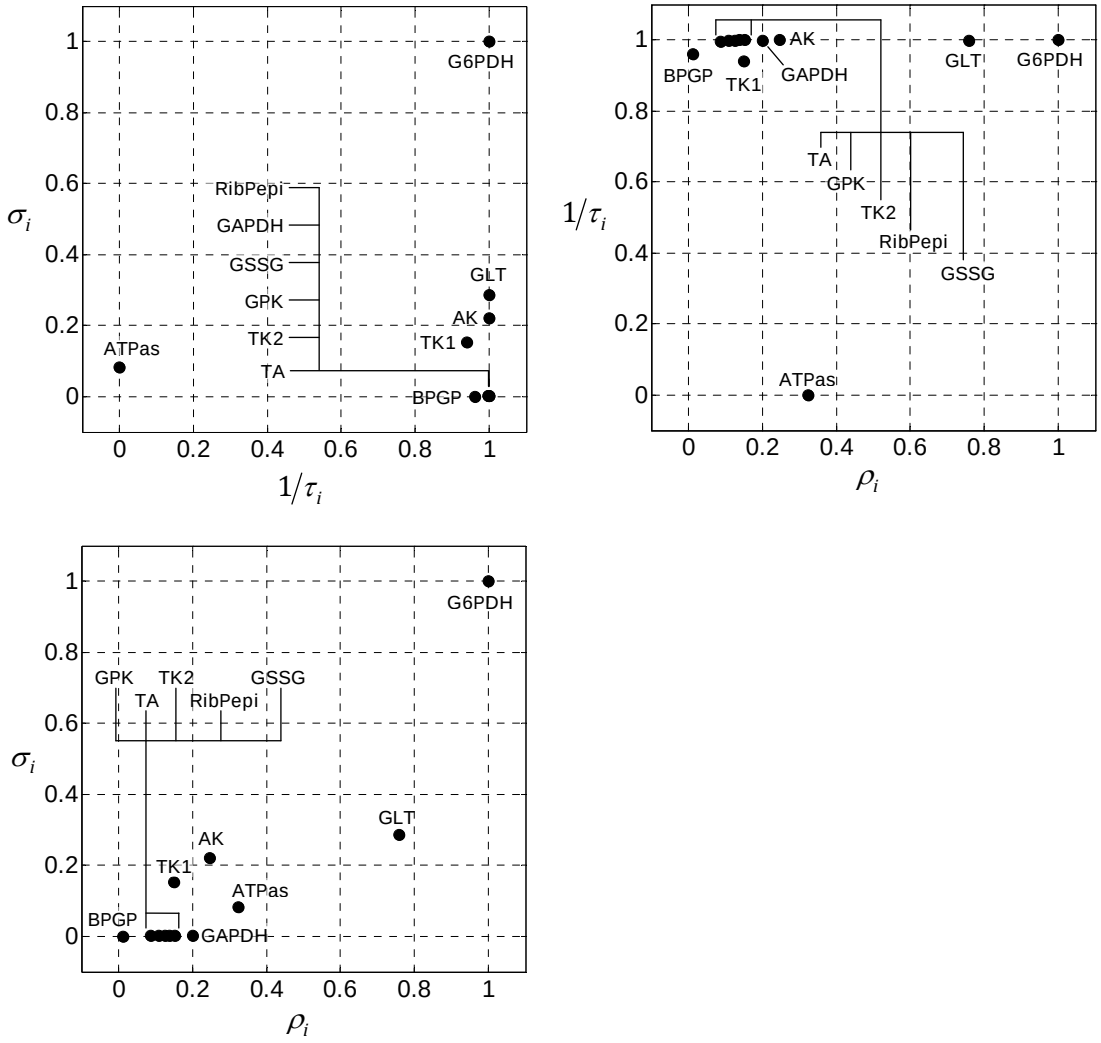

Figure 5 –Normalized selectivity, safety and reliability coefficients plotted versus each other – Strategy 2. *The value of the coefficients was evaluated with regard of the clinical strategy aimed to hinder ribose production.*

Fig. 5 shows that G6PDH is the best enzyme in terms of all the three criteria of selection. GLT might be also considered as a putative target, sharing with G6PDH (and other enzymes) the highest value of safety ($1/\tau_i$), and being the second best in terms of both selectivity and reliability. The central part of Table 2 shows the global score computed for the different enzymes through Eq.(18). The enzymes listed are the same shown in Fig. 5, *i.e.* enzymes with a positive average selectivity coefficient $S_i^{J_{target}}$.

STRATEGY 3: INHIBITING THE EXCRETION OF LACTATE

In a clinical strategy aimed to hinder the lactic acid efflux from cancer cells, one might consider to inhibit the activity of enzymes such as the lactate transporter (LCT) or the lactate dehydrogenases (denoted in Fig. 1 as LDH and LDHP). Paradoxically, in our study these enzymes show very low selective coefficients (Fig. 6.a-b-c), meaning that the control they exert on the lactic acid efflux is similar in cancer and normal cells. Although in general this represents a bad result in terms of drug selectivity, in the context of this specific clinical strategy it might not imply particularly harmful consequences. Normal cells, in fact, do not rely on the fermentative pathway of glycolysis (except for muscle cells under effort), hence a drug targeting one of these enzymes might elicit the desired effect on cancer cells without compromising the normal functioning of their normal counterparts. In this case, rather than the selectivity coefficient, one might be interested in the distribution of the pure flux control coefficients of LDH, LDHP and LCT. However, our data showed that for these enzymes not only the differential control (selectivity coefficient) was particularly low, but also the control coefficients corresponding to each of the metabolic states under comparison. In contrast with LCT, LDH and LDHP, a high selectivity coefficient is shown by the glucose transporter (Fig. 6.d). Interestingly, the values of the selectivity coefficient of GLT are higher for the efflux of lactate than for the

uptake of glucose. As a consequence of targeting GLT, then, one would expect an increasing in the ratio between the flux entering the TCA cycle and the glucose influx from the value of the unperturbed cancer phenotype.

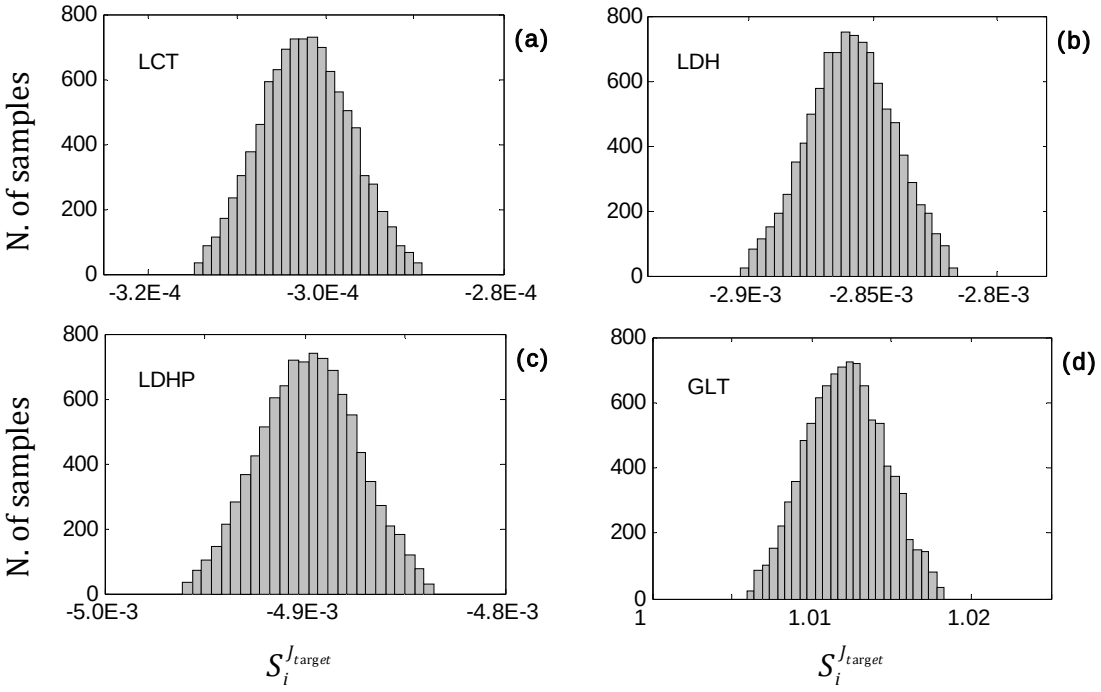

Figure 6 – Selectivity coefficient distributions – Strategy 3*. The selectivity coefficients are shown for LCT (a), LDH (b), LDHP (c) and GLT (d). The results refer to a clinical strategy aimed to inhibiting the excretion of lactate.*

From Fig. 7, GLT appears to be the best putative target in respect to the criteria of highest selectivity and reliability. It is also hardly toxic, as many of the other enzyme shown in the plot. The score of the enzymes in Fig. 8 is reported in the right part of Table 2.

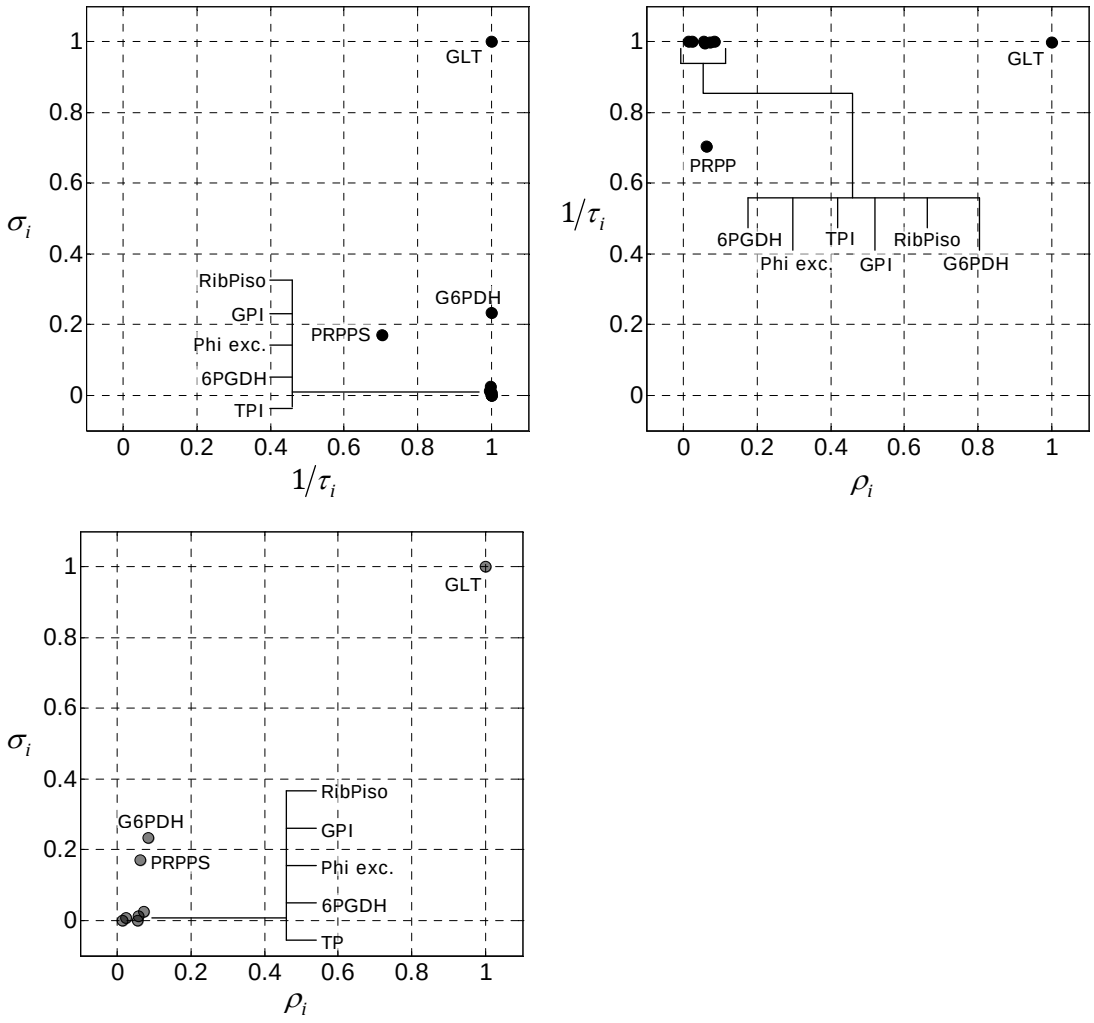

Figure 7 –Normalized selectivity, safety and reliability coefficients plotted versus each other – Strategy 3. *The value of the coefficients was evaluated with regard of the clinical strategy aimed to hinder the excretion of lactate.*

SAMPLING CONCENTRATIONS AND KINETIC PARAMETERS

The same analysis performed for the three clinical strategies illustrated above was repeated by including the kinetic parameters in the set of sampled quantities. In particular, in addition to the concentrations, we sampled parameters such as Michaelis-Menten constants of inhibition/activation constants, quantifying the strength of the interaction between the enzymes with their substrates and effectors.

