THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Constraint-based modeling of metabolism

- interpreting predictions of growth and ATP synthesis in human and yeast

AVLANT NILSSON

Constraint-based modeling of metabolism

- interpreting predictions of growth and ATP synthesis in human and yeast

Avlant Nilsson

Department of Biology and Biological Engineering
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden

Telephone +46(0) 31-772 1000

Constraint-based modeling of metabolism - interpreting predictions of growth and ATP synthesis in human and yeast
Avlant Nilsson
Department of Biology and Biological Engineering
Chalmers University of Technology

Abstract

Growth is the primary objective of the cell. Diseases arise when cells diverge from a healthy growth-pattern. An increased understanding of cellular growth may thus be translated into improved human health. The cell requires materials and free energy (in the form of ATP) in order to grow, metabolism supplies the cell with this. The rate of metabolism is ultimately constrained by the biophysical properties of the metabolic enzymes. Interactions between the constraints and the growth-objective gives rise to metabolic trade-offs, e.g. between ATP synthesis from respiration and fermentation. We can gain quantitative insight into these processes by simulating metabolism using mathematical models. In this thesis I simulated the metabolism of four biological systems: the infant, cancer, yeast and muscle. The simulations demonstrated how a shift in metabolic strategy may increase the rates of ATP synthesis and growth. These increased metabolic rates come at the expense of decreased resource efficiency, i.e. ATP produced per carbon consumed. The effect was primarily caused by the low catalytic efficiency of the respiratory enzyme complexes I and V. By shifting from respiratory to fermentative ATP synthesis, the cell was able to bypass these constraints. An intermediate strategy involved bypassing only complex I. The phenomenon was experimentally corroborated in the working muscle, and it is the native state of the yeast *Saccharomyces cerevisiae* (which lacks complex I). The differences in efficiency between the different metabolic pathways also explained why cells grow faster on some carbon sources, e.g. the specific growth rate for yeast is higher on glucose than on ethanol. These models were extended to predict the world-record running-speeds at different distances, by taking the sizes of the body's nutrient-deposits into account. A metabolic strategy employed by cancer cells involved excretion of the amino acid glutamate. The simulations showed a mechanistic relation to catabolism of branched-chain amino acids and the localization of amino acid metabolism to different cellular compartments. By experimentally inhibiting glutamate excretion using an off-the-shelf drug (sulfasalazine), the growth rate of a cancer cell line was reduced. The metabolic modeling involved integration of various types of data and thus demonstrated the potential to unify knowledge from different studies and domains. This exposed contradictory claims in literature and highlighted knowledge-gaps that need to be filled to further improve human health.

Vilkorsbaserad modellering av metabolismen – tolkning av prediktioner om tillväxt och ATP-syntes i människa och jäst
Avlant Nilsson
Institutionen för Biologi och bioteknik
Chalmers tekniska högskola

Sammanfattning

Cellens primära målsättning är att växa. Celler som avviker från ett hälsosamt tillväxtmönster kan ofta orsaka sjukdomar. Ökad kunskap om celltillväxt kan därmed omsättas i förbättrad hälsa. För att växa behöver cellen material och energi (i form av ATP), metabolismen förser cellen med detta. Takten på metabolismen bestäms i sista hand av de biofysiska egenskaperna hos dess enzymer. I interaktionen mellan de biofysiska villkoren och cellens tillväxtmål uppstår avvägningar mellan olika strategier för metabolismen, t.ex. mellan ATP syntes genom respiration eller fermentering. En kvantitativ inblick i dessa processer kan uppnås genom att simulera metabolismen med hjälp av matematiska modeller. I denna avhandling simulerade jag metabolismen hos fyra biologiska system: spädbarn, cancer, jäst och muskel. Simuleringarna visade hur ett byte av strategi kunde öka takten på ATP-syntesen och tillväxten. Ökningen kom dock på bekostnad av minskad resurseffektivitet, d.v.s. mängden ATP som producerades per förbrukat kol. Detta berodde främst på låg katalytisk effektivitet hos de respiratoriska enzymkomplexen I och V. Genom att byta strategi i ATP syntesen från respiration till fermentering så kunde cellen kringgå komplexen. En mellanliggande strategi innefattade att endast förbikoppla komplex I. Det fenomenet bekräftades experimentellt i en arbetande muskel, och det utgör normaltillståndet för jästen Saccharomyces cerevisiae (som helt saknar komplex I). Effektivitetsskillnader mellan de olika vägarna i metabolismen förklarade också varför celler växer fortare på vissa kolkällor än andra, t.ex. så växer jäst fortare på glukos än på alkohol. Modellerna utvidgades så att de kunde förklara skillnaderna i världsrekordhastighet i löpning på olika distanser, genom att ta hänsyn storleken på reserverna av olika näringsämnen. En strategi som används av cancerceller involverade utsöndring av aminosyran glutamat. Simuleringarna visade en mekanistisk relation med nedbrytningen av grenade aminosyror och med lokalisering av aminosyrametabolismen till olika delar av cellen. Genom att inhibera glutamatutsöndringen experimentellt med hjälp av ett läkemedel (sulfasalazin), kunde tillväxttakten reduceras i en cancercellinje. Olika typer av data integrerades med hjälp av modellen och påvisade därigenom dess potential i att förena kunskaper från olika studier och domäner. Modellerna exponerade motsägelsefulla påståenden i litteraturen och markerade de kunskapsluckor som måste fyllas för att ytterligare förbättra människors hälsa.

List of publications

This thesis is based on the work contained in the following papers and manuscripts:

Paper I: Predicting Growth of the Healthy Infant using a Genome Scale Metabolic Model
Nilsson A, Mardinoglu A and Nielsen J (2017) Npj Systems Biology and Applications.

Paper II: Liver cancer cells excrete glutamate of cytosolic but not mitochondrial origin
Nilsson A*, Haanstra JR*, Engqvist M, Gerding A, Bakker B, Klingmüller U, Teusink B and Nielsen J (2018) *Manuscript*.

Paper III: Metabolite Depletion Affects Flux Profiling of Cell Lines
Nilsson A, Haanstra JR, Teusink B, Nielsen J (2018) Trends in biochemical sciences.

Paper IV: Metabolic Trade-offs in Yeast are Caused by F1F0-ATP synthase
Nilsson A and Nielsen J (2016) Scientific Reports.

Paper V: Complex I is bypassed during high intensity exercise
Nilsson A, Björnson E, Flockhart M, Larsen F, Nielsen J (2018) *Manuscript.*

Paper VI: Metabolic Models of Protein Allocation Call for the Kinetome
Nilsson A, Nielsen J and Palsson BO (2017) Cell Systems

Additional publications during doctoral research not included in this thesis:

Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints
Sánchez BJ, Zhang C, Nilsson A, Lahtvee PJ, Kerkhoven EJ, Nielsen J (2017) Molecular systems biology

Genome scale metabolic modeling of cancer
Nilsson A, Nielsen J (2017) Metabolic engineering

Recon3D enables a three-dimensional view of gene variation in human metabolism
Brunk E, Sahoo S, Zielinski DC, Altunkaya A, Dräger A, Mih N, Gatto F, Nilsson A, Preciat Gonzalez GA, Aurich MK, Prlić A, Sastry A, Danielsdottir AD, Heinken A, Noronha A, Rose PW, Burley SK, Fleming RMT, Nielsen J, Thiele I, Palsson BO (2018) Nature biotechnology

*Authors contributed equally to this work

Contribution summary

My contributions to the papers are summarized in the table below (Table 1).

Table 1 Contribution summary.

	Paper I	II	III	IV	V	VI
Conceived the idea	White	White	Green	White	Yellow	Yellow
Developed the project	Green	Yellow	Green	Green	Yellow	Yellow
Conducted the experiments	Gray	White	Gray	Gray	White	Gray
Constructed the model	Green	Green	Gray	Green	Green	Gray
Performed simulations	Green	Green	Gray	Green	Green	Gray
Analyzed the data	Green	Green	Green	Green	Green	Gray
Drafted the paper	Green	Green	Green	Green	Green	Green

Key

Color	
Green	Performed the task
Yellow	Co-performed the task
Gray	Not applicable
White	Did not perform the task

Preface

This dissertation serves as partial fulfillment of the requirements to obtain the degree of Doctor of Philosophy at the Department of Biology and Biological Engineering at Chalmers University of Technology. The PhD studies were carried out between September 2014 and January 2019 at the division of Systems and Synthetic Biology under the supervision of Jens Nielsen. The project was co-supervised by Adil Mardinoglu and Martin Engqvist and examined by Stefan Hohmann. It was mainly funded by Bill and Melinda Gates foundation, and Vetenskapsrådet and Västra Götalands Regionen, which funded the ERA-NETproject IMOMESIC through the ERASysApp program.

Avlant Nilsson

January 2019

Table of content

Abbreviations

AKG   α-ketoglutarate
ALA   alanine
ALE   adaptive laboratory evolution
ASA   acute sensitivity analysis
ASP   aspartic acid
ATP   adenosine triphosphate
BCAA   branched chain amino acid
CO2   carbon dioxide
DNA   deoxyribonucleic acid
FBA   flux balance analysis
FBS   fetal bovine serum
gdw   gram cell dry weight
GEM   genome-scale metabolic model
HMR   human metabolic reaction
mRNA   messenger RNA
NADH   nicotinamide adenine dinucleotide (reduced)
O2   oxygen
pFBA   parsimonious flux balance analysis
PYR   pyruvate
QH2   ubiquinol
RNA   ribonucleic acid
RQ   respiration quotient
SSZ   sulfasalazine
TCA   tricarboxylic acid cycle
vCO2   carbon dioxide production rate
vO2   oxygen consumption rate

x

In loving memory of Marianne

Acknowledgments

Background

The cell is the fundamental unit of life. By improving our understanding of the cell, we may improve quality of life through a wide range of applications, from drug development to production of bio-based goods and services. The increasing availability of biological data allows us to ask more detailed scientific questions about the functions of cells under different conditions. Computer models help us to put the data in context of our collective knowledge and enable quantitative predictions of phenotypes through simulations. The models provide mechanistic descriptions of cellular behavior, founded in biophysical constraints and the theory of evolution. In this thesis, I developed metabolic models of yeast, infants, exercising humans and cancer. I use the models to simulate growth and ATP synthesis and interpret the predictions.

Understanding and improving life

We can improve life by modifying cells. All living organisms are made up of cells and are affected by changes at the cellular level. Cellular interventions may sometimes be desirable in order to cure a disease or improve the productivity of cells of industrial relevance. In the case of disease treatment, cells are reverted from a diseased to a healthy state through some type of intervention. Some diseases develop from malfunctioning cells, and may be treated with small molecule drugs that enter the cell and interact with its constituents (Adams et al. 2015). Others arise from malnutrition, and may be prevented by appropriate nutritional interventions (Prendergast & Humphrey 2014). Yet other diseases are caused by infections, and require the pathogen to be separated from the host (Bordbar & Palsson 2012). In the case of industrially relevant cells, they may be modified to increase productivity or to yield new types of products through metabolic engineering. An example is yeast, which traditionally has been used in baking, brewing, but has been modified to produce chemicals in several industrial fermentation processes (Nielsen 2015). To be able to modify cells with reliable results, a fundamental understanding of the cell is required.

Cells are at their core self-replicators that must divide to grow. The growth of cells is the result of complex interactions between a number of biological processes (Molenaar et al. 2009). Many of which may limit growth, e.g. DNA replication, synthesis of proteins and uptake of nutrients. Environmental factors such as substrate availability, temperature and stress factors may also play a role. In the case of multicellular organisms, the growth of cells is constrained by regulatory signals

(Hanahan & Weinberg 2011). Careful control of cell growth is required to maintain the organism in a healthy state.

Different diseases and syndromes are caused by malfunctioning growth processes. One example is cancer, it is a disease that involves uncontrolled cell growth and is estimated to cause 9.6 million deaths in 2018 (Bray et al. 2018). Mutations, and other genetic events, reprogram the cancer cells to avoid checkpoints in control of nutrient supply and growth (Hanahan & Weinberg 2011). Another example is the stunting syndrome, it involves reduced growth in children, and affects 165 million children younger than 5 years globally. It is associated with decreased cognition and health (Prendergast & Humphrey 2014). Since the human body is a complex system, examining cells in isolation may sometimes be more informative.

We can understand human cells better by studying other organisms. Due to the shared evolutionary history, similar principles may be used to understand disparate biological systems (Liu et al. 2014). The generalization of knowledge from one species to another may be illustrated by model organisms. These are less complex organisms that are easier to study, and for which more prior knowledge is available. Many discoveries in the model organism *Saccharomyces cerevisiae*, have later been translated to human (Hohmann 2016), e.g. discoveries related to cell division, protein transport and DNA protection by telomeres. A detailed understanding of the cell has emerged from studying, yeast and other model organisms.

Metabolism is the base, signaling the superstructure

Metabolism controls the cell. Cellular activities may be broadly divided into metabolism and signaling (Hyduke & Palsson 2010). Metabolism is the base and carries out the physical activities of the cell. It is central to their maintenance and growth. It converts nutrients into building blocks and free energy in the form of ATP through a vast network of enzyme mediated biochemical reactions. Signaling is the superstructure and carries out the decision-making activities of the cell. It regulates metabolism through the activities and abundances of enzymes. The signaling network responds to changes in the environment and to the new metabolic requirements that follow, e.g. depletion of a growth enhancing metabolite. It also carries out the internal dynamics of the cell, e.g. the progression through the cell cycle. The physical limitations of metabolism shape the signaling system, and the signaling system maintains metabolism in an appropriate form.

The cell must optimize its allocation of protein. Biological processes require enzymes and other proteins to function. The protein content of the cell is finite, and the cell must therefore optimize the allocation of proteins amongst these functions. This becomes a regulatory problem for the cell, i.e. the protein allocation problem. A concept that has been popularized through different versions of proteome fraction models (You et al. 2013; Hui et al. 2015; Mori et al. 2016). They divide the protein pool into different coarse-grained sectors such as core functions, the ribosomes and metabolism. The growth condition dictates the optimal allocation between these sectors. By definition, the core functions are assumed to cover housekeeping activities that are independent of growth. The demand for ribosomes is thought to be a simple linear function of the specific

growth rate. This leaves most of the dynamics to the metabolic domain. The protein allocation problem gives rise to trade-offs between different metabolic strategies.

Overflow metabolism is a trade-off between efficiency and rate. It denotes a shift in ATP synthesis strategy from respiration towards fermentation while under aerobic conditions (Molenaar et al. 2009). This results in increased consumption of substrate, e.g. glucose, and the excretion of fermentation products, and thus a lower yield of ATP per substrate. The catalytic rate of fermentation is however higher, and more ATP can thus be synthesized per enzyme mass. This makes overflow metabolism a favorable metabolic strategy under conditions when rapid ATP synthesis is required. The phenomenon occurs in many organisms under different names, *the Crabtree effect* in yeast, *the Warburg effect* in cancer, *the lactate threshold* in sports (Molenaar et al. 2009; Vazquez & Oltvai 2011). Analogous low-yield pathways also exist in many bacteria, e.g. *Escherichia coli* and *Bacillus subtilis* (van Hoek & Merks 2012). Cancer cells commonly gain access to overflow metabolism by dysregulating the cells oxygen sensor HIF-1α (Cairns et al. 2011). The resulting high glucose consumption rates can be exploited for tumor detection using positron emission tomography, i.e. PET scans.

Cancer cells modify metabolism to support rapid growth. There are several metabolic traits, aside from rapid glucose consumption, that are shared by many cancer cells. The amino acid glutamine is commonly consumed at high rates (Wise & Thompson 2010). Several other amino acids are instead excreted, e.g. glutamate, alanine and glycine (Zielinski et al. 2017; Jain et al. 2012). It remains unclear why cells excrete glutamate (Seidlitz et al. 2009; Stepulak et al. 2014), and despite many influential studies on the importance of metabolism for cancer cell proliferation, we are missing an overall view of how cancer cells coordinate acquisition of the diverse nutrients required for growth (Palm & Thompson 2017). This calls for an integrative study of metabolism in cancer using a systems biology approach.

Computer simulations in systems biology

Systems biology takes a holistic perspective on the cell (Kitano 2002). It focuses on the interactions between the cells components and on the complex behavior that emerges from these interactions. This is facilitated by computer models, which integrate knowledge and data into a structured framework (Figure 1A). The use of computer models in systems biology may be broadly divided into the top-down approach and the bottom-up approach (Bruggeman & Westerhoff 2007). Top-down systems biology maps experimental data to models to identify phenomenological facts, which may in turn be interpreted as molecular mechanisms. Bottom-up systems biology simulates molecular mechanisms *in silico,* i.e. in the computer, and generates quantitative and testable predictions of cell behavior.

Mechanistic relationships in reality may be inferred from the simulations. The simulated interaction between the components of a model can be interpreted as a mechanistic relationship *in silico*, e.g. knocking out a certain gene causes a reduction in cell growth. This relationship in the model may then be translated to statements about reality, if the correspondence between the model and reality is considered to be sufficiently good. The predictions do, however, always have an implicit
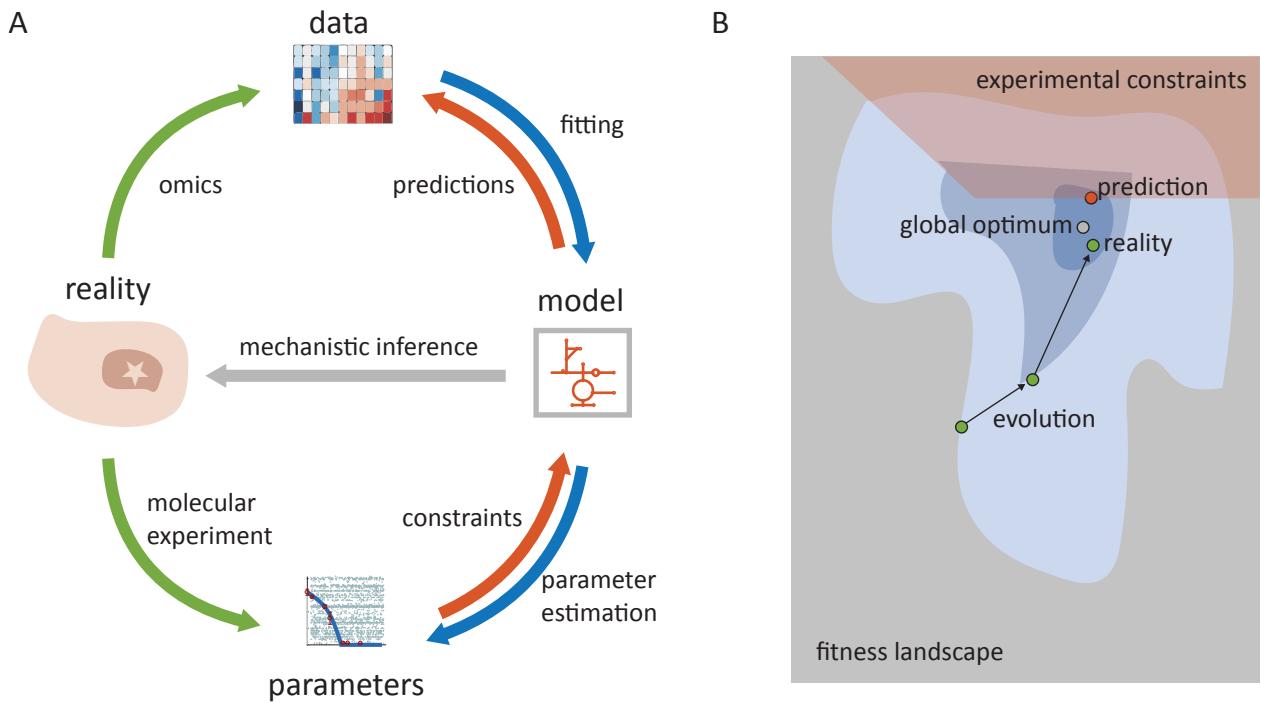
Figure 1. The relation between model and reality in systems biology. A) A good model has an injective relationship with reality, which allows mechanistic relationships in the model to be directly mapped to reality (mechanistic inference). Data is gathered about reality through experiments (green arrows). These may be aimed at identifying parameter values that are invariant to time and the condition. They may also be used to give an experimental account of the state of the cell. Bottom-up systems biology (orange arrows) constrains models with parameter values to make experimentally verifiable predictions. Top-down systems biology (blue arrows) integrates experimental data to fit a model. The fitted model may potentially be used to estimate parameters and thus allow bottom-up systems biology. B) The biophysical constraints on a cell generate a fitness landscape, which may have some global optimum. Evolutionary processes drive the cells towards optimality. Assuming that cells have reached near optimality, the state of the cell may be predicted from the fitness landscape. Experimentally derived constraints may be used to rule out possible but counterfactual states.

dependency on the accuracy of the model structure and the parameter values. In the absence of complete information, predictions can also be informative (Gutenkunst et al. 2007). They may show that a property is predictable in principle under more favorable research conditions, e.g. in the case of random sampling of parameter values and ensemble modeling (Morris et al. 2016; Tran et al. 2008; Beg et al. 2007; Steuer et al. 2006). In the absence of detailed parameter data only general conclusions may be drawn about a system of interest.