The effect of enlarging the set of sampled quantities consisted of spreading the distributions of the control properties over wider ranges of values. In the vast majority of cases (90%), the averages of the selectivity coefficient distributions maintained their sign. This fact implies that the main traits of the control properties of the system already emerge from the data defining the metabolic states under comparison. In particular, the qualitative behaviour of the enzymes, in terms of the differential response that their inhibition would elicit in the system between the two metabolic states, is statistically predicted to be the same independent of whether the kinetic parameters are included or not in the set of sampled quantities. However, the increased uncertainty in the definition of the model, due to the sampling of the parameters, might affect the results obtained previously in terms of the suitability of the different enzymes as putative drug targets. In Table 3 we provide the listing of the top scoring enzymes for the three different clinical strategies where the kinetic parameters were included in set of sampled quantities.

Strategy 1		**Strategy 2**		**Strategy 3**	
Reaction step	**Score**	**Reaction step**	**Score**	**Reaction step**	**Score**
GLT	1.00	G6PDH	1.00	GLT	1.00
Phi exc.	0.49	GLT	0.59	Phiexch	0.42
GPI	0.42	AK	0.48	GPI	0.41
ACO	0.38	TrKet1	0.40	G6PDH	0.34
G6PDH	0.35	GSSGRD	0.37	RibPiso	0.34
Phi trs.	0.35	GAPDH	0.36	TrAld	0.34
RibPiso	0.34	RibPepi	0.35	TrKet2	0.34
TPI	0.34	TrKet2	0.35	TrKet1	0.32
TrAld	0.34	TrAld	0.35		
TrKet2	0.34	PGK	0.35		
TrKet1	0.33	ATPase	0.09		

Table 3 – Suitability of the enzymes as drug targets. when both metabolite concentrations and kinetic parameters are sampled. *As in Table 2, except for the fact that the results are obtained by sampling both the metabolite concentrations of the cancerous phenotype and the kinetic parameters.*

DISCUSSION

In this paper we presented a study aimed to identify putative targets for a drug operating at the metabolic level and designed to attack breast cancer. The suitability of the different enzymes as putative targets was assessed with respect to criteria of maximal efficacy and minimal toxicity, both evaluated in terms of control coefficients. The criterion of low toxicity, in particular, requires that both the normal and the cancerous metabolic phenotypes are taken into account and their control profiles compared.

The two metabolic states (normal and cancer) and the dynamic properties of the system were described based on currently available literature data. Unknown quantities were sampled randomly and the control properties of the system evaluated at each sampling iteration. The search for putative targets was performed with regard to three possible clinical strategies: starvation of cancer cells through inhibition of glucose uptake; hindering of cell replication by inhibition of ribose production; prevention of tumour expansion via acidification of extracellular environment by inhibiting lactate excretion. The glucose transporter (GLT) and glucose-6-phosphate dehydrogenase (G6PDH) emerged as the two best putative targets in respect to all the three clinical approaches. In interpreting these results, however, some considerations have to be made. For example, the normal metabolic phenotype has been described in terms of the flux pattern and metabolite concentrations of different cell types. In particular, the concentrations of cytosolic metabolites and the ratio between the glycolytic flux and the PPP flux were taken from Schuster's model of human erythrocytes [24]; on the other hand the concentration of most metabolites in mitochondria and the portion of the glycolitic flux entering the TCA cycle were retrieved by experimental data referring to skeletal muscle cells under resting condition [25]. In doing so we intended to describe the normal metabolic phenotype based, to the wider possible extent, on experimental data, hence

using physiological values for the quantities defining the normal metabolic state. However, the heterogeneity of these data and the fact that they do not specifically refer to human breast cells' metabolism may affect the outcome of our study and the reliability of its predictions. For example, it is known that the glucose transporter do not exert a high control on the glucose uptake in human erythrocytes [10]. This might be (or have contributed to) the reason why GLT showed a particularly high differential control between the cancerous and the normal metabolic states over influx of glucose. The experimental characterization of the metabolic phenotype in non-neoplastic human breast cells would contribute to improve the reliability of our study, whose probabilistic nature remains particularly valuable in the abundant cases in which the value of the kinetic parameters are unknown or uncertain.

To further improve the reliability of our predictions, another aspect to be considered is the value of the equilibrium constants. In our analysis, the equilibrium constants were assumed to be known and a thermodynamically consistent set of values was obtained from [39]. We point out two possible limitations of such an assumption. In the first place, the values of the equilibrium constants are often retrieved using computational algorithms designed to provide reasonable approximations [61-63], and they do not necessarily reflect the value of the relevant *in vivo* conditions. Experimental quantification of the free energy changes occurring in different metabolic reactions is available on repositories such as Web GCM [64], but they are currently far to cover even just the central carbon metabolism as represented in Fig. 1. A second limitation is introduced by considering the same set of equilibrium constants for both the normal and the cancerous metabolic phenotypes. The value of the equilibrium constants varies depending on the specific physiological conditions, and factors such as the pH can affect it to a considerable extent [65]. Despite the internal pH of cancer cells is basically the same as in normal cells [33], the mitochondrial concentration of protons (which

is buffered by the bigger cytosolic volume) might significantly differ in the two phenotypes, resulting in a different set of values for the equilibrium constants of mitochondrial reactions.

Regarding specifically the third clinical approach, we notice that inhibiting the excretion of lactic acid has a double effect: on one hand it prevents cancer cells from creating the hyperacidic extracellular environment by means of which they invade the normal tissue; on the other hand it results in accumulation of intracellular lactic acid with toxic effect for neoplastic cells. Depending on the rationale driving this clinical strategy, different definitions of selectivity and toxicity coefficients may be provided. If the focus is on increasing the concentration of lactic acid in transformed cells above toxic levels (rather than decreasing the acidity of the tumour micromilieu), the criteria of selectivity and toxicity might be better defined in terms of concentrations coefficients rather than flux coefficients.

SUPPLEMENTARY MATERIAL

RATE EQUATIONS

Metabolic reactions

The rate laws used to describe the kinetics of the different reaction steps are of four different forms. All the metabolic reactions in our model follow kinetic laws that can be expressed in the following general form

$$v = V_{\max} \cdot \left(\Pi_k [A_k]^{\alpha_k} - \frac{\Pi_l [B_l]^{\beta_l}}{K_{eq}} \right) \cdot f(\mathbf{A},\mathbf{B},\mathbf{E},\mathbf{K}) \qquad \text{(S.1)}$$

where $[A_k]$ and $[B_l]$ denote the concentrations of substrates and products involved in the generic reaction $\alpha_1 A_1 + \alpha_2 A_2 + ... \rightarrow \beta_1 B_1 + \beta_2 B_2 + ...$, while α_k and β_l denote their corresponding stoichiometric coefficients. f is a function describing the interactions of the enzyme with substrates (**A**), products (**P**) and allosteric effector (**E**). The strength of such interactions depends on kinetic parameters (**K**) such as Michaelis-Menten constants of inhibition/activation constants. As stated in the main text, the benefit of using kinetic laws following the general form of Eq.(A.1) is that the thermodynamics of the reaction, expressed by the factor $\left(\Pi_k [A_k]^{\alpha_k} - \Pi_l [B_l]^{\beta_l} / K_{eq} \right)$, is maintained separated from the properties ascribable to the specific enzymatic mechanism governing the reaction kinetics and described by $f(\mathbf{A},\mathbf{B},\mathbf{E},\mathbf{K})$.

Carrier-mediated transport

The carrier mediated transport of a generic metabolite *S* from compartment *c1* to compartment *c2* was modelled through the general rate law proposed by Li et al. [25]:

$$v = T_{\max}\left(\frac{S_{c1}}{K_{c1}+S_{c1}} - \frac{S_{c2}}{K_{c2}+S_{c2}}\right) \tag{S.2}$$

where $T_{\max}$ is the maximal transport rate from compartment $c1$ to compartment $c2$, and K_M is the Michaelis-Menten constant.

A slightly modified version of Eq. (S.2) was used to describe the transport of cofactors between cytosol and mitochondria. In particular, the antitransport of ATP and ADP across the inner mitochondrial membrane was modelled through the following equation:

$$v = T_{MAX}\cdot\left(\frac{\mathrm{PS}_{\mathrm{cyt}}}{K^{\mathrm{cyt}}_{\mathrm{cyt}\to\mathrm{mit,PS}}+\mathrm{PS}_{cyt}} - \frac{\mathrm{PS}_{\mathrm{mit}}}{K^{\mathrm{mit}}_{\mathrm{cyt}\to\mathrm{mit,PS}}+\mathrm{PS}_{\mathrm{mit}}}\right) \tag{S.3}$$

where $\mathrm{PS}=[\mathrm{ATP}]/[\mathrm{ADP}]$ is the phosphorylation potential. A similar expression was used for the apparent transport of NADH from cytosol to mitochondria in exchange for NAD:

$$v = T_{MAX}\cdot\left(\frac{\mathrm{RS}_{\mathrm{cyt}}}{K^{\mathrm{cyt}}_{\mathrm{cyt}\to\mathrm{mit,RS}}+\mathrm{RS}_{cyt}} - \frac{\mathrm{RS}_{\mathrm{mit}}}{K^{\mathrm{mit}}_{\mathrm{cyt}\to\mathrm{mit,RS}}+\mathrm{RS}_{\mathrm{mit}}}\right) \tag{S.4}$$

where $\mathrm{RS}=[\mathrm{NADH}]/[\mathrm{NAD}]$ denotes the redox potential. We note that Eqs.(S.3)-(S.4) are characterized by two Michaelis-Menten constants, as these flux expressions describe the action of antiporters [25]. In particular, the

phenomenological Michaelis-Menten parameters include the effect of mitochondrial membrane potential ($\Delta\Psi$).

Passive diffusion

Passive diffusion was modelled through the following expression:

$$v = \lambda\left(S_{c1} - S_{c2}\right) \tag{S.5}$$

where λ denotes the permeability coefficient for diffusion from compartment $c1$ to compartment $c2$.

Electron transport and oxidative phosphorylation

The electron transport chain and oxidative phosphorylation were described through the following three lumped reactions, as in Li et al. [25]:

$$\mathrm{NADH(m) + H^+(m) + 0.5\ O_2(m) \rightarrow NAD(m) + H_2O + 10\ \Delta H^+}$$

$$\mathrm{FADH2(m) + 0.5\ O_2 \rightarrow FAD(m) + H_2O + 6\ \Delta H^+}$$

$$\mathrm{ADP(m) + Phi(m) + H^+(m) + 3\ \Delta H^+ \rightarrow ATP(m)}$$

The first two reactions require pumping protons from the mitochondrial matrix (m) into the inter membrane space, while the third reaction requires 3 of these protons to be pumped back into the mitochondrial matrix, providing the motive force for the phosphorylation of ADP into ATP. The kinetics of these three reaction were described through the following rate laws (taken from [25])

$$v = \left(\frac{\exp\left(-\frac{10\Delta G_{H^+}}{RT}\right) \cdot V_{MAX}^{f} \frac{[\mathrm{NADH}][\mathrm{O_2}]^{0.5}[\mathrm{H^+}]}{K^{f}} - V_{MAX}^{b} \frac{[\mathrm{NAD}]}{K^{b}}}{1 + \frac{[\mathrm{NADH}][\mathrm{O_2}]^{0.5}[\mathrm{H^+}]}{K^{f}} + \frac{[\mathrm{NAD}]}{K^{b}}} \right) \tag{S.6}$$

$$v = \left(\frac{\exp\left(-\frac{6\Delta G_{H^+}}{RT} \right) \cdot V_{MAX}^{f} \frac{[\mathrm{FADH_2}][\mathrm{O_2}]^{0.5}}{K^{f}} - V_{MAX}^{b} \frac{[\mathrm{FAD}]}{K^{b}}}{1 + \frac{[\mathrm{FADH_2}][\mathrm{O_2}]^{0.5}}{K^{f}} + \frac{[\mathrm{FAD}]}{K^{b}}} \right) \tag{S.7}$$

$$v = \left(\frac{\exp\left(+\frac{3\Delta G_{H^+}}{RT} \right) \cdot V_{MAX}^{f} \frac{[\mathrm{ADP}][\mathrm{Pi}][\mathrm{H^+}]}{K^{f}} - V_{MAX}^{b} \frac{[\mathrm{ATP}]}{K^{b}}}{1 + \frac{[\mathrm{ADP}][\mathrm{Pi}][\mathrm{H^+}]}{\mathrm{K^f}} + \frac{[\mathrm{ATP}]}{\mathrm{K^b}}} \right) \tag{S.8}$$

Eqs. (S.6)-(S.7)-(S.8) have the form of generalized Michaelis-Menten equations where the forward rate has been modified to depend on the proton motive force defined by ΔG_{H^+}. The proton motive force, in turn, can be expressed as:

$$\Delta G_{H^+} = F \cdot \Delta\Psi + RT \ln\left(\frac{\mathrm{H^+_{cyt}}}{\mathrm{H^+_{mit}}} \right) \tag{S.9}$$

where

- F is the Faraday's constant ($F = 2.3061 \times 10^{-2} \frac{\mathrm{cal}}{\mathrm{V \cdot mmol}}$);
- R is the gas constant ($R = 1.986 \cdot 10^{-3} \frac{\mathrm{cal}}{\mathrm{K \cdot mmol}}$);
- T the absolute temperature (in the human body $T \simeq 312\ \mathrm{K} = 37°\mathrm{C}$);
- $\Delta\Psi$ is the mitochondrial membrane potential. If the inner and outer surface of the mitochondrial membrane are thought as the plates of a capacitor [66], the membrane potential can be expressed as:

  $$\Delta\Psi = \frac{\mathrm{H^+_{cyt}} - \mathrm{H^+_{mit}}}{C} \tag{S.10}$$

  where C denote the capacitance. Empirical values can be found in the literature for C. We set $C = 6.75 \times 10^{-3} \frac{\mathrm{mM}}{\mathrm{V}}$ [67, 68].

List of the rate equations

The control coefficients were evaluated at each sampling iteration step based on a set of rate equations governing the kinetic of every reaction or metabolic process. Here we report the complete set of rate equations characterizing the dynamics of our reconstruction of central carbon metabolism. We also provide the parameter values that give rise to the normal metabolic steady-state (see next section). In particular:

- For thermodynamic consistency, all the equilibrium constants (K_{eq}) were retrieved from the same source [39];
- The maximal activities or rates ($V_{\max}, T_{\max}$ or λ) were evaluated to make the set of metabolite concentrations and the values of the other parameters consistent with the flux characterising the steady-state under consideration.
- For all the other parameters we kept the value used in the paper from which the rate equation was taken.

Reactions occurring in the cytoplasm

Hexokinase (HXK) – $GLC + MgATP \rightarrow G6P + MgADP + H^+$

$$v_{HXK} = \frac{V_{\max} \cdot \frac{[GLC]}{[GLC]+K^M_{GLC}} \cdot \frac{V_{M1}}{K^M_{MgATP}} \cdot \left([MgATP] + \frac{\frac{V_{M2}}{V_{M1}} \cdot [MgATP] \cdot [Mg]}{K_{Mg,ATPMg}} - \frac{[G6P] \cdot [MgADP]}{K_{eq}} \right)}{1 + \frac{[MgATP]}{K^M_{MgATP}} \cdot \left(1 + \frac{[Mg]}{K_{Mg,ATPMg}} \right) + \frac{[Mg]}{K^I_{Mg}} + \left(1.55 + \frac{G6P}{K^M_{G6P}} \right) \cdot \left(1 + \frac{[Mg]}{K^I_{Mg}} \right) + \frac{[G23P2]}{K^I_{G23P2}} + \frac{[Mg] \cdot [G23P2]}{K^I_{Mg} \cdot K^I_{MgG23P2}}}$$

$V_{\max}$=1.29E-01; K_{eq}=3.90E+03; V_{M1}=1.58E+01; V_{M2}=3.32E+01; K^M_{GLC}=1.00E+01; K^M_{G6P}=4.50E-03; K^I_{Mg}=1.03E+00; K^M_{MgATP}=1.44E+00; $K_{Mg,ATPMg}$=1.14E+00; K^I_{G23P2}=2.70E+00; $K^I_{MgG23P2}$=3.44E+00;

Glucosephosphate isomerise (GPI) – $G6P \rightarrow F6P$

$$v_{GPI} = \frac{V_{\max} \cdot \left(G6P - \frac{F6P}{K_{eq}} \right)}{G6P + K_{G6P}^{M} \cdot \left(1 + \frac{F6P}{K_{F6P}^{M}} \right)}$$

$V_{\max}$=1.20E+02; K_{eq}=3.93E-01; K_{G6P}^{M}=1.82E-01; K_{F6P}^{M}=7.10E-02;

Phosphofructokinase (PFK) – $F6P + MgATP \rightarrow F16P2 + MgADP + H^{+}$

$$v_{PFK} = \frac{V_{\max} \cdot \left([F6P] \cdot [MgATP] - \frac{[F16P2] \cdot [MgADP]}{K_{eq}} \right)}{\left(F6P + K_{F6P}^{M} \right) \cdot \left([MgATP] + K_{MgATP}^{M} \right) \cdot \left(1 + L_0 \cdot \left(\frac{\left(1 + \frac{[ATP]}{K_{ATP}^{M}} \right) \cdot \left(1 + \frac{[Mg]}{K_{Mg}^{I}} \right)}{\left(1 + \frac{[AMP_{TOT}]}{K_{AMP}^{A}} \right) \cdot \left(1 + \frac{[F6P]}{K_{F6P}^{M}} \right)} \right)^{4} \right)}$$

$V_{\max}$=1.49E+00; K_{eq}=1.00E+05; K_{F6P}^{M}=1.00E-01; K_{ATP}^{M}=1.00E-02; K_{AMP}^{A}=3.30E-02;
K_{MgATP}^{M}=6.80E-02; K_{Mg}^{I}=4.40E-01; L_0=1.07E-03;

Aldolase (ALD) – $F16P2 \rightarrow DHAP + GraP$

$$v_{ALD} = \frac{V_{\max} \cdot \frac{[F16P2]}{K_{F16P2}^{M}} \left(1 - \frac{\frac{[DHAP] \cdot [GraD]}{[F16P2]}}{K_{eq}} \right)}{1 + \frac{[F16P2]}{K_{F16P2}^{M}} + \frac{[GraP]}{K_{GraP}^{I}} + \frac{[DHAP] \cdot \left([GraP] + K_{GraP}^{M} \right)}{K_{DHAP}^{M} \cdot K_{GraP}^{I}} + \frac{[F16P2] \cdot [GraP]}{K_{F16P2}^{M} \cdot K_{GraP}^{II}}}$$

$V_{\max}$=1.27E+01; K_{eq}=1.14E-01; K_{F16P2}^{M}=7.10E-03; K_{DHAP}^{M}=3.64E-02; K_{GraP}^{M}=1.91E-01;
K_{GraP}^{I}=5.72E-02; K_{GraP}^{II}=1.76E-01;

Triosephosphate isomerase (TPI) – $DHAP \rightarrow GraP$

$$v_{TPI} = \frac{V_{\max} \cdot \frac{[DHAP]}{K_{DHAP}^{M}} \cdot \left(1 - \frac{[GraP]/[DHAP]}{K_{eq}} \right)}{1 + \frac{[GraP]}{K_{GraP}^{M}} + \frac{[DHAP]}{K_{DHAP}^{M}}}$$

$V_{\max}$=7.03E+02; K_{eq}=4.07E-02; K_{DHAP}^{M}=8.38E-01; K_{GraP}^{M}=4.28E-01;

Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) –

$GraP + Phi + NAD \rightarrow 13P2G + NADH + H^{+}$

$$v_{GAPDH} = \frac{V_{\max} \cdot \frac{[GraP] \cdot [NAD] \cdot [Phi]}{K^M_{GraP} \cdot K^M_{NAP} \cdot K^M_{Phi}} \cdot \left(1 - \frac{\frac{[13P2G] \cdot [NADH]}{[GraP] \cdot [NAD] \cdot [Phi]}}{K_{eq}}\right)}{\left(1 + \frac{[GraP]}{K^M_{GraP}}\right) \cdot \left(1 + \frac{[NAP]}{K^M_{NAP}}\right) \cdot \left(1 + \frac{[Phi]}{K^M_{Phi}}\right) + \left(1 + \frac{[13P2G]}{K^M_{13P2G}}\right) \cdot \left(1 + \frac{[NADH]}{K^M_{NADH}}\right) - 1}$$

$V_{\max}$=5.54E+02; K_{eq}=6.61E-05; K^M_{GraP}=5.00E-03; K^M_{NAD}=5.00E-02; K^M_{Phi}=1.13E+01; K^M_{13P2G}=3.50E-03; K^M_{Phi}=8.30E-03;

Phosphoglycerate kinase – (PGK) $13P2G + MgADP \rightarrow 3PG + MgATP$

$$v_{PGK} = \frac{\frac{V_{\max}}{K^M_{MgATP} \cdot K^M_{13P2G}} \cdot \left([MgADP] \cdot [13P2G] - \frac{[MgATP] \cdot [3PG]}{K_{eq}}\right)}{\left(1 + \frac{[MgADP]}{K^M_{MgADP}}\right) \cdot \left(1 + \frac{[13P2G]}{K^M_{13P2G}}\right) + \left(1 + \frac{[MgATP]}{K^M_{MgATP}}\right) \cdot \left(1 + \frac{[3PG]}{K^M_{3PG}}\right) - 1}$$

$V_{\max}$=6.44E+02; K_{eq}=1.46E+03; K^M_{MgADP}=3.50E-01; K^M_{MgATP}=4.80E-01; K^M_{13P2G}=2.00E-03; K^M_{3PG}=1.20E+00;

Bisphosphoglycerate mutase (BPGM) – $13G2P \rightarrow 23G2P$

$$v_{BPGM} = \frac{V_{\max} \cdot K^M_{13P2G} \cdot \left([13P2G] - \frac{[23P2G]}{K_{eq}}\right)}{1 + \frac{[23P2G]}{K^M_{23P2G}}}$$

$V_{\max}$=1.29E-01; K_{eq}=1.00E+05; K^M_{13P2G}=7.60E+04; K^M_{23P2G}=4.00E-02;

Bisphosphoglycerate phosphatase (BPGP) – $23P2G \rightarrow 3PG + Phi$

$$v_{BPGP} = \frac{V_{\max} \cdot \left([23P2G] - \frac{[3PG]}{K_{eq}}\right)}{[23P2G] + K^M_{23P2G}}$$

$V_{\max}$=6.82E-02; K_{eq}=1.00E+05; K^M_{23P2G}=2.00E-01;

Phosphoglycerate mutase (PGM) – $3PG \rightarrow 2PG$

$$v_{PGM} = \frac{V_{\max} \cdot \left([3PG] - \frac{[2PG]}{K_{eq}}\right)}{[3PG] + K^M_{3PG} \cdot \left(1 + \frac{[2PG]}{K^M_{2PG}}\right)}$$

$V_{\max}$=2.58E+02; K_{eq}=1.45E-01; K^M_{3PG}=5.00E+00; K^M_{2PG}=1.00E+00;

Enolase (ENO) – $2PG \rightarrow ENO$

$$v_{ENO} = \frac{V_{\max} \cdot \left([2PG] - \frac{[PEP]}{K_{eq}} \right)}{[2PG] + K_{2PG}^{M} \cdot \left(1 + \frac{[PEP]}{K_{PEP}^{M}} \right)}$$

$V_{\max}$=1.93E+02; K_{eq}=1.70E+00; K_{2PG}^{M}=1.00E+00; K_{PEP}^{M}=1.00E+00;

Pyruvate kinase (PYK) – $PEP + MgADP + H^{+} \rightarrow PYR + MgATP$

$$v_{PYK} = \frac{V_{\max} \cdot \left([PEP] \cdot [MgADP] - \frac{[PYR] \cdot [MgATP]}{K_{eq}} \right)}{\left([PEP] + K_{PEP}^{M} \right) \cdot \left([MgADP] + K_{MgADP}^{M} \right) \cdot \left(1 + \frac{L_0 \cdot \left(1 + \frac{[ATP_{TOT}]}{K_{ATP}^{I}} \right)^{4}}{\left(1 + \frac{[PEP]}{K_{PEP}^{M}} \right)^{4} \cdot \left(1 + \frac{[F16P2]}{K_{F16P2}^{A}} \right)^{4}} \right)}$$

$V_{\max}$=3.68E+01; K_{eq}=1.38E+04; L_0=1.9E+01; K_{PEP}^{M}=2.25E-01; K_{MgADP}^{I}=4.74E-01;
K_{ATP}^{I}=3.39E+00; K_{F16P2}^{I}=5.00E-03;