A good model should describe the quantitative state of a cell under a condition of interest. The state refers to the quantities of the components of interest, e.g. concentrations of metabolites and enzymes or metabolic fluxes. The state may have time dynamics or be in steady state. By definition a cell has a single state at a given point in time, but many states may exist within a population of cells (Labhsetwar et al. 2017). A common abstraction is to assume that all of the studied cells are in a single representative steady state. Some properties of the cells may be directly quantified experimentally, e.g. the concentration of proteins and metabolites. Other properties are abstractions which cannot be directly observed, e.g. the specific metabolic fluxes (Winter &

Krömer 2012). A well calibrated model overcomes these limitations, and allows a non-disruptive assessment of the state of the system. This may be particularly useful if the physical integrity of the system is important, as in the case of humans.

Constraint-based modeling of biology has been a fruitful paradigm (Orth et al. 2010; Palsson 2015). Biological knowledge can often be expressed in the form of constraints. The knowledge may be derived from experiments, e.g. the observed glucose uptake rate under a condition of interest. It may also be derived from biophysics, e.g. that the catalytic capacity of an enzyme does not exceed some value. The biophysical constraints form a fitness landscape for the cell (Figure 1B). The state of the cell may be predicted from the interaction between the biophysical constraints on the cell and the evolutionary drive towards optimality, the theory of dual causality (Bock 2017; Palsson 2015). Using this theory, the model may be able to explain the rationale behind an observed cellular phenomenon, e.g. the cell prefers to consume glucose because it can consume it faster than other sugars. It should be noted that many biophysical constraints arise from the physical properties of enzymes, and may be subject to change on an evolutionary time scale (Savir et al. 2010).

Constraint-based modeling of metabolism
Different types of metabolic models exist. They primarily fall into the categories of topological, stoichiometric and kinetic models (Steuer 2007). These describe the cell with increasing level of detail and hence with increasing demand for parameter values (Figure 2). Topological models require little experimental data, but are limited to qualitative predictions about the state of the cell, e.g. lethal or non-lethal gene deletions. Dynamic models have the potential to give a detailed description of the state of the cell, but due to data availability they are often limited to studying subsystems. In particular the enzyme capacity data commonly used as parameter in dynamic models ($v_{max}=[E]\cdot kcat$) has an implicit dependency on enzyme concentrations ($[E]$) that are specific to the studied condition (Müller et al. 2015). General conclusions transcending the experimentally determined state cannot easily be made. In this thesis I make use of stoichiometric models of metabolism.

A genome-scale metabolic model (GEM) is a mathematical representation of metabolism (Orth et al. 2010; Palsson 2015). It contains information about the stoichiometry of (nearly) all biochemical reactions in an organism, and links them to the genes that encode for the enzymes in question. Thus, genes, metabolites and reactions are joined to form a network (Figure 3A), expressed as a stoichiometric matrix (Figure 3B). For multi cellular organisms a multi tissue model can be developed by joining several models through a shared compartment (Bordbar et al. 2011). Flux balance analysis (FBA), is a method used to predict the metabolic fluxes of the reactions under a condition of interest (Figure 3C). Metabolites are assumed to be well mixed and at a constant concentration within each cellular compartment. This relies on the steady state assumption, i.e. that fluxes producing and consuming each metabolite cancel out, preventing buildup or depletion of internal metabolites.

FBA integrates constraints with the model. Experimental data can be integrated in the form of flux bounds and constraints, e.g. measurements of exchange fluxes and the accumulation of
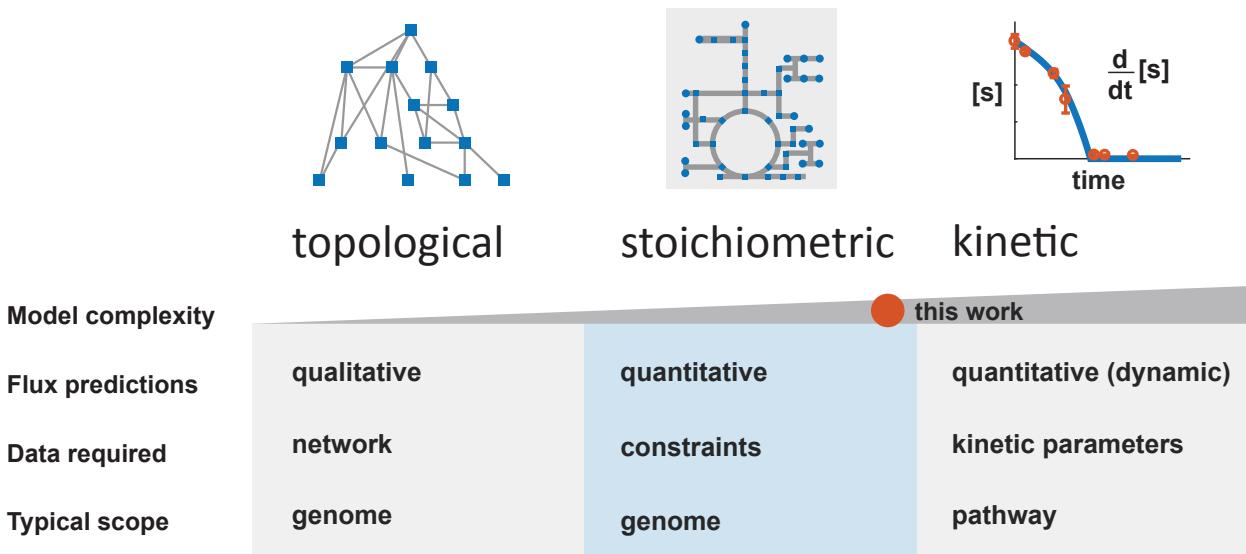
Figure 2 Different types of metabolic models. Different metabolic models may be suitable depending on the research question. Topological models require little experimental data and can give qualitative predictions about fluxes and essentiality of genes. Stoichiometric models require quantitative constraints to make quantitative flux predictions. Kinetic models give dynamic flux predictions, but rely on a large number of parameter values. This typically reduces the scope of the models to the level of pathways, and the parameters used are often specific to a particular context.

biomass. Alternatively, constraints may be added based on prior knowledge, e.g. enzyme pool constraints. Models constrained with experimental data are, by construction, limited to describe the state for which the data was collected. Models with biophysical constraints may be generalized to any condition for which the biophysical description is valid. The addition of experimental and biophysical constraints is seldom sufficient to uniquely determine the fluxes, resulting in a large solution space. For the biophysically constrained models this can be interpreted as a high degree of metabolic flexibility being retained by the cells.

Fluxes are predicted by solving an optimization problem. To identify a quantitative and unique flux prediction, the cells are assumed to (successfully) optimize some objective function (Orth et al. 2010). This is done by solving a linear programming problem that optimizes the flux through the affiliated reactions (Figure 3C). For growing cells, this may be to maximize the specific growth rate, and for non-growing cells, this may instead be to make efficient use of limited resources, by minimizing the substrate consumption. The living cell may have other priorities than maximizing growth. One proposed trade-off is between efficiency and the robustness to perturbations of metabolism (Segrè et al. 2002; Stelling et al. 2002). Another trade-off may be between accuracy and efficiency; for the ribosomes fast processing of mRNAs comes at the price of more mistranslation (Johansson et al. 2012). However, in some cases adapting the organism to the laboratory conditions, will drive the phenotype closer to the predicted optimum (Lewis et al. 2010). Multiple optimal solutions often remain after the first optimization, and a second optimization is often run to identify a minimal flux vector (Raman & Chandra 2009). This process may be viewed as a proxy for minimizing enzyme usage, but may not always be reliable, since all reactions carry equal weights.
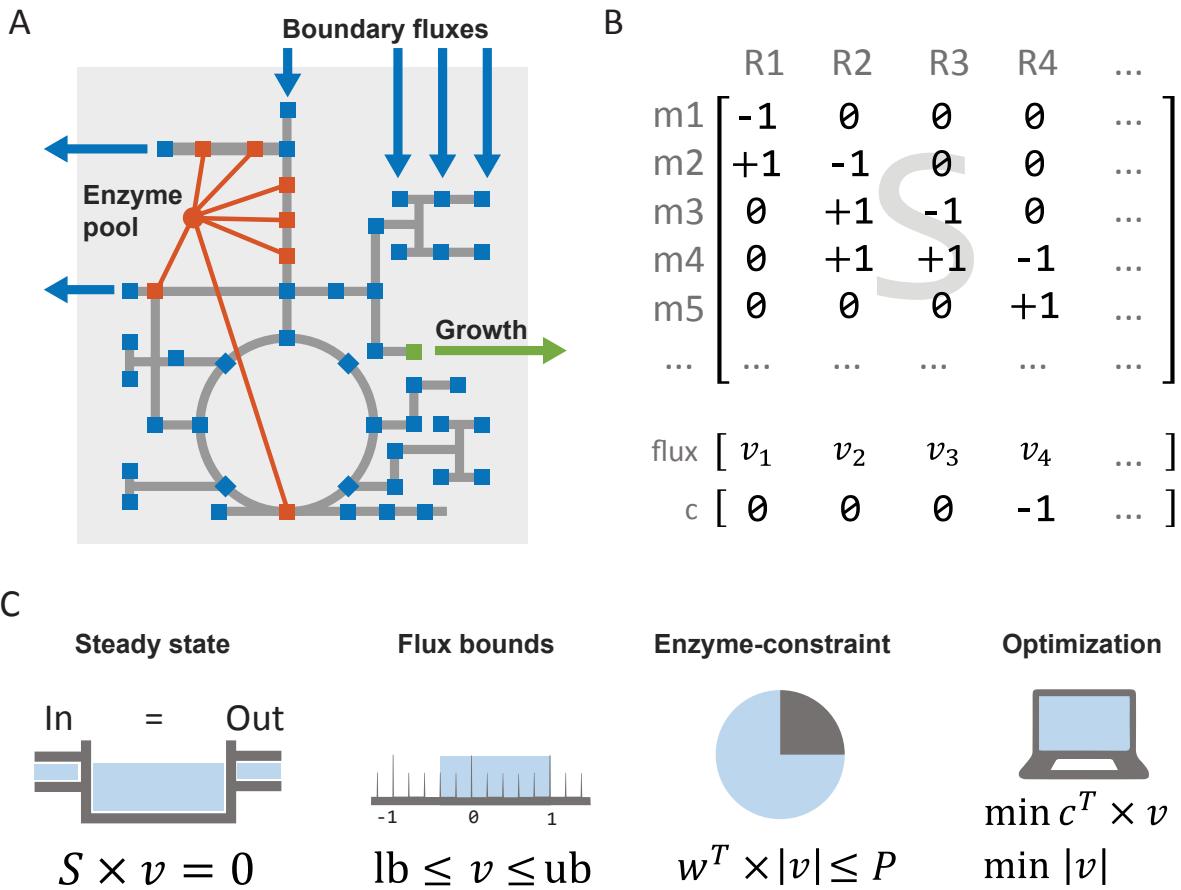
Figure 3. Genome-scale metabolic modeling using flux balance analysis. A) A metabolic network is reconstructed and experimental data on boundary fluxes are gathered, alternatively, parameter data is gathered for the relation between fluxes and enzymes. B) The network is represented by a stoichiometric matrix (*S*) where the rows correspond to metabolites and columns to reactions. Each reaction is represented by assigning the stoichiometric coefficients of the participating metabolites to the corresponding rows, where negative values indicate consumption and positive values production. C) Simulations are conducted to predict the fluxes (*v*), expressed in mmol per gram dry cell weight per hour (mmol/gdw/h). The simulations assume that metabolite concentrations are in steady state. Fluxes are constrained by upper and lower bounds (ub and lb) from experimentally determined intervals, when available. For irreversible reactions the lower bound is set to zero (lb=0). Alternatively or additionally, the fluxes are constrained through their use (*w*) of a constrained pool of total enzyme (*P*). A flux distribution is sought that optimizes the flux through one or more reactions (*c*), e.g. maximizes growth. There are often multiple flux distributions that result in the same optimal solution. The distributions often contain large predicted fluxes due to loops in the model. A second optimization is therefore often run to reduce unnecessarily large fluxes and remove loops. The absolute value of all fluxes is minimized, whilst maintaining the optimal solution. This results in a vector with the same flux through the objective function but with the lowest sum of absolute values of fluxes i.e. the minimum L1 norm. This is referred to as parsimonious FBA (pFBA).

Enzyme-constraints integrate FBA with the protein allocation problem. Enzyme-constrained metabolic models attempt to account for the protein investment required to sustain metabolic fluxes (Molenaar et al. 2009; Müller et al. 2015). They account for the constraints posed by limited enzyme capacities and have been used to model the processes underlying phenotypes in several organisms. The models have improved our understanding of constraints on the maximum specific growth rate for different organisms (Adadi et al. 2012), the diauxic shift (Beg et al. 2007) and overflow metabolism (Vazquez & Oltvai 2011; Shlomi et al. 2011). They come in the form of

constraints from macromolecular crowding (Beg et al. 2007), constraints on the sum of enzyme mass (Adadi et al. 2012) or on crowding of membranes (Zhuang et al. 2011; Szenk et al. 2017). Several metabolic models aspire to expand these constraints further to balance all of the resources in the cell, e.g. resource balance analysis (Goelzer et al. 2011), models linking metabolism to protein expression (O'Brien & Palsson 2015), models using proteome sectors (Mori et al. 2016) and whole cell models (Karr et al. 2012). We have only begun to discover the prediction power of resource-constrained models.

Aim and significance

Millions of people are affected by severe health issues. To reliably revert unhealthy cells to a healthy state, we must understand the healthy cell on a fundamental level. Growth is a core property of the healthy cell, and diseases often emerge when cells diverge from a healthy growth pattern. Metabolism supplies the materials and free energy required for growth. It is constrained by the biophysical properties of the enzymes involved, and complex behavior arises from the cells adaptations to the constraints. By simulating metabolism using mathematical models, we can gain an understanding of these processes. In this thesis I aimed to advance our fundamental understanding of cells, with implications for human health. I focused on mechanistic models of processes involved in growth and ATP synthesis. I was particularly interested in how biophysical constraints shape the fitness landscape of the cell and how phenotypes emerge from the cells' attempts to optimize metabolic fluxes. Much work has already been done in this area and preexisting methods and models were modified to improve their scope and quantitative accuracy.

In this thesis (**Paper I**), I conducted the first genome-scale metabolic simulations of growth in a healthy infant. Several mathematical models of infant growth have been developed in the past (Fjeld et al. 1989; Butte 2005; Jordan & Hall 2008). These models do, however, not explicitly model the metabolic fluxes of the cells. I took a data driven (top-down) approach to characterize the metabolic state of the infant. Various published datasets were integrated to estimate the metabolic fluxes and the growth rate. I found that growth was primarily constrained by the ATP synthesis requirements, which justifies the assumptions of the preexisting models.

In this thesis (**Paper II**), I conducted the first genome-scale metabolic simulations of a cancer cell line constrained with data from a dedicated experiment. There already exist several genome-scale metabolic models of cancer cells (Shlomi et al. 2011; Agren et al. 2012; Agren et al. 2014), as well as models of other mammalian cell types, e.g. hepatocytes (Mardinoglu et al. 2014), adipocytes (Mardinoglu et al. 2013), myocytes (Väremo et al. 2015), hybridoma cells (Sheikh et al. 2005) and CHO cells (Martínez et al. 2012). However, metabolic models of cancer that make use of experimental data have been limited in scope (Fan et al. 2013; Fan et al. 2014; Zielinski et al. 2017; Quek et al. 2014). There are several reasons for this. There are documented problems with model consistency for some genome-scale models (Swainston et al. 2016). Many popular algorithms and methods cannot be applied to genome-scale models (Erdrich et al. 2015). And it may be easier to interpret the results of a smaller model. Here I expanded the scope of cancer simulations to the genome-scale with a model that included multiple cellular compartments. This was instrumental to explain why cancer cells excrete glutamate.

When I integrated the experimental data on metabolite concentrations I noticed that some metabolites had been depleted before the experiment was completed. I therefore developed methods to account for depletion events in an appropriate way. I expected that other studies also may have encountered metabolite depletion. In this thesis (**Paper III**), I reanalyzed the data from a published high throughput flux characterization experiment (Jain et al. 2012). The reanalysis showed that growth enhancing metabolites had been depleted before the samples were taken in many of the cell cultures. This appears to have both qualitative and quantitative effects the characterized fluxes.

In this thesis (**Paper IV**), I developed the first fully parameterized enzyme-constrained (bottom-up) model of intermediary metabolism in yeast. Several metabolic models with enzyme constraints exist for *E.coli* (Beg et al. 2007; Adadi et al. 2012) and yeast (van Hoek & Merks 2012). However, these models either rely on randomly sampled parameter values, or are only partially parameterized with experimental values. By limiting the scope of the model to intermediary metabolism, and employing extensive search in literature and databases, it was possible to assemble a complete set of kinetic parameters. The model showed an increased use of glycolysis at high ATP synthesis rates, i.e. overflow metabolism, thus corroborating the hypothesis that fermentation has a higher catalytic efficiency than respiration. From the simulations it was possible to infer that the low efficiency of F_1F_O-ATP synthase was the most influential biophysical constraint behind the phenomenon. The model was later expanded to the genome-scale level in another project (Sánchez et al. 2017).

In this thesis (**Paper V**), I developed a fully parameterized enzyme-constrained (bottom-up) model of intermediary metabolism in human muscle. Several metabolic models with enzyme constraints exist for human (Shlomi et al. 2011; Vazquez & Oltvai 2011). However, these model are only partially parameterized. The fully parameterized model enabled a more detailed analysis of the constraints involved in ATP synthesis. Unlike yeast, human muscles do not have time to adjust their proteome during ATP synthesis. By integrating protein abundance measurements with the enzyme parameters, more realistic constraints were calculated for the fluxes. Phenotypes of multicellular organisms depend on the metabolic interaction between multiple cell types. I therefore expanded the model to a multi-tissue model. Previous multi tissue models of human (Bordbar et al. 2011) do not use enzyme-constraints, and were not used to study exercising humans.

The use of enzyme-constrained models relies heavily on parameter values from literature. In this thesis (**Paper VI**), I proposed that a systematic characterization of all kinetic parameter values, i.e. the kinetome, is necessary for enzyme-constrained models to reach their full potential.

Methods

Genome-scale metabolic modeling of growth using flux balance analysis is a well-established method (Palsson 2015; Orth et al. 2010). It relies on a metabolic network with a biomass equation. The use of interior constraints, e.g. enzyme capacity constraints, enables quantitative predictions under a condition of interest. If such constraints are not available, or are deemed too imprecise, experimentally derived boundary constraints may be used to predict the flux state of the cell. Sensitivity analysis may then be used to infer the relative importance of the fluxes and predict the effects of perturbations.

Network reconstructions and the biomass equation

The metabolic network is central to a GEM. Community efforts have resulted in several generic metabolic models of human and yeast (Mardinoglu et al. 2013; Brunk et al. 2018; Förster et al. 2003). These models can be reduced to a model of a cell type or a tissue by removing reactions without support in transcriptomics or proteomics data (Agren et al. 2014; Agren et al. 2012). Additional manual curation is often required, and the models are commonly reduced in size to give a more manageable set of reactions (Zielinski et al. 2017; Quek et al. 2014). For simulations of human in this thesis (**Paper I, III, V**), the networks originated from the generic human reconstruction, HMR 2.0 (Mardinoglu et al. 2013). A multi-tissue model (**Paper V**), was also constructed for different muscle tissue types, based on a reconstruction of the myocyte (Väremo et al. 2015). For simulations of yeast (**Paper IV**), I use a reduced version of the iFF708 reconstruction (Förster et al. 2003; Agren et al. 2013). The biomass equation is a crucial component of metabolic models designed for quantitative flux predictions.

The biomass equation contains all the components required to synthesize a cell (Figure 4A), e.g. proteins, lipids, nucleotides and glycans. It is an artificial reaction which acts as a sink for the fluxes and has a different unit (gdw/gdw/h) than the other fluxes (mmol/gdw/h). Its stoichiometric coefficients therefore contains an implicit change of units (gdw/mmol). The full biomass equation contains a large number of metabolites and it may therefore be convenient to split it into sub modules. For small-scale models, few of the biomass components are in the metabolite list. In this thesis (**Paper IV**) the drain of metabolites required for growth in a small-scale model was identified based on simulations in a genome-scale model (iFF708). The protein fraction is normally the dominating component in the biomass equation (by mass). The amino acid composition of the protein fraction must be known to accurately represent it. In this thesis (**Paper II**) this was estimated using protein abundance data and the amino acid frequencies of the proteins.

A cells composition depends on the growth rate. The biomass composition and the ATP requirement for growth are often taken as constant. However, in the living cell they commonly depend on the growth rate (Dikicioglu et al. 2015). This is thought to influence the model predictions. In this thesis, condition dependent biomass equations were used. For the infant simulations (**Paper I**), the age dependent fat fraction in newly formed biomass was estimated from body composition data (Figure 4B). The energy requirement for growth was dynamically calculated from the composition, 5.5 kcal/g protein and 1.6 kcal/g fat. For the yeast simulations (**Paper IV**),
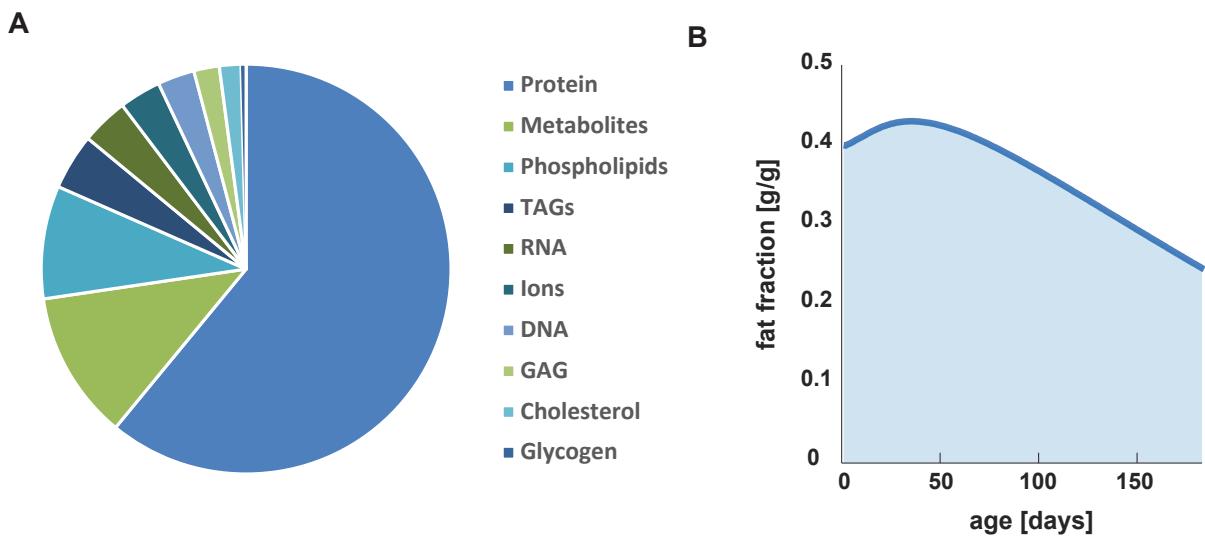
Figure 4. Composition and dynamics of the biomass equation. A) The biomass composition of a liver cancer cell line was assembled from literature sources. B) The fat fraction in newly synthesized infant tissue is a function of age.

experimental data was used to adjust the biomass equation to the growth rate. The ATP expenditure required for growth was also adjusted depending on the growth rate to stay consistent with the rates calculated from exchange fluxes.

Metabolites are diluted in growing cells. It is reasonable to assume that the metabolite concentrations of the cell fluctuate around some mean value. However, the steady state assumption neglects the dilution term associated with growth (Benyamini et al. 2010; Dikicioglu et al. 2015). When the mass of cells double, so must all metabolites to avoid depletion. Since the flux over most metabolite pools is much larger than the dilution of the pools, this assumption has limited effect for most predictions. But as the precision of simulations improve, the effect could become noticeable. The problem may be overcome by adding the metabolite pool to the biomass reaction, or adding dilution coefficients to each reaction. In this thesis I added the metabolite pool to the biomass reaction (**Paper II**) or neglected the dilution term altogether (**Paper I, IV**).

The wide range in concentrations of the biomass components may cause numerical issues for the computer. The optimization problem in FBA is solved by a linear programming algorithm. These algorithms commonly rely on floating point arithmetic (numbers are expressed as a significand and a base), which is a compromise between precision and range. Since the significand is expressed with a limited number of bits, rounding errors may occur when applying some arithmetic operations, e.g. division, and some values, e.g. 1/3, cannot be represented exactly. The use of LP solvers for FBA problems has been criticized, with the argument that the rounding errors may result in incorrectly predicted feasibility (Chindelevitch et al. 2014). An alternative is to use an algorithm with rationals (Applegate et al. 2007), but this is normally slower than to increase the precision of the solver. The magnitude of the (non-zero) values in the solution to a FBA problem may span 11 orders of magnitude. The upper bound of the fluxes is 10^3 by default and the lowest flux is typically given by the concentration of a metabolite in the biomass equation multiplied with

the growth rate. Metabolite concentrations may be as low as 10^{-6} (Park et al. 2016), and a typical growth rate is around 10^{-2}. In this thesis (**Paper II**), metabolites with a concentration lower than 10^{-4} were removed to retain fast solutions and avoid numerical issues.

Interior constraints, enzyme-constrained models

Interior flux constraints are useful to understand the behavior of cells. These constraints may come in the form of experimentally derived flux capacities ($v \leq v_{max}$). These capacities may be calculated ($v_{max} = [E] \cdot kcat$) from enzyme parameters and abundance data (Sánchez et al. 2017). It is possible to identify which of the interior constraints are active under the studied condition, i.e. which fluxes are equal to the constraint ($v_i = [v_{max}]_i$). These constraints are condition specific since they are experimentally derived, and may not necessarily reflect a global constraint. In this thesis (**Paper V**), I make use of interior flux constraints calculated from experimental data.

The internal fluxes may also be constrained by a limited enzyme pool (Figure 5). The enzyme requirement for each metabolic flux (v_i) is estimated by multiplication with a weight (w_i). Models with enzyme pool constraints enable simulations of resource allocation. The predicted reallocation of enzyme abundances between conditions is thus a predictor of metabolic regulation. Since the enzyme abundances are not fitted to a particular condition these models reflect global constraints on growth. Various nutritional conditions may be simulated by opening or closing the boundary fluxes. This results in different metabolic strategies (A, B, …), which may be referred to by their corresponding flux vector (v^A, v^B, ...). The catalytic efficiency of the strategies may be compared as $w^T(v^A - v^B) \leq 0$. By applying elasticity analysis to the weights, their relative contribution to the efficiency of a strategy may be investigated. In this thesis (**Paper IV, V**), I examined the constraints on ATP synthesis and growth using enzyme pool constraints.
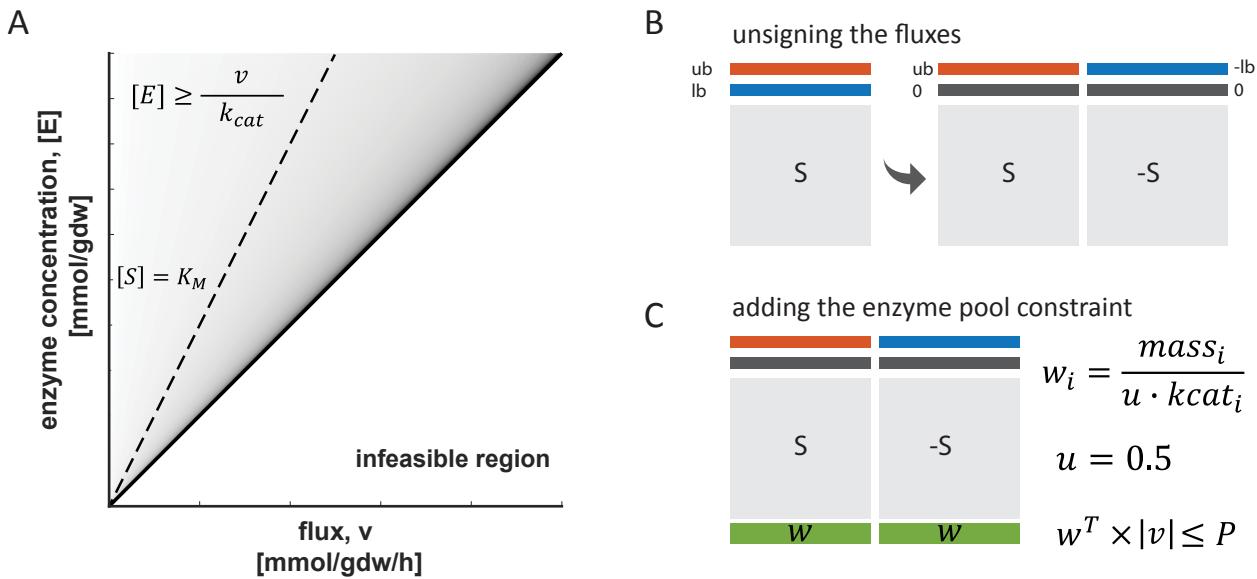
Figure 5. Enzyme-constrained models. A) The concentration of an enzyme ([E]) required for a flux (v) depends on the turnover rate of the enzyme (k_{cat}) and the degree of enzyme saturation. A direct relationship between flux and enzyme requirement can be calculated (dashed line), assuming Michaelis–Menten kinetics and a substrate concentration ([S]) equal to the Michaelis constant (K_M). B) The model must be converted to an unsigned format to ensure that all fluxes are positive. The matrix S is multiplied with -1 and concatenated with S. Lower bounds are set to zero for all fluxes and the upper bounds for the concatenated matrix are set to -1 times the lower bound. C) An enzyme pool constraint is added. A vector of weights (w) may be calculated for each reaction (i) to translate flux to the required enzyme mass concentration. To calculate the weights, the enzyme saturation (u) must be estimated.

Sensitivity analysis with boundary constraints

Metabolic models with boundary constraints make predictions about the interior fluxes. These constraints are empirical, not biophysical, and the observed boundary fluxes arise from interior constraints in one, several or all reactions. To make predictions about counterfactual states, e.g. the predicted maximum growth rate on a different carbon source, the fluxes must be translated to constraints. Elasticity analysis (Figure 6A) estimates the effect on growth from perturbations in a flux ($E_{v_i}^{\mu}$). The predicted flux is taken as reference, and the flux is scaled by some arbitrarily lower value (1-ε) and constrained. The growth in the perturbed state (μ_i) is compared with non-perturbed growth (μ).

$$E_{v_i}^{\mu} = \frac{1 - \mu_i/\mu}{\varepsilon}$$

Growth is expected to be sensitive to perturbations in reactions with active constraints. But a high predicted sensitivity does not necessarily imply an active constraint. The predicted flux may reflect parsimonious nutrient utilization by the cell and the enzyme may have plenty of excess capacity, or may be regulated to provide this.

The predicted sensitivity of many reactions is often low due to redundant pathways. These pathways may however not always be physiologically relevant, e.g. the flux capacity of the pentose phosphate pathway may be insufficient to circumvent a block in glycolysis. The predicted flux through a reaction under a favorable condition is a rough estimate of the effective v_{max} of the
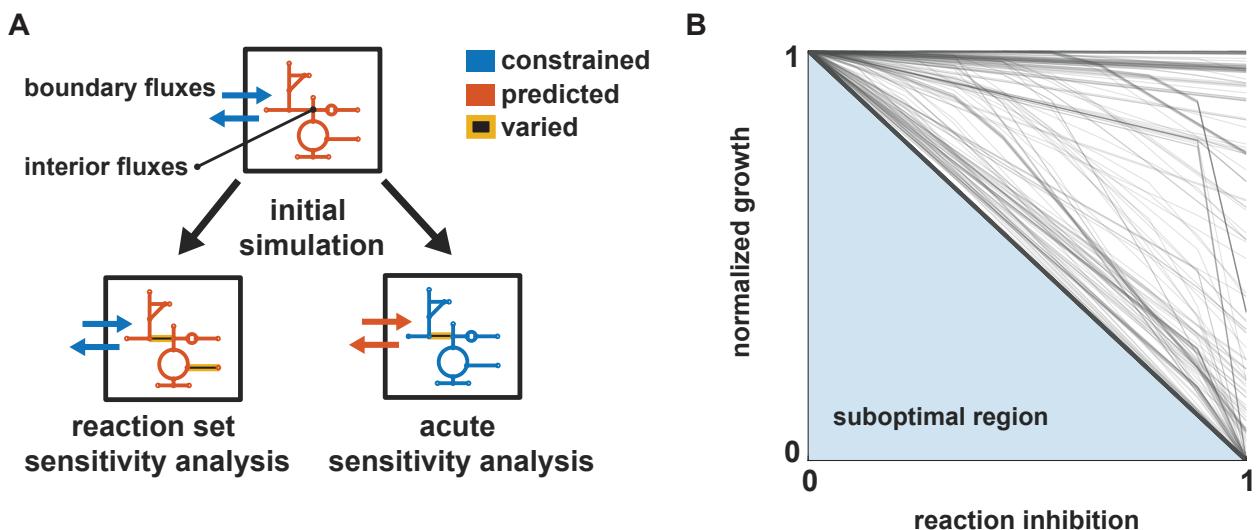
Figure 6. Assumptions about interior constraints for sensitivity analysis. A) The interior fluxes are estimated from boundary constraints. Interior fluxes are assumed to represent the effective v_{max} for a set of reactions or for all reactions. The elasticity of perturbations is then investigated. B) The linearity of the problem ensures that $E_{v_i}^{\mu} < 1$ for all reactions, and $E_{v_i}^{\mu} = 1$ for essential reactions. This assumes that there are no maintenance tasks. These may make the solution infeasible before a flux is completely inhibited and may thus allow $E_{v_i}^{\mu} > 1$.

reaction. A conservative approach may then be to use all predicted fluxes as capacity constraints (Figure 6). Although this is expected to prevent non-physiological rerouting, it may also prevent physiological rerouting. It can therefore be thought of as the acute adaptation to a perturbation, and later the flux landscape may be remodeled by transcriptional and regulatory changes. In this thesis (**Paper II**), this type of acute sensitivity analysis (ASA) was applied to study inhibition of glycolysis and other ATP synthesizing pathways.

From the optimality assumption, it is expected that growth should be sensitive to perturbations in all reactions. In practice this may not be the case since some fluxes are forced to fit experimental data. If the fluxes are small this may reflect experimental error, but if they are substantial, this may instead indicate an incomplete description of the objective of the cell. In this thesis (**Paper II**), I identify instances of this type of flux.

Results

In this thesis, I study ATP synthesis and growth in computer models of yeast and human. The models of infant and liver cancer used boundary constraints to predict the maximum specific growth rate and to estimate the interior fluxes. The models of yeast and muscle used interior constraints to study metabolic trade-offs and the maximum metabolic capacity on different substrates. The different models operated at specific growth rates and ATP synthesis rates that differed by orders of magnitude. This affected their responses to perturbations in energy metabolism. Nevertheless all models showed increased utilization of glycolysis at high ATP synthesis rates. Overflow metabolism from glycolytic substrates to fermentation products was found to be an efficient metabolic strategy. The cancer cells also excreted glutamate. The simulations related this to compartmentalization of glutamate metabolism.

Constraint-based models of metabolism

Simulations of metabolism were conducted in four different models, an infant, a cancer cell line, yeast and muscle. A model of high quality is a requisite for causal inferences. In this section I present the models together with some fundamental quality measures and principal results. In the next section I then proceed to analyze the phenotypic predictions arising from the models.

Growth of an infant on breastmilk (Paper I)

Healthy infants more than double their weight during their first 6 months in life. Growth may be challenged by complications in retention or availability of nutrients (Prendergast & Humphrey 2014). In severe cases, stunting syndrome may develop, i.e. very short length for the age. This primarily occurs in developing countries, and cost-efficient interventions would therefore be advantageous. Metabolic models provide insight on constraints on growth with reaction-level resolution. However, current models of infants (Fjeld et al. 1989; Butte 2005; Jordan & Hall 2008) do not explicitly model the metabolic fluxes. The relatively high growth rate, the uniform diet, i.e. breastmilk, and the abundance of experimental data make infants suitable for metabolic modeling.

A model of the growing infant was reconstructed. It was based on the HMR 2.0 model, which was modified to include age dependent reactions for biomass synthesis, energy expenditure and nutrient uptake. The biomass of the infant was broadly classified as fat and lean mass, which was represented by the free fatty acid composition and the amino acid composition of infant. The fat fraction in newly formed biomass decreased from around 40% during the first 3 months to around 25% at month 6 (Fields et al. 2011). The energy expenditure of the infant involved the requirements for growth, maintenance and physical activity and depended on body composition and age. The uptake of nutrients by the infant was calculated from the age dependent metabolite composition of breast milk and the daily intake rates (Butte et al. 2002; Hester et al. 2012; Yamawaki et al. 2005; Peng et al. 2007), and was used to constrain the boundary fluxes. Growth was then simulated by maximizing the synthesis of biomass.
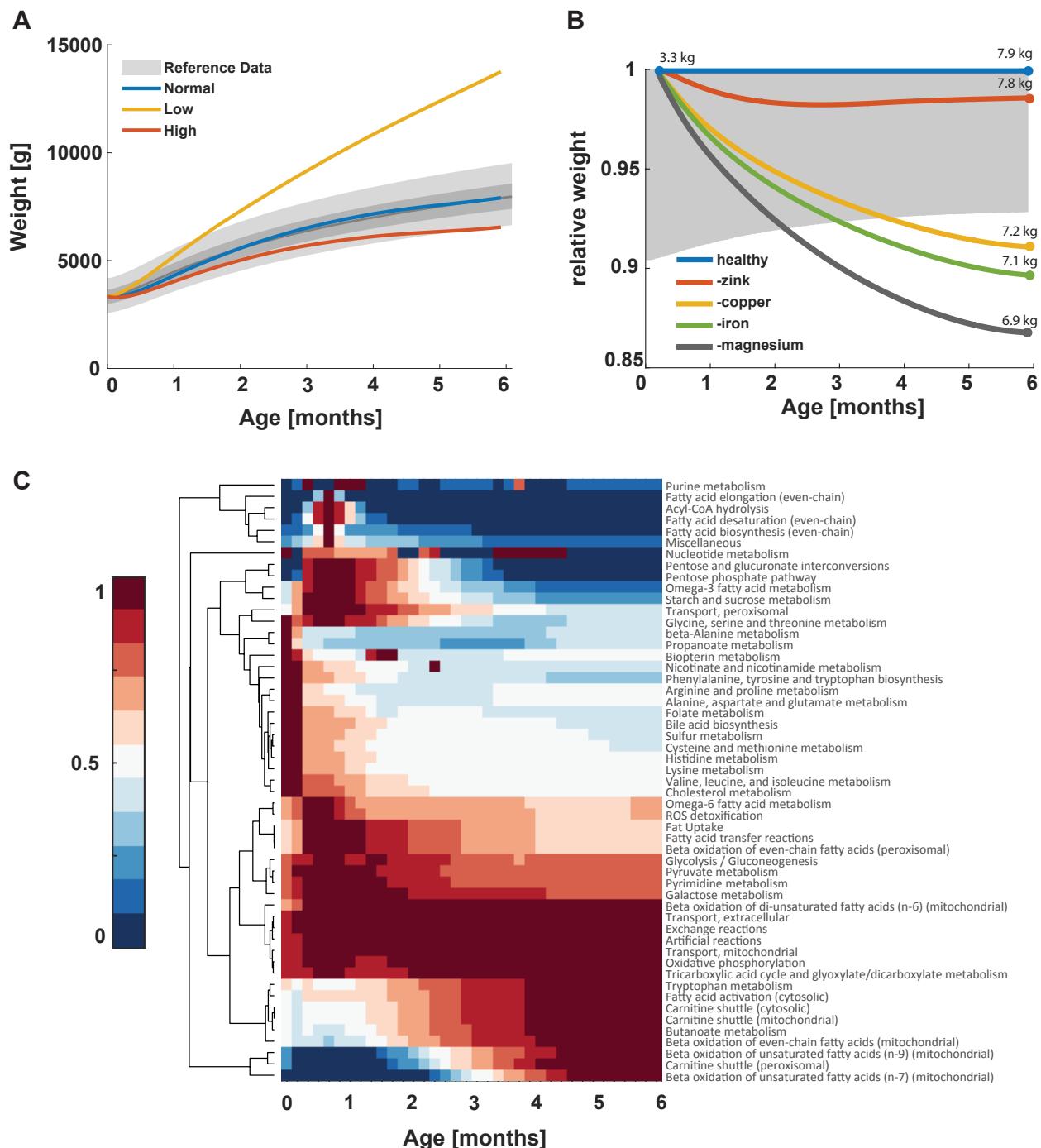
Figure 7. Growth of the infant during the first 6 months. A) The predicted growth for the infant model depends on the energy expenditure. Growth trajectories with normal, high and low energy expenditure was compared to reference data (World Health Organization 2009) at different percentiles (5,25,50,75,95). The simulations assumed that there was no compensatory increase in breastmilk intake. B) The predicted growth of infants was sensitive to mineral deficiencies. The growth effects were simulated for a 5% reduction in flux through reactions that depend on various mineral cofactors. Several predicted weights were below the 25-50th percentile in reference data (shaded gray area). C) The mean absolute flux in different pathways at increasing ages, normalized by the highest value in each pathway.