Lactate dehydrogenase (LDH) – $PYR + NADH + H^{+} \rightarrow LAC + NAD$

$$v_{LDH} = \frac{V_{\max} \cdot \frac{[PYR] \cdot [NADH]}{K_{PYR}^{M} \cdot K_{NADH}^{M}} \cdot \left(1 - \frac{[LAC] \cdot [NAD]}{[PYR] \cdot [NADH]} \right)}{\left(1 + \frac{[PYR]}{K_{PYR}^{M}} \right) \cdot \left(1 + \frac{[NADH]}{K_{NADH}^{M}} \right) + \left(1 + \frac{[LAC]}{K_{LAC}^{M}} \right) \cdot \left(1 + \frac{[NAD]}{K_{NAD}^{M}} \right) - 1}$$

$V_{\max}$=2.98E+02; K_{eq}=9.09E+03; K_{PYR}^{M}=3.98E-01; K_{NADH}^{M}=7.00E-03; K_{LAC}^{M}=8.00E+00
K_{NAD}^{M}=5.00E+00;

Lactate dehydrogenase 2 (LDHP) – $PYR + NADPH + H^{+} \rightarrow LAC + NADP$

$$v_{LDH} = \frac{V_{\max} \cdot \frac{[PYR] \cdot [NADPH]}{K_{PYR}^{M} \cdot K_{NADPH}^{M}} \cdot \left(1 - \frac{[LAC] \cdot [NADP]}{[PYR] \cdot [NADPH]} \right)}{\left(1 + \frac{[PYR]}{K_{PYR}^{M}} \right) \cdot \left(1 + \frac{[NADPH]}{K_{NADPH}^{M}} \right) + \left(1 + \frac{[LAC]}{K_{LAC}^{M}} \right) \cdot \left(1 + \frac{[NADP]}{K_{NADP}^{M}} \right) - 1}$$

$V_{\max}$=1.98E-01; K_{eq}=1.42E+04; K_{PYR}^{M}=3.98E-01; K_{NADPH}^{M}=7.00E-03; K_{LAC}^{M}=8.00E+00
K_{NADP}^{M}=5.00E+00;

Glucose 6-phosphate dehydrogenase (G6PDH) – $G6P + NADP \rightarrow 6PG + NADH + H^+$

$$v_{G6PDH} = \frac{V_{\max} \cdot \frac{[G6P] \cdot [NADP]}{K^M_{G6P} \cdot K^M_{NADP}} \cdot \left(1 - \frac{\frac{[6PG] \cdot [NADPH]}{[G6P] \cdot [NADP]}}{K_{eq}} \right)}{1 + \frac{[NADP] \cdot \left(1 + \frac{[G6P]}{K^M_{G6P}} \right)}{K^M_{NADP}} + \frac{[ATP_{TOT}]}{K^I_{ATP}} + \frac{[NADPH]}{K^M_{NADPH}} + \frac{[23P2G]}{K^I_{23P2G}}}$$

$V_{\max}$=2.09E+01; K_{eq}=2.00E+03; K^M_{G6P}=6.67E-02; K^M_{NADP}=3.67E-03; K^I_{NADPH}=3.12E-03; K^I_{ATP}=7.49E-01; K^I_{23P2G}=2.29E+00;

Phosphogluconate dehydrogenase (6PGDH) – $6PG + NADP \rightarrow Ru5P + NADPH + H^+$

$$v_{6PGDH} = \frac{V_{\max} \cdot \frac{[6PG] \cdot [NADP]}{K^M_{6PG} \cdot K^M_{NADP}} \cdot \left(1 - \frac{\frac{[Ru5P] \cdot [NADPH]}{[6PG] \cdot [NADP]}}{K_{eq}} \right)}{\left(1 + \frac{[NADP]}{K^M_{NADP}} \right) \cdot \left(1 + \frac{[6PG]}{K^M_{6PG}} + \frac{[23P2G]}{K^I_{23P2G}} \right) + \frac{[ATP_{TOT}]}{K^I_{ATP}} + \frac{[NADPH] \cdot \left(1 + \frac{[6PG]}{K^M_{6PG}} \right)}{K^M_{NADPH}}}$$

$V_{\max}$=2.03E+02; K_{eq}=1.42E+02; K^M_{6PG}=1.00E-02; K^M_{NADP}=1.80E-02; K^I_{NADPH}=4.50E-03; K^I_{ATP}=1.54E-01; K^I_{23P2G}=1.20E-01;

Glutathione reductase (GSSG) – $GSS + NADPH + H^+ \rightarrow 2\ GSH + NADP$

$$v_{GSSGR} = \frac{V_{\max} \cdot \left(\frac{[GSSG] \cdot [NADPH]}{K^M_{GSSG} \cdot K^M_{NADPH}} - \frac{\left(\frac{[GSH]}{K^M_{GSH}} \right)^2 \cdot \frac{[NADP]}{K^M_{NADP}}}{K_{eq}} \right)}{1 + \frac{[NADPH] \cdot \left(1 + \frac{[GSSG]}{K^M_{GSSG}} \right)}{K^M_{NADPH}} + \frac{[NADP]}{K^M_{NADP}} \cdot \left(1 + \frac{[GSH] \cdot \left(1 + \frac{[GSH]}{K^M_{GSH}} \right)}{K^M_{GSH}} \right)}$$

$V_{\max}$=1.16E+01; K_{eq}=1.04E+00; K^M_{GSSG}=6.52E-02; K^M_{NADPH}=8.52E-03; K^M_{GSH}=2.00E+01; K^I_{NADP}=7.00E-02;

Glutathione oxidation (GSHox) – $2\ GSH \rightarrow GSSG$

$$v_{GSHox} = K \cdot [GSH]$$

K=3.86E-03;

Phosphoribolose epimerase (EP) – $Ru5P \rightarrow X5P$

$$v_{EP} = \frac{V_{\max} \cdot \left([Ru5P] - \frac{[X5P]}{K_{eq}} \right)}{[Ru5P] + K^{M}_{Ru5P} \cdot \left(1 + \frac{X5P}{K^{M}_{X5P}} \right)}$$

$V_{\max}$=5.97E+02; K_{eq}=2.7E+00; K^{M}_{Ru5P}=1.90E-01; K^{M}_{X5P}=5.00E-01;

Ribose phosphate isomerise (KI) – $Ru5P \rightarrow R5P$

$$v_{KI} = \frac{V_{\max} \cdot \left([Ru5P] - \frac{[R5P]}{K_{eq}} \right)}{[Ru5P] + K^{M}_{Ru5P} \cdot \left(1 + \frac{R5P}{K^{M}_{R5P}} \right)}$$

$V_{\max}$=9.40E+01; K_{eq}=3.00E+00; K^{M}_{Ru5P}=7.80E-01; K^{M}_{R5P}=2.20E+00;

Phosphoribosylpyrophosphate synthetase (PRPPS) – $R5P + ATP \rightarrow PRPP + AMP$

$$v_{PRPPS} = \frac{V_{\max} \cdot \left([R5P] \cdot [MgATP] - \frac{[PRPP] \cdot [MgAMP]}{K_{eq}} \right)}{\left(K^{M}_{ATP} + [MgATP] \right) \cdot \left(K^{M}_{R5P} + [R5P] \right)}$$

$V_{\max}$=1.42E-01; K_{eq}=1.00E+05; K^{M}_{R5P}=5.70E-01; K^{M}_{ATP}=3.00E-20;

Transketolase 1 (TK1) – $X5P + R5P \rightarrow GraP + S7P$

$$v_{TK1} = \frac{V_{\max} \cdot \left([R5P] \cdot [X5P] - \frac{[GraP] \cdot [S7P]}{K_{eq}} \right)}{\left(K_1 + [R5P] \right) \cdot [X5P] + \left(K_2 + K_6 \cdot [S7P] \right) \cdot [R5P] + \left(K_3 + K_5 \cdot [S7P] \right) \cdot [GraP] + K_4 \cdot [S7P] + \ldots \\ \ldots K_7 \cdot [X5P] \cdot [GraP]}$$

$V_{\max}$=3.02E+00; K_{eq}=1.05E+00; K_1=4.18E-01; K_2=3.06E-01; K_3=1.24E+01; K_4=4.96E-03; K_5=4.11E-01; K_6=7.74E-03; K_7=4.88E+01;

Transaldolase (TA) – $GraP + S7P \rightarrow E4P + F6P$

$$v_{TA} = \frac{V_{\max} \cdot \left([S7P] \cdot [GraP] - \frac{[E4P] \cdot [F6P]}{K_{eq}} \right)}{\left(K_1 + [GraP] \right) \cdot [S7P] + \left(K_2 + K_6 \cdot [F6P] \right) \cdot [GraP] + \left(K_3 + K_5 \cdot [F6P] \right) \cdot [E4P] + K_4 \cdot [F6P] + \ldots \\ \ldots K_7 \cdot [S7P] \cdot [E4P]}$$

$V_{\max}$=3.50E+00; K_{eq}=1.05E+00; K_1=8.23E-03; K_2=4.77E-02; K_3=1.73E-01; K_4=6.10E-03; K_5=8.68E-01; K_6=4.66E-01; K_7=2.52E+00;

Transketolase 2 (TK2) – $X5P + E4P \rightarrow F6P + GraP$

$$v_{TK2} = \frac{V_{\max} \cdot \left([E4P] \cdot [X5P] - \frac{[GraP] \cdot [F6P]}{K_{eq}} \right)}{\left(K_1 + [E4P]\right) \cdot [X5P] + \left(K_2 + K_6 \cdot [F6P]\right) \cdot [E4P] + \left(K_3 + K_5 \cdot [F6P]\right) \cdot [GraP] + K_4 \cdot [F6P] + \ldots \\ \ldots K_7 \cdot [X5P] \cdot [GraP]}$$

$V_{\max}$=3.02E+00; K_{eq}=1.2E+00; K_1=1.84E-03; K_2=3.06E-01; K_3=5.48E-02; K_4=3.00E-04; K_5=2.87E-02; K_6=1.22E-01; K_7=2.15E-01;

Adenylate kinase (AK) – $ATP + AMP \rightarrow 2\,ADP$

$$v_{AK} = \frac{\frac{V_{\max}}{K^M_{ATP} \cdot K^M_{AMP}} \cdot \left([MgATP] \cdot [AMP] - \frac{[MgADP] \cdot [ADP]}{K_{eq}} \right)}{\left(1 + \frac{[MgATP]}{K^M_{ATP}}\right) \cdot \left(1 + \frac{[AMP]}{K^M_{AMP}}\right) + \frac{\left([ADP] + [MgADP]\right)}{K^M_{ADP}} + \frac{[MgADP] \cdot [ADP]}{\left(K^M_{ADP}\right)^2}}$$

$V_{\max}$=1.78E+02; K_{eq}=2.5E-01; K^M_{ATP}=9.00E-02; K^M_{ADP}=1.10E-01; K^M_{AMP}=8.00E-02;

ATPase – $ATP \rightarrow MgADP + Phi + H^+$

$$v_{ATPase} = K \cdot [MgATP]$$

K=2.13E+00;

Reactions occurring in mitochondria

Pyruvate dehydrogenase (PDH) – $PYR + NAD + CoA \rightarrow ACoA + NADH + CO_2 + H^+$

$$v_{PDH} = \frac{V_{\max} \cdot [PYR] \cdot [CoA] \cdot [NAD] \cdot \left(1 - \frac{[PYR] \cdot [CoA] \cdot [NAD]}{\frac{[ACoA] \cdot [CO_2] \cdot [NADH]}{K_{eq}}} \right)}{KMC \cdot \left(1 + \frac{[NADH]}{K^M_{NADH}}\right) \cdot [PYR] \cdot [CoA] + KMB \cdot \left(1 + \frac{[ACoA]}{K^M_{ACoA}}\right) \cdot [PYR] \cdot [NAD] + KMA \cdot [CoA] \cdot [NAD]}$$

$V_{\max} = 9.18\text{E-}01$; $K_{eq} = 1.00\text{E+}05$; $KMA = 3.83\text{E-}02$; $KMB = 9.90\text{E-}03$; $KMC = 6.07\text{E-}02$; $K^I_{ACA} = 4.02\text{E-}02$; $K^I_{NADH} = 4.00\text{E-}02$;

Citrate synthase (CIS) – $ACoA + OAA \rightarrow Cit + CoA + H^+$

$$v_{CIS} = \frac{V_{\max} \cdot [OAA] \cdot [ACoA] \cdot \left(1 - \frac{\frac{[CoA] \cdot [Cit]}{[OAA] \cdot [ACoA]}}{K_{eq}}\right)}{K_A^I \cdot K_B^M \cdot \left(1 + \frac{[Cit]}{K_{Cit}^I}\right) + K_A^M \cdot [ACoA] \cdot \left(1 + \frac{[Cit]}{K_{Cit}^I}\right) + [OAA] \cdot [ACoA] + \ldots}$$
$$\ldots [OAA] \cdot K_B^M \cdot \left(1 + \frac{[ATP_{TOT}]}{K_{ATP}^I} + \frac{[ADP_{TOT}]}{K_{ADP}^I} + \frac{[AMP_{TOT}]}{K_{AMP}^I} + \frac{[CoA]}{K_{CoA}^I} + \frac{[SCA]}{K_{SCA}^I}\right)$$

$V_{\max}$=2.69E+00; K_{eq}=1.00E+05; K_{ATP}^I=9.00E-01; K_{ADP}^I=1.80E+00; K_{AMP}^I=6.00E+00;
K_{Cit}^I=1.60E+00; K_{CoA}^I=6.75E-01; K_{SCA}^I=1.40E-01; K_A^I=3.33E-03; K_A^M=4.00E-03; K_B^I=7.00E-03;

Aconitase (ACO) – $Cit \rightarrow ICit$

$$v_{ACO} = \frac{V_{\max} \cdot \frac{[Cit]}{K_{Cit}^M} \cdot \left(1 - \frac{[ICit]/[Cit]}{K_{eq}}\right)}{1 + \frac{[Cit]}{K_{Cit}^M} + \frac{[ICit]}{K_{ICit}^M}}$$

$V_{\max}$=5.56E+01; K_{eq}=7.14E-02; K_{Cit}^I=1.10E+01; K_{ICit}^I=2.96E-01;

Isocitrate dehydrogenase – $ICit + NAD = AKG + NADH + CO_2 + H^+$

$$v_{IDH} = \frac{V_{\max} \cdot \left([NAD] \cdot [ICit]^3 - \frac{[ICit]^2 \cdot [AKG] \cdot [CO_2] \cdot [NADH]}{K_{eq}}\right)}{\left([NAD] + K_A^M\right) \cdot [ICit]^3 + \left([NAD] \cdot \left(K_B^M + K_A^M \cdot K_B^I\right) + K_A^M \cdot \frac{[NADH] \cdot [ICit]^3}{K_Q^I}\right) \cdot \ldots}$$
$$\ldots \left(1 + \frac{K_{ADP}^A}{[ADP_{TOT}]}\right) \cdot \left(1 + \frac{[ATP_{TOT}]}{K_{ATP}^I}\right)$$

$V_{\max}$=1.26E+03; K_{eq}=1.004E+00; K_A^M=7.40E-02; K_B^M=5.90E-02; K_{ATP}^I=9.10E-02;
K_B^I=7.66E-03; K_Q^I=2.90E-02;

α-ketoglutarate dehydrogenase (AKD) – $AKG + NAD + CoA \rightarrow SCA + NADH + CO_2$

$$v_{AKD} = \frac{V_{\max} \cdot \left(1 - \frac{[CO_2] \cdot [SCA] \cdot [NADH]}{[AKG] \cdot [CoA] \cdot [NAD]} \cdot \frac{1}{K_{eq}}\right)}{1 + \frac{K_A^M}{[AKG]} \cdot \left(1 + \frac{K_{ADP}^A}{[ADP_{TOT}]}\right) \cdot \left(1 + \frac{[ATP_{TOT}]}{K_{ADP}^I}\right) + \frac{K_B^M}{[CoA]} \cdot \left(1 + \frac{[SCA]}{K_Q^I}\right) + \frac{K_C^M}{[NAD]} \cdot \left(1 + \frac{[NADH]}{K_R^I}\right)}$$

$V_{\max} = 7.73\text{E+01};\ K_{eq} = 1.00\text{E+05};\ K_{ADP}^{A} = 1.00\text{E-01};\ K_{ATP}^{I} = 5.00\text{E-02};\ K_{A}^{M} = 8.00\text{E-02};$
$K_{B}^{M} = 5.50\text{E-02};\ K_{C}^{M} = 2.10\text{E-02};\ K_{Q}^{I} = 6.90\text{E-03};\ K_{R}^{I} = 4.50\text{E-03};$

Succinyl-CoA synthetase (SCAS) – $SCA + ADP + Phi \rightarrow Succ + ATP$

$$v_{SCAS} = \frac{V_{\max} \cdot \frac{[SCA] \cdot [ADP_{TOT}] \cdot [Phi]}{K_{SCA}^{M} \cdot K_{ADP}^{M} \cdot K_{Phi}^{M}} \cdot \left(1 - \frac{\frac{[Succ] \cdot [ATP_{TOT}] \cdot [CoA]}{[SCA] \cdot [ADP_{TOT}] \cdot [Phi]}}{K_{eq}} \right)}{\left(1 + \frac{[SCA]}{K_{SCA}^{M}} \right) \cdot \left(1 + \frac{[ADP_{TOT}]}{K_{ADP}^{M}} \right) \cdot \left(1 + \frac{[Phi]}{K_{Phi}^{M}} \right) + \left(1 + \frac{[Succ]}{K_{Succ}^{M}} \right) \cdot \left(1 + \frac{[ATP_{TOT}]}{K_{ATP}^{M}} \right) \cdot \ldots \\ \ldots \left(1 + \frac{[CoA]}{K_{CoA}^{M}} \right) - 1}$$

$V_{\max} = 7.10\text{E-01};\ K_{eq} = 2.00\text{E+00};\ K_{ADP}^{M} = 2.50\text{E-01};\ K_{SCA}^{M} = 4.10\text{E-02};\ K_{Phi}^{M} = 1.40\text{E+00};$
$K_{Succ}^{M} = 1.00\text{E+00};\ K_{ATP}^{M} = 1.60\text{E+00};\ K_{CoA}^{I} = 1.00\text{E-02};$

Succinate dehydrogenase (SDH) – $Succ + FAD \rightarrow Fum + FADH_2$

$$v_{SDH} = \frac{V_{\max} \cdot \frac{[Succ] \cdot [FAD]}{K_{Succ}^{E} \cdot K_{FAD}^{M}} \cdot \left(1 - \frac{\frac{[Fum] \cdot [FADH_2]}{[Succ] \cdot [FAD]}}{K_{eq}} \right)}{1 + \frac{[Succ]}{K_{Succ}^{E}} + \frac{[FAD]}{K_{FAD}^{M}} \cdot \frac{K_{Succ}^{M}}{K_{Succ}^{E}} + \frac{[Succ]}{K_{Succ}^{E}} \cdot \frac{[FAD]}{K_{FAD}^{M}} + \frac{[Fum]}{K_{Fum}^{E}} + \frac{[FADH_2]}{K_{FADH_2}^{M}} \cdot \frac{K_{Fum}^{M}}{K_{Fum}^{E}} + \frac{[Fum]}{K_{Fum}^{E}} \cdot \frac{[FADH_2]}{K_{FADH_2}^{M}}}$$

$V_{\max} = 3.30\text{E-01};\ K_{eq} = 1.00\text{E+05};\ K_{Succ}^{M} = 1.30\text{E-01};\ K_{FAD}^{M} = 3.00\text{E-04};\ K_{Fum}^{M} = 2.50\text{E-02};$
$K_{FADH_2}^{M} = 1.50\text{E-03};\ K_{Succ}^{E} = 1.00\text{E-02};\ K_{Fum}^{E} = 2.90\text{E-01};$

Fumarase (FUM) – $Fum \rightarrow Mal$

$$v_{FUM} = \frac{V_{\max} \cdot \frac{[Fum]}{K_{Fum}^{M}} \cdot \left(1 - \frac{[Mal]/[Fum]}{K_{eq}} \right)}{1 + \frac{[Fum]}{K_{Fum}^{M}} + \frac{[Mal]}{K_{Mal}^{M}}}$$

$V_{\max} = 3.58\text{E+00};\ K_{eq} = 5.00\text{E+00};\ K_{Fum}^{M} = 4.70\text{E-02};\ K_{Mal}^{M} = 1.70\text{E-02};$

Malate dehydrogenase (MDH) – $Mal + NAD \rightarrow OAA + NADH + H^{+}$

$$v_{MDH} = \frac{V_{\max} \cdot \frac{[Mal] \cdot [NAD]}{K_{Mal}^{M} \cdot K_{NAD}^{M}} \cdot \left(1 - \frac{\frac{[OAA] \cdot [NADH]}{[Mal] \cdot [NAD]}}{K_{eq}} \right)}{\left(1 + \frac{[Mal]}{K_{Mal}^{M}} \right) \cdot \left(1 + \frac{[NAD]}{K_{NAD}^{M}} \right) + \left(1 + \frac{[OAA]}{K_{OAA}^{M}} \right) \cdot \left(1 + \frac{[NADH]}{K_{NADH}^{M}} \right) - 1}$$

$V_{\max} = 8.03\text{E-}01$; $K_{eq} = 6.26\text{E+}03$; $K_{Mal}^{M} = 5.50\text{E-}01$; $K_{NAD}^{M} = 1.00\text{E-}01$; $K_{OAA}^{M} = 1.00\text{E-}02$; $K_{NADH}^{M} = 1.00\text{E-}02$;

Electron transport and oxidative phosphorylation

Complex I+III+IV – $NADH + H^{+} + 0.5\ O_2 \rightarrow NAD + H_2O + 10\Delta H^{+}$

$$v_{I,III,IV} = \frac{V_{MAX}^{f} \cdot \exp\left(-\frac{10\Delta G_{H^+}}{RT}\right) \cdot \frac{[\text{NADH}][\text{O}_2]^{0.5}[\text{H}^+]}{K^f} \cdot \left(1 - \frac{\frac{[NAD]}{[\text{NADH}][\text{O}_2]^{0.5}[\text{H}^+]} \cdot \exp\left(\frac{10\Delta G_{H^+}}{RT}\right)}{K_{eq}}\right)}{1 + \frac{[\text{NADH}][\text{O}_2]^{0.5}[\text{H}^+]}{K^f} + \frac{[\text{NAD}]}{K^b}}$$

$V_{\max}$=4.72E+04; K_{eq}=1.00E+12; K_f=2.05E-06; K_b=3.15E+00;

Complex II + III + IV – $FADH_2 + 0.5\ O_2 \rightarrow FAD + H_2O + 6\Delta H^{+}$

$$v_{II+III+IV} = \frac{V_{\max} \cdot \exp\left(-\frac{6\Delta G_{H^+}}{RT}\right) \cdot \frac{[\text{FADH}_2][\text{O}_2]^{0.5}}{K^f} \cdot \left(1 - \frac{\frac{[FAD]}{[\text{FADH}_2][\text{O}_2]^{0.5}} \cdot \exp\left(\frac{6\Delta G_{H^+}}{RT}\right)}{K_{eq}}\right)}{1 + \frac{[\text{FADH}_2][\text{O}_2]^{0.5}}{K^f} + \frac{[\text{FAD}]}{K^b}}$$

$V_{\max}$=2.89E+02; K_{eq}=1.00E+12; K_f=3.94E-02; K_b=2.12E+00;

Complex V – $ADP + Phi + H^{+} + 3\Delta H^{+} \rightarrow ATP$

$$v_V = \frac{V_{\max} \cdot \exp\left(+\frac{3\Delta G_{H^+}}{RT}\right) \cdot \frac{[\text{ADP}][\text{Pi}][\text{H}^+]}{K^f} \cdot \left(1 - \frac{\frac{[ATP]}{[\text{ADP}][\text{Pi}][\text{H}^+]} \cdot \exp\left(-\frac{3\Delta G_{H^+}}{RT}\right)}{K_{eq}}\right)}{1 + \frac{[\text{ADP}][\text{Pi}][\text{H}^+]}{\text{K}^{\text{f}}} + \frac{[\text{ATP}]}{\text{K}^{\text{b}}}}$$

$V_{\max}$=4.05E-01; K_{eq}=1.00E+12; K_f=2.71E-04; K_b=8.68E+00;

Exchange and translocation reactions

Glucose uptake (GLT) – $GLCo \rightarrow GLC(c)$

$$v_{GLT} = \frac{V_{\max} \cdot \frac{[GLCo]}{K^{M}_{GLCo}} \cdot \left(1 - \frac{[GLC]/[GLCo]}{K_{eq}}\right)}{1 + \frac{[GLCo]}{K^{M}_{GLCo}} + \frac{[GLC]}{K^{M}_{GLC}} + \frac{\alpha_0 \cdot [GLCo] \cdot [GLC]}{K^{M}_{GLCo} \cdot K^{M}_{GLC}}}$$

$V_{\max}$=4.32E+00; K_{eq}=1.00E+00; K^{M}_{GLCo}=1.70E+00; K^{M}_{GLC}=6.90E+00; α_0=5.40E-01;

Lactate exchange (LCT) – $LAC(c) \rightarrow LACo$

$$v_{LCT} = V_{\max} \cdot \left(LAC - \frac{LACo}{K_{eq}}\right)$$

$V_{\max}$=2.95E+02; K_{eq}=1.00E+00;