16

Growth was found to be energy limited. The predicted growth was in good agreement with reference data (Figure 7A). The growth rate was constrained by the intake of one or several different types of nutrients, the model was used to identify which was the active constraint. Some amino acids cannot be synthesized by the body, i.e. essential amino acids, and must be provided by the diet. Some fatty acids are also essential. The total intake of nitrogen must be sufficient to cover the synthesis of non-essential amino acids and nucleotides. Finally, the intake of nutrients for ATP synthesis may be limiting, i.e. calorie restriction. The sensitivity analysis revealed that growth was energy limited, this could be illustrated by removing the maintenance energy requirement. If other factors than energy were limiting, growth would not increase in this hypothetical scenario. However, the predicted weight more than doubled. Additionally, a small increase (20%) in maintenance expenditure was sufficient for the predicted weight to drop below the 5th percentile within 6 months. Other factors than nutritional may still influence growth.

Simulated mineral deficiencies decreased growth. Mineral deficiencies are known to cause growth defects in animals, and have been investigated as a cause of stunting (Yang et al. 2012; Castillo-Duran & Uauy 1988; Hill et al. 1983). I therefore investigated how mineral deficiencies could affect growth in the infant. Annotation data from a protein database (Bateman et al. 2017) was used to identify sets of reactions that shared the same mineral as a cofactor. The flux through these reactions was then reduced to 95% of the optimal flux. Cofactors that were involved in energy metabolism had the strongest predicted effects on growth (Figure 7B). This was an expected result based on the high sensitivity of growth to perturbations in energy metabolism at low specific growth rates. Reducing the flux further resulted in infeasible solutions at the later ages due to an insufficient capacity to synthesize ATP, to meet the maintenance requirement.

The model predicted the interior fluxes of the infant. The changes in flux in different subsystems were predicted over time (Figure 7C). During the first days they indicated catabolism of amino acids, due to insufficient milk intake. Over the first two months they shifted towards biosynthesis of fatty acid and metabolism of sugars. And during the last three months, metabolism was centered on beta oxidation of fatty acids. This illustrates the capacity of GEMs to provide high resolved predictions in time with regards to metabolism.

Growth of cancer cell lines on medium (Paper II)

An integrative understanding of metabolism may promote increased cancer survival. Cancer survival is currently increasing in developed countries due to implementation best practices and management (Bray et al. 2018). But future improvements depend on the development of new treatments (Adams et al. 2015). Many human diseases involve reprogramming energy metabolism, and it is considered an emerging hallmark of cancer (Hanahan & Weinberg 2011). Cancer metabolism is the subject of many models (Shlomi et al. 2011; Agren et al. 2012; Agren et al. 2014; Fan et al. 2014; Zielinski et al. 2017), however these are either small in scope or do not make use of experimental data. We still lack an integrative view on metabolism in cancer (Palm & Thompson 2017). Pending questions include: which metabolites are consumed, and at which rates, what is their metabolic fate once consumed and why are some metabolites excreted.
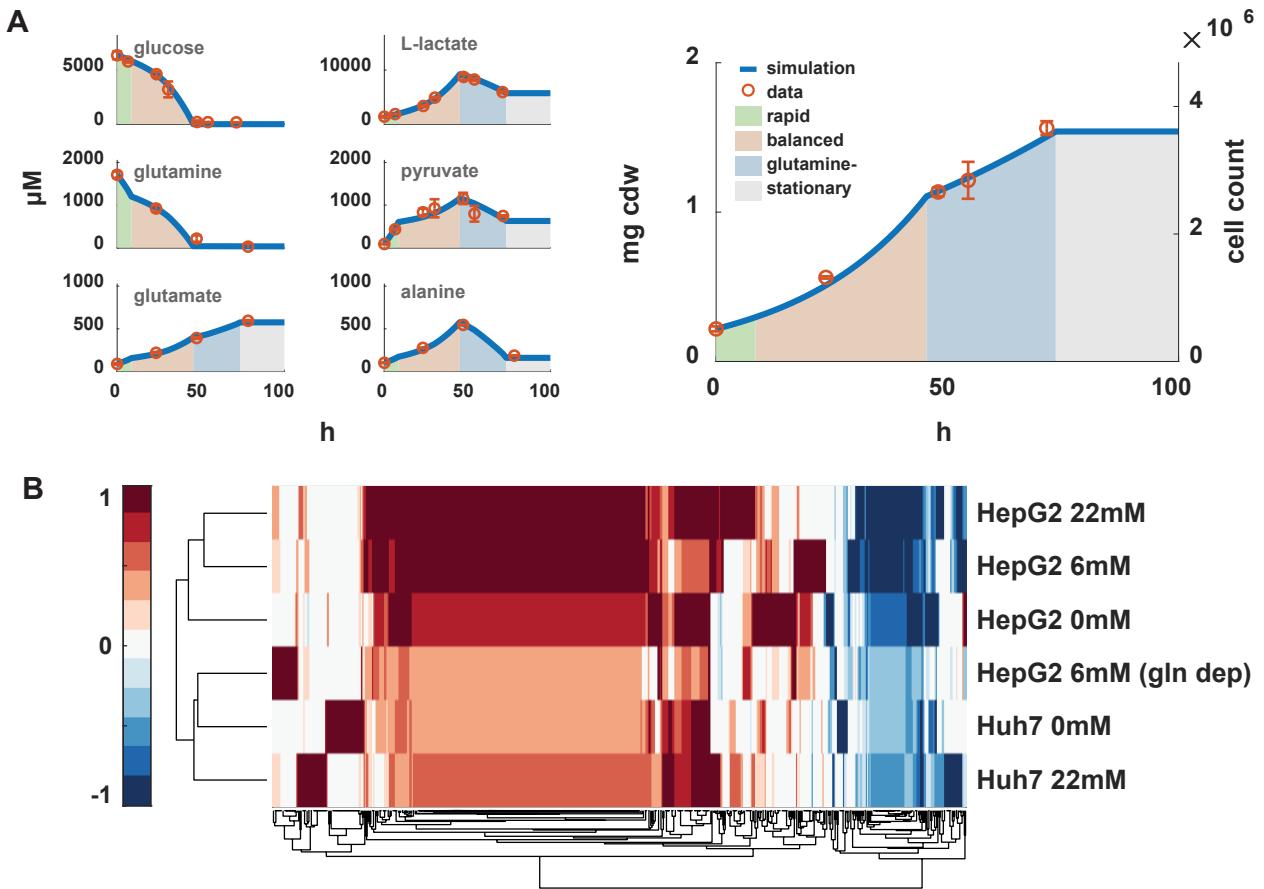
Figure 8. Growth of a liver cancer cell line (HepG2) on modified growth medium (DMEM). A) Experimental fits to the time dependent changes in metabolite concentrations and cell counts. The data was divided into four metabolic phases with different specific exchange fluxes, based on metabolite depletion events. B) The normalized predicted fluxes cluster by cell line and growth condition (0, 6 or 22 mM initial glucose) using hierarchical clustering.

A model of liver cancer was reconstructed. The model was based on HMR 2.0 with some additional manual curation and removal of reactions without support in mRNA sequencing data from a liver cancer cell line (HepG2). The biomass equation consisted of 60% protein, 15% phospholipids and fats, 12% metabolites, 6% RNA and DNA and 2% glycosaminoglycans. Data on protein abundances and amino acid frequencies was used to estimate the amino acid composition of the protein fraction. Metabolite abundance data was used for the composition of the metabolite fraction.

The exchange fluxes were charaterized. A liver cancer cell line (HepG2) was cultivated in standard growth medium (DMEM). Concentrations of glucose, pyruvate, lactate and most amino acids were measured over time. Specific exchange rates were fitted to the data and used as boundary constraints for the model. The specific growth rate was predicted by maximizing the biomass. The predicted growth was used to calculate metabolite trajectories, and exchange fluxes were refitted until both predicted cell counts and metabolite concentrations were in acceptable agreement with the experimental data (Figure 8A). Several metabolites were depleted over the course of the cultivation, which was expected to affect the magnitude of the fluxes. Unique sets of specific

exchange fluxes were therefore fitted for each growth phase between the depletion events. An additional growth phase was introduced at the first hours of cultivation to account for the rapid disappearance of several metabolites during this period. To fulfill the carbon balance under this phase a protein secretion flux was fitted as an aditional sink for amino acids.

The predicted fluxes followed an anticipated pattern. Hierarchical clustering (Figure 8B) of the predicted fluxes separated the experiments by cell lines and placed similar conditions adjacent to each other. Huh7 cells clustered together with the glutamine depleted simulations, reflecting a lower dependence on glutaminolysis in this cell line. The fluxes related to biomass synthesis constituted a large share of the fluxes and had a similar pattern for all cell lines reflecting the differences in specific growth rate. Most fluxes were strictly in the direction defined as forward in the model (76%) and only a small set of fluxes (7%) changed direction between conditions.

Enzyme constraints on growth of yeast (Paper IV)

Yeast is well suited for metabolic modeling. The yeast *S.cerevisiae* is a model organism, and can be grown under controlled conditions. It is thus well suited to study fundamental question about metabolism. There is a rich collection of experimental measurements in different setups available in literature, e.g. protein abundances, metabolite concentrations, growth rates and fluxes (de Godoy et al. 2008; E Postma et al. 1989; Larsson et al. 1993; Boer et al. 2010). Yeast undergoes the *Crabtree effect*, where cells growing above a critical growth rate switch from pure respiration to a mixture between respiration and fermentation. This drastically reduces the yield of biomass on glucose, i.e. the amount of biomass produced per glucose consumed. A model of yeast metabolism with randomly sampled parameter values had previously demonstrated that this may be related to enzyme-constraints (van Hoek & Merks 2012). It remained unclear if these results would persist using parameter values from literature, however previous studies in microbes and humans using such parameters suggested that predictions may improve (Adadi et al. 2012; Shlomi et al. 2011).

To study the Crabtree effect, an enzyme-constrained model of yeast was reconstructed. The model was based on a small-scale model of intermediary metabolism in yeast (Agren et al. 2013), which was reconstructed based on the iFF708 model (Förster et al. 2003). Enzyme constraints were incorporated into the model based on parameter values from literature. A vector of conversion factors (w) from reaction flux (v) to enzyme mass was constructed. The factors were calculated from the reported specific activity of the purified enzymes (μmol/mg protein/min) and an assumed saturation. In this case it was assumed that all enzymes operate at half their maximal rate. A constraint was introduced such that the sum of mass of the enzymes involved in the model ($w^T|v|$) was equal to or less than the observed sum of mass (P) in proteomics measurements, around 0.1g/gdw.

The simulation results were in good agreement with experimental data. The growth of yeast in a glucose limited chemostat was simulated (van Hoek et al. 1998). A chemostat is a device where cells are cultivated in a tank, fresh medium is supplied at a constant rate (F), and the culture volume (V) is held constant by removal of excess volume. The specific growth rate of the cells (μ) then becomes equivalent to the dilution rate ($\mu=D=F/V$) at steady state and can thus be experimentally controlled. The optimal fluxes at different specific growth rates were simulated by constraining
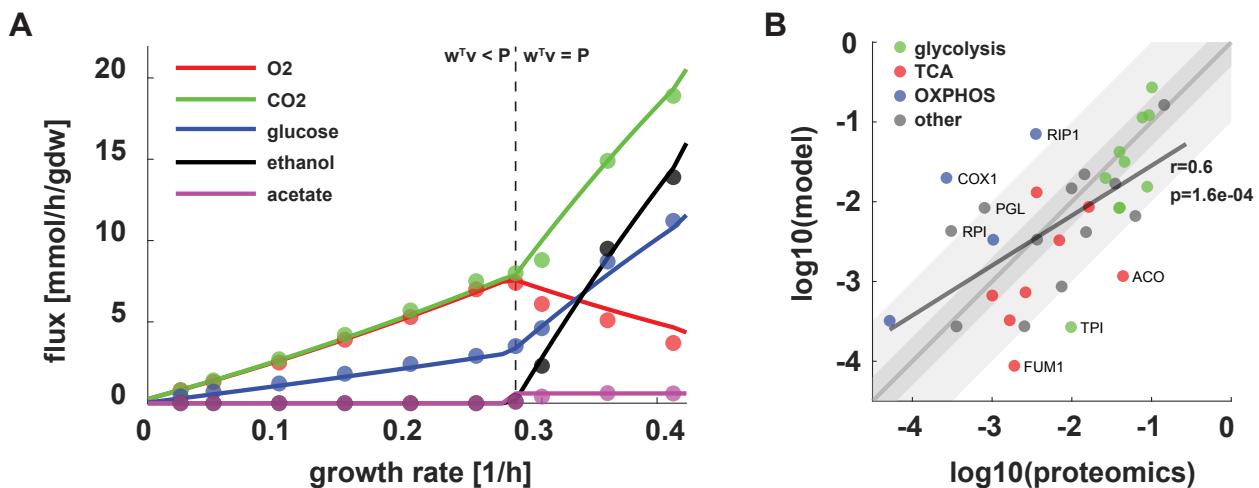
Figure 9. Predicted exchange fluxes and protein concentrations for yeast grown on glucose. A) Good agreement between predicted (lines) and observed (dots) exchange fluxes at different specific growth rates. The predicted sum of enzyme mass (w^Tv) reaches the protein constraint (P) at a specific growth rate of 0.28. B) Acceptable agreement (Pearson correlation 0.6) between predicted and observed enzyme concentrations, during the exponential growth phase.

growth and minimizing the specific glucose uptake rate. The predicted exchange fluxes were in good agreement with experimental values (Figure 9A). The model could accurately predict the shift from respiratory metabolism at low specific growth rates ($\mu < 0.28$) to respiro-fermentative metabolism at high specific growth rates. In the model this occurred when the predicted sum of enzyme mass reached the protein constraint. An empirical constraint on acetate production was required to fit the fluxes.

The models description of NADPH synthesis was incomplete. The excretion of acetate was predicted to originate from increased flux through aldehyde dehydrogenase (ALD6). The enzyme ALD6 consumes acetaldehyde to produce acetate and NADPH. The increased flux, reflects a shift in the strategy for synthesizing NADPH. The exchange of acetate was constrained to the highest observed secretion level, to prevent higher synthesis rates than observed. Unconstrained acetate production would only have had minor effects on the overall fluxes, but the requirement for an empirical constraint indicated that the model of NADPH synthesis at high growth rates was incomplete.

The predictions rely on approximated enzyme abundances. These have been experimentally determined during exponential growth in batch conditions (de Godoy et al. 2008). Under these conditions, nutrients are available in excess and cells grow at their maximum rate until some nutrient is depleted. I simulated exponential growth under batch conditions and the predicted enzyme abundances were compared with proteomics data (Figure 9B). The predictions were in reasonable agreement with the experimental values, given the crude underlying assumptions. Enzymes from the TCA cycle were generally under predicted and oxidative phosphorylation was in general over-predicted. This may reflect differences in enzyme saturation, e.g. higher saturation in glycolysis compared to the TCA cycle, due to it being a linear pathway. But it may also reflect

experimental challenges in accurately quantifying proteins that originate from different cellular compartments or are attached to membranes.

Enzyme constraints on the activity of muscle (Paper V)

Muscles are tailored towards ATP synthesis. In the context of human endurance sports, high sustained ATP production is one of the most important competitive advantages (Peronnet & Thibault 1989). The working muscle must adjust the rate of ATP synthesis to match the work rate. The metabolism of the muscle can thus be experimentally controlled by adjusting the resistance (Poole & Richardson 1997). At high exercise intensities the muscles begin to produce lactate and there is a trade-off between endurance and power output (Grassi et al. 1999; Peronnet & Thibault 1989). Previous metabolic models of muscles, make qualitative predictions about these phenomena (Vazquez & Oltvai 2011). However these models do not include parameter values for the individual enzymes.

An enzyme-constrained muscle model was reconstructed. A model of intermediary metabolism in a muscle fiber was reconstructed from a subset of HMR 2.0. The subset consisted of the reactions that remained after removing reactions that did not carry flux when simulating ATP synthesis under various conditions, e.g. absence of complex I and V and absence of oxygen. Protein mass abundance data (E) and the specific activity (a) of the purified enzymes were used to calculate the maximum theoretical flux for each enzyme mediated reaction in the model ($v_{max}=E \cdot a$). The reactions were constrained to half the maximum flux to reflect saturation effects.

The simulations were in good agreement with experimental data. The optimal fluxes for the muscle fiber were simulated at different ATP synthesis rates. This was done by constraining the ATP synthesis rate and minimizing the specific glycogen depletion rate. The predicted maximum oxygen uptake rate occurred when complex IV reached the saturation constraint. This was in good agreement with the maximum oxygen uptake rate observed in muscle fibers (Larsen et al. 2012). The model predicted a shift from a mode with full respiration to a mode where complex I was saturated and thus bypassed. The oxygen uptake rate where this occurred was in good agreement with observed oxygen uptake rates for muscle fibers restricted to NADH-producing substrates.

Fermentation increased the ATP synthesis rate. The predicted exchange fluxes (Figure 10A) agreed with known physiology (Grassi et al. 1999), where metabolism was fully respiratory at low ATP synthesis rates, and fermentative at high rates. The maximum ATP synthesis rate (19.9 mmol/gdw/h) occurred when GAPDH was saturated. This occurred at ATP synthesis rates around 3.6 times higher than the rate at which oxygen uptake was saturated. This is in good agreement with the relatively higher (3.2 times) maximum work rate (1130 W) compared with the maximum respiratory steady state rate (350 W) for exercise on cycle (Elmer et al. 2011). A trend towards higher predicted flux in reactions with high v_{max} was also noted (Figure 10B).
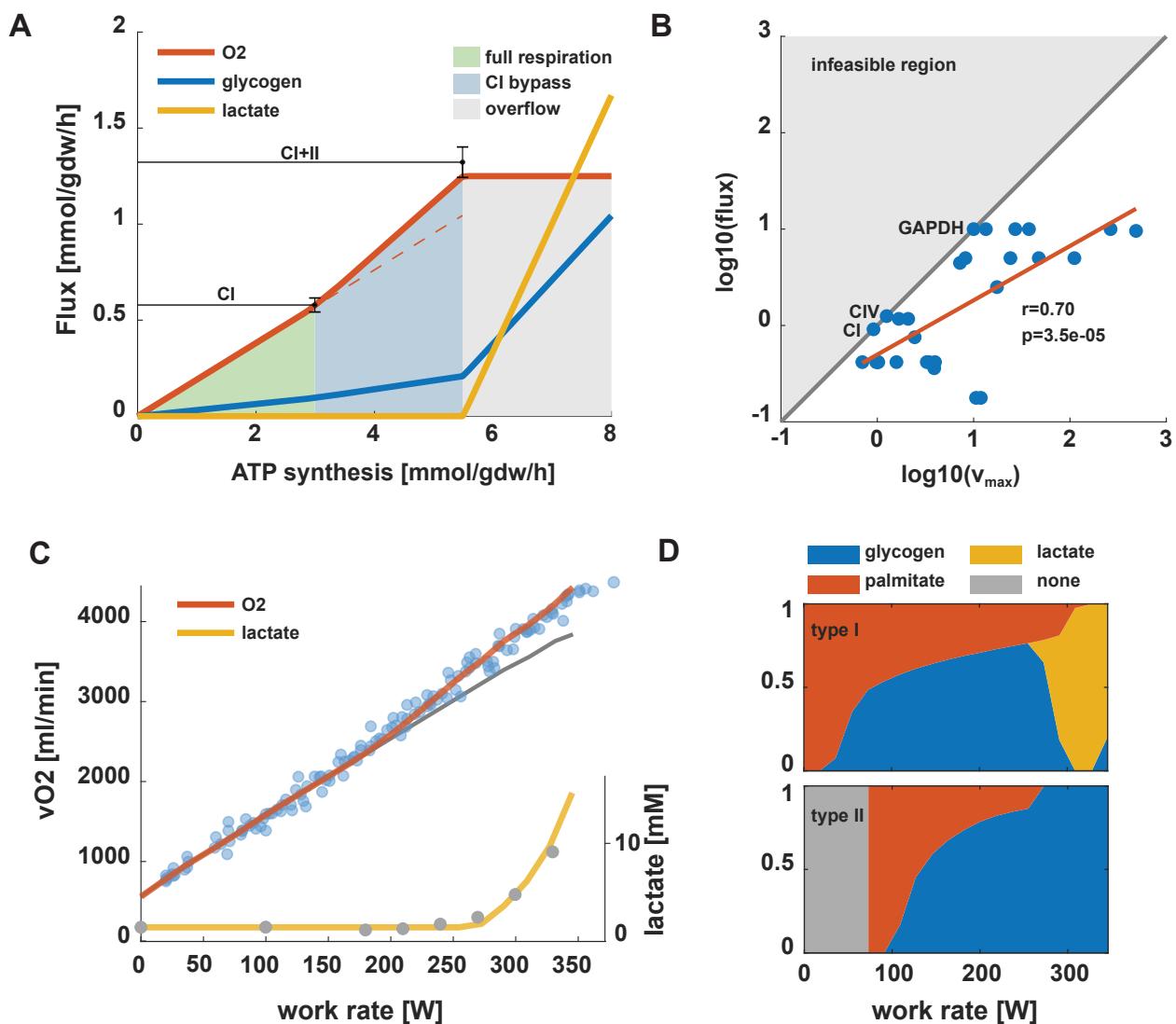
Figure 10. Predicted and experimental fluxes for a single muscle fiber and a human exercising on a bike. A) Predicted exchange fluxes at different ATP synthesis rates in a single muscle fiber. There was good agreement between predicted and observed maximum oxygen consumption rate in isolated muscle fibers on NADH substrates only (CI) and NADH and QH_2 substrates combined (CI+II). The predicted oxygen consumption would be underestimated in the absence of a complex I constraint (dashed line). B) Acceptable agreement (Pearson correlation 0.7) between maximum flux and the highest predicted flux amongst all ATP synthesis rates. Complex I (CI), complex IV (CIV) and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) were rate limiting. C) Predicted oxygen flux and lactate concentration (lines) were in good agreement with experimental data (points). The predicted oxygen flux without complex I bypass (gray line) shown for reference. D) There is a shift from fatty acid metabolism towards glycolysis at high work rates.

Prolonged fermentation at high rates is unsustainable, since lactate accumulates in the human body. However, at intermediate fermentation rates, lactate produced by one cell type may be used to fuel another (Brooks 2009). The metabolic interaction between multiple cell types can be studied using a multi tissue model (Bordbar et al. 2011; Shaw & Cheung 2018). Previous multi tissue models of human do not use enzyme-constraints, and have not been applied to study exercising humans. A model of the muscle system of the full body was reconstructed to study these dynamics. The model

consisted of 3 genome-scale models that were joined together by an additional compartment. This allowed exchange of extracellular metabolites between the sub-models, thus resembling the blood stream. The sub-models were parameterized to represent muscle fibers with high mitochondrial content (type I fibers) and with intermediary mitochondrial content (type II fibers). The third sub-model represented peripheral tissues and accounted for metabolic events that do not take place in the active muscle. Parameters were fitted to constrain the flux through complex I and complex IV, the reactions with active constraints in the muscle fiber model. A maintenance energy expenditure rate was also fitted based on the oxygen expenditure of the subject at rest.

The multi-tissue model described the experimental fluxes from an exercise bike. The optimal fluxes were simulated for a subject cycling on a stationary bike at increasing work rates. ATP synthesis was constrained and an objective function was minimized. The objective function reflected a tradeoff between minimization of oxygen uptake and glycogen expenditure. There was good agreement (Figure 10C) between the predicted exchange fluxes and experimental data. The model described a super-linearity in the oxygen expenditure, reflecting complex I bypass. It also described the shift from fatty acid synthesis towards utilization of glycogen and the excretion of lactate by type II fibers and metabolism of lactate by type I fibers (Figure 10D).

Interpretation of phenotypic predictions
Mechanistic relations are inferred from the models. Extracting knowledge from data is an emerging challenge. Data is seldom stored in a form that is easily accessible for humans. Simulation results and experimental datasets are consistently increasing in information content, due to improvements in computational power and data generation from high throughput experiments. In the previous section I presented the models and some of the predicted phenotypes. In this section I analyze the models and the simulation results, to infer mechanistic relations behind the predicted phenotypes.

Elasticity of growth with respect to the ATP synthesis rate
Biosynthesis depends on input of free energy in the form of ATP. Fast growing cells must therefore synthesize ATP at higher rates than slow growing cells. There is thus a near linear relationship between ATP synthesis and specific growth rate for fast growing cells. For slow growing cells the energy expenditure for maintenance dominates the energy budget, and factors other than specific growth rate explain the ATP synthesis rate. In the infant, such factors include changes in body composition and increased physical activity with age (Wells & Davies 1998). There is currently no model available to predict all ATP demanding processes associated with growth. The ATP cost of growth (G) is thus estimated from the relation between ATP synthesis (T), maintenance (M) and growth (μ).

$$T = G\mu + M$$

This expression can be rearranged to predict growth from ATP synthesis, an implicit assumption that explains much of the prediction power of GEMs (Durot et al. 2009). ATP synthesis and growth rates span orders of magnitude when comparing different organisms (Figure 11A). This may affect the simulations.
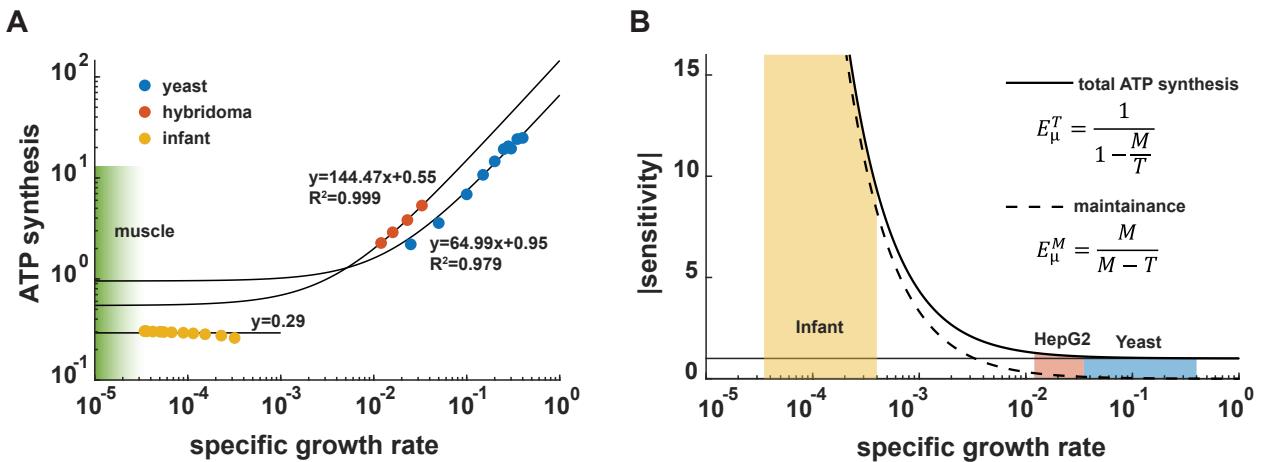
Figure 11. Growth rate dependent ATP synthesis rate and elasticity. A) Linear fits to literature data on growth and ATP synthesis rates for yeast cells, mammalian cells and infants (Hoek et al. 1998; van Hoek et al. 1998; Butte 2005). The range of steady state ATP synthesis rates in the exercising muscle as reference. B) Growth rate dependent elasticity to perturbations in total energy intake and maintenance energy expenditure for the model T=Gμ+M with M=0.5 and G=150. For mammalian cell lines and yeast cells the model predicts negligible effects of maintenance (M << T), and that growth is proportional to energy expenditure (M/T ≈ 0). For infants (M ≈ T) growth is hyper sensitive to changes in both maintenance and energy intake.

An elasticity analysis was performed on the linear model which underlies the growth predictions of GEMs (Figure 11B). The elasticity of growth to perturbations in total ATP synthesis and maintenance depends on the specific growth rate. At higher specific growth rates, the maintenance expenditure becomes asymptotically negligible, and growth becomes directly proportional to the ATP synthesis rate. At low specific growth rates, growth is hypersensitive to fluctuations in both maintenance energy and total ATP synthesis. In the context of GEMs this means that perturbations in fluxes supplying the maintenance reactions may have greater than proportional effects on growth.

The direct proportionality between ATP synthesis rate and growth emphasizes that increased ATP synthesis rates may be of great value to a growing organism. It should however be noted that the linear relation between energy expenditure and growth does not necessarily imply that growth is constrained by the ATP synthesis rate. It may also reflect parsimonious resource utilization where the cells synthesize no more ATP than what is required, even though the potential capacity may have been higher.

The hypersensitivity to perturbations in ATP synthesis and maintenance at low growth rates may be an underlying explanation for growth spurts, where the specific growth varies strongly from day to day (Dobbing & Sands 1973; Wales & Gibson 1994). It may also explain the prevalence of under and over weight in infants, as small variations in environmental factors may translate to large differences in outcome. This assumes that no compensatory mechanisms may act to dampen the effects. In living humans such compensatory mechanism are known to exist (Byrne et al. 2017).

NADH and overflow metabolism

There are several different ways to synthesize ATP. Cells synthesize ATP through substrate level phosphorylation in glycolysis and the TCA cycle and through oxidative phosphorylation (Figure 12A). Aside from ATP, glycolysis also generates pyruvate and NADH. The pyruvate may be further metabolized to CO_2 through the TCA cycle or converted to a fermentation product, e.g. lactate or ethanol, and released from the cell. The fermentation products lactate and ethanol dispose of both the carbon from pyruvate and the electrons from NADH. The free energy stored in NADH and QH_2 may also be converted to a proton gradient and then ATP by oxidative phosphorylation.

Overflow metabolism is a phenomenon where cells ferment a carbon source even in the presence of oxygen. Overflow metabolism appears to be the preferred metabolic strategy under conditions where high ATP synthesis rates are required and substrate availability is not limiting (Molenaar et al. 2009). Thus, overflow metabolism is observed in fast growing yeast, in the vigorously exercising muscle and in many cancer cells. Respiration has a lower catalytic efficiency (ATP synthesis/g protein) than fermentation (Figure 12B). The models of yeast and human predict that fermentation is 2-3.5 times more catalytically efficient than respiration. It is particularly complex I and complex V that make respiration inefficient (Figure 12C). The modeling suggest that cells may bypass complex I to achieve a higher catalytic efficiency of respiration, at the expense of a lower substrate efficiency (ATP/substrate). This occurs by default in *S. cerevisiae*, as it lacks a proton pumping complex I (Dashko et al. 2014). Uncoupling of the respiration chain through complex V bypass results in a lower catalytic activity than respiration, despite the low catalytic activity of complex V. This may not be surprising since it is responsible for the bulk of the ATP synthesis from respiration.

The synthesis of NADH is balanced under fully fermentative conditions. But in growing yeast cells the drain of pyruvate for biosynthesis and NADH production from other processes results in excess formation of NADH (van Dijken & Scheffers 1986). Because of the low efficiency of respiration this may become a burden to the cells. Excess NADH can be relieved through glycerol synthesis from glucose, which occurs under oxygen limited conditions. This disposes NADH more efficiently than respiration (NADH/g protein), but comes at a cost of ATP (Figure 12D). Uncoupling of respiration is 40% more efficient than glycerol synthesis and does not require an ATP investment. This suggests that yeast may use uncoupled respiration to manage excess NADH. Petite yeast cells lack mitochondria and thus cannot uncouple excess NADH. The specific growth rate of these yeasts is around 25% lower than that of normal yeast (De Deken 1966).
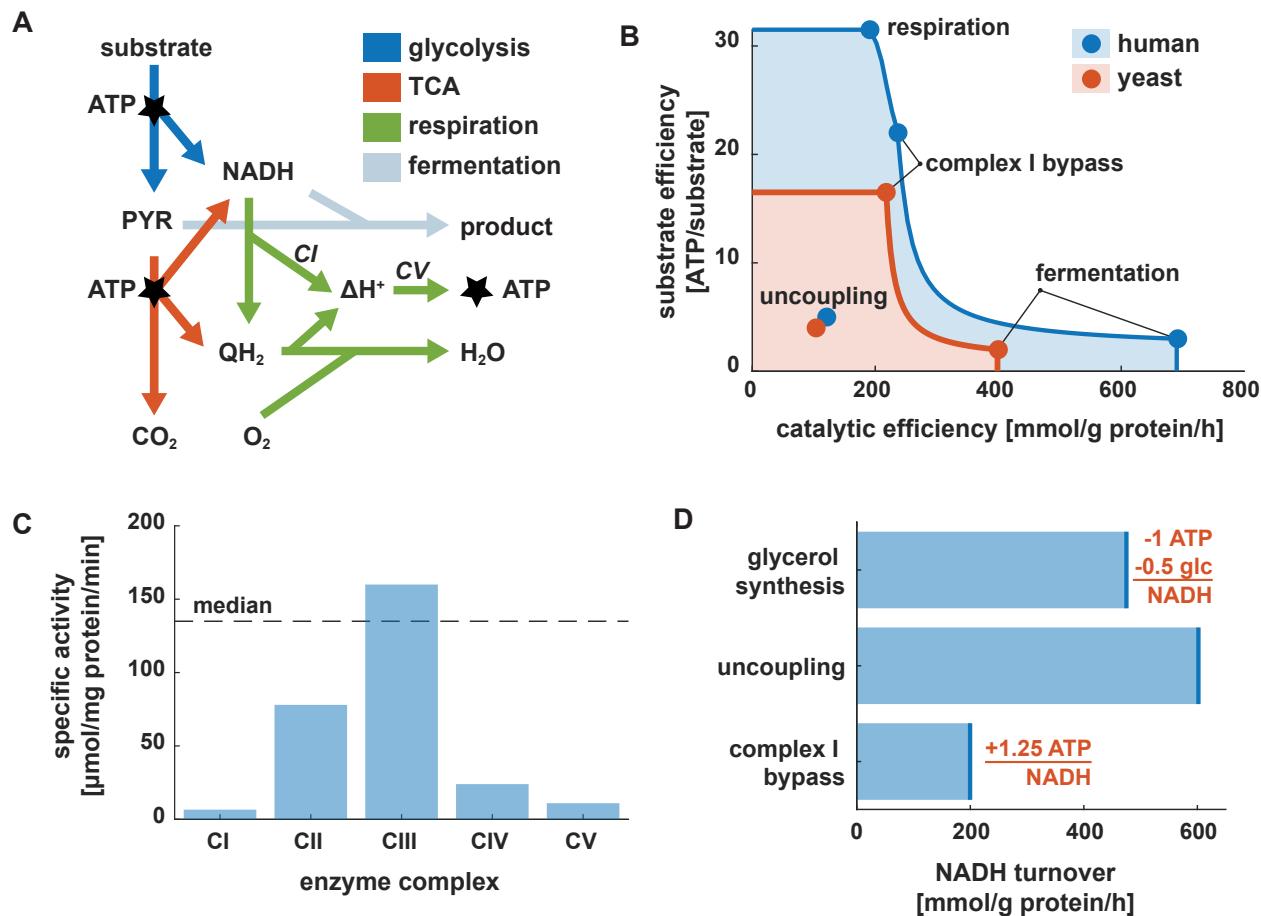
Figure 12. Trade-off between catalytic efficiency and substrate efficiency. A) Schematic representation of energy metabolism. B) Pareto front for maximizing the catalytic and substrate efficiency in muscle and yeast. C) Comparison of the specific activity of the enzyme complexes involved in oxidative phosphorylation in human, with the median values of intermediary metabolism as reference. D) Uncoupling is the most protein efficient strategy to handle NADH-imbalances. Abbreviations: complex I-V (CI-V), pyruvate (PYR), ubiquinol (QH_2) and proton gradient (ΔH^+).

The phenotypic effects of metabolite depletion

Metabolite depletion affects ATP synthesis and growth. Cells may adapt to depletion of a metabolite by reallocating their proteome to utilize another. Due to differences in catalytic efficiency of the metabolic pathways, the rate of ATP synthesis and thus growth may differ between metabolites. Metabolites with high growth rates are often consumed first, and metabolite depletion is hence associated with decreased proliferation rates. Metabolic models with interior constraints may provide insight into these dynamics.

The growth rate of yeast depends on the carbon source. The enzyme-constrained yeast model was used to predict the maximum growth rate on different substrates (Figure 13A). The predicted growth was higher on glycolytic substrates compared to non-fermentative due to the higher catalytic efficiency. In a follow up study using a GEM, these results were replicated for a large set of different metabolites (Sánchez et al. 2017).
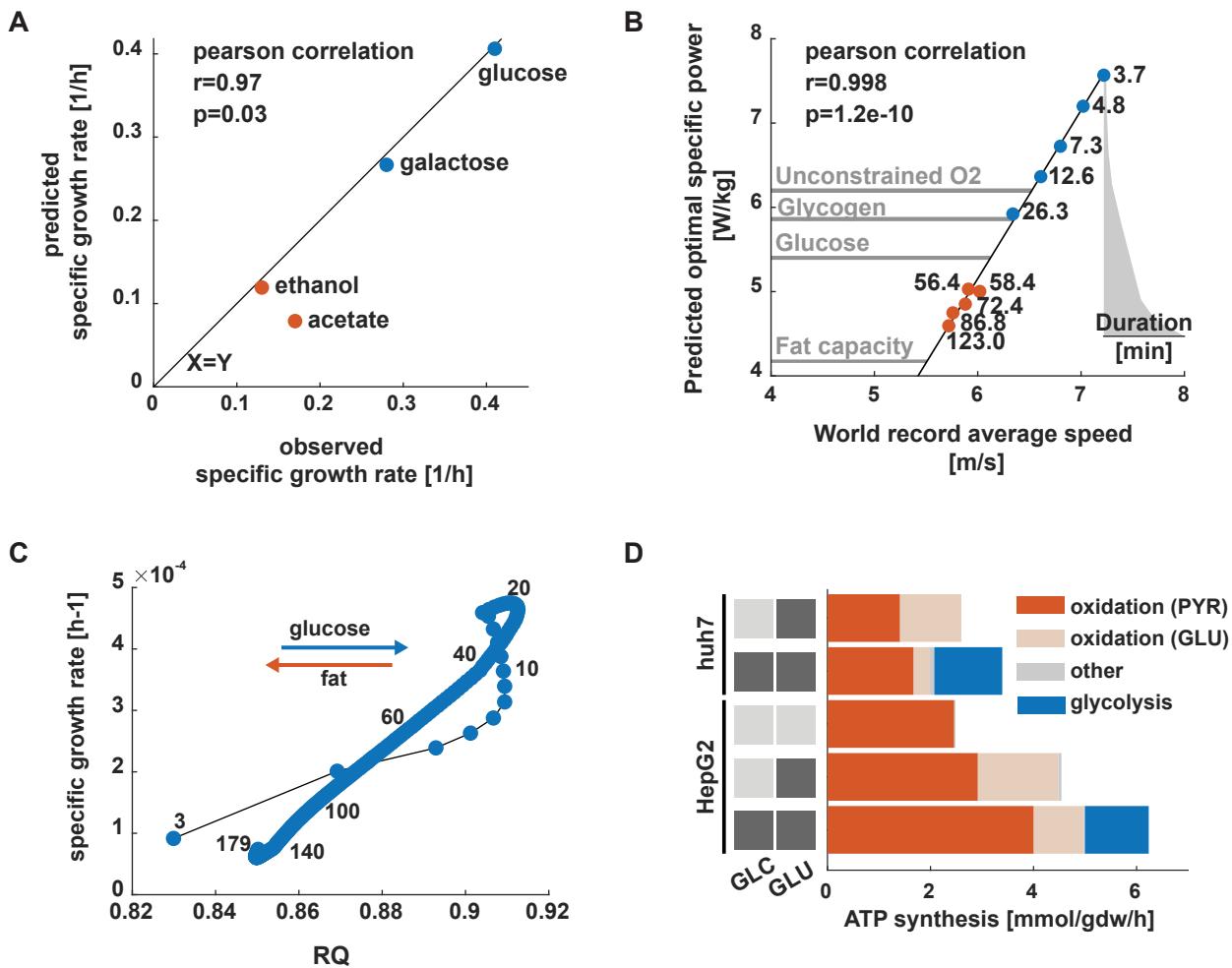
Figure 13. Phenotypic effects of metabolite depletion in yeast and human. A) The enzyme-constrained model of yeast was used to predict the maximum specific growth rate on two glycolytic substrates (blue dots) and two non-fermentative substrates (orange dots). B) The full body muscle model was used to predict the maximum work rate on different substrates (gray lines). The maximum sustainable work rate and optimal metabolic strategy, respiration (orange dots) or respiro-fermentation (blue dots) was predicted for increasing exercise durations (numbers and gray subplot). There was good agreement between predicted specific power output and the average speed, which is a function of power at different durations. C) High specific growth rates in the infant model were associated with high glucose utilization between the age 3-179 days. D) The ATP synthesis rates in the cancer model depend on the substrate availability. The results suggest modular contribution from the different subsystems.