Pyruvate Translocation (PYRT) – $PYR(c) \rightarrow PYR(m)$

$$v_{PYRT} = T_{\max} \cdot \left(\frac{[PYR(c)]}{[PYR(c)] + K^{M}_{PYR(c)\rightarrow(m)}} - \frac{[PYR(m)]}{[PYR(m)] + K^{M}_{PYR(c)\rightarrow(m)}}\right)$$

$T_{\max}$=9.92E-01; $K^{M}_{PYR(c)\rightarrow(m)}$=5.28E-02;

NADH/NAD translocation – $NADH(c) + NAD(m) \rightarrow NADH(m) + NAD(c)$

$$v_{NAD(H)\ Transp} = T_{MAX} \cdot \left(\frac{\left([\mathrm{NADH}]/[NAD]\right)_{\mathrm{cyt}}}{K^{\mathrm{cyt}}_{\mathrm{cyt}\rightarrow\mathrm{mit,PS}} + \left([\mathrm{NADH}]/[NAD]\right)_{\mathrm{cyt}}} - \frac{\left([\mathrm{NADH}]/[NAD]\right)_{\mathrm{mit}}}{K^{\mathrm{mit}}_{\mathrm{cyt}\rightarrow\mathrm{mit,PS}} + \left([\mathrm{NADH}]/[NAD]\right)_{\mathrm{mit}}}\right)$$

$T_{\max}$=3.87E-01; $K^{cyt}_{cyt\rightarrow mit}$=6.28E-04; $K^{mit}_{cyt\rightarrow mit}$=1.58E+01;

ATP/ADP translocation – $ATP_{TOT}(m) + ADP_{TOT}(c) \rightarrow ATP_{TOT}(c) + ADP_{TOT}(m)$

$$v_{AXP\ Transp} = T_{MAX} \cdot \left(\frac{\left([\mathrm{ATP}_{TOT}]/[ADP_{TOT}]\right)_{\mathrm{mit}}}{K^{\mathrm{mit}}_{\mathrm{mit}\rightarrow cyt} + \left([\mathrm{ATP}_{TOT}]/[ADP_{TOT}]\right)_{\mathrm{mit}}} - \frac{\left([\mathrm{ATP}_{TOT}]/[ADP_{TOT}]\right)_{\mathrm{cyt}}}{K^{\mathrm{cyt}}_{\mathrm{mit}\rightarrow cyt} + \left([\mathrm{ATP}_{TOT}]/[ADP_{TOT}]\right)_{\mathrm{cyt}}}\right)$$

$T_{\max}$=6.66E+00; $K^{mit}_{mit\rightarrow cyt}$=5.54E-01; $K^{cyt}_{mit\rightarrow cyt}$=6.66E+02;

Intermembrane protons balancing (Protons leaking) – $H^{+}(im) \rightarrow H^{+}(m)$

$$v_{H,im\rightarrow m} = \lambda \cdot \left([H^{+}(im)] - [H^{+}(m)]\right)$$

λ =1.56E+00;

Mitochondrial protons balancing – $H^{+}(c) \rightarrow H^{+}(m)$

$$v_{H^+,\ cyt\to mit} = T_{MAX}\cdot\left(\frac{[H^+]_{cyt}}{K^M_{cyt\to mit}+[H^+]_{cyt}} - \frac{[H^+]_{mit}}{K^M_{cyt\to mit}+[H^+]_{mit}}\right)$$

$T_{\max}$=4.53E+01; $K^M_{cyt\to mit}$=2.75E-05;

Oxygen exchange between cytosol and mitochondria – $O_2(c)\to O_2(m)$

$$v_{O_2,\ c\to m} = \lambda\cdot\left([O_2(c)]-[O_2(m)]\right)$$

λ=1.29E+02;

Oxygen exchange between cytosol and extra-cellular environment –

$O_2(ex)\to O_2(c)$

$$v_{O_2,\ ex\to c} = \lambda\cdot\left([O_2(ex)]-[O_2(c)]\right)$$

λ=8.67E+00;

Phosphate exchange between cytosol and mitochondria – $Phi(c)\to Phi(m)$

$$v_{Phi,\ c\to m} = \lambda\cdot\left([Phi(c)]-[Phi(m)]\right)$$

λ=2.87E+00;

Phosphate exchange between cytosol and extra-cellular environment –

$Phi(ex)\to Phi(c)$

$$v_{Phi,\ ex\to c} = \lambda\cdot\left([Phi(ex)]-[Phi(c)]\right)$$

λ=1.29E+01;

Carbon Dioxide exchange – $CO_2(m)\to CO_2(ex)$

$$v_{CO_2,\ m\to ex} = \lambda\cdot\left([CO_2(m)]-[CO_2(ex)]\right)$$

λ=1.00E+00;

NORMAL METABOLIC PHENOTYPE

Reaction	Flux	Reaction	Flux
HXK	1.95E-01	PDH	2.90E-01
GPI	1.83E-01	CIS	2.90E-01
PFK	1.89E-01	ACO	2.90E-01
ALD	1.89E-01	IDH	2.90E-01
TPI	1.89E-01	AKD	2.90E-01
GAPDH	3.80E-01	SCAS	2.90E-01
PGK	3.17E-01	SDH	2.90E-01
BPGM	6.36E-02	FUM	2.90E-01
BPGP	6.36E-02	MDH	2.90E-01
PGM	3.80E-01	Complex I,III,IV	1.46E+00
ENO	3.80E-01	Complex II,III,IV	2.90E-01
PK	3.80E-01	Complex V	4.10E+00
LDH	7.71E-02	GLT	1.95E-01
LDHP	1.29E-02	LCT	9.00E-02
G6PDH	1.25E-02	PYR translocation	2.90E-01
PGLDH	1.25E-02	NADH/NAD transloc.	3.03E-01
GSSGRD	1.20E-02	ATP/ADP transloc.	4.39E+00
GSHox	1.20E-02	Phi exchange (ext→cyt)	9.99E-03
EP	6.08E-03	Phi transolc. (cyt→mit)	4.39E+00
KI	6.37E-03	O_2 diffusion (cyt→mit)	8.77E-01
PPRPPS	3.33E-03	O_2 diffusion (ext→cyt)	8.77E-01
TK1	3.04E-03	H+ diffusion (cyt→mit)	4.40E+00
TA	3.04E-03	H^+ exchange (cyt→ext)	6.00E-01
TK2	3.04E-03	H^+ leaking (im→mit)	4.10E+00
AK	3.33E-03	CO_2 diffus. (mit→ext)	8.71E-01
ATPase	4.69E+00		

Table S.1 – List of metabolic fluxes in the normal metabolic phenotype

Metabolite	Conc. [Ref.]	Metabolite	Conc. [Ref.]
GLCo	5.00E+00 [25]	Suc	9.50E-01 [25]
GLC	4.57E+00 [39]	Fum	9.50E-01
G6P	3.95E-02 [39]	Mal	9.50E-01 [25]
F6P	1.53E-02 [39]	OAA	2.40E-03 [25]
F16P2	9.68E-03 [39]	FAD	2.12E+00 [25]
GraP	6.06E-03 [39]	$FADH_2$	2.40E-01 [25]
DHAP	1.49E-01 [39]	NADPH	4.89E-03 [39]
13P2G	4.81E-04 [39]	NADP	2.34E-05 [39]
23P2G	2.75E+00 [39]	NAD(c)	6.53E-02 [39]
3PG	6.58E-02 [39]	NADH(c)	1.56E-04 [39]
2PG	8.44E-03 [39]	NAD(m)	3.15E+00 [25]
PEP	1.09E-02 [39]	NADH(m)	4.98E-01 [25]
PYR(c)	8.40E-02 [39]	ATP(c)*	1.60E+00 [39]
LAC	1.68E+00 [39]	ADP(c)*	3.24E-01 [39]
LACo	1.68E+00 [39]	AMP(c)*	7.47E-02 [39]
6PG	2.51E-02 [39]	ATP(m)	8.68E+00 [25]
Ru5P	4.72E-03 [39]	ADP(m)	7.84E+00 [25]
GSSG	1.87E-04 [39]	AMP(m)	0.00E+00 [25]
GSH	3.11E+00 [39]	Phi(c)	2.90E+00 [25]
X5P	1.27E-02 [39]	Phi(m)	1.38E+00 [25]
R5P	1.40E-02 [39]	Phi out	2.90E+00 [25]
S7P	1.54E-02 [39]	O_2(c)	3.38E-02 [25]
E4P	6.27E-03 [39]	O_2(m)	2.70E-02 [25]
PRPP	1.00E+00 [39]	O_2 out	1.35E-01 [25]
PYR(m)	2.50E-02 [25]	H^+(c)	7.94E-05 [25]
ACoA	2.00E-02 [25]	H^+(m)	5.01E-05 [25]
CoA	4.00E-02 [25]	H^+(im)	= H^+(m)
Cit	9.50E-01 [25]	H^+ out	4.79E-05 [25]
ICit	6.70E-02	CO_2(m)	1.53E+00 [25]
AKG	1.25E-01 [25]	CO_2 out	1.22E+00 [25]
SCA	1.25E+00 [25]		

Table S.2 –List of metabolic fluxes in the normal metabolic phenotype

* These concentrations refer to the total amount of ATP, ADP and AMP in cytoplasm. The free and Mg-bound portions of each of these cofactors were obtained assuming that the association/dissociation reactions involving Mg and AXP are much faster than the other ones (and hence carrying a null flux at steady-state). The concentration of free Mg was set to 5.56E-01 [39] and was treated a parameter rather than a system variable.

CANCER METABOLIC PHENOTYPE

Reaction	Flux	Reaction	Flux
HXK	1.70E+00	PDH	3.85E-01
GPI	1.20E+00	CIS	3.85E-01
PFK	1.51E+00	ACO	3.85E-01
ALD	1.51E+00	IDH	3.85E-01
TPI	1.51E+00	AKD	3.85E-01
GAPDH	3.18E+00	SCAS	3.85E-01
PGK	2.76E+00	SDH	3.85E-01
BPGM	4.24E-01	FUM	3.85E-01
BPGP	4.24E-01	MDH	3.85E-01
PGM	3.18E+00	Complex I,III,IV	2.82E+00
ENO	3.18E+00	Complex II,III,IV	3.85E-01
PK	3.18E+00	Complex V	7.63E+00
LDH	1.90E+00	GLT	1.70E+00
LDHP	8.95E-01	LCT	2.80E+00
G6PDH	5.00E-01	PYR translocation	3.85E-01
PGLDH	5.00E-01	NADH/NAD transloc.	1.28E+00
GSSGRD	1.05E-01	ATP/ADP transloc.	8.01E+00
GSHox	1.05E-01	Phi exchange (ext→cyt)	8.71E-02
EP	3.14E-01	Phi transolc. (cyt→mit)	8.01E+00
KI	1.86E-01	O_2 diffusion (cyt→mit)	1.60E+00
PPRPPS	2.90E-02	O_2 diffusion (ext→cyt)	1.60E+00
TK1	1.57E-01	H+ diffusion (cyt→mit)	8.91E+00
TA	1.57E-01	H^+ exchange (cyt→ext)	3.09E+00
TK2	1.57E-01	H^+ leaking (im→mit)	7.63E+00
AK	2.90E-02	CO_2 diffus. (mit→ext)	1.15E+00
ATPase	1.07E+01		

Table S.3 – List of metabolic fluxes in the cancer metabolic phenotype

SAMPLING METABOLITE CONCENTRATIONS

To infer the control properties of the system under study, we used a random sampling approach where the unknown or uncertain quantities are sampled and the control coefficient evaluated subsequently (see *Methods* in the main text). Some of the metabolite concentrations defining the cancer metabolic state were among these quantities. In sampling the metabolite concentrations one has to consider the thermodynamic constraints they have to satisfy. For example, for any reaction governed by a kinetics of the form of Eq.(S.1), the following constraint can be derived

$$\frac{\prod_r A_r^{\alpha_r}}{\prod_p B_p^{\beta_p}} \lessgtr K_{eq} \tag{S.11}$$

where the direction of the inequality depends on the sign of the flux (the "less than" sign corresponds to a positive flux, while the "greater than" sign correspond to a negative flux). Similar constraints can also be derived for reactions governed by the other kinetics laws (see previous section). By taking their logarithmic form, most of these constraints can be rewritten as linear inequalities. Eq.(S.11), for example, becomes:

$$\sum_r \alpha_r \ln(A_r) - \sum_p \beta_p \ln(B_p) \lessgtr K_{eq} \tag{S.12}$$