The finite metabolite deposits in the human body explain differences in running-speed. The human body contains reserves of different metabolites, around 0.5 kg of glycogen in muscle tissues, 0.1 kg of glycogen in the liver, and large reserves of fat. Due to the difference in maximum specific ATP synthesis rates on different substrates, the exercising human must economize these resources to achieve high sustained ATP synthesis rates. At high work rates, lactate accumulates, as it is produced at higher rates than it can be cleared. The maximum lactate concentration that can be tolerated becomes an additional constraint. For longer exercise durations, the sustainable specific consumption rates are lower, and the predicted maximum specific power output is thus also lower. Assuming that the metabolite pool is depleted at the end of the exercise, then the maximum steady state flux is given by the pool size divided by the duration. With these modeling assumptions a

good agreement was found between world-record running-speeds and the maximum predicted power output for different exercise durations (Figure 13B).

Infants shift metabolic strategy with age. In humans, it is possible to determine which carbon source is utilized from non-invasive respiratory measurements. The ratio of CO_2 eliminated to O_2 consumed ($RQ=vCO_2/vO_2$) tells if glucose ($RQ=1$), fat ($RQ=0.7$) or a mixture is used ($0.7<RQ<1$). The predicted RQ was compared with the predicted specific growth rate in the infant model (Figure 13C) it showed an association between growth and high glucose utilization. This may reflect the higher energy synthesis capacity on glucose, but could also be an indirect effect of increased fat storage rates, and thus lower fat usage, at high growth rates.

ATP synthesis rates correlate with cellular growth in cancer. ATP synthesis rates have been shown to correlate with growth in mammalian cells (Xie & Wang 1996). The ATP synthesis rates were characterized for the cancer cell lines under various nutrient conditions. A modular contribution from the different ATP synthesis pathways was found, consistent with enzyme capacity constraints. Cells cultivated in the presence of glucose had a higher ATP synthesis rate than cells cultured without (Figure 13D). Cells cultured under glutamine depleted conditions had yet lower levels. The contribution from various sources made ATP supply robust, and growth proceeded also in the absence of several growth enhancing metabolites. The sensitivity of growth was investigated for partial and complete inhibition of these pathways (Figure 14A) using acute sensitivity analysis. The predicted growth reduction was in agreement with the reduction that could be expected based on the pathways relative contribution to the total ATP synthesis flux. The predicted effects at full inhibition were in reasonable agreement with the predicted growth under nutritional conditions where the flux through these pathways was zero. Sensitivity analysis was also applied to investigate increased flux through the maintenance reaction. Under some conditions the growth response was greater than what would be anticipated from a linear analysis (Figure 14B).

Compartmentalization and overflow of glutamate

Many cancer cells excrete glutamate. The excretion rates correlate with growth rate and aggressiveness of the cancer (Stepulak et al. 2014). The precursor of glutamate, glutamine, is often consumed at high rates by cancer cells (Wise & Thompson 2010). Glutamate is the first intermediate in the metabolism of glutamine (Figure 15A) towards several important downstream products, e.g. α-ketoglutarate (AKG), aspartate (ASP), alanine (ALA) and pyruvate (PYR). Glutamate arises as a byproduct from nitrogen donation by glutamine in several biosynthetic processes, e.g. the synthesis of nucleotides. It is also a byproduct of deamination of branched chain amino acids (BCAA). Many cancer cells overexpress the first step in BCAA metabolism, the BCAT1 enzyme. The *de novo* synthesis of serine by phosphohydroxythreonine aminotransferase (PSAT) consumes glutamate and releases AKG.
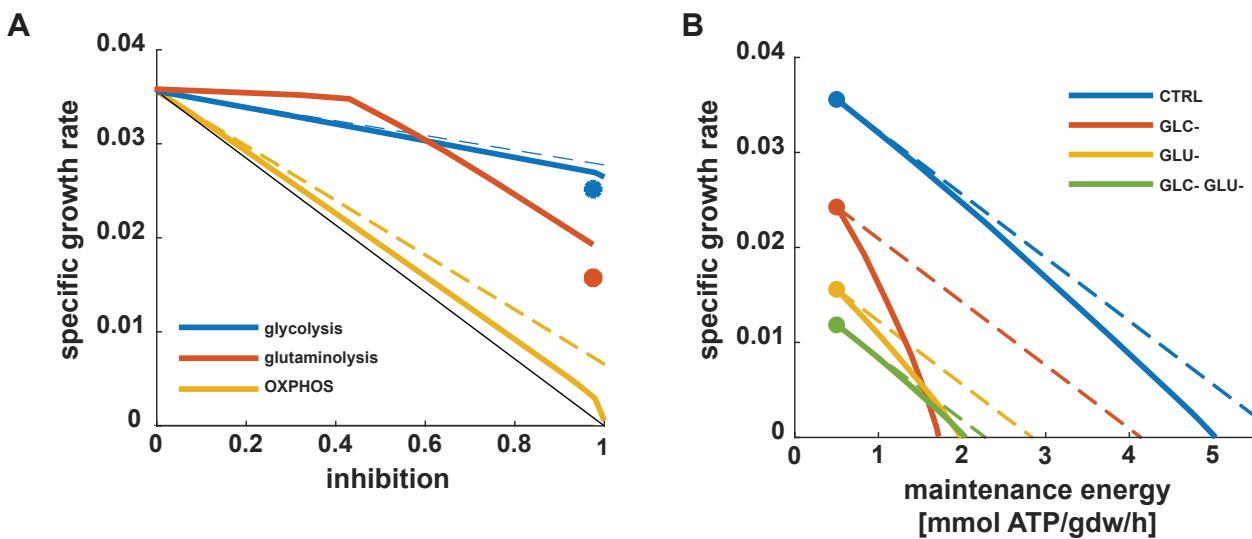
Figure 14. Predicting the effect of perturbations using interior constraints derived from estimated fluxes. A) Acute sensitivity analysis was applied to the cancer model to predict growth effects from reduced flux through glycolysis, glutaminase and oxidative phosphorylation (solid lines). For glycolysis and oxidative phosphorylation the predictions are compared with the expected effect (dashed line) based on the pathways' contribution to ATP synthesis and the model $T=150\mu+0.5$. The predictions at full inhibition are compared with the specific growth rate under glucose depleted (blue point) and glutamine depleted (orange point) conditions. B) The growth effects are predicted for increased maintenance energy expenditure using ASA. Model predictions (solid lines) are compared to the expected effects (dashed lines) calculated from the model $T=150\mu+0.5$.

Due to the physical structure of the cell, excreted glutamate must come from the cytosol. The interior fluxes surrounding cytosolic glutamate were investigated using the compartmentalized model of liver cancer cell metabolism (Figure 15B). The production of cytosolic glutamate were found to outweigh the consumption, so that excess glutamate must be transported to another compartment to ensure mass balance. Rather than transporting glutamate to the mitochondria to be further metabolized, the cells excreted it.

It is not clear why the cells elect to excrete glutamate. The mitochondria clearly has capacity to metabolize glutamate, as it consumes glutamine, and thus glutamate, at high rates. It could be due to an unfavorable concentration gradient, however glutamate appears to be readily transported into the mitochondria as part of the malate-aspartate shuttle, which moves the charge of NADH in the cytosol to NADH in the mitochondria. The malate aspartate shuttle is located in the intermembrane space, and may potentially not be accessible for cytosolic glutamate. It may also be due to the glutamate transporter being an antiporter with aspartate (Figure 15C). If the cell solved the imbalance of glutamate by transporting it to the mitochondria, it may instead have to solve an imbalance in aspartate, since all compartments are mass balanced.

I hypothesized that blocking glutamate export would affect fitness negatively. With the excretion option blocked, the cell would need to reduce the fluxes through BCAA metabolism or biosynthesis to uphold flux balance. An inhibitor of glutamate transport (sulfasalazine) was added to the medium. It reduced both glutamate production and the proliferation rate, and allowed the hypothesis to be validated (Figure 15D).
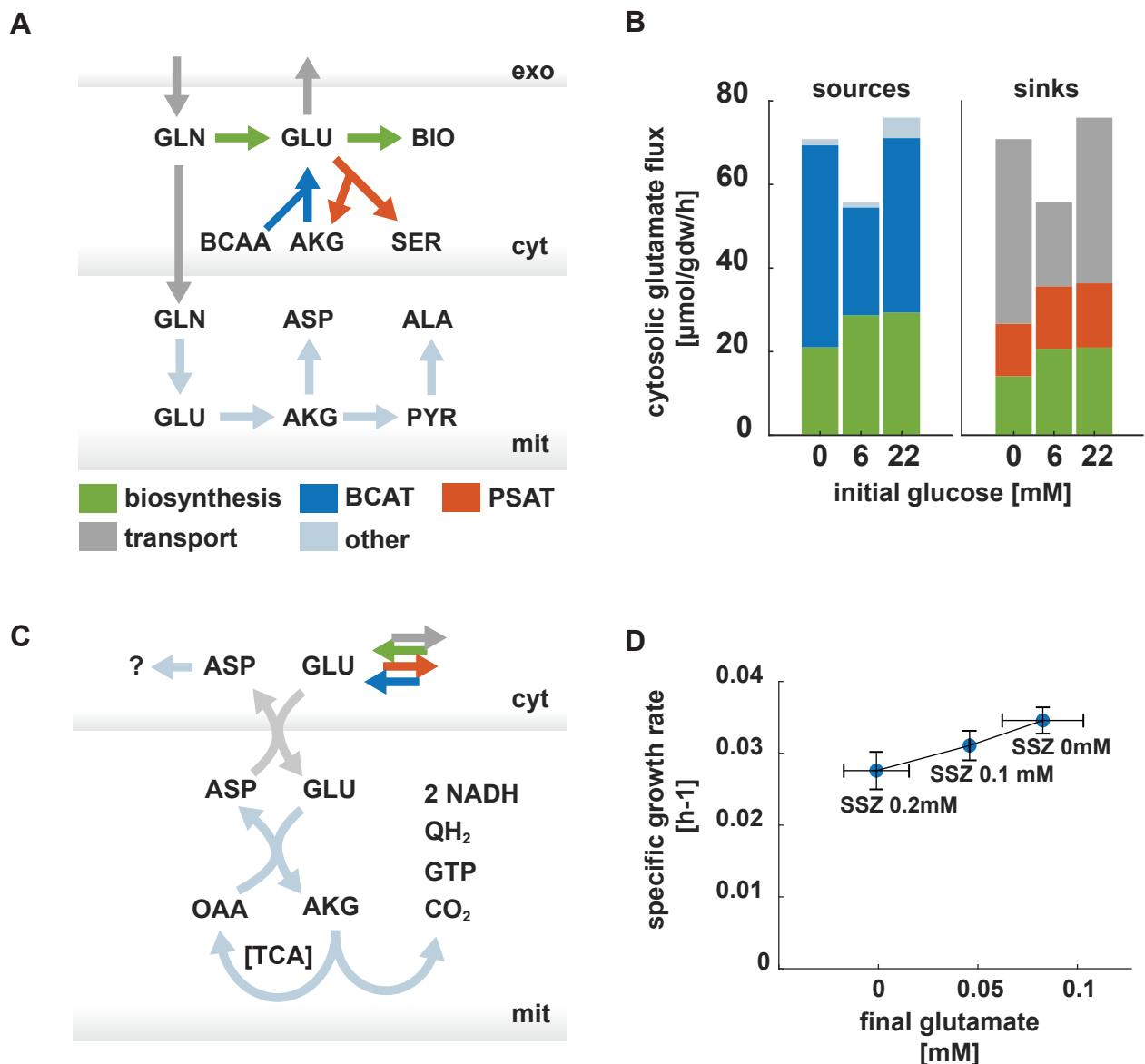
Figure 15. Effects of the imbalance in the cytosolic glutamate flux. A) Schematic representation of the fluxes surrounding cytosolic and mitochondrial glutamate. B) The sources and sinks of cytosolic glutamate are imbalanced under several different conditions. Excess glutamate is transported to other compartments, primarily the extracellular compartment. C) The influx of glutamate into the mitochondria must be balanced by an efflux of aspartate due to the glutamate-aspartate antiporter. D) Increasing concentrations of sulfasalazine (SSZ) inhibited glutamate export and reduced the specific growth rate of the cells.

Discussion

In this thesis, metabolic models of human and yeast were developed. Models with boundary constraints were used to predict internal fluxes, and models with interior constraints were used to simulate the trade-off between different metabolic strategies. Preexisting methods were modified to improve model accuracy and better interpret the constraints on growth and ATP synthesis. I aimed to understand how biophysical constraints shape the fitness landscape of the cell, and how phenotypes emerge from optimization of protein allocation. For cancer cells, the development towards optima appears to be incomplete, several regulatory constraints from the cell of origin appear to remain. Some phenotypes of the cell lines may be remnants of social interactions from the in vivo condition. Various data types are integrated through modeling, thus biological models have potential to unify knowledge from different studies and domains. This may expose contradictory claims and highlight what remains unknown.

Development of new metabolic models and simulations

In this thesis, I developed mechanistic models of processes involved in growth and ATP synthesis. Much work has already been done in this area and preexisting methods and models were modified to improve their scope and quantitative accuracy.

In this thesis, the first GEM of an infant was developed. The whole body of the infant was modeled as if it consisted of a single cell type with accesses to all biochemical reactions encoded by the human genome. In the reconstruction the HMR 2.0 model was used, but it was possible to port the simulations to the Recon3D model (Brunk et al. 2018), yielding similar results. Earlier studies of infants had shown that changes in body composition and energy expenditure were of importance (Wells & Davies 1998; Butte 2005). A modular biomass composition was thus introduced as a function of age, based on the experimentally determined body composition. By integrating age dependent data on organ sizes, the hypothesis that age dependent changes in maintenance energy could be fully explained by changes in body composition (Wang 2012), was corroborated. The simulations showed that energy was the growth-limiting factor for breastfed infants. This justified the assumptions of the preexisting models (Fjeld et al. 1989; Butte 2005; Jordan & Hall 2008), which relied on converting all nutritional input to their corresponding energy content (kcal), using the Atwater factors.

There already exist several GEMs of cancer cells, as well as other mammalian cell types (Shlomi et al. 2011; Agren et al. 2012; Agren et al. 2014; Sheikh et al. 2005; Martínez et al. 2012). The existing metabolic models of cancer that make use of experimental data are however small-scale models. In this thesis, I developed the first GEM of a cancer cell line constrained with data from a dedicated experiment. Previous models (Zielinski et al. 2017) had used published flux data (Jain et al. 2012) to show a correlation between metabolite excretion, and excessive uptake of amino acids. However, by using the full genome-scale model together with experimentally determined fluxes I could show a mechanistic relation between glutamate, branched-chain amino acids and cellular compartmentalization. The use of a genome-scale model, in favor of a small-scale model, exposed a high degree of variability for many of the fluxes. In general, only fluxes in close proximity to the

biomass equation and the exchange fluxes could be estimated with any certainty. The uncertainty for the remaining reactions may to some degree be a consequence of metabolic loops and synonymous submodules without effects on the overall analysis (Kelk et al. 2012). But it should be noted that many fluxes may differ from their predicted values, which limits the set of claims that can be made with certainty.

There are previous metabolic models with enzyme-constraints for both *E.coli*, yeast and human. Some of these models used parameters derived from literature in favor of randomly sampled parameters, which increases the accuracy of the predictions (Adadi et al. 2012). I developed the first fully parameterized enzyme-constrained model of intermediary metabolism in yeast. In this work I introduced the idea to use proteomics data to constrain the total pool of enzymes to the experimentally observed levels, the P-parameter. This parameter attempted to address the issue that the model only covers a part of the activities of the cell, and prevented the predicted enzyme concentrations from becoming unrealistically high. An enzyme saturation parameter was also introduced. This parameter attempted to capture the difference between the effective flux in the cell and the maximum flux of an enzyme under optimized conditions. In this thesis the parameter was set to 0.5 for all enzymes. Whilst this assumption may be correct on average, the accuracy of the enzyme concentration predictions was modest in most cases. New types of models are under development (Noor et al. 2016), which allow the saturation of each enzyme to be estimated. They predict the optimal enzyme concentrations and substrate concentrations of the cell simultaneously, by minimizing the enzyme cost.

There are two previous models of ATP synthesis in muscle fibers using enzyme constraints (Vazquez et al. 2010; Vazquez & Oltvai 2011). The constraints of these models were however not on the level of individual enzymes, but were based on the assumed crowding capacity of different compartments. In this thesis, I reconstructed the first enzyme-constrained model of intermediary metabolism in muscle fibers with parameter values from literature. These parameters were integrated with proteomics data to estimate the effective v_{max} of each enzyme-constrained reaction. The reactions with active constraints were identified and constrained in a GEM. For this I developed the first multi tissue GEM of the human muscle system. It was used to study the lactate shuttle and the maximal fat oxidation capacity of cells. It was also used to study the phenomenon that the running speed in exercising humans decreases as a function of the length/duration of the race. Whilst this has been studied before using mathematical models of world record performance (Peronnet & Thibault 1989), a more detailed mechanistic description was achieved using the GEM. The differences in ATP yield from different carbon sources together with limited pool sizes allowed the optimal work output to be predicted for different exercise durations.

Constraint-based modeling at low specific growth rates

Survival may be even more fundamental than growth. The mechanisms for survival are normally modeled as growth-independent maintenance fluxes, e.g. a constant expenditure of ATP. At low specific growth rates, the maintenance reactions drive the fluxes. This makes the growth predictions sensitive to perturbations in energy metabolism and to nutrient intake. For most applications the predicted growth does not decrease more than proportionally to a reduction in

any flux, since FBA models are linear equation systems. But at low predicted growth rates the maintenance fluxes introduce an exception to this and also a small decrease in ATP synthesis markedly reduces the predicted growth rate. If the rate of ATP synthesis is insufficient to meet the maintenance requirement, the simulations become infeasible. This may be interpreted as death, since the organism is not able to meet the requirements for survival. In the living cell, maintenance fluxes are often more plastic than in the simulations, and cells have been shown to survive at ATP synthesis rates below the maintenance requirements found in growing cells (Vos et al. 2016). Infeasibility often occurred when perturbing the simulations of infant growth (**Paper I**). This was resolved by introducing a second optimization where ATP synthesis from the fatty acid stores was minimized to fulfill the maintenance requirement.

Despite the methodological challenges at low growth rates, maximization of growth was still a viable objective function. This can conversely be viewed as parsimonious energy intake of the infant. In the absence of interior constraints on growth, the model would predict unrestricted growth for unrestricted milk intake. This is clearly not physiological, but was not a practical concern since infants did not consume milk in these quantitates (Butte et al. 2002). If the model were to be adapted to simulate child obesity, appropriate interior constraints on the synthesis of lean mass would have been required to ensure the accumulation of fat tissue.

FBA simulations assume that growth is exponential. An implicit assumption of FBA is that cells grow exponentially with some rate µ (1/h), which depends on the flux of biomass (X) over the boundary. This assumption arises from the use of mass specific fluxes (mmol/gdw/h).

$$\frac{dX}{dt} = \mu \ \Rightarrow X_t = X_0 e^{\mu t}$$

For some applications this assumption may not be suitable, e.g. for cells growing in monolayers, in tumors where only a fraction of the cells are actively growing (Murphy et al. 2016) or for yeast cells growing as stalks (Vulin et al. 2014). Assuming exponential growth for linearly growing cell cultures may result in marked underestimation of the specific growth rate and uptake fluxes in the growing portion. Whilst at the same time fluxes may be overestimated for the quiescent cells. At short time scales or low specific growth rates the linear and exponential models become nearly equivalent (by tailor expansion to the first order).

$$X_0 e^{t\mu} \approx X_0 (1 + t\mu)$$

The growth rate of an infant declines over time and a linear growth model may be more appropriate than an exponential. Due to the low specific growth rate and high time resolution (day-to-day resolution) used in this thesis, the differences between the two models is expected to be low.

Modeling of mineral deficiencies

A method to model the effects of mineral deficiencies in a cell was introduced. The method was conceptually similar to earlier sensitivity tests. It is common to constrain flux through sets of reactions, e.g. all reactions sharing a gene or all reactions sharing a metabolite (Agren et al. 2014; O'Brien et al. 2015). The concept was expanded to the set of reactions that share a certain mineral

cofactor. The approach has clear potential for generalization and the recent cross-referencing of human metabolic networks with protein structural databases and cheminformatics platforms (Brunk et al. 2018), opens up many possibilities to construct and simulate informative reaction sets. The simulated reduction in flux is intended to mimic a reduced concentration of active enzyme due to insufficient mineral concentrations. In principle it should be possible to obtain a rough estimate of the mineral demand of a cell directly from the abundance and stoichiometry of mineral binding enzymes. These enzyme abundances could either be derived from enzyme estimates by the models or more directly from proteomics. In the future, model driven estimates of mineral requirements may be able to replace the current empirical recommendations for daily intake of mineral and vitamins.

Genome-scale metabolic modeling of a cell culture

Cancer cells consume many different metabolites simultaneously. A challenge for modeling growth of mammalian cells is the broad range of metabolites that are simultaneously consumed and must be quantified. In the cancer study (**Paper II**) most amino acids were measured (20/22), as well as glucose, lactate and pyruvate. Still additional assumptions were required about non-quantified exchange fluxes, e.g. tryptophan, CO_2 and vitamins. Fetal bovine serum (FBS) is an undefined component of the culture medium. It contains metabolites that were not quantified, but which may potentially have been consumed, e.g. fatty acids. Many metabolites are essential for growth and have direct control over the growth predictions. The existence of multiple substrates made it less clear how to conduct simulations with fixed growth and minimization of substrate uptake. Growth maximization was therefore the only objective function used in these simulations.

Metabolites are often depleted over the course of the cultivation. Metabolites depletion made the analysis more complicated. The time points of the depletion events were generally unknown and must therefore be identified by extrapolation of the metabolite consumption curves. The depletion events were also expected to affect the exchange fluxes and they therefore had to be refitted for each event. Most data was thus used to fit new sets of exchange fluxes, rather than to improve the accuracy and estimate the error of the fits. Not accounting for the depletion events would however have had an even more adverse effect on the flux predictions, as was seen in a high throughput study (**Paper III**). In addition to the metabolite depletion events, an initial metabolic phase was identified, where amino acids were consumed at markedly higher rates than at later time points. More frequent sampling and quantification of the protein secretion rates may be required to examine this phenomenon further.

The uncertainty of the exchange fluxes was not assessed. Due to limitations in the experimental setup it was not possible to properly assess the experimental uncertainty of the fitted exchange fluxes. Whilst potentially important, this is commonly neglected in simulations using GEMs (Sánchez et al. 2014). Experience from this study (**Paper II**) stresses the value of identifiability analysis and dedicated experimental design. But it should also be noted that even if the complete flux distribution were to be estimated with high confidence, this distribution would likely only be one of the possible states for the cell.

From the large number of possible reactions (>8000) in the models, only a small fraction (<500 reactions) carried flux. This may have several explanations. Many reactions are involved in biosynthesis of complex biomolecules which are not represented with this level of detail in the biomass equation, e.g. lipids (Sanchez et al. 2018). There may be nutrient sources that the cells do not encounter in the experimental setup. Protein modifications and tRNA loading (Feizi et al. 2013) involve reactions that are not included in the models. Metabolic modeling with optimality principles will not predict the occurrence of undesired spontaneous reactions, but in the living cell, non-enzymatic reactions and enzyme promiscuity may result in the buildup metabolites that must be cleared with dedicated detoxification pathways, e.g. reactive oxygen species and the oncometabolite 2-hydroxyglutarate (Xu et al. 2011; Piedrafita et al. 2015). However, the quantitative effects of these reactions are likely low. Different maintenance tasks, such as upholding osmotic balance may also be influential. Finally, proteins that take part in signaling processes are often enzymes (Hyduke & Palsson 2010) and these are not expected to carry flux in the simulations. More refined simulations are required to cover all the activities of the cell.

It was challenging to extrapolate the predictions of models with boundary constraints. The absence of internal capacity constraints made it unclear which fluxes were constraining growth and which fluxes were accidental. The introduction of acute sensitivity analysis (ASA), allowed some degree of extrapolation. It prevented uncontrolled growth in the absence of boundary constraints, and it appeared that the fluxes in one condition could be generalized as capacity constraints in another, in particular for reactions relating to ATP synthesis.

Metabolic trade-offs in ATP synthesis

There are multiple lines of evidence that limited ATP synthesis rate is a prevalent constraint in biology. Limited ATP synthesis may arise from capacity constraints in the ATP-synthesizing enzymes or constrained uptake of carbon sources and thus limited downstream flux.

The major trade-off in ATP synthesis is between fermentation and respiration. It is driven by the low catalytic efficiency of respiration. Using sensitivity analysis, enzyme-parameters with a big effect on the predictions were identified; in the case of yeast this was primarily ATP synthase (complex V). Yeast cells ferment at high growth rates and they therefore grow slower on non-fermentative carbon sources where respiration is the only option. Reduced glucose uptake rates through deletion of the low-affinity/high-capacity glucose transporter HXK2 in yeast makes fermentation non-viable and reduces growth (Schuurmans et al. 2008). Likewise reduced growth was observed in an adaptive laboratory evolution (ALE) aimed at removing the fermentative phenotype in yeast (Dai et al. 2018). In the carbon limited chemostat, growth aligns with the dilution rate at steady state. The high cell density at low dilution rates reduces the extracellular glucose concentration to near undetectable levels (Postma et al. 1989). This affects glucose uptake rates, and in turn ATP synthesis, and is likely the mechanistic explanation for the reduced growth rate in the chemostat. A minor metabolic trade-off is the bypass of complex I. This strategy is more catalytically efficient than complete respiration. The flux through complex I is constrained in muscle fibers, and was experimentally detectable as a reduced ATP yield per oxygen consumed in exercising humans.

The role of NADH in energy metabolism depends on the ATP synthesis strategy. The free energy carrying metabolite NADH is the source of most ATP under respiratory conditions, but the low efficiency of respiration makes NADH a liability under conditions where ATP synthesis needs to be maximized. If it cannot be balanced by consumption in the synthesis of fermentation products enzymes must be invested into clearing excess NADH. Uncoupled respiration is a strategy for clearing NADH predicted by the modeling. In yeast the permeability transition pore allows uncoupling of respiration (Beutner et al. 2017). In in human several dedicated mitochondrial uncoupling proteins (UCP1-4) exist. Under high intensity exercise in a small muscle group it was predicted that uncoupling can mitigate the accumulation of lactate. The increasing difference between the pulmonary oxygen expenditure and the oxygen expenditure in the leg at high exercise intensities may indicate that such a process takes place in humans (Poole & Richardson 1997).

Interpretation of the constraints on ATP synthesis and growth

ATP synthesis may be constrained in liver cancer cell lines. The predicted growth rate of the HepG2 cancer cell line was constrained by the uptake rates of several essential amino acids as well as synthesis of ATP. The balance between uptake and biosynthesis is consistent with parsimonious substrate utilization and may not be constraining the cell. The excretion of lactate suggests that ATP synthesis from respiration was an active constraint. It however appears that partial fermentation was sufficient (lactate/glucose < 2) to supply the required ATP, similar to what is observed for yeast at intermediary growth rates ($0.28 < \mu < 0.42$) in the chemostat. This suggests a glucose uptake rate constraint as in the chemostat, or that some other constraint on growth limits the demand for further ATP synthesis. In support of a constrained ATP synthesis by respiration, the predicted growth was lower in the absence of glucose. Glutamine supported ATP synthesis, and a reduction in growth was observed after glutamine was depleted. This growth reduction was however likely due to the biosynthetic role of glutamine, since reduced glucose uptake flux and reduced lactate production also was observed, indicating a decreased demand for ATP. Our simulations and experiments suggest that cells prefer to not uptake cytosolic glutamate into the mitochondria, even when this would seemingly be beneficial. This suggests that mitochondrial uptake is constrained in the living cell.

The growth of infants was limited by the ATP synthesis rate in the simulations. The growth rate was sensitive to perturbations in energy metabolism. It is thought that the high energy expenditure and low growth rate in infants may reflect an evolutionary trade-off between rapid maturity and a large brain (Kuzawa et al. 2014). The prevalence of growth spurts suggests that there may be sufficient enzymatic capacity to sustain higher growth than what is observed (Wales & Gibson 1994) and thus that the constrained ATP synthesis rate arises from limited nutrient uptake rates. The degree of hunger and intake of breastmilk may in turn be driven by evolutionary pressure towards parsimonious energy usage. Regulatory constraints on hunger from growth hormones may be the fundamental explanatory factor behind observed growth rates under normal conditions.

Regulatory constraints on metabolism

Enzyme concentrations are optimized by evolution. At long time scales evolutionary pressure optimizes the kinetic parameters of enzymes towards what is biophysically feasible under the given

biological conditions. On shorter evolutionary time scales, kinetic parameters may not be dramatically improved. Evolution instead acts on the regulation of the cell to provide the most optimal enzyme concentrations for the encountered conditions. Solving an optimization problem in the computer thus mimics regulation, and is a model of the biophysical constraints on regulation. ALE experiments (Yu et al. 2018) have revealed that regulation may be optimized to a specific condition, which reflects fine-tuning of the cell's protein allocation to the selective constraints.

Genetic deletions modify the fitness landscape. The dependence on regulation to achieve an optimal protein allocation may constrain the cell to less than optimal performance under many circumstances. There is a dual causality, where biophysics shapes the fitness landscape, and regulation places the cell therein. An interesting case study is the combination of genetic engineering with ALE (Yu et al. 2018). The fitness landscape was reshaped by genetic deletion events that could not be overcome by evolution in the short term. In the background of these deletions, fatty acid secretion became the new optima. By evolving the cells, the regulatory program was allowed to identify this option, and high fatty acid synthesis rates were achieved. The short term evolution towards optima in response to a genetic events has potential to make ALE an interesting experimental setup to study rewiring of the regulatory network in cancer.

Cancer cells evolve on a short timescale. Healthy cells have specific tasks within an organism and their growth is carefully regulated. Genetic events, e.g. mutations, deletions and copy number alterations, may affect the cell's signaling system and allow them to avoid checkpoints in control of nutrient supply, growth and dissemination, to become cancer cells (Hanahan & Weinberg 2011). Actively growing cancer cells undergo evolution due to acquired genomic instability. This allows them to adapt to new environments. Yet on an evolutionary timescale, cancer cells have only a short time to evolve. This suggests that the cells may not reach optimality before they become fatal or are treated. Thus the phenotypes of cancer cells may best be understood as a transition between the cell type of origin and some more optimized state.

There are many strategies that may increase growth. By comparing the phenotype of many cancer cell lines we may gain an understanding of evolutionary conserved strategies. Fermentation of glucose and glutamine addiction appear to be two such strategies. Excretion of glutamate may also constitute such a strategy. These may be achieved through various regulatory paths, e.g. HIF1-α or truncation of oxidative phosphorylation through succinate dehydrogenase (Yizhak et al. 2015). But there are also examples of increased ATP synthesis through increased respiration. It is debated whether the changes observed in cancer metabolism are driving (supply model) the cancer progression or if they are driven (demand model) by increased consumption rates (Ward & Thompson 2012).

Metabolism does not only support biosynthesis. Cancer cells evolve to evade the immune system. It has been proposed that some of the metabolic activities of cancer cells take part in communication with the immune system (Adams et al. 2015). This may be an attractive explanation for metabolic phenotypes for which a biosynthetic explanation cannot be found. An example may be the release of ornithine seen in many cancer cells (Jain et al. 2012). The overconsumption of branched chain amino acids could not be explained by any biosynthetic requirement. This may

indicate a role in signaling or communication. It has been proposed that metabolism of branched chain metabolism affects fermentation signaling through HIF1-α, through metabolic interaction with α-ketoglutarate (Raffel et al. 2017). It was assumed that the cancer cell line secreted proteins, and secretome analysis shows high abundances of native proteins (Caruso et al. 2017), which may indicate incomplete optimization. But the secretome also shows high levels of metalloproteases, which may be used to modify the extracellular matrix. The secretion of these proteins drained resources without any apparent benefit for the cells under the laboratory conditions, but may be an adaptation to increase growth *in vivo*.

Ethical and environmental risks and benefits

The use of computer models in this thesis allowed experiments to be performed at high pace and at a low economic and environmental cost. The computer models largely relied on prior knowledge and parameter values from literature. There is a risk that errors may propagate from these sources to incorrect model predictions. Non-canonical components may not be represented in the model and this may thus bias predictions in favor of prior assumptions.

Knowledge was generated about biological systems of high public interest, which has potential to improve both health and economy. With this comes the ethical risk that the results may be over- or misinterpreted. This could result in the use of inappropriate strategies in treatment of cancer, faulty nutritional management of infants, ineffective training regimes for sportsmen and wasted economic or environmental resources in the biotechnological industry. The risk of these scenarios was assessed to be minor since the results are on the level of basic research and the methods are relatively straight-forward to interpret. In this thesis I attempted to keep the titles of the publications short, to facilitate their impact (Letchford et al. 2015). However, a short title will in general contain less information than a longer title and may increase the risk for misinterpretation.

The pharmaceutical industry is an external stakeholder in this project, since it benefits from an improved understanding of biology. A closer relationship with the pharmaceutical industry could have been advantageous for the project. The loose connection to industry may prevent the results from coming to practical use, either due to lack of interest, or challenges in their application. Involving the stakeholders closer in defining the objectives and requirements may have decreased the risk. However, the research and development in the pharmaceutical industry is currently struggling with an innovation crisis (Munos 2009), and too close of a connection may have threatened the scientific independence of this work. However, the predictions can further be validated by pre-clinical researchers and may then be translated into the clinic with the help of pharmaceutical companies.

Perspective, towards a standard model of biology

The standard model of particle physics has emerged as a unifying theory of fundamental forces and elementary particles (Ellis 1986). The model came out of a drive to integrate all physics into a theory of everything, a set of self-consistent equations from which the behavior of all matter and energy could (in principle) be derived. The standard model is not a theory of everything, e.g. it does not describe the gravitational force. But the model has been instrumental to systemizing physics and many predictions made by the model have later been experimentally verified, including the existence of several elementary particles.

Analogously a unified model of biology may be envisioned. Such a model should be able to account for all biological observations, and could become a scaffold onto which all the knowledge of the field may be attached. Ideally new data should be tested against the theory and inconsistencies in the model or the data would thereby be exposed. The requirement to pre-register trials has been successful in the pharmaceutical industry where it appears to have reduced publication bias (Kaplan & Irvin 2015). Attempting to model all of these experiments could help to accentuate where the blank spaces in our knowledge reside. The amount of data generated in biology will only increase and the storage of knowledge in the form of a models is compact and persistent. It also allows new work to be built on earlier without detailed understanding of its origin.

There are several ongoing attempts at generating unifying models of biology, e.g. the whole cell model (Karr et al. 2012) and the protein sector models (Klumpp et al. 2013). And the whole field of systems biology may be seen to have an integrative agenda (Brigandt 2013). FBA simulations with a GEM is a simplification of the more general problem to identify the dynamics for all metabolite and enzyme concentrations of the cell (Müller et al. 2015). Enzyme-constrained models constitute a significant advancement toward integrating GEMs with kinetic modeling and the fields of structural biology and evolutionary biology (**Paper VI**). The integration of kinetic parameter values in metabolic models link advancements in deriving enzyme function from sequence directly to the phenotype level. GEMs are moving from predictions about growth rates for single cells towards simulations of organs, multicellular organisms and interactions between hosts and microbial communities (Thiele et al. 2018). Enzyme-constrained models have potential to provide these simulations with a solid foundation in biophysics, but more reliable kinetic parameter values are required to move ahead. There is a need for a standardized characterization of all enzymes, the kinetome.

The human genome project cost 3 billion U.S. dollars and lasted for 15 years (Collins et al. 2003). It is a successful example of a "big science" project in biology. The next logical step is a "functional" genome project, which would systematically characterize the functional properties of the proteins expressed by the genes. The results of such a characterization would form the basis for a bottom-up whole body model, and pave the way for a standard model of biology.

Conclusion

Growth is the primary objective of the cell. Growing cells require ATP and metabolites that is supplied by metabolism. Metabolic phenotypes, as observed in the fluxes over the boundary of the cell, originate from interior constraints. Mechanistic relations in the living cell can be inferred from quantitative simulations of metabolism using mathematical models. These constraints are provided directly by enzyme-constrained metabolic models. They may also be inferred from analysis of the fluxes in boundary-constrained models. In this thesis I aimed to advance our fundamental understanding of growing cells, with implications for human health. By applying different types of metabolic models I was able to: identify the growth limiting factor for breastfed infants (the energy content); estimate the catalytic advantage of fermentation over respiration (2-3.5 times more efficient); identify an intermediate metabolic mode between respiration and fermentation (complex I bypass); quantitatively predict the growth rate for yeast grown on different carbon sources; and explain the excretion of glutamate in cancer cells (compartmentalization). This thesis demonstrates the value of systematic integration of knowledge using mathematical models.

References

Adadi, R. et al., 2012. Prediction of Microbial Growth Rate versus Biomass Yield by a Metabolic Network with Kinetic Parameters. *PLoS Computational Biology*, 8, p.e1002575.

Adams, J.L. et al., 2015. Big opportunities for small molecules in immuno-oncology. *Nature Reviews Drug Discovery*, 14, p.603.

Agren, R. et al., 2014. Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. *Molecular Systems Biology*, 10.

Agren, R. et al., 2012. Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT. *PLoS Computational Biology*, 8.

Agren, R. et al., 2013. The RAVEN Toolbox and Its Use for Generating a Genome-scale Metabolic Model for Penicillium chrysogenum. *PLoS Computational Biology*, 9(3), p.e1002980.

Applegate, D.L. et al., 2007. Exact solutions to linear programming problems. *Operations Research Letters*, 35(6), pp.693–699.

Bateman, A. et al., 2017. UniProt: The universal protein knowledgebase. *Nucleic Acids Research*, 45(D1), pp.D158–D169.

Beg, Q.K. et al., 2007. Intracellular crowding defines the mode and sequence of substrate uptake by Escherichia coli and constrains its metabolic activity. *Proceedings of the National Academy of Sciences of the United States of America*, 104, pp.12663–12668.

Benyamini, T. et al., 2010. Flux balance analysis accounting for metabolite dilution. *Genome biology*, 11, p.R43.

Beutner, G. et al., 2017. The Mitochondrial Permeability Transition Pore and ATP Synthase BT - Pharmacology of Mitochondria. In H. Singh & S.-S. Sheu, eds. Cham: Springer International Publishing, pp. 21–46.

Bock, W.J., 2017. Dual Causality and the Autonomy of Biology. *Acta Biotheoretica*, 65(1), pp.63–79.

Boer, V.M. et al., 2010. Growth-limiting intracellular metabolites in yeast growing under diverse nutrient limitations. *Molecular biology of the cell*, 21, pp.198–211.

Bordbar, A. et al., 2011. A multi-tissue type genome-scale metabolic network for analysis of whole-body systems physiology. *BMC Systems Biology*, 5, p.180.