The only exceptions are represented by Eqs. (S6)-(S7)-(S8), where the exponential factor used to account for the mitochondrial membrane potential does not allow to linearize their corresponding thermodynamic constraints. This problem was circumvented by replacing the exponential factors with appropriate values, such that the sampled concentrations were compliant with the original thermodynamic constraint derived from these kinetic laws. For

example, the constraints imposed by Eq.(S.7), assuming the flux of the corresponding reaction being positive, is:

$$\exp\left(-\frac{6\Delta G_{H^+}}{RT}\right)\cdot V_{MAX}^{f}\frac{[\mathrm{FADH_2}][\mathrm{O_2}]^{0.5}}{K^{f}}-V_{MAX}^{b}\frac{[\mathrm{FAD}]}{K^{b}}>0 \qquad \text{(S.13)}$$

and can be rewritten as

$$\frac{[\mathrm{FADH_2}][\mathrm{O_2}]^{0.5}}{[\mathrm{FAD}]}>\frac{V_{MAX}^{b}}{V_{MAX}^{f}}\cdot\frac{K^{f}}{K^{b}}\cdot\exp\left(\frac{6\Delta G_{H^+}}{RT}\right) \qquad \text{(S.14)}$$

By replacing $\exp\left(\frac{6\Delta G_{H^+}}{RT}\right)$ with the maximum value (MV) that this factor can assume based on the sampling interval of H^{+}_{mit} concentrations, the constraint expressed through Eq.(S.14) can be rewritten as

$$\frac{[\mathrm{FADH_2}][\mathrm{O_2}]^{0.5}}{[\mathrm{FAD}]}>\frac{V_{MAX}^{b}}{V_{MAX}^{f}}\cdot\frac{K^{f}}{K^{b}}\cdot\mathrm{MV} \qquad \text{(S.15)}$$

We note that Eq.(S.15) represents a more strict constraint with respect to Eq.(S.14). The advantage of using Eq.(S.15), however, is that it can be linearized by taking its logarithmic form:

$$\ln[\mathrm{FADH_2}]+0.5\cdot\ln[\mathrm{O_2}]-\ln[\mathrm{FAD}]>\ln\left(\frac{V_{MAX}^{b}}{V_{MAX}^{f}}\cdot\frac{K^{f}}{K^{b}}\cdot\mathrm{MV}\right) \qquad \text{(S.16)}$$

The same procedure was applied to linearize the constraints derivable from Eqs.(S.6)-(S.8).

The entire set of linear inequalities derived from the *n* different reaction steps in the system can be expressed as follows

$$\sum_{i=1}^{m} \lambda_{ij} X_i \lessgtr b_j \quad \text{with } j = 1,2,...,n \tag{S.17}$$

where X_i is the logarithmic concentration of metabolite S_i and λ_{ij} is the stoichiometric coefficient of S_i in reaction *j*. The direction of the inequality depends on the sign of the flux of reaction *j* and the specific rate law from which the constraint is derived.

Eq.(S.17) define a convex hull in the *m* dimensional space of the logarithmic metabolite concentrations. Fig.S.1 shows an exemplifying representation of a convex hull in a three dimensional space, corresponding to a scenario where only three metabolite concentrations are considered (one for each dimension).

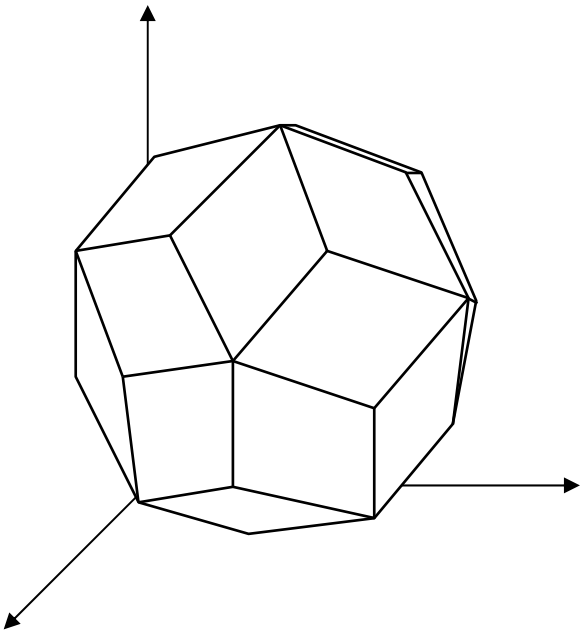

Figure S.1 – Schematic representation of a convex hull in three dimensions*. Each dimension represents the logarithmic concentration of a metabolite.*

For the metabolite concentrations to be thermodynamically meaningful, their logarithmic values have to satisfy simultaneously all the *n* linear constraints of Eq. (S.17), *i.e.* they need to be sampled within the convex hull defined by this constraints. In order to do so, we used the known property according to which, given a set of solutions $\left\{\mathbf{X}^{(1)},\mathbf{X}^{(2)},...,\mathbf{X}^{(K)}\right\}$ of Eq.(S.17), any linear combination of the form

$$\frac{\sum_{k=1}^{K}\varphi_k\mathbf{X}^{(k)}}{\sum_{k=1}^{K}\varphi_k} \tag{S.18}$$

is also a solution of the same set of inequalities. This means that, once an representative initial set of solutions is found, a thermodynamically compliant way to sample the metabolite concentrations consists of combining these solutions linearly with random coefficients φ_k. To find a first representative set of solutions, we used a linear programming approach. In particular we applied the algorithm proposed by Lee et al. [38] to find the corner points of the convex hull. These corner points were then used as the initial set of solution from which to sample thermodynamically compliant sets of metabolite concentrations through Eq.(S18).

REFERENCES

1. Warburg, O., F. Wind, and E. Negelein, *The Metabolism of Tumors in the Body.* Journal of General Physiology, 1927. **8**(6): p. 519-530.
2. Gottlieb, E., *The fat and the furious.* NATURE, 2009. **461**: p. 44-45.
3. Bowden, A.C., *Metabolic control analysis in biotechnology and medicine.* Nature Biotechnology, 1999. **17**(7): p. 641-643.
4. Kitano, H., et al., *Metabolic Syndrome and Robustness Tradeoffs* Diabetes, 2004. **53**(3): p. S6-S15.
5. Radisavljevic, Z., *Locus of fragility in robust breast cancer system.* Journal of Cellular Biochemistry, 2004. **92**(5): p. 1020-1024.
6. Chu, L.-H. and B.-S. Chen, *Comparisons of Robustness and Sensitivity between Cancer and Normal Cells by Microarray Data.* Cancer Informatics, 2008. **6**: p. 165-181.
7. Hornberg, J.J., et al., *Metabolic control analysis to identify optimal drug targets*, in *Systems Biological Approaches in Infectious Diseases*. 2007, Birkhäuser Basel. p. 171-189.
8. Cascante, M., et al., *Metabolic control analysis in drug discovery and disease.* Nature Biotechnology, 2002. **20**: p. 243-249.
9. Cascante, M., et al., *The evolution of molecular biology into systems biology.* Nature Biotechnology, 2002. **20**: p. 243-249.
10. Bakker, B.M., et al., *Metabolic control analysis of glycolysis in trypanosomes as an approach to improve selectivity and effectiveness of drugs.* Molecular and Biochemical Parasitology, 2000. **106**: p. 1-10.
11. Schellenberger, J. and B.Ø. Palsson, *The use of randomized sampling for analysis of metabolic networks.* Journal of Biological Chemistry, 2008. **284**(9): p. 5457–5461.
12. Wang, L. and V. Hatzimanikatis, *Metabolic engineering under uncertainty. I: Framework development.* Metabolic Engineering, 2006. **8**: p. 133-141.
13. Wang, L. and V. Hatzimanikatis, *Metabolic engineering under uncertainty—II: Analysis of yeast metabolism.* Metabolic Engineering, 2006. **8**: p. 142-159.
14. Tran, L.M., M.L. Rizk, and J.C. Liao, *Ensemble Modeling of Metabolic Networks.* Biophysical Journal, 2008. **95**(12): p. 5606–5617.
15. Steuer, R., *Computational approaches to the topology, stability and dynamics of metabolic networks.* Phytochemistry, 2007. **68**(16-18): p. 2139-2151.

16. Steuer, R., et al., *Structural kinetic modeling of metabolic networks.* Proceedings of the National Academy of Sciences, 2006. **103**(32): p. 11868–11873.

17. Grimbs, S., et al., *The stability and robustness of metabolic states: identifying stabilizing sites in metabolic networks.* Molecular Systems Biology, 2007. **3**(146).

18. Murabito, E., et al., *A probabilistic approach to identify putative drug targets in biochemical networks.* Submitted to the Journal of the Royal Society Interface, 2010.

19. Richardson, A.D., et al., *Central carbon metabolism in the progression of mammary carcinoma.* Breast Cancer Research and Treatment, 2008. **110**: p. 297–307.

20. Cohen, J.S., et al., *Differences in Phosphate Metabolite Levels in Drug-sensitive and -resistant Human Breast Cancer Cell Lines Determined by 31P Magnetic Resonance Spectroscopy.* Cancer Research, 1986. **46**: p. 4087-4090.

21. Heinrich, R. and T.A. Rapoport, *A Linear Steady-State Treatment of Enzymatic Chains. General Properties, Control and Effector Strength.* European Journal of Biochemistry, 1974. **42**(1): p. 89-95.

22. Fell, D.A., *Metabolic Control Analysis: a survey of its theoretical and experimental development.* Biochem. J., 1992. **286**: p. 313-330.

23. Reder, C., *Metabolic control theory: a structural approach.* Journal of Theoretical Biology, 1988. **135**(2): p. 175-201.

24. Schuster, R. and H.G. Holzhütter, *Use of mathematical models for predicting the metabolic effect of large-scale enzyme activity alterations Application to enzyme deficiencies of red blood cells.* European journal of biochemistry / FEBS, 1995. **229**(2): p. 403-418.

25. Li, Y., et al., *Role of NADH/NAD+ transport activity and glycogen store on skeletal muscle energy metabolism during exercise: in silico studies.* American Journal of Physiology. Cell Physiology, 2009. **296**: p. C25-C46.

26. Wu, F., et al., *Computer Modeling of Mitochondrial TCA Cycle, Oxidative Phosphorylation, Metabolite Transport, and Electrophysiology.* The Journal of Biological Chemistry, 2007. **282**(34): p. 24525–24537.

27. Mogilevskaya, E., O. Demin, and I. Goryanin, *Kinetic Model of Mitochondrial Krebs Cycle: Unraveling the Mechanism of Salicylate Hepatotoxic Effects.* Journal of Biological Physics, 2006. **32**(3-4): p. 245–271.

28. Liebermeister, W. and E. Klipp, *Bringing metabolic networks to life: convenience rate law and thermodynamic constraints.* Theoretical Biology and Medical Modelling, 2006. **3**(41).

29. deGraaf, A.A., et al., *Lactate Imaging of the Human Brain at 1.5 T Using a Double-Quantum Filter.* Magnetic Resonance in Medicine, 1993. **30**(2): p. 231-235.

30. He, Q., et al., *Single-scan in vivo lactate editing with complete lipid and water suppression by selective multiple-quantum-coherence transfer (Sel-MQC) with application to tumors.* Journal of Magnetic Resonance. Series B, 1995. **106**(3): p. 203-2011.

31. Hurd, R.E. and D.M. Freeman, *Metabolite specific proton magnetic resonance imaging.* Proceedings of the National Academy of Sciences, 1989. **86**(12): p. 4402-4406.

32. Schupp, D.G., et al., *Localized detection of glioma glycolysis using edited 1H MRS.* Magnetic Resonance in Medicine, 2005. **30**(1): p. 18-27.

33. Griffiths, J.R., *Are cancer cells acidic?* British journal of cancer, 1991. **64**(3): p. 425-427.

34. Gillies, R.J., Z. Liu, and Z. Bhujwalla, *31P-MRS measurements of extracellular pH of tumors using 3-aminopropylphosphonate.* The American Journal of Physiology, 1994. **267**: p. C195-C203.

35. Martin, G.R. and R.K. Jam, *Noninvasive Measurement of Interstitial pH Profiles in Normal and Neoplastic Tissue Using Fluorescence Ratio Imaging Microscopy.* Cancer Research, 1994. **54**: p. 5670-5674.

36. Stubbs, M., et al., *Metabolic Consequences of a Reversed pH Gradient in Rat Tumors.* Cancer Research, 1994. **54**: p. 4011-4016.

37. Wright, A.J., et al., *Ex-vivo HRMAS of adult brain tumours: metabolite quantification and assignment of tumour biomarkers.* Molecular Cancer, 2010. **9**(66).

38. Lee, S., et al., *Recursive MILP model for finding all the alternate optima in LP models for metabolic networks.* Computers & Chemical Engineering, 2000. **24**(2-7): p. 711-716.

39. Holzhütter, H.-G., *The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks.* European Journal of Biochemistry, 2004. **271**: p. 2905–2922.