Bordbar, A. & Palsson, B.O., 2012. Using the reconstructed genome-scale human metabolic network to study physiology and pathology. In *Journal of Internal Medicine*. pp. 131–141.

Bray, F. et al., 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 0(0).

Brigandt, I., 2013. Systems biology and the integration of mechanistic explanation and mathematical explanation. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 44(4, Part A), pp.477–492.

Brooks, G.A., 2009. Cell-cell and intracellular lactate shuttles. *Journal of Physiology*, 587(23), pp.5591–5600.

Bruggeman, F.J. & Westerhoff, H. V, 2007. The nature of systems biology. *Trends in Microbiology*, 15(1), pp.45–50.

Brunk, E. et al., 2018. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nature Biotechnology*, 36, p.272.

Butte, N.F., 2005. Energy requirements of infants. *Public health nutrition*, 8, pp.953–967.

Butte, N.F., Lopez-Alarcon, M.G. & Garza, C., 2002. *Nutrient adequacy of exclusive breast feeding for the term infant during the first six months of life*, Geneva : World Health Organization.

Byrne, N.M. et al., 2017. Intermittent energy restriction improves weight loss efficiency in obese men: the MATADOR study. *International Journal Of Obesity*, 42, p.129.

Cairns, R., Harris, I. & Mak, T., 2011. Regulation of cancer cell metabolism. *Nature Reviews Cancer*, 11(2), pp.85–95.

Caruso, M.B. et al., 2017. Proteomic analysis of the secretome of HepG2 cells indicates differential proteolytic processing after infection with dengue virus. *Journal of Proteomics*, 151, pp.106–113.

Castillo-Duran, C. & Uauy, R., 1988. Copper deficiency impairs growth of infants recovering from malnutrition. *The American Journal of Clinical Nutrition*, 47(4), pp.710–714.

Chindelevitch, L. et al., 2014. An exact arithmetic toolbox for a consistent and reproducible structural analysis of metabolic network models. *Nature communications*, 5, p.4893.

Collins, F.S., Morgan, M. & Patrinos, A., 2003. The Human Genome Project: Lessons from Large-Scale Biology. *Science*, 300(5617), p.286 LP-290.

Dai, Z. et al., 2018. Global rewiring of cellular metabolism renders Saccharomyces cerevisiae Crabtree negative.

Nature communications, 9(1), p.3059.

Dashko, S. et al., 2014. Why, when, and how did yeast evolve alcoholic fermentation? *FEMS Yeast Research*, 14(6), pp.826–832.

De Deken, R.H., 1966. The Crabtree effect: a regulatory system in yeast. *Journal of general microbiology*, 44, pp.149–156.

van Dijken, J.P. & Scheffers, W.A., 1986. Redox balances in the metabolism of sugars by yeasts. *FEMS Microbiology Letters*, 32, pp.199–224.

Dikicioglu, D., Kirdar, B. & Oliver, S.G., 2015. Biomass composition: the elephant in the room of metabolic modelling. *Metabolomics*, 11(6), pp.1690–1701.

Dobbing, J. & Sands, J., 1973. Quantitative growth and development of human brain. *Archives of disease in childhood*, 48, pp.757–767.

Durot, M., Bourguignon, P.-Y. & Schachter, V., 2009. Genome-scale models of bacterial metabolism: reconstruction and applications. *FEMS Microbiology Reviews*, 33(1), pp.164–190.

Ellis, J., 1986. The superstring: theory of everything, or of nothing? *Nature*, 323, p.595.

Elmer, S.J. et al., 2011. Joint-Specific Power Production during Submaximal and Maximal Cycling. *Medicine & Science in Sports & Exercise*, 43(10).

Erdrich, P., Steuer, R. & Klamt, S., 2015. An algorithm for the reduction of genome-scale metabolic network models to meaningful core models. *BMC systems biology*, 9(1), p.48.

Fan, J. et al., 2013. Fatty acid labeling from glutamine in hypoxia can be explained by isotope exchange without net reductive isocitrate dehydrogenase (IDH) flux. *Journal of Biological Chemistry*, 288(43), pp.31363–31369.

Fan, J. et al., 2014. Quantitative flux analysis reveals folate-dependent NADPH production. *Nature*, 510, pp.298–302.

Feizi, A. et al., 2013. Genome-Scale Modeling of the Protein Secretory Machinery in Yeast. *PLOS ONE*, 8(5), p.e63284.

Fields, D.A. et al., 2011. Longitudinal body composition data in exclusively breast-fed infants: a multicenter study. *Obesity (Silver Spring, Md.)*, 19, pp.1887–1891.

Fjeld, C.R., Schoeller, D.A. & Brown, K.H., 1989. A new model for predicting energy requirements of children during catch-up growth developed using doubly labeled water. *Pediatric research*, 25, pp.503–508.

Förster, J. et al., 2003. Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. *Genome research*, 13, pp.244–253.

de Godoy, L.M.F. et al., 2008. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*, 455(7217), pp.1251–1254.

Goelzer, A., Fromion, V. & Scorletti, G., 2011. Cell design in bacteria as a convex optimization problem. *Automatica*, 47, pp.1210–1218.

Grassi, B. et al., 1999. Blood lactate accumulation and muscle deoxygenation during incremental exercise. *Journal of Applied Physiology*, 87(1), pp.348–355.

Gutenkunst, R.N. et al., 2007. Universally Sloppy Parameter Sensitivities in Systems Biology Models. *PLOS Computational Biology*, 3(10), p.e189.

Hanahan, D. & Weinberg, R.A., 2011. Hallmarks of cancer The next generation. *Cell*, 144, pp.646–674.

Hester, S.N. et al., 2012. Is the Macronutrient Intake of Formula-Fed Infants Greater Than Breast-Fed Infants in Early Infancy? *Journal of Nutrition and Metabolism*, 2012, p.891201.

Hill, G.M. et al., 1983. A Copper Deficiency in Neonatal Pigs Induced by a High Zinc Maternal Diet. *The Journal of Nutrition* , 113(4), pp.867–872.

van Hoek, M.J. & Merks, R.M., 2012. Redox balance is key to explaining full vs. partial switching to low-yield metabolism. *BMC Systems Biology*, 6, p.22.

van Hoek, P., van Dijken, J.P. & Pronk, J.T., 1998. Effect of specific growth rate on fermentative capacity of baker's yeast. *Appl.Environ.Microbiol.*, 64, pp.4226–4233.

Hoek, P.I.M.V. a N., Dijken, J.P.V. a N. & Pronk, J.T., 1998. Effect of Specific Growth Rate on Fermentative Capacity of Baker's Yeast. *Society*, 64, pp.4226–4233.

Hohmann, S., 2016. Nobel Yeast Research. *FEMS Yeast Research*, 16(8), pp.fow094-fow094.

Hui, S. et al., 2015. Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Molecular Systems Biology*, 11(2), p.784.

Hyduke, D.R. & Palsson, B.O., 2010. Towards genome-scale signalling-network reconstructions. *Nature Reviews Genetics*.

Jain, M. et al., 2012. Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. *Science*

(New York, N.Y.), 336(6084), pp.1040–4.

Johansson, M., Zhang, J. & Ehrenberg, M., 2012. Genetic code translation displays a linear trade-off between efficiency and accuracy of tRNA selection. *Proceedings of the National Academy of Sciences*, 109(1), p.131 LP-136.

Jordan, P.N. & Hall, K.D., 2008. Dynamic coordination of macronutrient balance during infant growth: insights from a mathematical model. *The American journal of clinical nutrition*, 87(3), pp.692–703.

Kaplan, R.M. & Irvin, V.L., 2015. Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased over Time. *PLOS ONE*, 10(8), p.e0132382.

Karr, J.R. et al., 2012. A Whole-Cell Computational Model Predicts Phenotype from Genotype. *Cell*, 150, pp.389–401.

Kelk, S.M. et al., 2012. Optimal flux spaces of genome-scale stoichiometric models are determined by a few subnetworks. *Scientific Reports*, 2, p.580.

Kitano, H., 2002. Computational systems biology. *Nature*, 420, p.206.

Klumpp, S. et al., 2013. Molecular crowding limits translation and cell growth. *Proceedings of the National Academy of Sciences of the United States of America*, 110, pp.16754–9.

Kuzawa, C.W. et al., 2014. Metabolic costs and evolutionary implications of human brain development. *Proceedings of the National Academy of Sciences*, 111(36), pp.13010–13015.

Labhsetwar, P. et al., 2017. Population FBA predicts metabolic phenotypes in yeast. *PLOS Computational Biology*, 13(9), p.e1005728.

Larsen, S. et al., 2012. Biomarkers of mitochondrial content in skeletal muscle of healthy young human subjects. *Journal of Physiology*, 590(14), pp.3349–3360.

Larsson, C. et al., 1993. Growth and metabolism of Saccharomyces cerevisiae in chemostat cultures under carbon-, nitrogen-, or carbon- and nitrogen-limiting conditions. *Journal of Bacteriology*, 175, pp.4809–4816.

Letchford, A., Moat, H.S. & Preis, T., 2015. The advantage of short paper titles. *Royal Society Open Science*, 2(8).

Lewis, N.E. et al., 2010. Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models. *Molecular systems biology*, 6, p.390.

Liu, J.K. et al., 2014. Reconstruction and modeling protein translocation and compartmentalization in Escherichia coli at the genome-scale. *BMC systems biology*, 8(1), p.110.

Mardinoglu, A. et al., 2014. Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. *Nature communications*, 5(May 2013), p.3083.

Mardinoglu, A. et al., 2013. Integration of clinical data with a genome-scale metabolic model of the human adipocyte. *Molecular systems biology*, 9, p.649.

Martínez, V.S. et al., 2012. Flux balance analysis of CHO cells before and after a metabolic switch from lactate production to consumption. *Biotechnology and Bioengineering*, p.n/a-n/a.

Molenaar, D. et al., 2009. Shifts in growth strategies reflect tradeoffs in cellular economics. *2*, 5, p.323.

Mori, M. et al., 2016. Constrained Allocation Flux Balance Analysis. *PLoS Computational Biology*.

Morris, M.K. et al., 2016. Systematic Analysis of Quantitative Logic Model Ensembles Predicts Drug Combination Effects on Cell Signaling Networks. *CPT: Pharmacometrics & Systems Pharmacology*, 5(10), pp.544–553.

Müller, S., Regensburger, G. & Steuer, R., 2015. Resource allocation in metabolic networks: kinetic optimization and approximations by FBA. *Biochemical Society Transactions*, 43(6), pp.1195–1200.

Munos, B., 2009. Lessons from 60 years of pharmaceutical innovation. *Nature Reviews Drug Discovery*, 8, p.959.

Murphy, H., Jaafari, H. & Dobrovolny, H.M., 2016. Differences in predictions of ODE models of tumor growth: a cautionary example. *BMC cancer*, 16, p.163.

Nielsen, J., 2015. Yeast cell factories on the horizon. *Science*, 349(6252), p.1050 LP-1051.

Noor, E. et al., 2016. The Protein Cost of Metabolic Fluxes: Prediction from Enzymatic Rate Laws and Cost Minimization. *PLOS Computational Biology*, 12(11), p.e1005167.

O'Brien, E.J., Monk, J.M. & Palsson, B.O., 2015. Using Genome-scale Models to Predict Biological Capabilities. *Cell*, 161(5), pp.971–987.

O'Brien, E.J. & Palsson, B.O., 2015. Computing the functional proteome: Recent progress and future prospects for genome-scale models. *Current Opinion in Biotechnology*, 34, pp.125–134.

Orth, J.D., Thiele, I. & Palsson, B.Ø.O., 2010. What is flux balance analysis? *Nature biotechnology*, 28(3), pp.245–248.

Palm, W. & Thompson, C.B., 2017. Nutrient acquisition strategies of mammalian cells. *Nature*, 546(7657), pp.234–242.

Palsson, B., 2015. *Systems biology : constraint-based reconstruction and analysis*, Cambridge, United Kingdom:

Cambridge University Press.

Park, J.O. et al., 2016. Metabolite concentrations, fluxes and free energies imply efficient enzyme usage. *Nature Chemical Biology*, 12(7), pp.482–489.

Peng, Y.M. et al., 2007. Fatty acid composition in breast milk and serum phospholipids of healthy term Chinese infants during first 6 weeks of life. *Acta Pædiatrica*, 96(11), pp.1640–1645.

Peronnet, F. & Thibault, G., 1989. Mathematical analysis of running performance and world running records. *Journal of Applied Physiology*, 67(1), pp.453–465.

Piedrafita, G., Keller, A.M. & Ralser, M., 2015. The Impact of Non-Enzymatic Reactions and Enzyme Promiscuity on Cellular Metabolism during (Oxidative) Stress Conditions. *Biomolecules* , 5(3).

Poole, D.C. & Richardson, R.S., 1997. Determinants of oxygen uptake: Implications for exercise testing. *Sports Medicine*, 24(5), pp.308–320.

Postma, E. et al., 1989. Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of Saccharomyces cerevisiae. *Applied and environmental microbiology*, 55, pp.468–477.

Postma, E., Alexander Scheffers, W. & van Dijken, J.P., 1989. Kinetics of growth and glucose transport in glucose-limited chemostat cultures of Saccharomyces cerevisiae CBS 8066. *Yeast*, 5(3), pp.159–165.

Prendergast, A.J. & Humphrey, J.H., 2014. The stunting syndrome in developing countries. *Paediatrics and International Child Health*, 34(4), pp.250–265.

Quek, L.-E. et al., 2014. Reducing Recon 2 for steady-state flux analysis of HEK cell culture. *Journal of Biotechnology*, 184, pp.172–178.

Raffel, S. et al., 2017. BCAT1 restricts αkG levels in AML stem cells leading to IDHmut-like DNA hypermethylation. *Nature*, 551(7680), pp.384–388.

Raman, K. & Chandra, N., 2009. Flux balance analysis of biological systems: Applications and challenges. *Briefings in Bioinformatics*, 10(4), pp.435–449.

Sanchez, B.J. et al., 2018. SLIMEr: probing flexibility of lipid metabolism in yeast with an improved constraint-based modeling framework. *bioRxiv*.

Sánchez, B.J. et al., 2017. Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Molecular Systems Biology*, 13(8).

Sánchez, B.J., Pérez-Correa, J.R. & Agosin, E., 2014. Construction of robust dynamic genome-scale metabolic model structures of Saccharomyces cerevisiae through iterative re-parameterization. *Metabolic Engineering*, 25, pp.159–173.

Savir, Y. et al., 2010. Cross-species analysis traces adaptation of Rubisco toward optimality in a low-dimensional landscape. *Proceedings of the National Academy of Sciences of the United States of America*, 107(8), pp.3475–3480.

Schuurmans, J.M. et al., 2008. Effect of hxk2 deletion and HAP4 overexpression on fermentative capacity in Saccharomyces cerevisiae. *FEMS Yeast Research*, 8(2), pp.195–203.

Segrè, D., Vitkup, D. & Church, G.M., 2002. Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(23), pp.15112–7.

Seidlitz, E.P. et al., 2009. Cancer cell lines release glutamate into the extracellular environment. *Clinical & Experimental Metastasis*, 26(7), p.781.

Shaw, R. & Cheung, C.Y.M., 2018. A Dynamic Multi-Tissue Flux Balance Model Captures Carbon and Nitrogen Metabolism and Optimal Resource Partitioning During Arabidopsis Growth. *Frontiers in plant science*, 9, p.884.

Sheikh, K., Förster, J. & Nielsen, L.K., 2005. Modeling hybridoma cell metabolism using a generic genome-scale metabolic model of Mus musculus. *Biotechnology Progress*, 21(1), pp.112–121.

Shlomi, T. et al., 2011. Genome-Scale Metabolic Modeling Elucidates the Role of Proliferative Adaptation in Causing the Warburg Effect. *PLoS Comput Biol*, 7(3), p.e1002018.

Stelling, J. et al., 2002. Metabolic network structure determines key aspects of functionality and regulation. *Nature*, 420, p.190.

Stepulak, A. et al., 2014. Glutamate and its receptors in cancer. *Journal of neural transmission (Vienna, Austria : 1996)*, 121(8), pp.933–44.

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

Steuer, R. et al., 2006. *Structural kinetic modeling of metabolic networks.*,

Swainston, N. et al., 2016. Recon 2.2: from reconstruction to model of human metabolism. *Metabolomics*, 12(7),

pp.1–7.

Szenk, M., Dill, K.A. & de Graff, A.M.R., 2017. Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the Membrane Real Estate Hypothesis. *Cell Systems*, 5(2), pp.95–104.

Thiele, I. et al., 2018. When metabolism meets physiology: Harvey and Harvetta. *bioRxiv*, p.255885.

Tran, L.M., Rizk, M.L. & Liao, J.C., 2008. Ensemble Modeling of Metabolic Networks. *Biophysical Journal*, 95(12), pp.5606–5617.

Väremo, L. et al., 2015. Proteome- and Transcriptome-Driven Reconstruction of the Human Myocyte Metabolic Network and Its Use for Identification of Markers for Diabetes. *Cell Reports*, 11(6), pp.921–933.

Vazquez, A. et al., 2010. Catabolic efficiency of aerobic glycolysis: the Warburg effect revisited. *BMC systems biology*, 4, p.58.

Vazquez, A. & Oltvai, Z.N., 2011. Molecular crowding defines a common origin for the warburg effect in proliferating cells and the lactate threshold in muscle physiology. *PLoS ONE*, 6(4), p.e19538.

Vos, T. et al., 2016. Maintenance-energy requirements and robustness of Saccharomyces cerevisiae at aerobic near-zero specific growth rates. *Microbial Cell Factories*, 15(1), p.111.

Vulin, C. et al., 2014. Growing yeast into cylindrical colonies. *Biophysical Journal*, 106(10), pp.2214–2221.

Wales, J.K. & Gibson, A.T., 1994. Short-term growth: rhythms, chaos, or noise? *Archives of disease in childhood*, 71(1), pp.84–89.

Wang, Z., 2012. High ratio of resting energy expenditure to body mass in childhood and adolescence: A mechanistic model. *American Journal of Human Biology*, 24, pp.460–467.

Ward, P.S. & Thompson, C.B., 2012. Metabolic Reprogramming: A Cancer Hallmark Even Warburg Did Not Anticipate. *Cancer Cell*, 21(3), pp.297–308.

Wells, J.C. & Davies, P.S., 1998. Estimation of the energy cost of physical activity in infancy. *Archives of disease in childhood*, 78(2), pp.131–6.

Winter, G. & Krömer, J.O., 2012. Fluxomics – connecting 'omics analysis and phenotypes. *Environmental Microbiology*, 15(7), pp.1901–1916.

Wise, D.R. & Thompson, C.B., 2010. Glutamine addiction: a new therapeutic target in cancer. *Trends in Biochemical Sciences*, 35(8), pp.427–433.

World Health Organization, 2009. *WHO Child Growth Standards: Growth Velocity Based on Weight, Length and Head Circumference*, World Health Organization.

Xie, L. & Wang, D.I.C., 1996. Energy metabolism and ATP balance in animal cell cultivation using a stoichiometrically based reaction network. *Biotechnology and Bioengineering*, 52(5), pp.591–601.

Xu, W. et al., 2011. Oncometabolite 2-Hydroxyglutarate Is a Competitive Inhibitor of α-Ketoglutarate-Dependent Dioxygenases. *Cancer Cell*, 19(1), pp.17–30.

Yamawaki, N. et al., 2005. Macronutrient, mineral and trace element composition of breast milk from Japanese women. *Journal of Trace Elements in Medicine and Biology*, 19, pp.171–181.

Yang, W. et al., 2012. Anemia, malnutrition and their correlations with socio-demographic characteristics and feeding practices among infants aged 0-18 months in rural areas of Shaanxi province in northwestern China: a cross-sectional study. *BMC Public Health*, 12, p.1127.

Yizhak, K. et al., 2015. Modeling cancer metabolism on a genome scale. *Molecular Systems Biology*, 11(6), pp.817–817.

You, C. et al., 2013. Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature*, 500, p.301.

Yu, T. et al., 2018. Reprogramming Yeast Metabolism from Alcoholic Fermentation to Lipogenesis. *Cell*, 174(6), p.1549–1558.e14.

Zhuang, K., Vemuri, G.N. & Mahadevan, R., 2011. Economics of membrane occupancy and respiro-fermentation. *Molecular systems biology*, 7, p.500.

Zielinski, D.C. et al., 2017. Systems biology analysis of drivers underlying hallmarks of cancer cell metabolism. *Scientific Reports*, 7, p.41241.