40. Medina, R.A. and G.I. Owen, *Glucose transporters: expression, regulation and cancer.* Biological Research, 2002. **35**(1): p. 9-26.

41. Harmon, A.W. and Y.M. Patel, *Naringenin inhibits glucose uptake in MCF-7 breast cancer cells: a mechanism for impaired cellular proliferation.* Breast Cancer Research and Treatment, 2004. **85**(2): p. 103-110.

42. Horecker, B.L., G. Domagk, and H.H. Hiatt, *A comparison of C14-labeling patterns in deoxyribose and ribose in mammalian cells.* Archives of Biochemistry and Biophysics, Vol. 78, No. 2. (December 1958), pp. 510-517., 1958. **78**(2): p. 510-517.

43. Macallan, D.C., et al., *Measurement of cell proliferation by labeling of DNA with stable isotope-labeled glucose: Studies in vitro, in animals, and in humans.* Proceedings of the National Academy of Sciences, 1998. **95**(2): p. 708-713.

44. Boren, J., et al., *Metabolic Control Analysis aimed at the ribose synthesis pathways of tumor cells: a new strategy for antitumor drug development.* Molecular biology Reports, 2002. **29**(7-12).

45. Gatenby, R.A. and E.T. Gawlinski, *A Reaction-Diffusion Model of Cancer Invasion.* Cancer Research, 1996. **56**(24): p. 5745-5753.

46. Annibaldi, A. and C. Widmann, *Glucose metabolism in cancer cells.* Current Opinion in Clinical Nutrition & Metabolic Care, 2010. **13**(4): p. 466-470.

47. Shime, H., et al., *Tumor-Secreted Lactic Acid Promotes IL-23/IL-17 Proinflammatory Pathway.* The Journal of Immunology, 2008. **180**: p. 7175-7183.

48. Belt, J.A., et al., *Inhibition of Lactate Transport and Glycolysis in Ehrlich Ascites Tumor Cells by Bioflavonoidst.* American Chemical Society, 1979. **18**(16): p. 3506-3511.

49. Kim, J.H., et al., *Quercetin, an Inhibitor of Lactate Transport and a Hyperthermic Sensitizer of HeLa Cells.* Cancer Research, 1984. **44**: p. 102-106.

50. Mathupala, S.P., et al., *Lactate and malignant tumors: A therapeutic target at the end stage of glycolysis.* Journal of Bioenergetics and Biomembranes, 2007. **39**(1): p. 73-77.

51. Mathupala, S.P., Y.H. Ko, and P.L. Pedersen, *Hexokinase II: Cancer's double-edged sword acting as both facilitator and gatekeeper of malignancy when bound to mitochondria.* Oncogene, 2006. **25**(34): p. 4777-4786.

52. Brown, R.S., et al., *Expression of hexokinase II and Glut-1 in untreated human breast cancer.* Nuclear Medicine and Biology, 2002. **29**(4): p. 443-453.

53. Shreeve, C.M., et al. *Correlation of hexokinase-2 expression with overall survival and potential as a therapeutic target in the treatment of breast cancer brain metastases.* in *ASCO Annual Meeting* 2008.

54. Torres, N.V., et al., *Kinetics of metabolic pathways. A system in vitro to study the control of flux.* Biochemical Journal, 1986. **234**(1): p. 169-174.

55. Bos, R., et al., *Biologic correlates of (18)fluorodeoxyglucose uptake in human breast cancer measured by positron emission tomography.* Journal of clinical oncology, 2002. **20**(2): p. 379-387.

56. Brown, R.S. and R.L. Wahl, *Overexpression of glut-1 glucose transporter in human breast cancer an immunohistochemical study.* Cancer, 1993. **72**(10): p. 2979-2985.

57. Macheda, M.L., S. Rogers, and J.D. Best, *Molecular and cellular regulation of glucose transporter (GLUT) proteins in cancer.* Journal of cellular physiology, 2005. **202**(3): p. 654-662.

58. Brown, R.S. and R.L. Wahl, *Overexpression of Glut-1 glucose transporter in human breast cancer. An immunohistochemical study.* Cancer, 1993. **72**(10): p. 2979-2985.

59. Younes, M., et al., *Wide Expression of the Human Erythrocyte Glucose Transporter Glut1 in Human Cancers.* Cancer Research, 1996. **56**(1164-1167).

60. Reuss, M. *Multiscale Modeling in Cancer Therapy – Synergistic Interplay of Target Identification and Drug Delivery*. in *Metabolic Engineering VIII Conference*. 2010. Jeju Island, Korea.

61. Jankowski, M.D., et al., *Group contribution method for thermodynamic analysis of complex metabolic networks.* Biophysical Journal, 2008. **95**(3): p. 1487-1499.

62. Alberty, R.A., *The role of water in the thermodynamics of dilute aqueous solutions.* Biophysical Chemistry, 2003. **100**(1-3): p. 183-192.

63. Henry, C.S., et al., *Genome scale thermodynamic analysis of Escherichia coli metabolism.* Biophysical Journal, 2006. **90**(4): p. 1453-1461.

64. Jankowski, M.D., et al., *Group Contribution Method for Thermodynamic Analysis of Complex Metabolic Networks.* Biophysical Journal, 2008. **95**(3): p. 1487–1499.

65. Alberty, R.A. and A. Cornish-Bowden, *The pH dependence of the apparent equilibrium constant, K', of a biochemical reaction.* Trends in Biochemical Sciences, 1993. **18**(8): p. 288-291.

66. Magnus, G. and J. Keizer, *Minimal model of β-cell mitochondrial Ca^{2+} handling.* American Journal of Physiology, Endocrinology and Metabolism, 1997. **273**(2): p. C717-C733.

67. Gunter, T.E. and D.R. Pfeiffer, *Mechanisms by which mitochondria transport calcium.* American Journal of Physiology, Endocrinology and Metabolism, 1990. **258**(5): p. C755-C786.

68. Wojtczak, L., A. Żółkiewska, and J. Duszyński, *Energy-storage capacity of the mitochondrial proton-motive force.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1986. **851**(2): p. 313-321.

CHAPTER 5

CONCLUSIONS

SYNOPSIS

In the quest for anti-cancer drugs, cancer metabolism has increasingly been a focus of interest in clinical research. Enhanced glycolysis and robust production of lactate constitute a characteristic trait that discriminates many cancerous cells from their normal counterparts. As one of the main problems in defeating cancer is the genetic heterogeneity of the altered cells constituting the malignancy, the near-universality of these metabolic features can provide researchers with a handle on such a complex disease, regardless of the specific genotype of the single cells.

In this thesis we have developed and applied analytical approaches, mainly drawn from the field of metabolic control analysis (MCA), to the study of cancer metabolism. In particular we aimed to assess whether and to what extent the metabolic features of cancer cells can be exploited to attack the malignancy more specifically than through traditional approaches, such as chemotherapy or radiotherapy. The underlying idea consists of identifying enzymes that represent points of fragility that specifically characterise the cancerous metabolic phenotype. These enzymes are such that an alteration in their activity (due for example to the action of an anticancer drug) would elicit the desired response in cancer cells, without affecting their normal counterparts. MCA can help to identify such points of fragility by assessing how the control over important cellular functions is distributed amongst the different enzymes. Enzymes which exert a strong control over a fundamental biological property in the cancer metabolic phenotype and a low control over that same property in the normal phenotype can be considered good candidate targets for a drug aimed at eliciting a high differential response between neoplastic and normal cells.

The application of MCA relies on a mathematical representation of the system under study. Creating such a model is often hampered by the lack of data about the precise kinetic laws governing the different reaction steps and the value of their corresponding parameters. One of the most important results presented in this thesis is that the metabolic quantities defining the normal and cancer phenotypes (such as fluxes and metabolite concentrations), together with heuristic assumptions about the properties of typical enzyme-catalyzed reactions, already allow for a fast and efficient way to explore the effectiveness of putative drug targets. This result was obtained by integrating a random sampling approach into an MCA framework. In particular, we showed that if the quantities defining a metabolic phenotype are known, it is possible to gain a probabilistic insight on how the control over relevant systemic properties is distributed amongst the different enzymatic steps, although only partial knowledge may be available with respect to the kinetic properties of the system. The comparison of the probabilistic control profile of the system between the two metabolic phenotypes (normal vs. cancer, or host vs. pathogen) allows for a first screening of the different enzymes as suitable target for drugs with high efficacy and low toxicity operating at the metabolic level. The relevance of this result lies in the fact that metabolomics and fluxomics studies are now standard techniques in the analysis of cellular metabolism, and the quantities defining a metabolic phenotype are experimentally more accessible than the kinetic parameters of the different enzymatic steps in the system.

We applied our methodology in a case study aimed to identify putative molecular targets for drugs designed to attack human breast cancer. We based our study on fluxomic and metabolomic data currently available in literature, and made use, where possible, of actual kinetic equations in the attempt to minimize the uncertainty introduced in the description of the system. Interestingly the results obtained were in accordance with previous experimental work, highlighting the central role of certain enzymes such as GLT

and G6PDH in controlling glycolysis and the pentose phosphate pathway respectively in some kinds of cancer.

CRITICISM

One possible criticism that could be levelled at the work presented in this thesis is the limitations inherent to the probabilistic nature to our sampling-based analytical approach. Indeed, biology seems to work contrary to randomness and selective pressure may have confined the parameter values to a tiny region of the entire parameter space. Such a subspace might restrict the behaviour of the system to an extent that it is not reflected by the "most probable" outcome. Consequently, the results obtained through a random sampling approach in the investigation of the control properties of a system might face some scepticism as they would seem to suggest that the most probable values are indeed (close to) the value that one would observe in the real system. However, the real system might be found in the "tail" of the distribution of the possible outcomes. A possible answer to this criticism is that what defines the tail of a distribution depends on the average and width of the distribution itself. In this respect, one has to consider that the reliability of any probabilistic prediction depends on the quality of its underlying assumptions. In our work we constrained the sampling interval within reasonable range of values in order to limit the possible responses of the system to what is reasonably expectable. On this basis it is more reasonable to be guided by the most probable result, rather than the less likely, in an attempt to infer the control properties of the system. We also underline that a common practice in modelling biological systems consists of "adjusting" the parameters in order to make the model fit with the experimentally observed behaviour. Although such a model might reproduce some known characteristic of the system under study, many other possible choices for the parameter values are able to reproduce that same behaviour. Sampling the parameters, then, corresponds to evaluating an entire population

of models, each of them, in absence of any prior knowledge about the parameter values, equally eligible to represent the system under study. In our case, all the models of the population (differing from each other by their specific choice of parameters) were able to reproduce the observed metabolic phenotypes under comparison (normal and cancer) in terms of fluxes and metabolite concentrations at steady-state. We suggest that this approach is preferable in the abundant cases where the parameter values of the relevant *in vivo* conditions are unknown.

A different criticism might be put forward when considering the specific choice of the metabolic pathways taken into account as representative of the system under study. Glycolysis, the pentose phosphate pathway and the TCA cycle might not account for all the relevant metabolic changes that cells undergo during carcinogenesis. In this case further refinement of the model in terms of actual kinetic equations and parameter values would not necessarily lead to more realistic results. This issue may be partly addressed through more extensive experimental quantification of the metabolic changes occurring in cancer. From a modelling perspective, the application of flux balance analysis could also contribute in identifying a scaffold of reactions which are supposedly hosting the main metabolic changes occurring in transforming cells. In particular, FBA could be used to investigate the motive force that pushes many cancer cells to (partially or totally) switch from respiration to fermentation, by considering possible scenarios where the functioning of cancer cells is mainly driven by different biological needs. In so doing, FBA solutions reproducing the observed pattern of fluxes in well studied metabolic pathways could be used to infer and predict other regions of metabolism where cancer and normal cells differ. Identifying the biological functions driving and shaping the metabolic phenotype characteristic of many cancer cells could also provide further criteria to target the malignancy, based on the reprioritization of the systemic properties underlying the functioning of cancer cells.

FURTHER WORK

The results obtained during this project suggest possible directions to further investigate cancer metabolism in a clinical context. Because the control profile of a metabolic system seems to be determined to a good extent by the metabolic state under consideration, it would be sensible to dedicate resources to retrieve a detailed characterization of such metabolic states. This would be done for both control and cancer cells of a specific type of neoplasia. Human breast cancer is a natural choice to focus the experimental quantifications of these metabolic profiles, since it has already been partly characterized. Measured fluxes and concentrations would be integrated in our MCA-based analytical approach to predict the suitability of the different enzymes as molecular targets, and the validity of these predictions tested experimentally. In particular, siRNA and transient transfection methods could be employed to down-regulate the expression of the key enzymes highlighted in the study as best putative targets, and the differential response of cancer versus control cells assessed.